Stockholms matematiska cirkel

Elliptiska Kurvor

Lukas Gustafsson
Ludvig Olsson

Institutionen för matematik, KTH och
Matematiska institutionen, Stockholms universitet
2023–2024

Stockholms matematiska cirkel genom tiderna
(tidigare KTH:s matematiska cirkel)

2023-2024	Elliptiska Kurvor
2022-2023	Variationskalkyl
2021-2022	Matematik och AI
2020-2021	Musik och matematik
2019-2020	Datorernas matematik
2018-2019	Grafteori med inriktning på färgläggning
2017-2018	Geometriska konstruktioner
2016-2017	Vad är ett tal?
2015-2016	Fraktaler
2014-2015	Polytoper
2013-2014	Grupper, mönster och symmetrier
2012-2013	Den matematiska analysens grunder
2011-2012	Diofantiska ekvationer
2010-2011	Polynom
2009-2010	Hyperbolisk geometri
2008-2009	Talteori
2007-2008	Sannolikhetsteori
2006-2007	Gruppteori
2005-2006	Vad är ett tal?
2004-2005	Integraler
2003-2004	Linjär algebra och bioinformatik
2002-2003	Algebra och kryptografi
2001-2002	Analysens grunder
2000-2001	Talföljder, rekursioner och iterationer
1999-2000	Linjära avbildningar

Innehåll

Lista över grekiska alfabetet

A	α	alfa
B	β	beta
Γ	γ	gamma
Δ	δ	delta
E	ε	epsilon
Z	ζ	zeta
H	η	eta
Θ	θ	theta
I	ι	iota
K	κ	kappa
Λ	λ	lambda
M	μ	my
N	ν	ny
Ξ	ξ	xi
O	o	omikron
Π	π	pi
P	ρ	rho
Σ	σ	sigma
T	τ	tau
Υ	υ	ypsilon
Φ	ϕ	fi
X	χ	chi
Ψ	ψ	psi
Ω	ω	omega

Lista över mängdsymboler

$x \in X$	x är ett element i mängden X.
$X \subset Y$	X är en delmängd till Y.
$\forall x \in X$	för varje x som är ett element i mängden X.
$\exists x \in X \ldots$	existerar ett element x i mängden X sådant att $\ldots$
$f \colon X \to Y$	f är en funktionen med definitionsmängd X och målmängd Y.

Några ord på vägen

Detta kompendium är skrivet för att användas som kurslitteratur till Stockholms matematiska cirkel under läsåret 2023–2024 och består av åtta kapitel.

Kompendiet är inte tänkt att läsas enbart på egen hand, utan ska ses som ett skriftligt komplement till undervisningen. Alla elever rekommenderas att läsa igenom varje kapitel själv innan föreläsningen. Det är inte nödvändigt att förstå alla detaljer vid den första genomläsningen.

Som de flesta matematiska skrifter på högre nivå är kompendiet kompakt skrivet. Detta innebär att man i allmänhet inte kan läsa det som en vanlig bok. Istället bör man pröva nya satser och definitioner genom att på egen hand exemplifiera. Därmed uppnår man oftast en mycket bättre förståelse av vad dessa satser och deras bevis går ut på.

Till varje kapitel finns ett antal övningsuppgifter. De udda övningarna har lösningar längst bak i kompendiet. Syftet med dessa är att eleverna ska kunna lösa dem och på egen hand kontrollera att de förstått materialet. Övningar med jämna nummer saknar facit och kan användas som examination. Det rekommenderas dock att man försöker lösa dessa uppgifter även om man inte examineras på dem.

Om man kör fast kan man alltid fråga en kompis, en lärare på sin skola eller någon av författarna. Under årets gång kommer det att finnas övningstillfällen där eleverna kan jobba med uppgifterna, själva eller i grupp, och få hjälp av oss.

De övningsuppgifter som är något svårare markeras med en stjärna ($\star$). Uppgifter som är extra utmanande markeras med två stjärnor ($\star\star$).

Vissa övningar kan ha flera lösningar och det som står i facit bör i detta fall endast ses som ett förslag.

Målet med årets kurs är att introducera konceptet *elliptiska kurvor* och att ge en introduktion till de verktyg som används för att göra algebra inom matematik. Innan vi gör det så studerar och repeterar vi de grundläggande principerna i matematik i kapitel 1. Sedan introducerar vi grupper i kapitel 2 som är en viktig typ av matematisk struktur. Kapitel 3 behandlar ett viktigt exempel av en grupp, nämligen $\mathbb{Z}/(n)$, och introducerar andra strukturer som ringar och kroppar. I kapitel 4 lär vi oss mer om ringar och introducerar polynom över en godtycklig kropp. Kapitel 5 är första gången vi introducerar konceptet elliptiska kurvor, som kommer att utgöra huvudtemat för kursen. I kapitel 6 diskuterar vi hur elliptiska kurvor används för att göra kryptografi, och vad som menas med en elliptisk kurva över en ändlig kropp. I kapitel 7 bevisar vi en mycket central sats gällande elliptiska kurvor över de rationella talen. Slutligen i kapitel 8 går vi igenom elliptiska kurvor över de komplexa talen.

Några ord om cirkeln

Stockholms matematiska cirkel är en kurs för matematikintresserade gymnasieelever, som arrangeras av Kungliga Tekniska högskolan och Stockholms universitet. Cirkeln startade 1999. Vid starten hade den namnet KTH:s matematiska cirkel och hölls i KTH:s ensamma regi. Ambitionen med cirkeln är att sprida kunskap om matematiken och dess användningsområden utöver vad eleverna får genom gymnasiekurser, och att etablera ett närmare samarbete mellan gymnasieskolan och högskolan. Cirkeln ska särskilt stimulera elevernas matematikintresse och inspirera dem till fortsatta naturvetenskapliga och matematiska studier.

Till varje kurs skrivs ett kompendium som distribueras gratis till eleverna. Detta material, föreläsningsschema och övrig information om Stockholms matematiska cirkel finns tillgängligt på

https://www.math-stockholm.se/cirkel

Cirkeln godkänns ofta som en gymnasiekurs eller som matematisk breddning på gymnasieskolorna. Det är upp till varje skola att godkänna cirkeln som en kurs och det är lärarna från varje skola som sätter betyg på kursen. Lärarna är självklart också välkomna till cirkeln och många har kommit överens med sin egen skola om att få cirkeln godkänd som fortbildning eller som undervisning.

Vi vill gärna understryka att föreläsningarna är öppna för alla gymnasieelever, lärare eller andra matematikintresserade.

Vi har avsiktligt valt materialet för att ge eleverna en inblick i matematisk teori och tankesätt och presenterar därför både några huvudsatser inom varje område och bevisen för dessa resultat. Vi har också som målsättning att bevisa alla satser som används om de inte kan förutsättas bekanta av elever från gymnasiet.

Författarna, sommaren 2023

1 Matematik

Temat för årets matematiska cirkel är Elliptiska kurvor. Detta kapitel är en introduktion i den matematiska metoden och ett antal grundbegrepp som vi kommer använda oss av i kursen. Under föreläsningen för detta kapitel kommer vi att fokusera på avsnitten 1.5–1.6 medan läsaren uppmanas att själv läsa genom de första avsnitten mer noggrant.

1.1 Definition, axiom, sats och bevis

I detta avsnitt ska vi beskriva den matematiska metoden utifrån fyra begrepp: *definition*, *sats*, *bevis* och *axiom*.

En *definition* bestämmer vad en term betyder så att man kan arbeta matematiskt med den. Till exempel kan vi definiera udda och jämna tal på följande sätt.

Definition 1.1.1. Ett heltal n är *udda* om det finns ett heltal k som uppfyller att $n = 2k + 1$. △

Definition 1.1.2. Ett heltal n är *jämnt* om det finns ett heltal k som uppfyller att $n = 2k$. △

Ofta har man en intuition om vad en term betyder redan innan man definierar den. Läsaren hade till exempel säkert en uppfattning om vad udda och jämna tal är innan vi definierade dem. Syftet med en definition är att precisera detta.

När definitionen är gjord, så överger man sina tidigare uppfattningar om vad termen betyder och utgår endast ifrån definitionen. Man säger att definitionen är *stipulativ*. En definition är alltså inte rätt eller fel, utan bara mer eller mindre användbar och intuitiv.

Definitioner bygger ofta på begrepp som läsaren är bekant med. Till exempel utgår Definition 1.1.1 och 1.1.2 från att läsaren redan vet vad ett heltal är.

En *sats* är ett påstående som bevisats vara sant. Varje sats hör samman med ett *bevis*: ett argument för att påståendet är sant.

Sats 1.1.3. *Om n är udda, så är $n + 1$ jämnt.*

Bevis. Om n är udda så finns det ett heltal k så att $n = 2k + 1$. Då gäller att

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

Eftersom $k + 1$ är ett heltal, så är $n + 1$ ett jämnt tal. □

Bevisen kombinerar definitioner och olika logiska slutledningsregler för att nå den önskade slutsatsen. Sats 1.1.3 har en syskonsats. Beviset är mer eller mindre identiskt, och lämnas som övning.

Sats 1.1.4. *Om n är jämnt, så är $n + 1$ udda.*

En sats vars främsta syfte är att användas i beviset av en annan sats kallas för en *hjälpsats* eller ett *lemma*. En sats som följer omedelbart ur en annan sats, till exempel som ett specialfall, kallas för en *följdsats* eller ett *korollarium*.

Ett påstående måste vara bevisat för att få kallas för en sats. Om man har goda skäl att tro att ett påstående är sant men inte formellt bevisat det kallas påståendet för en *förmodan*, eller *hypotes*. Två exempel är *Riemannhypotesen* och *primtalstvillingsförmodan*.

En förmodan kan förbli obevisad i hundratals år. Ett berömt exempel är *Fermats stora sats*, som formulerades av Pierre de Fermat (1607–1665) år 1637 men bevisades först av Andrew Wiles år 1995. Riemannhypotesen, som ännu är obevisad, formulerades 1859 av Bernard Riemann (1826–1866).

Eftersom bevisen utgår ifrån definitionen, och inte vår intuition, så behöver man ibland bevisa saker som känns uppenbara. Läsaren vet till exempel att

(i) alla tal är antingen udda eller jämna, och

(ii) ett tal kan inte vara udda och jämnt samtidigt.

Men om man läser Definition 1.1.1 och 1.1.2 så ingår inte dessa påståenden. Kan man inte tänka sig tal som varken är udda eller jämnt? Eller tal som är både och?

Bevis utgår ifrån antaganden och tidigare kända satser. Dessa tidigare satser måste också bevisas innan de kan anses giltiga. Men dessa bevis måste också bygga på antaganden och satser, som också måste bevisas, och så vidare.

För att undvika en oändlig kedja av bevis, eller ett cirkulärt bevis (ett bevis som använder sig av det man försöker bevisa) så måste man göra grundantaganden som inte behöver bevisa. Dessa kallas för *axiom*. Exempel på axiom är att mängden av heltal existerar och att addition uppfyller

- (Associativitet) För alla heltal n, m, p gäller $(n + m) + p = n + (m + p)$.
- (Identitet) Det finns ett element 0, som vi kallar *nollan*, sådan att för alla heltal n så gäller $0 + n = n + 0 = n$.
- (Inverser) För varje heltal n existerar ett heltal $(-n)$, som vi kallar för *minus* n, sådan att $0 = n + (-n) = (-n) + n$.
- (Kommutativitet) För alla heltal n, m så gäller $n + m = m + n$.

1.2 Bevistekniker

Ett bevis för en sats är ett argument som förklarar varför satsen är sann. Vi har redan sett ett exempel när vi bevisade Sats 1.1.3. I detta avsnitt ska vi gå igenom tre tekniker för att bevisa matematiska satser: direkta bevis, motsägelsebevis och induktionsbevis.[1]

[1] Ibland förekommer termen *indirekt bevis*. Vissa använder det som synonym till motsägelsebevis, andra som en synonym till bevisregeln *modus tollens*. Vi undviker den helt.

Ett *direkt bevis* utgår ifrån satsens antaganden och definitioner och bevisar satsen rakt på, så att säga. Beviset av Sats 1.1.3 är ett exempel på direkt bevis. Ett annat är följande sats.

Sats 1.2.1. *Antalet funktioner från en mängd A med n element till en mängd B med m element är m^n.*

Bevis. Varje funktion från A till B kan beskrivas som en tabell där varje element i A motsvaras av precis ett element i B. Listan innehåller totalt n platser, och på varje plats kan vi välja bland m element att välja bland. Alltså finns det totalt

$$\underbrace{m \cdot m \cdots m \cdot m}_{n \text{ stycken}} = m^n$$

olika funktioner. □

Ibland går det inte att använda direkta bevis, till exempel när man ska bevisa att något inte är fallet. Då kan det vara enklare att anta att det man vill bevisa är falskt, och visa att detta leder till en motsägelse. Om alla steg i beviset är korrekta så måste det ursprungliga antagandet vara fel. Detta kallas för ett *motsägelsebevis*.

Sats 1.2.2. *Ett tal kan inte vara udda och jämnt samtidigt.*

Bevis. Antag att n är ett tal som är båda udda och jämnt. Då finns det två heltal, k och l, så att $n = 2k$ och $n = 2l + 1$. Då gäller att

$$2k = n = 2l + 1 \implies 2k - 2l = 1 \implies 2(k - l) = 1.$$

Med andra ord finns det heltal $m = k - l$ så att $2m = 1$. Kan det finns ett sådant tal? Det finns två fall.

(i) Om $m \leqslant 0$, så är $1 = 2m \leqslant 0$. Motsägelse!

(ii) Om $m \geqslant 1$ så är $2m \geqslant 2 > 1$. Motsägelse! □

Ett berömt motsägelsebevis är följande.

Sats 1.2.3. *Talet $\sqrt{2}$ är irrationellt.*

Bevis. Antag motsatsen, det vill säga att $\sqrt{2} = a/b$ för några heltal a och b. Antag att a och b är förkortade så långt som möjligt. Då kan endast en av a eller b vara jämn, eftersom om båda är jämna kan vi skriva

$$\sqrt{2} = \frac{a}{b} = \frac{2c}{2d} = \frac{c}{d}$$

och då var inte a och b förkortade så långt som möjligt.

Av definitionen av $\sqrt{2}$ får vi att

$$\sqrt{2}^2 = 2 = \frac{a^2}{b^2} \implies 2b^2 = a^2.$$

Den sista ekvationen säger att a^2 är jämn. Eftersom kvadrater av udda tal är udda (se Övning 1.25), så måste a vara ett jämnt tal, det vill säga $a = 2k$ för något heltal k. Då får vi att

$$2b^2 = (2k)^2 = 4k^2 \implies b^2 = 2k^2.$$

Eftersom b^2 är jämnt, så måste b vara jämnt. Men nu har vi bevisat att både a och b är jämna, vilket var omöjligt eftersom vi hade förkortat bråket så långt som möjligt. Detta är en motsägelse. □

Att bevis inkluderar antaganden och små hjälpsatser som inte nämns är snarare regel än undantag, till exempel lägger vi nästan aldrig tid på att visa $2 > 0$ i mitten av ett bevis. Ifall man bevisade precis vartenda antagande och påstående utifrån axiomen varje gång skulle bevisen bli väldigt långa och komplicerade. Läsaren förväntas själv fylla i de luckor som uppstår.

Det händer dock att uppenbara antaganden är mycket svåra, till och med omöjliga, att bevisa utifrån definitionerna. Historien är fylld av matematiker som gjort till synes självklara antaganden som sedan visat sig vara svåra att bevisa.

Beviset av Sats 1.2.3 är ett exempel på det. Vi antar att ett bråk kan förkortas så långt som möjligt. Detta är inte självklart, utan bygger i själva verket på aritmetikens fundamentalsats som vi inte kommer att behandla i kursen.

Den tredje bevistekniken som finns kallas för *induktionsbevis* men den kommer vi inte behöva i den här kursen.

1.3 Olika satser och hur man bevisar dem

I föregående avsnitt diskuterade vi olika bevistekniker. Men vilka tekniker är lämpliga för vilka typer av satser?

- **Implikation:** Man säger att P implicerar Q om Q är sant när P är det. Ett exempel är Sats 1.1.4, som säger att om ett tal n är jämnt, så är talet $n + 1$ udda. Man brukar beteckna implikationer med en tjock pil $\implies$, så att P medför Q skrivs

  $$P \implies Q.$$

  En implikation kan bevisas med ett direkt bevis. Då antar man att P är sant, och sedan visar man att Q också måste vara sant (det är så vi bevisar Sats 1.1.4). Man kan också använda ett motsägelsebevis. Då antar man att P är sann och att Q är falsk, och bevisar en motsägelse.

  Ett tredje sätt att bevisa att P implicerar Q är att bevisa att om Q är falsk, så är P falsk. Detta kallas för *omvändningen* av en implikation.

- **Ekvivalens:** En ekvivalens är när två påståenden P och Q implicerar varandra, alltså att om P så Q, och om Q så P. Man brukar använda

frasen P om och endast om Q. Man använder tjocka dubbelpilar för att beteckna ekvivalenser, så att P om och endast om Q skrivs som

$$P \iff Q.$$

Ekvivalenser bevisas genom att första visa att P implicerar Q, och sedan att Q implicerar P.

- **Universalsats:** En universalsats säger att alla n i en mängd M uppfyller något villkor P. Universalsatser kan bevisas som implikationer, genom att omformulera universalsatsen som att om n ligger i mängden M, så uppfyller n villkoret P, det vill säga

  $$n \in M \implies n \text{ uppfyller } P$$

  Man kan även bevisa en universalsats genom ett motsägelsebevis. Då antar man att det finns ett n i M som inte uppfyller P, och bevisar att det är omöjligt.

- **Existenssats:** En existenssats säger att finns ett objekt n som har egenskapen P. Den typiska existenssatsen är ekvationslösning. Att $x^2 = 3$ har en lösning är en existenssats, och kan omformuleras som att det finns ett tal x så att $x^2 = 3$.

  Ett sätt att bevisa en existenssats är att konstruera det sökta objektet utifrån objekt man redan vet finns. Till exempel så kan man bevisa att det finns ett udda kvadrattal genom att notera att $3^2 = 9$ är udda och ett kvadrattal.

  Man kan också använda ett motsägelsebevis. Då antar man att det inte existerar någon objekt med egenskapen P och visar att det leder till en motsägelse. Dessa bevis har fördelen att vi inte behöver beskriva hur objektet konstrueras. I gengäld kan bevisen vara mycket komplicerade.

En variant på universalsatsen är att inget n i M uppfyller P. Den kan omformuleras som att alla n i M saknar egenskapen P. För dessa typer av satser är motsägelsebevis ofta smidiga: man antar att det finns ett n i M som uppfyller P och härleder en motsägelse.

Universal- och existenssatser är duala till varandra, i bemärkelsen att om du ska bevisa en existenssats med hjälp av ett motsägelsebevis så antar du en universalsats, och vice versa, se bevisen av 1.2.2 och 1.2.3.

1.4 Mängder

En *mängd* är en samling objekt. Man kan samla nästan vad man vill i en mängd: tal, katter, och andra mängder.[2] Det viktiga är att man alltid kan avgöra ifall ett objekt tillhör mängden eller inte. De objekt som ligger i mängden kallas för *element*.

[2]Vi skriver *nästan* av en anledning. Det finns samlingar av objekt som kan beskrivas men som inte utgör en mängd. Detta kallas *Russells paradox*, efter Bertrand Russell (1872–1970). Russells exempel är samlingen av alla mängder som inte innehåller sig själv.

Det lättaste sättet att beskriva en mängd är att räkna upp elementen som ingår i den. För att markera att objekten ligger i en mängd, så omger man listan med *mängdklamrar* { och }. Mängden som innehåller 1, 2 och 3 skrivs alltså som

$$\{1, 2, 3\}.$$

Två mängder A och B är lika om de innehåller samma element, vilket skrivs $A = B$. Det spelar ingen roll i vilken ordning man skriver elementen eller hur många gånger de listas. Därför gäller att

$$\{1, 1, 2, 3\} = \{1, 2, 3\} = \{2, 3, 1\}.$$

Om ett element x tillhör en mängd A brukar man skriva $x \in A$, vilket uttalas som x *tillhör* A. Om x inte tillhör A skriver man $x \notin A$. Antalet element som tillhör en mängd A brukar betecknas med $|A|$ eller $\#A$. Eftersom vi inte vill blanda ihop detta begrepp med absolutbeloppet av ett tal som introduceras senare i detta kapitel använder vi $\#$.

Mängden på formen $\{\}$ innehåller inte några element alls och kallas den *tomma mängden*. Den brukar betecknas med $\emptyset$ och är unik i aspekten att den saknar element, vi skriver alltså $\#\emptyset = 0$.

En mängd kan innehålla andra mängder som element. Mängden

$$A = \{\{1, 2\}, 3\}$$

har två element vilket skrivs som $\#A = 2$. Dess element är: mängden $\{1, 2\}$ och talet 3. Mängden $\{1, 2\}$ innehåller i sin tur elementen 1 och 2. Däremot innehåller A varken 1 eller 2, det vill säga

$$\{1, 2\} \in A \qquad \text{men } 1 \notin A.$$

Att mängder kan innehålla andra mängder kan ha paradoxala konsekvenser. Till exempel kan vi lägga den tomma mängden i en mängd, och bilda mängden av den tomma mängden.

$$A = \{\emptyset\} = \{\{\}\}$$

Mängden A innehåller ett element, den tomma mängden, och är därför inte tom. Mängden av den tomma mängden är alltså inte lika med den tomma mängden.

Detta verkar motsägelsefullt. Den tomma mängden är ju tom, så mängden av den tomma mängden borde ju också vara tom? Tricket är att skilja på mängden och elementen i mängden. Den tomma mängden är ju ett element i sig, även om den inte innehåller några element, precis som att 0 är ett tal, trots att representerar ett antal som inte finns. Man kan tänka sig att en påse som innehåller en annan tom påse, inte är tom.

Det finns ingen begränsning på hur stor en mängd kan vara, och de flesta mängder man studerar innehåller oändligt många element. Dessa mängder kan naturligtvis inte skrivas ut som en lista. Istället beskriver man dem med *mängdbyggaren*, som har följande allmänna form $\{x \mid \text{villkor på } x\}$. Den här mängden består av alla element som uppfyller villkoret. Ett exempel är mängden

$$\{n \mid n \text{ är jämnt}\} = \{\ldots, -4, -2, 0, 2, 4, \ldots\}$$

som innehåller alla jämna tal.

En mängd B är en *delmängd* av en mängd A om alla element som tillhör B även tillhör A. Man skriver detta som $B \subset A$. Till exempel så är $\{1, 2\}$ en delmängd av $\{1, 2, 3\}$, eftersom 1 och 2 är element i båda mängderna. Om två mängder är delmängder av varandra så är de lika.

En icke-tom mängd har alltid minst två delmängder: sig själv och den tomma mängden. En delmängd B av A är *äkta* om $B \neq A$.

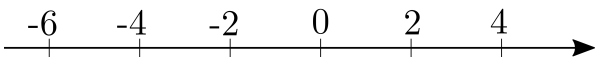
Det är lätt att blanda ihop element och delmängder. Det beror på att mängder kan innehålla andra mängder, så att en delmängd av en mängd kan vara ett element i mängden. Mängden $A = \{\emptyset\}$ är ett bra exempel. Den tomma mängden är både ett element i och en delmängd av A.

I mängden $A = \{1, 2, \{1, 2\}\}$ är $\{1, 2\}$ både en delmängd och ett element. Däremot så är $\{1\}$ enbart en delmängd av A, medan 1 enbart är ett element.

De olika talsystemen kan ses som mängder av tal, och har fått egna beteckningar. De *naturliga talen* betecknas med $\mathbb{N}$ och består av talen 0, 1, 2, 3, och så vidare.[3]

Naturliga tal kan adderas och multipliceras utan problem. Resultatet är alltid ett nytt naturligt tal. För att subtrahera behöver vi införa de negativa talen -1, -2, och så vidare. De naturliga talen tillsammans med de negativa talen kallas för *heltalen*, och betecknas med $\mathbb{Z}$ (av tyskans *Zahl* = tal).

Heltal kan adderas, subtraheras och multipliceras. Man kan däremot inte dividera dem med varandra. För detta krävs *rationella tal*. De definieras som alla kvoter a/b, där a och b är heltal och b är skilt från 0. Mängden av alla rationella tal betecknas med $\mathbb{Q}$.

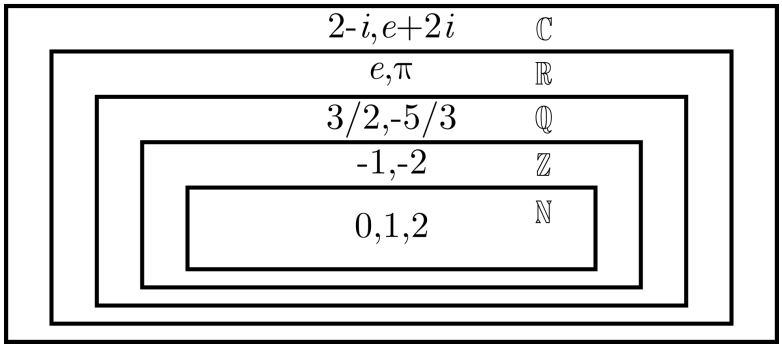


Figur 1.1: Tallinjen runt 0.

De rationella talen ligger på den så kallade *tallinjen*, som går från negativa tal till vänster och till positiva tal till höger (se Figur 1.1). Det finns dock tal som inte är rationella, men som ändå ligger på tallinjen. Ett exempel är $\sqrt{2}$, som är längden på diagonalen i en kvadrat med sidan 1. Lägger man till dessa tal får de *reella talen*, som betecknas med $\mathbb{R}$.[4] Reella tal som inte är rationella kallas för *irrationella*.

[3] Vissa exkluderar 0 från de naturliga talen. Att inkludera 0 har dock fördelar. Om man börjar räkna från 0 och går ett steg i taget kommer man ha gått n steg när man räknat till n. Exempel: om vi räknar till 3 från 0 så får vi $0 \to 1 \to 2 \to 3$, vilket är 3 steg. Om vi börjar från 1 får vi istället $1 \to 2 \to 3$, vilket är 2 steg.

[4] Reella tal är mycket mystiska. Den matematiska cirkeln 2016–2017, *Vad är ett tal?*, handlade om hur man kan definiera dem i termer av rationella tal. Den intresserade läsaren uppmanas att söka upp kompendiet på Cirkelns hemsida: https://www.math-stockholm.se/samverkan/cirkel/



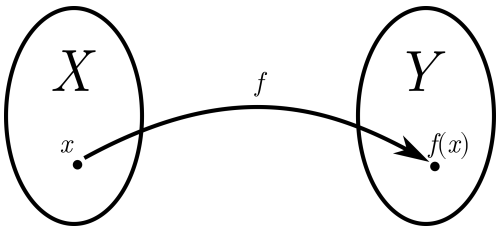
Figur 1.2: De olika talsystemen från $\mathbb{N}$ till $\mathbb{C}$.

De reella talen kan utvidgas ytterligare till de *komplexa talen*, som betecknas med $\mathbb{C}$, genom att lägga till ett tal i som uppfyller $i^2 = -1$.

1.5 Funktioner

En funktion f från en mängd X till en mängd Y beskriver hur man parar ihop element i en mängd X med element i en mängd Y. Det brukar skrivas som $f\colon X \to Y$. Mängden X kallas för *definitionsmängd* och mängden Y kallas för *målmängd*. Man kan se f som en process som tar ett element i mängden X och avger ett element som ligger i mängden Y. När man tillämpar en funktion på ett element x i X så kallas x för funktionens *argument*. Mängden av alla värden en funktion i praktiken antar kallas för funktionens *värdemängd*, och denna betecknas ofta med V_f. Värdemängden är då alltså en delmängd av målmängden, och kan beskrivas som $V_f = \{f(x) \mid x \in X\}$. Till exempel är $\sin(x)$ en funktion från $\mathbb{R}$ till $\mathbb{R}$, så funktionens målmängd är alltså $\mathbb{R}$, men värdemängden är $[-1, 1]$.

Två funktioner är lika när de har samma definitionsmängd, samma målmängd och de är lika på alla element i definitionsmängden. Definitions- och målmängden är alltså en del av funktionen.



Figur 1.3: En funktion f från X till Y.

Funktioner beskrivs ofta med formler. Exempelvis så kan funktionen $f\colon \mathbb{N} \to \mathbb{N}$ som tar ett naturligt tal och returnerar dess kvadrat beskrivas som $f(n) = n^2$. Alla polynom kan ses som en funktion från $\mathbb{R}$ till $\mathbb{R}$, som beräknas genom att man sätter in talet x i uttrycket. En funktion måste dock inte ges av en formel. Det enda som krävs är att funktionen är definierad för alla element i definitionsmängden, och att den alltid ger samma svar. Vi ger nu ett par

exempel på detta och hur man istället kan beskriva en funktion.

Exempel 1.5.1. Vårt första exempel är absolutbeloppet $|x|$ av ett reellt tal x, som definieras som avståndet på tallinjen från x till origo. Detta är en funktion vars definitionsmängd och målmängd är $\mathbb{R}$. Man kan beräkna den genom att man tar bort eventuella minustecken framför talet, det vill säga

$$|x| = \begin{cases} x & \text{om } x \geqslant 0 \\ -x & \text{om } x < 0. \end{cases}$$

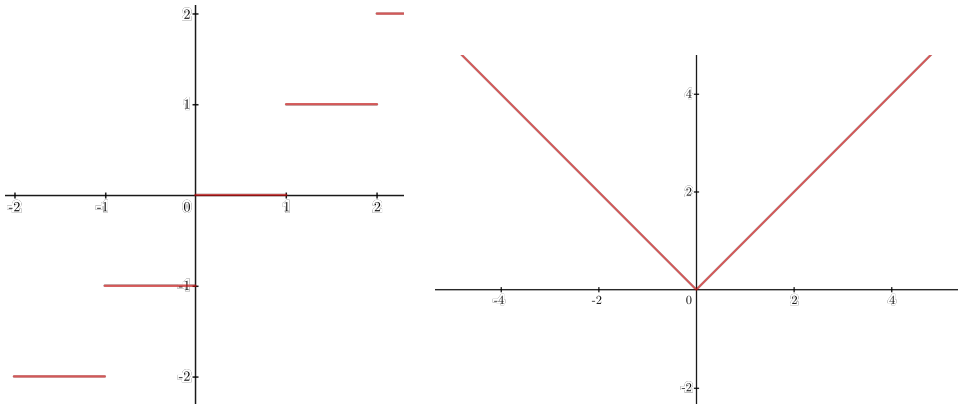
Exempelvis så gäller $|-3| = -(-3) = 3$ och $|2| = 2$. ▲

Exempel 1.5.2. *Golvfunktionen* är funktionen som avbildar reella tal x på heltalet man får när man avrundar x nedåt. Vi benämner den funktionen med $\lfloor \ldots \rfloor$. För att förtydliga hur denna och andra funktioner används brukar vi skriva på följande sätt

$$\begin{aligned} \lfloor \ldots \rfloor \colon \mathbb{R} &\to \mathbb{Z} \\ x &\mapsto \lfloor x \rfloor. \end{aligned}$$

eftersom det kanske inte är uppenbart för läsaren att man ska skriva $\lfloor x \rfloor$ när man vill referera till funktionens värde vid punkten x istället för $\lfloor \ldots \rfloor_x$ eller $\lfloor \ldots \rfloor(x)$. Exempelvis så gäller $\lfloor \pi \rfloor = 3$. ▲

En funktion $f \colon \mathbb{R} \to \mathbb{R}$ kan beskrivas genom sin *graf*, som definieras som mängden av punkter i planet på formen $(x, f(x))$.



(a) Grafen av $f(x) = \lfloor x \rfloor$ **(b)** Grafen av funktionen $f(x) = |x|$

Figur 1.4: Golv- och absolutbeloppsfunktionen.

Om man istället betraktar en funktion med ändlig definitionsmängd kan man beskriva den med en tabell.

Exempel 1.5.3. Betrakta funktionen från mängden $\{1, 2, 3\}$ till mängden $\{2, 3, 4, 5\}$ där $f(1) = 3$ och $f(2) = 4$ och $f(3) = 3$. Vi kan beskriva denna med en tabell eftersom vi bara ändligt många element i vår definitionsmängd,

x	f(x)
1	3
2	4
3	3

Vi kan också beskriva den med en formel eftersom definitionsmängden och målmängden är delmängder till de reella talen $\mathbb{R}$ så till exempel gäller

$$f(x) = 4 - (x-2)^2. \qquad \blacktriangle$$

Exempel 1.5.4. Betrakta funktionen från mängden $\{\text{hund}, \text{katt}\}$ till mängden $\{0, 1\}$ där $f(\text{hund}) = 0$ och $f(\text{katt}) = 1$. Vi kan beskriva denna med en tabell eftersom vi bara har ändligt många element i vår definitionsmängd,

x	f(x)
hund	0
katt	1

.

Det finns dock inte någon vettig formel som beskriver f eftersom det inte finns något välkänt sätt att 'addera' eller 'multiplicera' orden hund och katt på ett sätt som ger ett tal. $\blacktriangle$

Definition 1.5.5. Givet två funktioner sådana att den enas definitionsmängd är den andras målmängd, $f\colon Y \to Z$ och $g\colon X \to Y$ kan vi definiera deras *sammansättning* $f \circ g\colon X \to Z$ enligt regeln

$$(f \circ g)(x) = f(g(x)). \qquad \triangle$$

Definition 1.5.6. Givet en funktion $f\colon X \to Y$ och en delmängd $A \subset X$ så kallar vi mängden $\{f(x) \mid x \in A\}$ för *bilden av* A och den betecknas $f(A)$. Om x är ett element i X brukar vi även kalla $f(x)$ för *bilden* av x. $\triangle$

Definition 1.5.7. Givet en funktion $f\colon X \to Y$ och en delmängd $B \subset Y$ så kallar vi mängden $\{x \mid f(x) \in B\}$ för *urbilden av* B. Om y är ett element i Y brukar vi kalla urbilden $\{x \mid f(x) = y\}$ för *fibern* av y. Givet ett element x i fibern av y kommer vi i den här kursen beteckna fibern av y med $[x]$. $\triangle$

Definition 1.5.8. Vi säger att en funktion $f\colon X \to Y$ är *surjektiv* om alla element i Y är bilden av något element i X, det vill säga om värdemängden överensstämmer med målmängden. En funktion sägs vara *injektiv* om varje element i värdemängden är bilden av exakt ett element i definitionsmängden. Om en funktion är både surjektiv och injektiv säger vi att funktionen är *bijektiv*. $\triangle$

Notera att en bijektiv funktion är *inverterbar*: Det vill säga, det existerar en funktion som kallas *den inversa funktionen* eller f^{-1} med beskrivningen $f^{-1}\colon Y \to X$ och $f^{-1}(f(x)) = x$ för alla $x \in X$ och $f(f^{-1}(y)) = y$ för alla $y \in Y$. Den inversa funktionen, om den finns, är alltså den funktion som för varje element $y \in Y$ ger det unika elementet $x \in X$ som har egenskapen att $f(x) = y$. I övning 1.14 ser vi exempelvis att en strikt växande funktion $f\colon \mathbb{R} \to \mathbb{R}$ är injektiv. Det är även värt att notera att en injektiv funktion blir bijektiv om vi byter ut målmängden mot värdemängden.

Huruvida en funktion är inverterbar eller ej beror inte bara på regeln som beskriver funktionen, utan på definitionsmängd och målmängden.

Exempel 1.5.9. Betrakta funktionen $f(x) = x^2$ som tar ett tal och kvadrerar det.

Betraktat som funktion från $\mathbb{R}$ till $\mathbb{R}$ är den inte inverterbar. Den är inte surjektiv eftersom det till exempel inte finns något $x \in \mathbb{R}$ så att $x^2 = -1$. Problemet här är alltså att funktionens målmängd inte överensstämmer med dess värdemängd, det vill säga att funktionen inte är surjektiv.

Detta problem kan lösas genom att begränsa målmängden till värdemängden, men även om vi betraktar f som en funktion från $\mathbb{R}$ till dess värdemängd, det vill säga mängden av icke-negativa tal $[0, \infty)$, så är funktionen inte inverterbar. Vi har fortfarande problemet att funktionen inte är injektiv. Till exempel så har vi $(-1)^2 = 1^2 = 1$ och eftersom det finns två element som avbildas på 1, kan vi inte entydigt definiera en invers funktion.

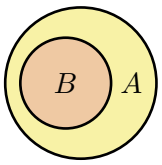
Om vi däremot betraktar f som en funktion från $[0, \infty)$ till $[0, \infty)$ så är den både injektiv och surjektiv, det vill säga bijektiv, och således har funktionen en väldefinierad invers, nämligen $f^{-1}(x) = \sqrt{x}$. Till exempel så är i detta fall $f^{-1}(1) = 1$ eftersom vi bara får välja den positiva roten.

Vi har en liknande situation med $\sin(x)$ och $\cos(x)$. Betraktade som funktioner från $\mathbb{R}$ till $\mathbb{R}$ är dessa funktioner varken injektiva eller surjektiva, men genom att betrakta $\sin(x)$ som en funktion från $[-\pi/2, \pi/2]$ till $[-1, 1]$ och $\cos(x)$ som en funktion från $[0, \pi]$ till $[-1, 1]$ erhåller vi bijektiva, och således inverterbara funktioner. ▲

1.6 Mängdoperationer

I detta avsnitt ska vi beskriva ett antal olika sätt från en eller flera mängder skapa en ny mängd. Vi kallar dessa för *mängdoperationer*. I nästa kapitel kommer vi även att se hur vissa av dessa mängdoperationer går att generalisera till andra situationer.

För att illustrera mängder använder man ibland *Venndiagram*, efter matematikern John Venn (1834–1923). Där representeras mängder som enkla former, oftast cirklar, och formernas förhållanden till varandra motsvarar mängdernas. Till exempel kan man illustrera att B är en delmängd av A genom att rita dem som två cirklar, där B ligger inuti A.



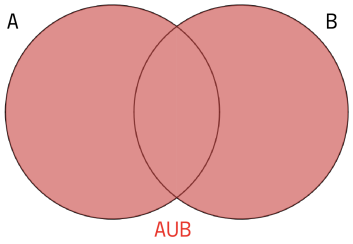
Figur 1.5: Venndiagram för $B \subset A$.

Unionen

Unionen av mängderna A och B är mängden som består av alla element som ligger i A *eller* i B. Den betecknas med $A \cup B$ och definieras som

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

Ett exempel är $\{1, 2, 3\} \cup \{2, 3, 4\} = \{1, 2, 3, 4\}$.



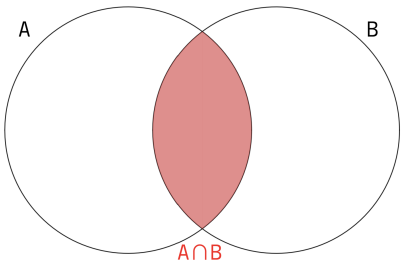
Figur 1.6: Venndiagram för $A \cup B$.

Snittet

Snittet av mängderna A och B är mängden som består av alla element som ligger i A *och* i B. Den betecknas med $A \cap B$ och definieras som

$$A \cap B = \{x \mid x \in A \text{ och } x \in B\}.$$

Ett exempel är $\{1, 2, 3\} \cap \{2, 3, 4\} = \{2, 3\}$.



Figur 1.7: Venndiagram för $A \cap B$.

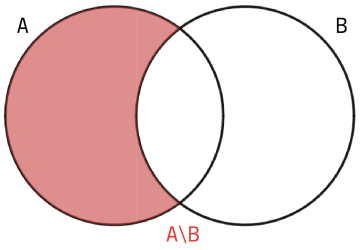
Två mängder A och B sägs vara *disjunkta* om de inte har några gemensamma element, det vill säga om $A \cap B = \emptyset$.

Differensen och komplementet

Differensen av en mängd A och B är mängden som består av alla element som ligger i A men inte i B. Den betecknas med $A \setminus B$ och definieras som

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

Ett exempel är $\{1, 2, 3\} \setminus \{2, 3, 4\} = \{1\}$.



Figur 1.8: Venndiagram för $A \setminus B$.

Notera att $A \setminus B$ inte är lika med $B \setminus A$, exempelvis gäller

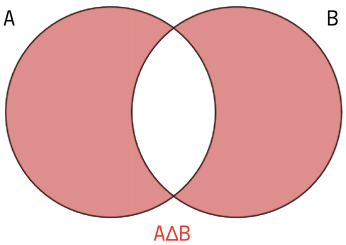
$$\{2, 3, 4\} \setminus \{1, 2, 3\} = \{4\} \neq \{1\} = \{1, 2, 3\} \setminus \{2, 3, 4\}.$$

Om alla uppträdande mängder är delmängder av en viss mer eller mindre underförstådd grundmängd M talar man ofta om $M \setminus A$ som *komplementet* till A (med avseende på M). Vi kommer att beteckna komplementet till A med A^c. Om vi till exempel pratar om mängder av heltal, och vi till exempel betraktar mängden $A = \{1, 2, 3\}$, så avser komplementet till A mängden av alla heltal *förutom* $1, 2$ och 3.

Symmetrisk differens

Den *symmetriska differensen* av två mängder A, B definieras som mängden av element som ligger *bara i* A eller *bara i* B. Vi skriver detta som

$$A \Delta B = \{x \mid x \in (A \cup B) \text{ och } x \notin (A \cap B)\}.$$



Figur 1.9: Venndiagram för $A \Delta B$.

Exempelvis gäller $\{2,3,4\}\Delta\{1,2,3\} = \{1,4\}$.

Den kartesiska produkten

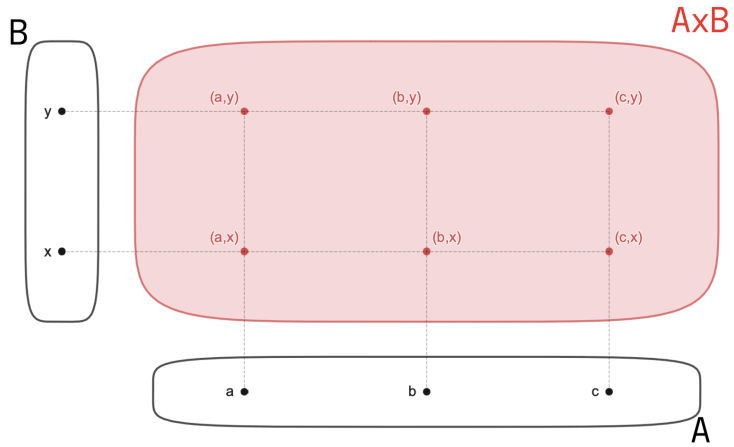
Den kartesiska produkten av två mängder A och B är mängden som består av alla par av element så att det första ligger i A och det andra i B. Den betecknas med $A \times B$ och definieras som

$$A \times B = \{(x,y) \mid x \in A \text{ och } y \in B\}.$$

Två olika exempel är

$$\{a,b,c\} \times \{x,y\} = \{(a,x),(a,y),(b,x),(b,y),(c,x),(c,y)\}$$
$$\{1,2,3\} \times \{2,3,4\} = \{(1,2),(1,3),(1,4),(2,2),(2,3),(2,4),(3,2),(3,3),(3,4)\}.$$

Notera att ordningen är viktig, exempelvis gäller $(1,2) \in \{1,2,3\} \times \{2,3,4\}$ men $(2,1) \notin \{1,2,3\} \times \{2,3,4\}$.



Figur 1.10: Venndiagram för $A \times B$.

När vi tar den kartesiska produkten av en mängd med sig själv brukar vi använda notationen $A^2 = A \times A$. Mer allmänt skriver vi

$$A^n = \underbrace{A \times \ldots \times A}_{n \text{ gånger}} = \{(a_1,\ldots,a_n) \mid \text{ alla } a_i \in A\}.$$

Till exempel är det reella talplanet $\mathbb{R} \times \mathbb{R}$ vilket vi brukar beteckna med $\mathbb{R}^2$.

Mängden av funktioner

Mängden av funktioner från mängden A till mängden B är mängden som består av alla funktioner från A till B. Den betecknas med B^A och definieras som

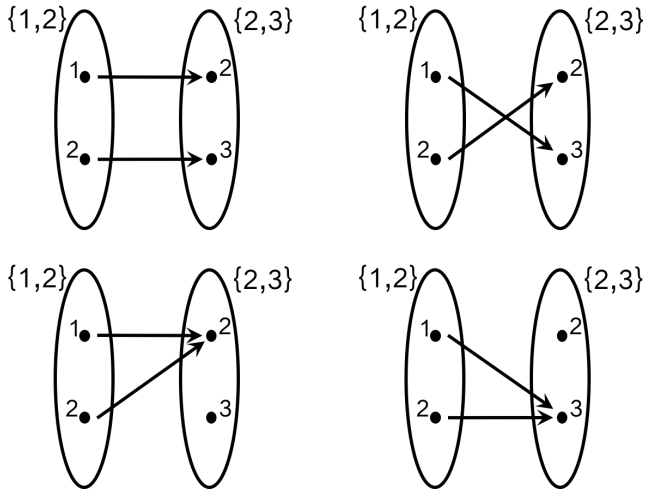
$$B^A = \{f \mid f\colon A \to B\}.$$

Ett exempel är

$$\{2,3\}^{\{1,2\}} = \left\{ \begin{array}{|c|c|}\hline \text{x} & \text{f(x)} \\ \hline 1 & 2 \\ \hline 2 & 2 \\ \hline \end{array}, \begin{array}{|c|c|}\hline \text{x} & \text{f(x)} \\ \hline 1 & 2 \\ \hline 2 & 3 \\ \hline \end{array}, \begin{array}{|c|c|}\hline \text{x} & \text{f(x)} \\ \hline 1 & 3 \\ \hline 2 & 2 \\ \hline \end{array}, \begin{array}{|c|c|}\hline \text{x} & \text{f(x)} \\ \hline 1 & 3 \\ \hline 2 & 3 \\ \hline \end{array} \right\}$$

vilken också kan skrivas som

$$\{2,3\}^{\{1,2\}} = \{f(x) = 2, f(x) = x + 1, f(x) = 4 - x, f(x) = 3\}.$$



Figur 1.11: Illustration av $\{2,3\}^{\{1,2\}}$.

Potensmängd

Vi definierar *potensmängden* av en mängd som mängden av alla delmängder till en mängd A. Den betecknas med

$$2^A = \{B \mid B \subset A\}.$$

Ett exempel är

$$2^{\{a,b,c\}} = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a,b\}, \{a,c\}, \{b,c\}, \{a,b,c\}\},$$

Kvotmängd (överkurs)

Kvotmängden av A kan definieras på många olika sätt. Vi väljer att definiera kvotmängden av A givet en funktion $f\colon A \to B$ som mängden vars element är fibrerna till f. Elementen i kvotmängden är alltså delmängder till A på formen $\{x \in A\colon f(x) = b\}$. För varje $a \in A$ använder vi notationen

$$[a] = \{x \in A\colon f(x) = f(a)\}.$$

Vi kallar i denna kontext mängden $[a]$ för en *ekvivalensklass* och vi kallar a för en *representant* för ekvivalensklassen $[a]$. Olika element $a \neq b \in A$ kan vara

representanter för samma ekvivalensklass, vi säger att a och b är *ekvivalenta* när $[a] = [b]$. Detta inducerar en s.k. *ekvivalensrelation* mellan elementen i A som vi betecknar med $\sim$.

Vi definierar kvotmängden

$$A/\sim \quad = \{[a] \mid a \in A\}.$$

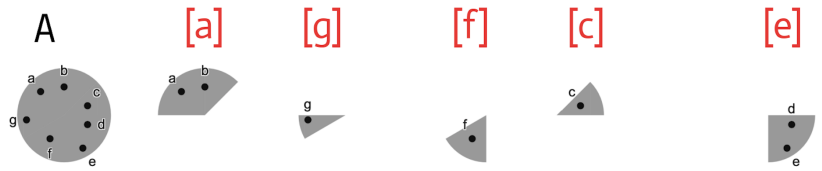
Låt oss demonstrera detta med ett par exempel.

Exempel 1.6.1. Låt $A = \{a, b, c, d, e, f, g\}$ och

$$\begin{aligned} &\phi\colon A \to \{0, 1, 2, 3, 4\} \\ &\quad \phi(a) = \phi(b) = 0 \quad \phi(g) = 1 \quad \phi(f) = 2 \quad \phi(c) = 3 \quad \phi(d) = \phi(e) = 4. \end{aligned}$$

och då blir

$$A/\sim \quad = \{\{a, b\}, \{g\}, \{f\}, \{c\}, \{d, e\}\}.$$



Figur 1.12: Illustration av hur elementen i $A/\sim$ är delmängder till A. ▲

Exempel 1.6.2. Ifall $f\colon \mathbb{Z} \to \{0, 1\}$ definieras som

$$f(n) = \begin{cases} 0 \text{ om } n \text{ jämnt} \\ 1 \text{ om } n \text{ udda.} \end{cases}$$

Då kan vi se att

$$\mathbb{Z}/\sim \quad = \{[0], [1]\}$$

är en mängd med två element, $[0]$ som är mängden av de jämna talen och $[1]$ som är mängden av de udda talen. Till exempel gäller det att $[1] = [3] = [51]$, eftersom talen $1, 3, 51$ kan alla väljas som representanter för de udda talen i detta exempel. ▲

Övningar

Övning 1.1. Lista elementen i följande mängder.

(i) $A = \{n \in \mathbb{N} \mid k < 5\}$.

(ii) $B = \{1, 2, \{2, 3\}\}$.

(iii) $C = \{k \in \mathbb{Z} \mid k^2 < 16\}$.

(iv) $A \cap B$.

(v) $(C \setminus A) \cup B$.

Övning 1.2. Lista elementen i följande mängder.

(i) $A = \{x \in \mathbb{Q} \mid x^2 = 2\}$.

(ii) $B = \{0, 1, 2, 3\}$.

(iii) $C = \{p/q \mid p, q \in \mathbb{N}, 0 \leqslant p < 3 \text{ och } 1 \leqslant q < 3\}$.

(iv) $(B \cup A) \cap C$.

(v) $(C \cap B) \setminus A$.

Övning 1.3. För nedanstående par av mängder A och B, avgör om A och B är lika, disjunkta, någon av dem är en äkta delmängd av den andra eller ingetdera.

(i) $A = \{1, 2, 3\}$ och $B = \{1, 1, 2\}$.

(ii) $A = \{0, 1, 2\}$ och $B = \{n \in \mathbb{N} \mid n^2 < 9\}$.

(iii) $A = \{\{\}\}$ och $B = \{x \in \mathbb{N} \mid 2x = -2\}$.

(iv) $A = \{x \in \mathbb{R} \mid |x| < 1\}$ och $B = \{x \in \mathbb{R} \mid |x - 1| < 1\}$.

(v) $A = \{x \in \mathbb{Q} \mid x^2 = 2\}$ och $B = \{x \in \mathbb{R} \mid x^2 = 2\}$.

Övning 1.4. För nedanstående par av mängder A och B, avgör om A och B är lika, disjunkta eller någon av dem är en äkta delmängd av den andra.

(i) $A = \{-2, 0, 2\}$ och $B = \{x \in \mathbb{Z} \mid |x| < 3 \text{ och } x \text{ är jämnt}\}$.

(ii) $A = \{x \in \mathbb{R} \mid x^2 < 2\}$ och $B = \{x \in \mathbb{Q} \mid x^2 \geqslant 2\}$.

(iii) $A = \{x \in \mathbb{Z} \mid x \text{ är jämnt}\}$ och $B = \{x \in \mathbb{Z} \mid x \text{ är kvadrattal}\}$.

(iv) $A = \{x \in \mathbb{Z} \mid 2x = -2\}$ och $B = \{x \in \mathbb{N} \mid 2x = 2\}$.

(v) $A = \{\emptyset, \{\emptyset\}\}$ och $B = \{\emptyset\}$.

Övning 1.5. Använd mängdbyggaren för att definiera följande mängder.

(i) Mängden av jämna, positiva heltal.

(ii) Mängden av rationella tal r så att $2r$ är ett heltal.

(iii) Mängden av irrationella tal som ligger inom avstånd 1 från origo.

Övning 1.6. Använd mängdbyggaren för att definiera följande mängder.

(i) Mängden av alla kvadrattal som är större än 2.

(ii) Mängden av rationella lösningar till $x^4 + x^2 - 1 = 0$.

(iii) Mängden av rationella tal som är volymen av en kub med rationella sidor.

Övning 1.7. Ange möjlig definitionsmängd och målmängd för följande funktioner.

(i) Funktionen som ger det n:te kvadrattalet.

(ii) Funktionen som beräknar arean av triangel.

(iii) Funktionen beräknar derivatan av ett andragradspolynom.

Övning 1.8. Ange möjlig definitionsmängd och målmängd för följande funktioner.

(i) Funktionen som ger arean av cirkel med radie r.

(ii) Funktionen som ger avståndet mellan 1 och ett tal r på tallinjen.

(iii) Funktionen som ger de rationella nollställena till ett förstagradspolynom med rationella koefficienter.

Övning 1.9. Är följande funktioner eller inte? Om inte, motivera varför.

(i) $f : \mathbb{R} \to \mathbb{R}$ där
$$f(x) = \begin{cases} 1 \text{ om } x \in \mathbb{Q} \\ 0 \text{ om } x \notin \mathbb{Q}. \end{cases}$$

(ii) $f : \mathbb{N} \to \mathbb{Q}$ där $f(n) = \sqrt{n}$.

(iii) $f : \mathbb{R} \to \mathbb{R}$ så att $f(x) = 0$ med sannolikhet $1/2$ och $f(x) = 1$ med sannolikhet $1/2$.

(iv) $f : \{0\} \to \mathbb{R}$ där $f(0) = 1$ om ordet Balkong börjar på B.

Övning 1.10. Är följande funktioner eller inte? Om inte, motivera varför.

(i) $f : \mathbb{Z} \to \mathbb{N}$, där $f(n)$ är siffersumman i det vanliga (decimala) talsystemet. Obs, alla *siffror* är icke-negativa.

(ii) $f : \mathbb{R} \to \mathbb{Q}$, där $f(x) = x/2$.

(iii) $f : \mathbb{N} \to \mathbb{R}$, där $f(n) = \sqrt[n+1]{n+1}$.

(iv) $f : \mathbb{Q} \to \mathbb{Z}$, där $f(p/q) = p$.

Övning 1.11. Avgör om följande funktioner är lika eller inte? Motivera varför.

(i) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$ och $g : \mathbb{R} \to \mathbb{R}$, $g(x) = \sqrt{x^2}$.

(ii) $f : \mathbb{Z} \to \mathbb{Q}$, $f(n) = 1/n$ och $g : \mathbb{N} \to \mathbb{Q}$, $g(m) = 1/m$.

(iii) $f : \mathbb{N} \to \mathbb{Q}$, $f(n) = n/(n+1)$ och $g : \mathbb{N} \to \mathbb{R}$, $g(z) = z/(z+1)$.

Övning 1.12. Avgör om följande funktioner är lika eller inte? Om inte, motivera varför.

(i) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2 + 2x + 1$ och $g : \mathbb{R} \to \mathbb{R}$, $g(x) = (x+1)^2$.

(ii) $f : \mathbb{Z} \to \mathbb{Z}$, $f(x) = x^2$ och $g : \mathbb{Z} \to \mathbb{Z}$, $g(x) = |x|^2$.

(iii) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \text{frac}(x)$ och $g : \mathbb{Q} \to \mathbb{R}$, $g(x) = x - \lfloor x \rfloor$.

Övning 1.13. Vad är värdemängden i exemplen 1.5.3 och 1.5.4? Avgör i vilka av de två exemplen som f är injektiv och i vilka som f är surjektiv.

Övning 1.14. Bevisa att en strängt växande funktion, alltså en funktion med egenskapen

$$x_1 < x_2 \implies f(x_1) < f(x_2)$$

är injektiv.

Övning 1.15. Visa att A och B är disjunkta om och endast om $A \setminus B = A$.

Övning 1.16. Visa att för två mängder A, B så uppfyller operationerna $\cup, \cap, A^C$ (där A^C betecknar mängdkomplement) följande räkneregler,

$$(A \cup B)^C = A^C \cap B^C \qquad (A \cap B)^C = A^C \cup B^C.$$

Övning 1.17. Visa att följande räkneregler håller

$$A \setminus B = A \setminus (A \cap B) \qquad (A \cup B) \setminus C = (A \setminus C)$$

Övning 1.18. Visa att

$$A \Delta B = (A \setminus B) \cup (B \setminus A).$$

Ledtråd: Använd övningen 1.17.

Övning 1.19. Visa att sammansättningen av två bijektioner är en bijektion.

Övning 1.20. Visa att

$$A/\sim \quad \subset 2^A.$$

Övning 1.21. Visa att antalet element i $A \times B$ är produkten av antalet element i A och antalet element i B.

Övning 1.22. Visa att den kartesiska produkten av två delmängder är delmängd till den kartesiska produkten av grundmängderna. Alltså om $B \subset A$ och $Y \subset X$ gäller det att

$$B \times Y \subset A \times X.$$

Övning 1.23. Visa att potensmängden $2^A = \{B \mid B \subset B\}$ är ett specialfall av mängdoperationen Y^A för ett par av mängder A, Y. Gör detta genom att visa att följande funktion är en bijektion,

$$\begin{aligned} \chi : & 2^A \to \{0,1\}^A \\ & B \mapsto \chi_B. \end{aligned}$$

Här definierar vi *indikatorfunktionen av* B som

$$\begin{aligned} \chi_B : A &\to \{0,1\} \\ a &\mapsto \begin{cases} 1 \text{ om } a \in B \\ 0 \text{ om } a \notin B. \end{cases} \end{aligned} .$$

Dra slutsatsen att 2^A har $2^{(\#A)}$ antal element.

Övning 1.24 ($\star$)**.** Använd ett direkt bevis för att bevisa att om n är ett jämnt tal så är $n+1$ ett udda tal (detta är Sats 1.1.4).

Övning 1.25 ($\star$)**.** Använd ett direkt bevis för att bevisa att om n är udda så är n^2 udda.

Övning 1.26 ($\star$)**.** Använd ett motsägelsebevis för att bevisa att summan av ett irrationellt tal och ett rationellt tal är irrationellt.

Övning 1.27 ($\star$)**.** Använd ett motsägelsebevis för att bevisa att om $a+b \geqslant c$ så är antingen $a \geqslant c/2$ eller $b \geqslant c/2$

Övning 1.28 ($\star\star$)**.** Bevisa att om $ab = c$ så är $a \geqslant \sqrt{c}$ och $b \leqslant \sqrt{c}$, eller tvärtom.

Övning 1.29 ($\star$)**.** Bevisa att det finns två irrationella tal a och b så att a^b är rationellt. Tips: använd att $\sqrt{2}$ är irrationellt.

Övning 1.30 ($\star\star$)**.** Bevisa att det finns alla rationella tal kan skrivas som en produkt av två irrationella tal.

Övning 1.31 ($\star$)**.** Visa följande påståenden för heltalen (du bör använda det du bevisat i ena för att bevisa nästa).

(i) Multiplikation med ett positivt tal bevarar olikheter. Alltså för $m > 0$ gäller

$$r \leqslant n \implies mr \leqslant mn.$$

(ii) Ett positivt heltal kan inte dela ett mindre positivt heltal. Alltså

$$0 < r < n \implies n \text{ delar inte } r.$$

(iii) Differensen av två heltal delbara med n är också delbar med n.

2 Grupper

I detta kapitel introducerar vi ett nytt koncept som vi kallar för en matematisk *grupp*. I talspråk kan man säga att en grupp är en mängd av objekt som har ett recept för hur man kan kombinera par av element för att skapa ett tredje element av samma sort. Exempelvis så är heltalen en grupp, eftersom vi kan addera ett par av heltal för att skapa ett tredje heltal. Ett annat exempel på en grupp är mängden av positiva reella tal, detta eftersom produkten av två positiva reella tal också är ett positivt reellt tal. Som tredje exempel och utblick så är huvudtemat för denna kurs, de *elliptiska kurvorna*, också grupper. Konceptet av en grupp är grundstenen för vad vi idag kallar för *abstrakt algebra*.

För att betona betydelsen av detta kapitel så är grupper alltså mängder där vi kan utöva algebra och de generaliserar den algebra som vi vanligtvis utövar på tallinjen. Vi kommer i senare kapitel behöva definitionen av en grupp för att sätta ord på hur varje elliptisk kurva har en väldigt intressant och rik algebraisk struktur.

Definition 2.0.1. En grupp är ett par $(G, *)$ där G är en mängd och $*$ är en funktion

$$\begin{aligned} *\colon G \times G &\to G \\ (x, y) &\mapsto x * y \end{aligned}$$

som vi kallar för grupp-operationen och den uppfyller *gruppaxiomen*:

(i) (Associativitet) För alla element $x, y, z \in G$ gäller $(x * y) * z = x * (y * z)$.

(ii) (Identitet) Det finns ett element $e \in G$, som vi kallar *identitetselement*, sådan att för alla $x \in G$ gäller $e * x = x * e = x$.

(iii) (Inverser) För varje element $x \in G$ existerar ett element $x^{-1} \in G$, som vi kallar för *inversen av* x, sådan att $e = x * x^{-1} = x^{-1} * x$.

En grupp $(G, *)$ kallas *abelsk* om den dessutom uppfyller ett fjärde villkor:

(iv) (Kommutativitet) För alla $x, y \in G$ så gäller $x * y = y * x$. △

Notera att gruppoperationen inte måste heta $*$, den kan också heta $\star, +, \cdot, \Delta$ och så vidare. Beroende på vilken grupp man jobbar med, såsom heltalen eller permutationsgruppen framgår det ofta från kontexten vilken gruppoperation man jobbar med. Man brukar alltså ofta bara skriva 'G' när man syftar på en grupp, där gruppoperation antingen framgår från kontexten eller där det inte spelar någon roll vad för namn man ger gruppoperationen. På samma sätt måste identiteten inte heta just e, antingen har den ofta ett speciellt namn för våra favoritgrupper, eller så spelar det inte så stor roll vad man benämner den. Identitetselementet kan till exempel heta $\mathrm{id}, i, u, 0, 1$ eller något dylikt.

I denna kurs kommer vi att fokusera på abelska grupper. Vi ger nu två fundamentala exempel på abelska grupper.

Exempel 2.0.2. Låt oss beteckna mängden av positiva reella tal med $\mathbb{R}_{>0}$. Ett första exempel på en grupp är de positiva reella talen med gruppoperationen multiplikation $\cdot$ vilket vi betecknar med $(\mathbb{R}_{>0}, \cdot)$. Detta kan vi bekräfta eftersom för positiva reella tal uppfyller gruppaxiomen:

(Associativitet) För alla element $x, y, z \in \mathbb{R}_{>0}$ gäller $(x \cdot y) \cdot z = x \cdot (y \cdot z)$.

(Identitet) För elementet $1 \in \mathbb{R}_{>0}$, som vi kallar *ettan*, gäller för alla $x \in \mathbb{R}_{>0}$ att $1 \cdot x = x \cdot 1 = x$.

(Inverser) För varje element $x \in \mathbb{R}_{>0}$ existerar ett element $\frac{1}{x} \in \mathbb{R}_{>0}$, som vi kallar för *ett över* x, sådan att $1 = x \cdot \frac{1}{x} = \frac{1}{x} \cdot x$.

(Kommutativ) För alla $x, y \in \mathbb{R}_{>0}$ så gäller $x \cdot y = y \cdot x$. ▲

Exempel 2.0.3. Ett andra exempel på en grupp är heltalen $\mathbb{Z}$ med addition som gruppoperation. Denna grupp som vi betecknar med $(\mathbb{Z}, +)$ uppfyller gruppaxiomen

(Associativitet) För alla element $n, m, p \in \mathbb{Z}$ gäller $(n + m) + p = n + (m + p)$.

(Identitet) För elementet $0 \in \mathbb{Z}$, som vi kallar *nollan*, gäller för alla $n \in \mathbb{Z}$ att $0 + n = n + 0 = n$.

(Inverser) För varje element $n \in \mathbb{Z}$ existerar ett element $(-n) \in \mathbb{Z}$, som vi kallar för *minus* n, sådan att $0 = n + (-n) = (-n) + n$.

(Kommutativitet) För alla $n, m \in \mathbb{Z}$ så gäller $n + m = m + n$.

På samma sätt är de rationella talen $(\mathbb{Q}, +)$ och de reella talen $(\mathbb{R}, +)$ grupper med addition som gruppoperation. ▲

2.1 Ett första exempel på en ändlig grupp

Låt oss nu studera ett annat exempel på en grupp vars gruppoperation är betydligt mer annorlunda från addition på tallinjen. Gruppen i fråga kommer att bara ha ändligt många element. När en grupp har ändligt många element kan vi beskriva dess gruppoperation med en så kallad *grupptabell* som vi kommer att se i detta utförliga exempel. En grupptabell är en tabell som har en rad och en kolumn för varje element i gruppen. Så givet två element $x, y \in G$ så finns det en position i tabellen på samma rad som x och samma kolumn som y och där skriver vi ned $x * y$. Denna tabell kommer alltså att innehålla all information som finns att hämta om gruppen.

Sats 2.1.1. *Definiera mängden av* permutationer av n element *som*

$$S_n = \{f \colon \{0, 1, \ldots, n-1\} \to \{0, 1 \ldots, n-1\} |\, f \text{ är en bijektion}\}.$$

Paret $(S_n, \circ)$*, där* $\circ$ *är sammansättning av funktioner, är en grupp.*

Bevis. Vi börjar med att anmärka att sammansättningen av två bijektioner också är en bijektion (se övning 1.19). Vi verifierar nu gruppaxiomen,

(i) (Associativitet). För alla element $f, g, h \in S_n$ gäller $(f \circ g) \circ h = f \circ (g \circ h)$. Kom ihåg att två funktioner är lika om och endast om de alltid har samma funktionsvärde. Vi kan då verifiera att för alla i $x \in X$ gäller

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

(ii) (Identitet) Vi kallar funktionen som skickar alla element till sig själva för identitetsfunktionen,

$$\begin{aligned} \mathrm{id} \colon \{0, 1, \ldots, n-1\} &\to \{0, 1 \ldots, n-1\} \\ m &\mapsto m. \end{aligned}$$

Identitetsfunktionen är en bijektion och det gäller för alla $f \in S_n$ att $\mathrm{id} \circ f = f \circ \mathrm{id} = f$.

(iii) (Inverser) För en bijektiv funktion f existerar en unik invers funktion (som vi diskuterat i kapitel 1) det vill säga att det existerar ett element $f^{-1} \in S_n$, som per definition uppfyller $\mathrm{id} = f \circ f^{-1} = f^{-1} \circ f$.

□

Exempel 2.1.2. Låt oss studera gruppen S_3 och namnge dess 6 olika element. Notera att det finns just 6 element enligt övningen 2.15.

$S_3 =$

x	id(x)
0	0
1	1
2	2

,

x	$\tau(x)$
0	1
1	2
2	0

,

x	$\pi(x)$
0	2
1	0
2	1

,

x	$f(x)$
0	0
1	2
2	1

,

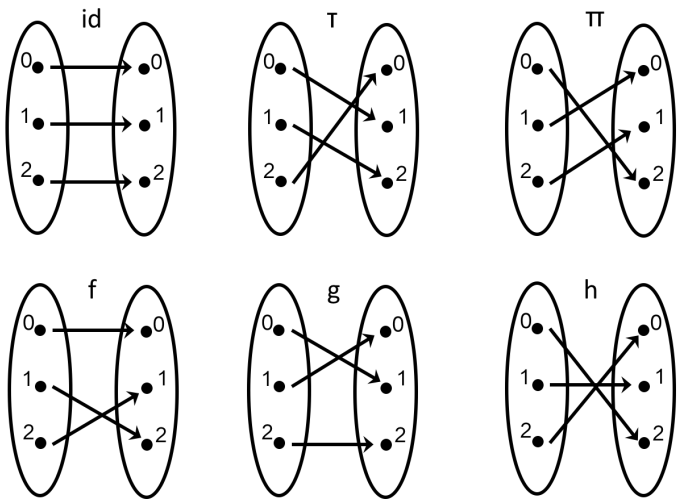
x	$g(x)$
0	1
1	0
2	2

,

x	$h(x)$
0	2
1	1
2	0

.

För att få en intuitiv förståelse av gruppen så kan följande figur vara till hjälp.

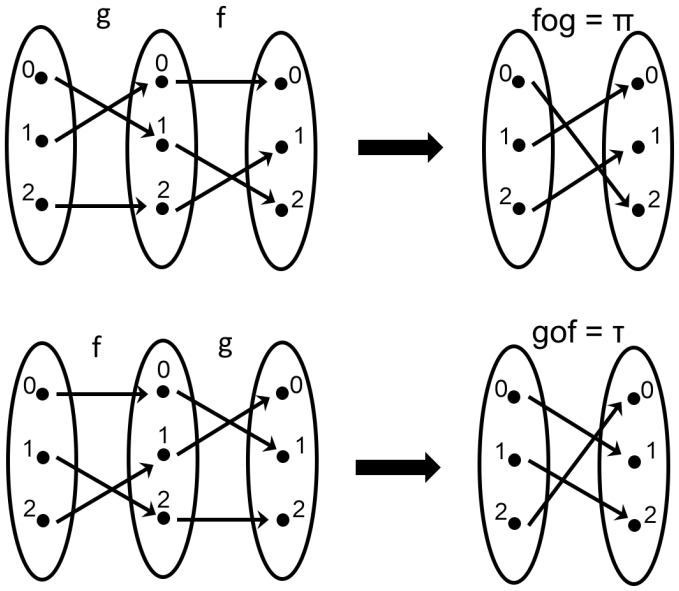


Figur 2.1: Illustration av S_3.

Notera hur dessa funktioner, de så kallade *permutationerna* byter plats på (alternativt kör hela havet stormar med) elementen i mängden $\{0, 1, 2\}$. Vi

observerar att denna grupp *inte* är en abelsk grupp. Som exempel, notera att $f \circ g \neq g \circ f$ i följande exempel

x	f	g	$f \circ g$	$g \circ f$
0	0	1	2	1
1	2	0	0	2
2	1	2	1	0



Figur 2.2: Illustration av hur S_3 *inte* är abelsk.

Slutligen kan vi skriva ned den så kallade *grupptabellen* för S_3.

$$G = \{\mathrm{id}, \tau, \pi, f, g, h\}$$

$\circ$	id	τ	π	f	g	h
id	id	τ	π	f	g	h
τ	τ	π	id	g	h	f
π	π	id	τ	h	f	g
f	f	h	g	id	π	τ
g	g	f	h	τ	id	π
h	h	g	f	π	τ	id

Tabell 2.3: Grupptabellen för S_3.

Högst upp i högra hörnet brukar vi skriva antingen namnet på gruppoperationen, $\circ$, eller namnet på gruppen, i detta fall S_3. ▲

2.2 Grundläggande gruppteori

I detta avsnitt ser vi till att övertyga oss själva om att grupper inte är så farliga som de verkar. Exempelvis så finns det bara ett unikt identitetselement i varje grupp och inversen till inversen av ett element är elementet självt. Dessutom introducerar vi konceptet av *grupphomomorfi* (även kallat *gruppavbildning*) som kommer att låta oss konstatera att två till synes olika grupper egentligen kan vara 'samma', vilket vi kallar för att grupperna är *isomorfa*.

Till exempel så kan vi betrakta en grupp som vi kallar $\mathbb{Z}/(2)$, den har elementen $\{0, 1\}$ beskrivs av följande grupptabell som vi ställer sida vid sida med grupptabellen för S_2

$\mathbb{Z}/(2)$	0	1
0	0	1
1	1	0

S_2	id	τ
id	id	τ
τ	τ	id

Tabell 2.4: Grupptabellen för S_2 och $\mathbb{Z}/(2)$ är identiska utöver att man bytt namn på elementen id, τ till $0, 1$. τ är permutationen som byter plats på 0 och 1.

När grupptabellerna ställs sida vid sida kan vi se att de är väldigt lika, man har i princip bara bytt namn på de två gruppelementen. Detta är ett första exempel på två *isomorfa* grupper. Vi avslutar kapitlet med att introducera konceptet delgrupper och kartesiska produkten av grupper.

Hjälpsats 2.2.1. *En grupp $(G, *)$ kan inte ha två identitetselement och inget element $x \in G$ har två stycken inverser.*

Bevis. För båda påståenden i satsen gör vi ett motsägelsebevis. Antag att det finns två olika identitetselement e och e' som båda uppfyller (ii), då gäller

$$e = e * e' = e'$$

vilket då motsäger gruppaxiomen. För det andra påståendet, antag att det finns ett element x med två olika inverser x^{-1} och x_{-1} som uppfyller (iii), då gäller

$$x^{-1} \overset{(ii)}{=} x^{-1} * e = x^{-1} * (x * x_{-1}) \overset{(i)}{=} (x^{-1} * x) * x_{-1} = e * x_{-1} \overset{(ii)}{=} x_{-1}$$

vilket då motsäger gruppaxiomen. □

Hjälpsats 2.2.2. *Givet en grupp $(G, *)$ och ett par av element $x, y \in G$ gäller*

$$(x * y)^{-1} = y^{-1} * x^{-1}.$$

Bevis. Vi beräknar helt enkelt

$$(x*y)*(y^{-1}*x^{-1}) = x*(y*(y^{-1}*x^{-1})) = x*((y*y^{-1})*x^{-1}) = x*(e*x^{-1}) = x*x^{-1} = e.$$

Repetera argumentet från andra hållet och så är beviset klart. □

Grupphomomorfier

Definition 2.2.3. Givet två grupper $(G, *)$ och $(H, \star)$ och en funktion $\phi\colon G \to H$ så säger vi att f är en *grupphomomorfi* om och endast om för alla par $x, y \in G$

$$\phi(x * y) = \phi(x) \star \phi(y).$$

Vi skriver då $\phi\colon (G, *) \to (H, \star)$. △

Hjälpsats 2.2.4. *En grupphomomorfi $\phi\colon (G, *) \to (H, \star)$ måste skicka identiteten i G till identiteten i H och $\phi(x^{-1}) = \phi(x)^{-1}$ för alla x.*

Bevis. Låt e vara identiteten i G och u vara identiteten i H. Vi börjar med att notera

$$\phi(e) = \phi(e * e) = \phi(e) \star \phi(e).$$

Då ser vi att

$$u \overset{(iii)}{=} \phi(e)^{-1} \star \phi(e) = \phi(e)^{-1} \star (\phi(e) \star \phi(e)) \overset{(i)}{=} (\phi(e)^{-1} \star \phi(e)) \star \phi(e) = u \star \phi(e) = \phi(e).$$

Det bevisar det första påståendet. För det andra påståendet så ser vi att

$$\phi(x) \star \phi(x^{-1}) = \phi(x * x^{-1}) = \phi(e) = u.$$

Om man repeterar argumentet för $\phi(x^{-1}) \star \phi(x)$ har vi nu visat att $\phi(x^{-1})$ är en invers till $\phi(x)$ och då ger hjälpsatsen 2.2.1 att det är den unika inversen. □

Exempel 2.2.5. Låt $\mathbb{R}_{>0}$ beteckna de positiva reella talen. Då är logaritmen en grupphomomorfi

$$\log\colon (\mathbb{R}_{>0}, \cdot) \to (\mathbb{R}, +).$$

Detta är på grund av den välkända logaritmlagen $\log(xy) = \log(x) + \log(y)$. ▲

Exempel 2.2.6. Betrakta grupphomomorfin $\phi\colon (\mathbb{Z}, +) \to (\mathbb{Z}, +)$ som definieras enligt

$$\phi(n) = 3n.$$

Funktionen ϕ är en grupphomomorfi eftersom

$$\phi(n + m) = 3(n + m) = 3n + 3m = \phi(n) + \phi(m).$$ ▲

Exempel 2.2.7. Betrakta grupphomomorfin $\phi\colon G \times G \to G$ som definieras enligt

$$\phi(x, y) = x.$$

Den glömmer helt enkelt bort en koordinat bara. Funktionen ϕ är en grupphomomorfi eftersom

$$\phi\left((x, y) * (a, b)\right) = \phi\left((x * a, y * b)\right) = x * a = \phi(x, y) * \phi(a, b).$$ ▲

Sats 2.2.8. *Sammansättningen av grupphomomorfier är en grupphomomorfi.*

Bevis. Låt $\phi\colon (G,*) \to (H,\star)$ och $\psi\colon (H,\star) \to (K,\times)$ vara grupphomomorfier. Då gäller det att

$$\psi(\phi(x*y)) = \psi(\phi(x)\star\phi(y)) = \psi(\phi(x)) \times \psi(\phi(y))$$

alltså har vi $(\psi\circ\phi)\colon (G,*)\to(K,\times)$. □

Hjälpsats 2.2.9. *Om en grupphomomorfi ϕ är bijektiv är dess inversfunktion också en grupphomomorfi. Vi kallar en sådan funktion för gruppisomorfi.*

Bevis. Låt $\phi\colon (G,*)\to(H,\star)$ vara bijektiv. Då existerar det en inversfunktion ϕ^{-1}. Vi ser att

$$\phi(\phi^{-1}(x\star y)) = x\star y = \phi(\phi^{-1}(x))\star\phi(\phi^{-1}(y)) = \phi(\phi^{-1}(x)*\phi^{-1}(y)).$$

Eftersom ϕ^{-1} är bijektiv kan vi nu applicera den på båda sidor av likheten ovan vilket resulterar i den önskade likheten

$$\phi^{-1}(x\star y) = \phi^{-1}(x)*\phi^{-1}(y).$$ □

Delgrupp

För att en delmängd till en grupp G ska kunna ärva egenskapen att vara en grupp med samma operation som G måste vi kräva ett par saker av den, vi beskriver dessa villkor i följande definition.

Definition 2.2.10. Givet en grupp $(G,*)$ och en delmängd $H\subset G$ säger vi att H är en *delgrupp* till G om och endast om $e\in H$ och om det gäller för alla $x,y\in H$ att $x^{-1}\in H$ och $x*y\in H$. Då är $(H,*)$ en grupp. △

Exempel 2.2.11. Givet ett positivt heltal n så är mängden

$$n\mathbb{Z} = \{m\in\mathbb{Z} \mid m \text{ är en heltalsmultipel av } n\}$$

en delgrupp av $(\mathbb{Z},+)$. Vi ser detta då 0 alltid är $0\cdot n$, om $m=kn$ så är $-m=(-k)\cdot n$ och summan av två tal delbara med n är också delbart med n. ▲

Kartesiska produkten av två grupper

Definition 2.2.12. *Den kartesiska produkten* av två grupper $(G,*)$ och $(H,\star)$ är mängden som består av alla par av element så att det första ligger i G och andra i H. Den betecknas med $G\times H$. Vi brukar utelämna att namnge den naturliga gruppoperationen, som definieras enligt

$$\begin{aligned}(G\times H)\times(G\times H) &\to G\times H\\ ((a,x),(b,y)) &\mapsto (a*b, x\star y).\end{aligned}$$ △

Exempel 2.2.13. Ett exempel är den kartesiska produkten av $(\mathbb{Z},+)$ med sig själv. Vi skriver då $\mathbb{Z}\times\mathbb{Z}$ som är mängden av punkter i talplanet med heltalskoordinater och kallar dess gruppoperation $+$. Exempelvis så har vi att

$$(5,7)+(2,-3) = (5+2,7-3) = (7,4).$$ ▲

Ordning av ett element

Definition 2.2.14. Låt $(G, *)$ vara en grupp och g vara ett element i G. Vi använder följande notation

$$g^0 = e, \quad g^1 = g, \quad g^n = \underbrace{g * \ldots * g}_{n \text{ gånger}}.$$

Om det finns ett n så att $g^n = e$, säger vi att det minsta sådana positiva n är g:s ordning. Om det inte finns ett sådant n, säger vi att g har oändlig ordning. △

Exempel 2.2.15. Betrakta elementet -1 i $(\mathbb{R} \setminus \{0\}, *)$. Vi har att

$$(-1)^2 = 1.$$

Alltså har -1 ordning 2. Om vi istället betraktar 2 så har vi att

$$2^2 = 4, \quad 2^3 = 8 \quad ...$$

Då resultatet bara växer finns inget n så att $2^n = 1$, alltså har 2 oändlig ordning. ▲

2.3 Analoger till mängdoperationer och funktioner

Grupper kan ses som mängder med 'extra struktur' i meningen att vi kan utföra algebra i mängden. På samma sätt är en grupphomomorfi en funktion med 'extra struktur' i det att den respekterar gruppoperationerna. Det är naturligt att fråga vad som händer med denna extra struktur när man utför diverse mängdoperationer på grupper och grupphomomorfier. Förlorar vi den eller bevaras den? I följande tabell listar vi olika operationer som tidigare definierats för ett par av mängder eller en funktion och ifall resultatet av operationen blir en grupp ifall man utgick ifrån ett par av grupper respektive en grupphomomorfi. Vi lämnar bevisen som övningar i slutet av kapitelet.

Mängdoperation	Gruppstruktur bevaras	Motivering
Union $\cup$	Nej	Övning 2.9
Delmängd $\subset$	Ibland	Definition 2.2.10
Snitt $\cap$	Ja	Övning 2.20
Differens $\setminus$ och Δ	Nej	Övning 2.14
Cartesisk produkt $\times$	Ja	Definition 2.2.12
Funktionsmängd G^X	Ja	Övning 2.21
Kvotmängd $G/\sim$	Ibland	$\mathbb{Z}/n\mathbb{Z}$ (bevis överkurs).

Funktionsoperation	Gruppstruktur bevaras	Motivering
Värdemängd	Ja	Övning 2.22
Urbild	Ja	Övning 2.22
Sammansättning	Ja	Sats 2.2.8
Inversfunktion	Ja	Hjälpsats 2.2.9

Exempel 2.3.1 (Snitt av delgrupper)**.** Ett exempel på ett snitt av delgrupper av $(\mathbb{Z}, +)$ är $2\mathbb{Z}, 3\mathbb{Z} \subset \mathbb{Z}$ där

$$(2\mathbb{Z}) \cap (3\mathbb{Z}) = \{n \mid n \text{ delbart med 2 och 3}\} = 6\mathbb{Z}.$$

▲

Exempel 2.3.2 (Gruppen av funktioner)**.** Betrakta gruppen $(\mathbb{R}, +)$ och $(\mathbb{R}^{\{1,2\}}, +)$. Ett par av element i $(\mathbb{R}^{\{1,2\}}, +)$ är funktionerna $f, g \colon \{1, 2\} \to \mathbb{R}$

x	$f(x)$
1	π
2	-1

x	$g(x)$
1	7
2	$\sqrt{2}$

Då får vi

x	$(f+g)(x)$
1	$\pi + 7$
2	$\sqrt{2} - 1$

▲

Exempel 2.3.3 (Gruppen av funktioner)**.** Betrakta gruppen $(\mathbb{R}, +)$ och $(\mathbb{R}^{\mathbb{R}}, +)$. Ett par av element i $(\mathbb{R}^{\mathbb{R}}, +)$ är funktionerna $f, g \colon \mathbb{R} \to \mathbb{R}$

$$f(x) = \sin(x) \qquad g(x) = 2x - 7x^8.$$

Då gäller det att

$$(f + g)(x) = \sin(x) + 2x - 7x^8.$$

▲

Exempel 2.3.4 (Värdemängden av en grupphomomorfi)**.** Betrakta grupphomomorfin

$$\begin{aligned} \phi \colon \mathbb{Z} &\to \mathbb{Z} \\ n &\mapsto 3n. \end{aligned}$$

Värdemängden för ϕ är alla tal delbara på tre som vi betecknar med $3\mathbb{Z}$ vilket är en delgrupp, vilket vi såg i exempel 2.2.11. ▲

Exempel 2.3.5 (Urbilden av grupphomomorfier)**.** Betrakta grupphomomorfin

$$\begin{aligned} \phi \colon \mathbb{Z} &\to \mathbb{Z} \\ n &\mapsto 3n. \end{aligned}$$

Funktionen ϕ är en grupphomomorfi eftersom

$$\phi(n + m) = 3(n + m) = 3n + 3m = \phi(n) + \phi(m).$$

Betrakta urbilden av delgruppen $6\mathbb{Z}$, alltså

$$\{n \mid 3n \in 6\mathbb{Z}\} = \{n \mid n \in 2\mathbb{Z}\} = 2\mathbb{Z}.$$

Urbilden av $6\mathbb{Z}$ är alltså gruppen av jämna heltal i detta fall. ▲

2.4 Notation

Vi har varit noggranna genom texten att skriva $(G, *)$ när vi refererar till en grupp, men i praktiken skriver man ofta G när man refererar till en grupp. Man brukar använda symbolen $\cdot$ för att beskriva multiplikation, eller inte skriva något alls, så att xy är produkten av x och y. Det här är för att göra matematisk text mer läsbar, och det är ofta tydligt från sammanhanget vad multiplikationen innebär.

Övningar

Övning 2.1. Är följande exempel grupper eller inte? Vi betecknar multiplikation med $\cdot$ och addition med $+$.

i) $(\mathbb{N}, +)$

ii) $(\mathbb{Z}, +)$

iii) $(\mathbb{Z} \setminus \{0\}, \cdot)$

iv) $(\{q \in \mathbb{Q} \mid q \geqslant 0\}, \cdot)$

v) $(\mathbb{C} \setminus \{0\}, \cdot)$

Övning 2.2. Visa att för alla element $\sigma \in S_3$ gäller det att

$$\underbrace{\sigma \circ \ldots \circ \sigma}_{6 \text{ gånger}} = \mathrm{id}.$$

Övning 2.3. Betrakta följande mängd med 2 element

$$B = \{\text{sant}, \text{falskt}\}.$$

Dessa två värden kallas för *booleans* inom programmering. Definiera gruppoperationen som vi kallar för *exclusive or* (förkortas till xor)

$$\begin{aligned} \text{xor} : & B \times B \to B \\ & (p, q) \to \text{antingen } p \text{ eller } q \text{ är 'sant' men inte båda.} \end{aligned}$$

Vi kan också beskriva gruppoperationen med en grupptabell

xor	falskt	sant
falskt	falskt	sant
sant	sant	falskt

(i) Visa att B är en grupp.

(ii) Visa att grupperna S_2, $\mathbb{Z}/(2)$ från tabell 2.4 och B alla är isomorfa. (Många programmeringsspråk tycker att 0 och 'falskt' är samma sak pga detta.)

Övning 2.4. Visa att $n\mathbb{Z} \cap m\mathbb{Z} = \text{mgm}(n, m)\mathbb{Z}$. Här står mgm för den *minsta gemensamma multipeln*. Talet $\text{mgm}(n, m)$ kan definieras som det tal som delar alla tal som delas av både n och m.

Övning 2.5. Är följande en delgrupp till S_3?

$$K = \{f \in S_3 \mid f(\{1,2\}) = \{1,2\} \quad f(\{3,4\}) = \{3,4\}\}.$$

(kom ihåg notationen för bilden, $f(A)$, av en delmängd A från kapitel 1).

Övning 2.6. Är följande exempel giltiga grupptabeller eller inte? Om exemplet är en grupp, är den abelsk eller inte?

$G = \{\text{id}, \sigma, \tau, \pi\}$

$*$	id	σ	τ	π
id	id	σ	τ	π
σ	σ	id	π	τ
τ	τ	π	id	σ
π	π	τ	σ	id

$H = \{e, i, j\}$

$\star$	e	i	j
e	e	i	j
i	i	i	e
j	j	e	j

Ledtråd: Du kan jämföra G med grupptabellen för $\mathbb{Z}/(2\mathbb{Z}) \times \mathbb{Z}/(2\mathbb{Z})$ där $\pi = (1, 1)$.

Övning 2.7. Givet en grupp $(G, *)$ med identitet e.

(a) Visa att identiten är sin egen invers, alltså $e = e^{-1}$.

(b) Visa att varje element x är inversen till sin invers, alltså $(x^{-1})^{-1} = x$.

Övning 2.8. Visa att för en abelsk grupp $(G, *)$ gäller det att inversen $f(x) = x^{-1}$ är en grupphomomorfi från G till sig själv.

Övning 2.9 ($\star$). Givet två delgrupper sådan att ingen innehåller den andra, visa att unionen aldrig är en delgrupp.

Övning 2.10. Visa att delmängden $\{e\}$ som bara har identiten som element, alltid är en delgrupp.

Övning 2.11. Verifiera att följande tabell gör paret $(\{0, 1, 2\}, \star)$ blir en abelsk grupp.

$\star$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

Övning 2.12. Studera mängden $G = \{e, a, b, c, d, f\}$ med en gruppoperation vars grupptabell är på formen

$\star$	e	a	b	c	d	f
e	e	a	b	c	d	f
a	a	e	d	f	b	c
b	b	f	e	d	c	a
c	c	d	f	e	a	b
d	d	c	a	b	f	e
f	f	b	c	a	e	d

Tabell 2.5: Grupptabellen för G

Låt oss verifiera (utöver associativitet) att detta ger en icke-abelsk grupp genom att motivera följande påståenden.

(i) (Associativ för mycket jobb) Men verifiera åtminstone att

$$(a \star d) \star c = a \star (d \star c).$$

(ii) Elementet e är identiteten.

(iii) Alla element har en invers.

(iv) Gruppen är icke-abelsk.

Övning 2.13. Givet en grupp $(G, *)$ kan vi definiera en 'ny' grupp $(G, *_{op})$ (*op* för opposite) som har samma element men vars gruppoperation är

$$x *_{op} y = y * x.$$

a) Visa att $(G, *_{op})$ är en grupp.

b) Om G har ändligt många element, hur förhåller sig grupptabellen för $(G, *_{op})$ och $(G, *)$.

c) Visa att $(G, *)$ är en abelsk grupp om och endast om $* = *_{op}$ som funktioner $G \times G \to G$.

Övning 2.14. Visa att mängddifferensen och den symmetriska differensen av två delgrupper aldrig är en delgrupp.

Övning 2.15. Visa att S_n har $n!$ antal element. Detta kallas för *fakulteten* av n,

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

Exempelvis är $\#(S_3) = 6 = 3 \cdot 2 \cdot 1$ från exempel 2.1.2.

Övning 2.16 ($\star$)**.** Visa att om $m \leqslant n$ är heltal så finns det en delgrupp $G \subset S_n$ sådan att G är isomorf med S_m.

Övning 2.17 ($\star$)**.** Visa att mängden av alla grupphomomorfier $\phi : (G, *) \to (H, \star)$ är en delgrupp till H^G om H är abelsk.

Övning 2.18 ($\star$)**.** Visa att om $(G, *)$ är en grupp (identiteten kallas e) med ändligt många element, så finns det ett heltal k sådant att för alla element $x \in G$ gäller det att

$$\underbrace{x * \ldots * x}_{k \text{ gånger}} = e.$$

Anmärkning: Det visar sig att $k = \#G$ funkar men det behöver du inte visa.

Övning 2.19 ($\star$)**.** En funktion $f : \mathbb{R} \to \mathbb{R}$ kallas jämn om $f(-x) = f(x)$ respektive udda om $f(-x) = -f(x)$. Låt $(\mathbb{R}[x], +)$ beteckna gruppen av polynom i en variabel och vars koefficienter är reella tal med addition som gruppoperation. Låt även J och U vara delmängderna av **j**ämna respektive **u**dda polynom. Definiera funktionen

$$\begin{aligned} \phi :&\mathbb{R}[x] \to J \times U \\ &\phi(f) = \left(\frac{f(x) + f(-x)}{2}, \frac{f(x) - f(-x)}{2} \right). \end{aligned}$$

(i) Skriv ned ett element i J och ett annat element i U.

(ii) Visa att J och U är delgrupper till $(\mathbb{R}[x], +)$.

(iii) Visa att ϕ är en grupphomomorfi.

(iv) Visa att ϕ har en invers och dra slutsatsen att ϕ är bijektiv.

(v) Vad är $J \cap U$?

Övning 2.20. Visa att om H, K är delgrupper så är $H \cap K$ en delgrupp.

Övning 2.21. Givet en mängd X och en grupp $(G, *)$ så är paret $(G^X, *)$ en grupp. Gruppoperation definieras enligt $(f * g)(x) := f(x) * g(x)$.

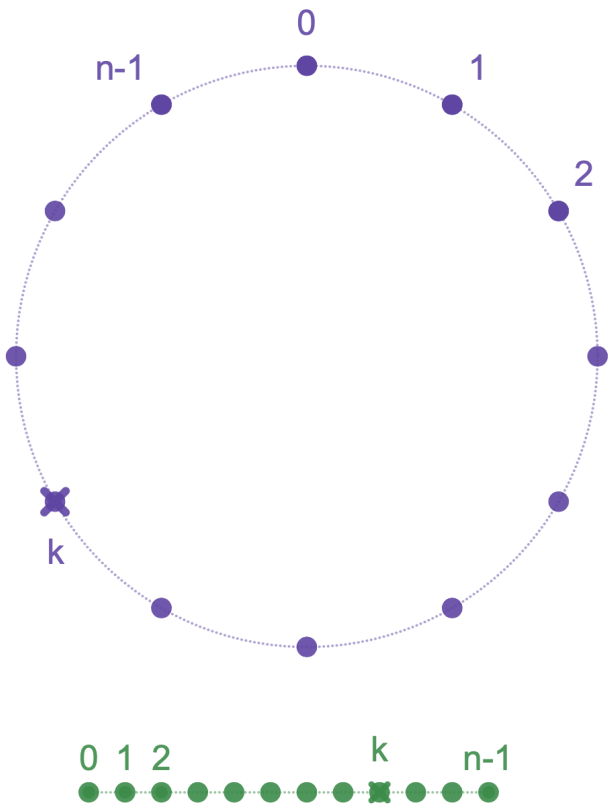
Övning 2.22. Låt $\phi : (G, *) \to (H, \star)$ vara en grupphomomorfi. Visa att

(i) värdemängden av ϕ är en delgrupp av H.

(ii) urbilden av en delgrupp är en delgrupp.

3 Modulär aritmetik och ringar

I detta kapitel kommer vi att spendera mycket tid med en av de viktigaste grupperna i matematiken, $\mathbb{Z}/(n)$, heltalen modulo n. Vi kommer även se att det går att bygga vidare på konceptet av en matematisk grupp. Exempelvis så har heltalen $\mathbb{Z}$ mer algebraisk struktur än det som garanteras av gruppaxiomen för $+$. Vi har nämligen en andra operation, multiplikationsoperationen, som också uppfyller en annan uppsättning regler, de så kallade *ringaxiomen*. Detta gäller även för heltalen modulo n. Vi kommer sedan spendera den resterande delen av kapitlet för att bekanta oss med heltalen modulo n, varför de är viktiga och hur ringar beter sig när man utför mängdoperationer på dem.



Figur 3.1: Illustration av $\mathbb{Z}/(n)$ för $n = 12$ som en 'klocka' eller som punkter på en linje.
(Länk till animation i Desmos)

3.1 Gruppen $\mathbb{Z}/(n)$

Vi ska i detta avsnitt definiera en samling väldigt speciella grupper. Mer specifikt så kan vi för varje heltal heltal n definiera gruppen som vi kallar *heltalen modulo n* . Vi betecknar heltalen modulo n med $\mathbb{Z}/(n)$. Att utföra beräkningar

i heltalen modulo n kallas för *modulär aritmetik*. Den modulära aritmetiken är väldigt viktigt eftersom det är med den som datorer utför sina beräkningar.

Exempel 3.1.1. En analog klocka har 12 stycken olika timslag. Om vi väntar 5 timmar och sedan 8 timmar till, så kommer avståndet mellan timvisarens start och slutpunkt vara detsamma som om du bara väntade en timme. Alltså skulle man kunna tänka sig att en klocka tycker att $'5+8=1'$, vilket kan låta absurt. Detta är vårt första exempel på modulär aritmetik där vi utfört en beräkning modulo 12. Vi ska i detta kapitel utveckla och försöka formalisera detta koncept. Här ger vi också två olika illustrationer av hur man kan tänka på heltalen modulo n.

▲

Definition 3.1.2. Givet ett heltal $n > 1$ så definierar vi mängden

$$n\mathbb{Z} = \{m | \, m \text{ är en heltalsmultipel av } n\}. \qquad \triangle$$

Definition 3.1.3. Givet ett positivt heltal n så definierar vi *heltalsresten vid division med* n som

$$\begin{aligned} \%n \colon \mathbb{Z} &\to \{0, 1, \ldots, n-1\} \\ k &\mapsto \begin{cases} 0 \text{ om } k \in n\mathbb{Z} \\ 1 \text{ om } k-1 \in n\mathbb{Z} \\ \vdots \\ n-1 \text{ om } k-(n-1) \in n\mathbb{Z} \end{cases} . \end{aligned}$$

Alternativt så kan $k\%n$ definieras som det unika tal i mängden $\{0, 1, \ldots, n-1\}$ som uppfyller $k - k\%n \in n\mathbb{Z}$. Vi definierar två heltal a, b att vara *kongruenta modulo* n om och endast om $a\%n = b\%n$ och vi skriver då

$$a \equiv_n b. \qquad \triangle$$

Resten av x vid division med n har många namn, ett annat exempel är x mod n. Läsaren som behöver ett bevis för att $\%n$ är väldefinierad hänvisas till övningen 1.31.

Exempel 3.1.4. Vi har att

$$\begin{aligned} \%2 \colon \mathbb{Z} &\to \{0, 1\} \\ k &\mapsto \begin{cases} 0 \text{ om } k \text{ jämnt} \\ 1 \text{ om } k \text{ udda} \end{cases} \end{aligned}$$

vilket exempelvis ger att $17\%2 = 1$ och $100\%2 = 0$. Alternativt så skriver man $17 \equiv_2 1$ eller $100 \equiv_2 0$. Ett annat exempel är

$$\begin{aligned} \%3 \colon \mathbb{Z} &\to \{0, 1, 2\} \\ k &\mapsto \begin{cases} 0 \text{ om } k \text{ delbart med } 3 \\ 1 \text{ om } k-1 \text{ delbart med } 3 \\ 2 \text{ om } k-2 \text{ delbart med } 3. \end{cases} \end{aligned}$$

där vi har använt att $k-2$ är delbart med 3 om och endast om $k-2+3 = k+1$ är det. Vi kan då beräkna

$$5\%3 = 2 \qquad 10\%3 = 1 \qquad 21\%3 = 0.$$

Alternativt så skriver man $5 \equiv_3 2$, $10 \equiv_3 1$ eller $21 \equiv_3 0$. ▲

Hjälpsats 3.1.5. *Om $m \in \{0, \ldots, n-1\}$ så gäller det att $m\%n = m$ och för två heltal $a, b \in \mathbb{Z}$ gäller*

$$a \equiv_n b \iff (a-b) \in n\mathbb{Z}.$$

Bevis. Det första påståendet följer direkt från faktumet att $0 \in n\mathbb{Z}$ oavsett n. För det andra påståendet visar vi först implikationen åt vänster. Antag att $(a-b) \in n\mathbb{Z}$, gäller

$$b - a\%n = (a + (b-a)) - a\%n = \underbrace{(a - a\%n)}_{\in n\mathbb{Z}} + \underbrace{(a-b)}_{\in n\mathbb{Z}} \in n\mathbb{Z}.$$

Alltså har $a\%n$ den unika egenskapen som $b\%n$ har, att när man subtraherar den från b får man något i $n\mathbb{Z}$. Därmed måste $a\%n$ och $b\%n$ vara samma, vilket vi skriver som $a \equiv_n b$. Avslutningsvis så ser vi att implikationen åt höger ges av att om vi låter $a\%n = b\%n$ så gäller

$$(a-b) = (a-b)+0 = (a-b)+(b\%n - a\%n) = \underbrace{(a - a\%n)}_{\in n\mathbb{Z}} - \underbrace{(b - b\%n)}_{\in n\mathbb{Z}} \in n\mathbb{Z}. \ \square$$

Definition 3.1.6. För ett heltal $n > 1$, definiera funktionen

$$\begin{aligned} +_n \colon \mathbb{Z} \times \mathbb{Z} &\to \{0, 1, \ldots, n\} \\ (a, b) &\mapsto (a+b)\%n. \end{aligned}$$

Alternativt, skriver vi $a +_n b = (a+b)\%n$. Denna operation kallas för addition modulo n. △

Denna operation gör *inte* heltalen $\mathbb{Z}$ till en grupp. Detta då det inte kan finnas något identitetselement. Varför inte? Vi gör ett kort motsägelseargument: om det fanns ett sådant heltal e skulle vi ha att $(n+1)\%e = n+1$ men $n+1$ ligger ju inte i värdemängden för $+_n$, så det går inte. Hur ska vi då använda denna operation för att skapa en grupp? Jo, om vi begränsar definitionsmängden!

Sats 3.1.7. *För ett heltal $n > 1$ så är paret $(\{0, 1, \ldots, n-1\}, +_n)$ är en abelsk grupp som kallas för $\mathbb{Z}/(n)$.*

Bevis. Det är klart att vi har kommutativitet från definitionen

$$a +_n b = (a+b)\%n = (b+a)\%n = b +_n a.$$

Nu visar ett påstående som kommer hjälpa se varför vi har associativitet. Vi visar att

$$(a +_n b) +_n c = (a+b+c)\%n \iff (a +_n b) + c \equiv_n a + b + c.$$

Det gör vi genom att använda hjälpsats 3.1.5 och visa att differensen är delbar på n. Differensen ses vara delbar på n genom att

$$(a+b+c)-((a+_n b)+c) = (a+b)-(a+b)\%n \in n\mathbb{Z}.$$

Nu kan vi verifiera gruppaxiomen.

(i) (Associativitet) För alla heltal a, b, c

$$(a+_n b)+_n c = (a+b+c)\%n = (b+c+a)\%n = (b+_n c)+_n a = a+_n(b+_n c).$$

(ii) (Identitet) 0 är identiteten eftersom för $a \in \{0, 1, \ldots, n-1\}$ enligt hjälpsats 3.1.5 att

$$a +_n 0 = (a+0)\%n = a\%n = a.$$

(iii) (Inverser) Inverselementet till $a \in \{0, 1, \ldots, n-1\}$ ges av $n-a$ eftersom

$$a +_n (n-a) = n\%n = 0.$$ □

Exempel 3.1.8. Låt oss skriva ned grupptabellerna för

$+_2$	0	1
0	0	1
1	1	0

$+_3$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$+_4$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

▲

Tabell 3.2: Grupptabellen för $\mathbb{Z}/(2), \mathbb{Z}/(3)$ och $\mathbb{Z}/(4)$.

Hjälpsats 3.1.9. *I gruppen $\mathbb{Z}/(n)$ gäller det att*

$$a = \underbrace{1 +_n \ldots +_n 1}_{a \text{ gånger}}.$$

Bevis. Detta följer direkt från hjälpsatsen 3.1.5 eftersom att så % inte gör något när man adderar två positiva tal vars summa är mindre än n. □

Att diskutera generella grupper är ofta lite för brett, även generella abelska grupper. Det kan därmed vara bättre att specialisera oss på grupper av en viss typ.

Definition 3.1.10. En abelsk grupp $(G, *)$ är *ändligt genererad* om det finns en ändlig delmängd $\{g_1, \ldots, g_k\} \subset G$, vars element kallas *generatorer*, sådan att alla element i G kan skrivas på formen

$$g = \underbrace{(g_1 * \ldots * g_1)}_{a_1 \text{ gånger}} * \ldots * \underbrace{(g_k * \ldots * g_k)}_{a_k \text{ gånger}}$$

för några heltal $a_1, \ldots, a_k$. Vi skriver då $G = \langle g_1, \ldots, g_k \rangle$. Notera att om $a_1 < 0$ så ska man istället ta g_1^{-1} gånger sig självt $|a_1|$ gånger istället och detsamma gäller de andra elementen i G. △

Abelska grupper som är ändligt genererade har en särskild form. Följande sats som vi inte bevisar ger oss exakt den form abelska grupper har i allmänhet.

Sats 3.1.11 (Klassifikation av ändligt genererade grupper)**.** *Låt G vara en ändligt genererad abelsk grupp. Då finns heltal $r, d_1, d_2, ...d_k$ så att*

$$G \cong \mathbb{Z}^r \times \mathbb{Z}/(d_1) \times \mathbb{Z}/(d_2) \times ... \times \mathbb{Z}/(d_k).$$

Talet r kallas för rangen av G.

Definition 3.1.12. En abelsk grupp $(G, *)$ är *cyklisk* om den har en enda generator g, alltså att alla element är en 'multipel' av g vilket skrivs som $G = \langle g \rangle$. △

Hjälpsats 3.1.13. *Gruppen $(\mathbb{Z}, +)$ är cyklisk.*

Bevis. Alla heltal kan skrivas på formen $1 + \ldots + 1$ eller $(-1) + \ldots + (-1)$ vilket betyder att 1 genererar $\mathbb{Z}$. □

Hjälpsats 3.1.14. *Gruppen $(\mathbb{Z}/(n), +_n)$ är cyklisk.*

Bevis. Alla heltal $m < n$ kan skrivas på formen

$$m = \underbrace{1 +_n \ldots +_n 1}_{m \text{ gånger}}$$

Så 1 genererar $(\mathbb{Z}/(n), +_n)$. □

3.2 Index

Vi definierar i detta avsnitt sidoklasser i grupper, för att lättare förså förhållandet mellan en grupp och dess delgrupper.

Definition 3.2.1. Låt $(G, +)$ vara en abelsk grupp, och $H \subset G$ vara en delgrupp. Om $g \in G$, så definierar vi *sidoklassen* av g som mängden

$$g + H \coloneqq \{g + h \mid h \in H\}.$$ △

Det visar sig att två olika sidoklasser är helt disjunkta, för om $g_1 + H \cap g_2 + H \neq 0$ finns det $h_1, h_2 \in H$ så att $g_1 + h_1 = g_2 + h_2$, och då är $g_1 - g_2 = h_2 - h_1 \in H$. Då är $g_1 + H = g_2 + (g_1 - g_2) + H = g_2 + H$. För att få den sista likheten använder jag att $h + H = H$ för $h \in H$.

Alla $g \in G$ är i någon sidoklass eftersom $g \in g + H$, då H innehåller identitetselementet. Alltså är sidoklasserna disjunkta mängder som tillsammans täcker G. Vi gör följande definition

Definition 3.2.2. Låt G vara en abelsk grupp, och $H \subset G$ en delgrupp. Antalet olika sidoklasser kallas för H:s index i G, skrivet $(G : H)$. Om det finns oändligt många säger vi att indexet är oändligt.

△

Exempel 3.2.3. Låt $G = \mathbb{Z}$, och $H = 2\mathbb{Z}$, delgruppen av jämna tal. Då är mängder på formen $n + 2\mathbb{Z}$ antingen alla udda tal eller alla jämna tal beroende på om n är ett jämnt eller udda tal. Alltså är $(\mathbb{Z} : 2\mathbb{Z}) = 2$.

Följande sats följer som en konsekvens av egenskaperna hos sidoklasser.

Sats 3.2.4 (Lagranges sats)**.** *Om G är en ändlig abelsk grupp och H är en delgrupp så gäller det att*

$$(\#H) \mid (\#G)$$

det vill säga att att antalet element i H delar antalet element i G.

Bevis. Alla sidoklasser av H i G har storlek exakt $\#H$, eftersom de är på formen $\{g + h \mid h \in H\}$. De är dessutom disjunkta och täcker G. Om n är antalet sidoklasser gäller då att

$$n \cdot (\#H) = G. \qquad \square$$

3.3 Ringar

Definition 3.3.1. En ring är en triplett $(R, +, \cdot)$ där R är en mängd och $+, \cdot$ är två olika funktioner $R \times R \to R$ kallas för addition och multiplikation som uppfyller de så kallade *ringaxiomen*

(i) (Addition) Paret $(R, +)$ är en abelsk grupp. Identiteten kallas 0 och inversen för x kallas för $-x$.

(ii) ($\cdot$ associativ) För alla För alla $a, b, c \in R$ gäller $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

(iii) (Distributivitet) För alla $a, b, c \in R$ gäller

$$a \cdot (b + c) = a \cdot b + a \cdot c$$
$$(a + b) \cdot c = a \cdot c + b \cdot c.$$

Om $\cdot$ dessutom är kommutativ och har ett identitetselement (som kallas 1) säger vi att $(R, +, \cdot)$ är en *kommutativ ring med etta*. $\triangle$

Exempel 3.3.2. Det första exemplet på en ring är $(\mathbb{Z}, +, \cdot)$ där vi har de vanliga operationerna 'plus' och 'gånger'. Vi vet från exempel 2.0.3 att $(\mathbb{Z}, +)$ är en grupp och från grundskolan att $\cdot$ är associativ och distributiv. $\blacktriangle$

Exempel 3.3.3. Nästa exempel på en ringar är

$$(\mathbb{Q}, +, \cdot) \qquad (\mathbb{R}, +, \cdot) \qquad (\mathbb{C}, +, \cdot). \qquad \blacktriangle$$

Ringarna $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ har ännu en egenskap som inte heltalen har, där kan man invertera tal med avseende på multiplikation. Det motiverar följande definition.

Definition 3.3.4. En *kropp* är en kommutativ ring med etta $(K, +, \cdot)$ sådan att K har minst två element och den extra egenskapen att $(K \setminus \{0\}, \cdot)$ är en abelsk grupp med 1 som identitet och där inversen för x kallas x^{-1}. $\triangle$

Exempel 3.3.5. Vi har tre exempel på kroppar,

$$(\mathbb{Q}, +, \cdot) \qquad (\mathbb{R}, +, \cdot) \qquad (\mathbb{C}, +, \cdot).$$

Vi visar det för $\mathbb{Q}$ och samma argument gäller för $\mathbb{R}$ och $\mathbb{C}$. Vi vet från exemplet precis före detta att $(\mathbb{Q}, +, \cdot)$ är en ring. Det vi måste visa är att $(\mathbb{Q} \setminus \{0\}, \cdot)$ har en identitet, har inverser och kommuterar. Vi vet från grundskolan att multiplikation är kommutativ, 1 agerar som identitet för $\cdot$ och att alla nollskilda $x \in \mathbb{Q}$ har egenskapen att $1/x \in \mathbb{Q}$. ▲

3.4 Ringen $\mathbb{Z}/(n)$

Definition 3.4.1. För ett heltal $n > 1$, definiera funktionen

$$\begin{aligned} \cdot_n \colon \mathbb{Z} \times \mathbb{Z} &\to \{0, 1, \ldots, n\} \\ (a, b) &\mapsto (ab)\%n. \end{aligned}$$

Alternativt, skriver vi $a \cdot_n b = (ab)\%n$. Denna operation kallas för multiplikation modulo n. △

Sats 3.4.2. *För varje heltal $n > 1$ så är $(\{0, 1, \ldots, n-1\}, +_n, \cdot_n)$ en kommutativ ring med etta 1 som vi kallar $\mathbb{Z}/(n)$.*

Bevis. Vi vet redan att vi har en grupp med avseende på $+_n$ enligt sats 3.1.7. Det är tydligt att multiplikationen är kommutativ med en etta som är elementet 1 enligt hjälpsats 3.1.5. Vi börjar med att visa att multiplikationen är associativ. För att göra det visar vi först att för alla heltal a, b, c har vi att $(a \cdot_n b) \cdot_n c \equiv_n abc$ genom att skriva

$$\begin{aligned} abc - (a \cdot_n b) \cdot_n c &= \Big(abc - (a \cdot_n b) \cdot c\Big) + \Big((a \cdot_n b) \cdot c - (a \cdot_n b) \cdot_n c\Big) \\ = \underbrace{\Big(ab - (a \cdot_n b)\Big)}_{\in n\mathbb{Z}} \cdot c &+ \underbrace{\Big((a \cdot_n b) \cdot c - ((a \cdot_n b) \cdot c)\%n\Big)}_{\in n\mathbb{Z}} \end{aligned}$$

Vi får då enligt hjälpsats 3.1.5 associativitet för alla heltal a, b, c

$$(a \cdot_n b) \cdot_n c = (abc)\%n = (bca)\%n = (b \cdot_n c) \cdot_n a = a \cdot_n (b \cdot_n c).$$

Till sist så måste vi visa att multiplikationen är distributiv, för detta behöver vi att $(a +_n b) \cdot c \equiv_n (a + b)c$. Detta ser vi eftersom

$$(a + b)c - (a +_n b) \cdot c = \underbrace{((a + b) - (a + b)\%n)}_{\in n\mathbb{Z}} \cdot c.$$

Vi behöver även att $a +_n b = a +_n (b\%n)$ vilket är sant ty

$$a +_n b = (a +_n b) +_n 0 = a +_n (b +_n 0) = a +_n (b\%n).$$

Då får vi alltså att

$$\begin{aligned} (a +_n b) \cdot_n c \equiv_n (a +_n b) \cdot c \equiv_n (a + b)c = ac + bc \\ \equiv_n ac +_n bc = (ac +_n bc) +_n 0 = a \cdot_n c +_n b \cdot_n c. \end{aligned}$$

Alltså är $(a +_n b) \cdot_n c \equiv_n a \cdot_n c +_n b \cdot_n c$ men båda är mellan 0 och $n - 1$ så de måste vara samma. □

Ringstrukturen för en ring R med ändligt många element kan beskrivas med en additions- och multiplikationstabell.

Exempel 3.4.3. Vi kan definiera $\mathbb{Z}/(3\mathbb{Z})$ rakt av genom att specificera addition- och multiplikationstabellerna

$+_3$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$\cdot_3$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Tabell 3.3: Tabellerna för $+$ och $\cdot$ för $\mathbb{Z}/(3)$.

▲

Exempel 3.4.4. Låt oss skriva ned tabellerna för $+$ och $\cdot$ modulo 4.

$+_4$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$\cdot_4$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

Tabell 3.4: Tabellerna för $+$ och $\cdot$ för $\mathbb{Z}/(4)$.

▲

Hjälpsats 3.4.5. *För ett givet primtal p så kan* alla *nollsklijda element i den cykliska gruppen $(\mathbb{Z}/(p), +_p)$ användas som generator (inte bara ettan).*

Bevis. Låt $a \in \mathbb{Z}/(p) \setminus \{0\}$. Studera sekvensen av element

$$0, a, 2 \cdot_n a, 3 \cdot_n a, \ldots, (p-1) \cdot_n a.$$

Detta är en sekvens av längd p. Vi gör ett motsägelsebevis där vi antar att något element i $\mathbb{Z}/(p)$ inte förekommer i denna sekvens. Notera att $\mathbb{Z}/(p)$ har p stycken element så att om något element i $\mathbb{Z}/(p)$ inte förekommer, måste ett annat förekomma två gånger (på grund av Dirichlets lådprincip för den som vill vara noggrann). Vi har då alltså två *olika* heltal $0 \leqslant b < c < p$ sådana att

$$a \cdot_n b = a \cdot_n c.$$

Detta betyder i sin tur enligt lemma 3.1.5 att

$$(c - b) \cdot a \in p\mathbb{Z}.$$

Detta betyder att p måste vara en primtalsfaktor i antingen a eller $c - b$ men detta är en motsägelse eftersom både a och $c - b$ är positiva heltal som är *mindre* än p. □

Sats 3.4.6. *Ringen $\mathbb{Z}/(n)$ är en kropp om och endast om $n > 1$ är ett primtal.*

Bevis. Vi börjar med att visa att om n inte är ett primtal så är $\mathbb{Z}/(n)$ inte en kropp. Ifall $n = m \cdot k$ för positiva heltal $1 < m \leqslant k < n$ så har vi att $m \cdot_n k = n\%n = 0$. Vi gör ett motsägelse bevis genom att anta att det finns ett naturligt tal $l < p$ sådan att

$$k \cdot_n \ell = 1.$$

Detta skulle isåfall innebära att

$$0 = 0 \cdot_n \ell = (m \cdot_n k) \cdot_n \ell = m \cdot_n (k \cdot_n \ell) = m \cdot_n 1 = m.$$

Men detta är en motsägelse eftersom $m \in \{2, \ldots, n-1\}$ och enligt hjälpsatsen 3.1.5 så är $m \neq 0$.

Om $n = p$ är ett primtal så måste vi visa att för varje nollskilt element a måste det existera ett tal a^{-1}. Nu använder vi hjälpsatsen 3.4.5 där vi ser att alla tal kan skrivas på formen $k \cdot_p a$ för något tal $k \in \mathbb{Z}/(p)$. Alltså finns det ett k sådant att $k \cdot_p a = 1$ och alltså är $k = a^{-1}$. Det avslutar beviset. □

Funktioner mellan ringar

Definition 3.4.7. En *ringhomomorfi* är en funktion mellan ringar $\phi\colon R \to S$ sådan att för alla par av element $x, y \in R$

$$\phi(x + y) = \phi(x) + \phi(y) \qquad \phi(x \cdot y) = \phi(x) \cdot \phi(y).$$

Samma definition kan användas för en funktion mellan kroppar när man lägger till antagandet att ettan skickas på ettan. Då kallar vi det för en homomorfi mellan kroppar. △

Det finns två olika operationer som båda kallas + i denna definition, + för R och + för S men det framgår från kontext vilken vi syftar på.

Hjälpsats 3.4.8. *Funktionen*

$$\begin{aligned} \%n\colon \mathbb{Z} &\to \mathbb{Z}/(n) \\ a &\to a\%n \end{aligned}$$

är en ringhomomorfi. Alltså har vi för alla heltal a, b att

$$\begin{aligned} (a + b)\%n &= (a\%n) +_n (b\%n) \\ (a \cdot b)\%n &= (a\%n) \cdot_n (b\%n). \end{aligned}$$

Bevis. Vi kan utnyttja kommutativitet och associativitet för $+_n$ och $\cdot_n$ gäller för alla heltal (inte bara de mellan 0 och $n-1$) enligt beviset av sats 3.1.7 och 3.4.2. Då får vi att

$$\begin{aligned} (a + b)\%n = a +_n b = ((a +_n b) +_n 0) +_n 0 = (a +_n 0) +_n (b +_n 0) \\ =(a\%n) +_n (b\%n) \end{aligned}$$

och på samma sätt

$$\begin{aligned} (a \cdot b)\%n = a \cdot_n b = ((a \cdot_n b) \cdot_n 1) \cdot_n 1 = (a \cdot_n 1) \cdot_n (b \cdot_n 1) \\ =(a\%n) \cdot_n (b\%n). \end{aligned}$$

□

Hjälpsats 3.4.9. *Om en ringhomomorfi ϕ är bijektiv är dess inversfunktion också en ringhomomorfi. Vi kallar en sådan funktion för ringisomorfi och ringarna för* isomorfa.

Bevis. Låt $\phi\colon (G, *, \cdot) \to (H, \star, \cdot)$ vara en bijektiv ringhomomorfi. Då existerar det en inversfunktion ϕ^{-1}. Vi vet från 2.2.9 att inversfunktionen respekterar $+$. Vi ser att

$$\phi(\phi^{-1}(x \cdot y)) = x \cdot y = \phi(\phi^{-1}(x)) \cdot \phi(\phi^{-1}(y)) = \phi(\phi^{-1}(x) \cdot \phi^{-1}(y)).$$

Eftersom ϕ^{-1} är bijektiv kan vi nu applicera den på båda sidor av likheten ovan vilket resulterar i den önskade likheten

$$\phi^{-1}(x \cdot y) = \phi^{-1}(x) \cdot \phi^{-1}(y). \qquad \square$$

3.5 Analoger till mängdoperationer och funktioner

Grupper kan ses som mängder med 'extra struktur' i meningen att vi kan utföra algebra i mängden och grupphomomorfier kan ses som funktioner med extra struktur. På samma sätt är en ring en mängd med 'extra mycket struktur' och ringhomomorfier är funktioner med 'extra mycket struktur' i och med att de respekterar ringoperationerna. Det är naturligt att fråga vad som händer med dessa extra struktur när man utför diverse mängdoperationer på ringar och ringhomomorfier. Förlorar vi de eller bevaras de? I följande tabell listar vi olika operationer som tidigare definierats för ett par av grupper eller en grupphomomorfi och ifall resultatet av operationen blir en ring då man utgår ifrån ett par av ringar/kroppar respektive en ringhomomorfi. Sedan ger vi en mer detaljerad förklaring av hur de olika operationerna bevarar ringstrukturen när det händer och ger övningar i slutet av kapitlet som visar vad som kan gå fel när det inte händer.

Notera att vi inte jämför med alla möjliga mängdoperationer eftersom att alla kroppar är ringar, alla ringar är grupper och alla grupper är ringar. Det betyder att om operationen inte bevarar gruppstruktur kan den omöjligt bevara ringstruktur och så vidare. Därför listar vi bara de operationer som vi vet redan fungerar för grupper.

Mängdoperation	Ringstruktur	Kroppstruktur	Motivering
Delmängd $\subset$	Ibland	Ibland	Definition 3.5.1
Snitt $\cap$	Ja	Ja	Övning 3.8
Cartesisk produkt $\times$	Ja	Nej	Övning 3.15
Funktionsmängd R^X	Ja	Nej	Övning 3.16
Kvotmängd $G/\sim$	Ibland	Nej	Överkurs

Funktionsoperation	Ring	Kropp	Motivering
Värdemängd	Ja	Ja	Övning 3.14
Urbild	Ja	Ja	Övning 3.14
Sammansättning	Ja	Ja	Övning 3.13
Inversfunktion	Ja	Ja	Hjälpsats 3.4.9

Definition 3.5.1. En delring till en ring $(R, +, \cdot)$ är en delmängd Q sådan att

- $(Q, +)$ är en delgrupp till $(R, +)$
- För alla $x, y \in Q$ gäller $x \cdot y \in Q$

En delring är en ring. Om K är en kropp så kallas en delmängd $F \subset K$ för en *delkropp* om

- F är en delring till K
- $(F \setminus \{0\}, \cdot)$ är en delgrupp till $(K \setminus \{0\})$. △

Hjälpsats 3.5.2. *För varje heltal $n > 1$ är mängden $n\mathbb{Z}$ en delring till $\mathbb{Z}$.*

Bevis. Vi vet redan att $n\mathbb{Z}$ är en delgrupp med avseende på addition. Vi behöver bara kontrollera att produkten av två heltal delbara med n också är delbar med n. Det är klart på grund av den associativa egenskapen. Låt $a, b \in \mathbb{Z}$, då gäller

$$(a \cdot n) \cdot (b \cdot n) = n \cdot (a \cdot n \cdot b). \qquad \square$$

Exempel 3.5.3. Bråktalen $(\mathbb{Q}, +, \cdot)$ är en delkropp till de reella talen $(\mathbb{R}, +, \cdot)$som är en delkropp till de komplexa talen $(\mathbb{C}, +, \cdot)$. ▲

Exempel 3.5.4 (Snittet av två delringar)**.** Ett exempel på ett snitt av delringar av $(\mathbb{Z}, +)$ är $2\mathbb{Z}, 3\mathbb{Z} \subset Z$ där

$$(2\mathbb{Z}) \cap (3\mathbb{Z}) = \{n \mid n \text{ delbart med 2 och 3}\} = 6\mathbb{Z}. \qquad ▲$$

Exempel 3.5.5 (Kartesiska produkten av ringar)**.** Ett exempel är den kartesiska produkten av $(\mathbb{Z}, +, \cdot)$ med sig själv. Vi skriver då $\mathbb{Z} \times \mathbb{Z}$ som är mängden av punkter i talplanet med heltalskoordinater och kallar dess operationer $+$ och $\cdot$. Exempelvis så har vi att

$$(5, 7) + (2, -3) = (5 + 2, 7 - 3) = (7, 4) \tag{3.1}$$
$$(5, 7) \cdot (2, -3) = (5 \cdot 2, 7 \cdot (-3)) = (10, -21) \qquad ▲$$

Exempel 3.5.6 (Ringen av funktioner)**.** Mängden av alla funktioner $\mathbb{R} \to \mathbb{R}$ betecknas $\mathbb{R}^{\mathbb{R}}$ och den är en ring eftersom det är väldefinierat hur man adderar och multiplicerar funktioner. Mängden av polynom är en delring till denna ring. Exempelvis så kan vi skriva

$$x + (4x - x^3) = 5x - x^3.$$

Summan av två funktioner blev en ny funktion. Samma sak gäller när man tar produkter. ▲

3.6 Notation

Precis som innan för grupper kommer vi i praktiken inte skriva $(R, +, \cdot)$ för en ring, utan snarare R för en ring. Vi kommer inte heller att skriva $+_n$ för addition modulo n, utan $+$. Vi förtydligar istället att additionen görs i $\mathbb{Z}/n$.

Övningar

Övning 3.1. Beräkna

(i) $1 +_n 1$

(ii) $1 +_n 1 +_n 1$

(iii) $2 +_n 5$

(iv) $2 \cdot_n 5$

(v) $0 \cdot_n 4$

(vi) $2 \cdot_n (5 +_n 7)$

(vii) $5 \cdot_n (3 +_n 9)$

(viii) $-_n 2$ (den additiva inversen för 2 i $\mathbb{Z}/(n)$).

(ix) $8 -_n 2$

för $n = 3$ och $n = 11$.

Övning 3.2. Beräkna den multiplikativa inversen för alla nollskilda element i $\mathbb{Z}/(3), \mathbb{Z}/(5), \mathbb{Z}/(5)$.

Övning 3.3. Skriv ned grupptabellen för $(\mathbb{Z}/(5\mathbb{Z}), +)$.

Övning 3.4. Skriv ned grupptabellen för $(\mathbb{Z}/(2\mathbb{Z}) \times (\mathbb{Z}/(3\mathbb{Z})$ med addition modulo n som operation.

Ledning: För att svaret ska stämma överens med facit sortera elementen i ordningen $(0,0), (0,1), (0,2), (1,0), (1,1), (1,2)$.

Övning 3.5. Verifiera att mängden av polynom i en variabel med koefficienter i $\mathbb{R}$ är en ring.

Övning 3.6. Verifiera att för en kropp K så är $0 \neq 1$.

Övning 3.7. Visa att produkten av två nollskilda element i en kropp K är nollskild.

Övning 3.8. Visa att snittet av två delringar är en delring och att snittet av två delkroppar är en delkropp.

Övning 3.9. Visa att i en ring gäller det för alla element x att $x \cdot 0 = 0$.

Övning 3.10. Låt -1 beteckna den additiva inversen för 1 i en kropp. Visa att för alla element i kroppen så erhålls den additiva inversen genom multiplikation med -1, alltså:

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

Ledtråd: Använd övningen 3.9

Övning 3.11. Låt -1 beteckna den additiva inversen för 1 i en kropp. Visa att

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

Övning 3.12. Visa att identitsfunktionen $x \mapsto x$ alltid är en ringhomomorfi från en ring R till sig själv.

Övning 3.13. Visa att sammansättningen av ringhomomorfier ger en ny ringhomomorfi.

Övning 3.14. Visa bilden av en ringhomomorfi $\phi : (R, +, \cdot) \to (S, +, \cdot)$ är en delring. Visa även att urbilden av en delring av S är en delring av R. Slutligen visa att samma påståenden gäller när R och S är kroppar så länge ettan $1_S \in S$ skickas på ettan $1_R \in R$.

Övning 3.15. Den *kartesiska* produkten av två ringar R, S är ringen vars ringstruktur ges av komponentvis addition och multiplikation.

$$\begin{aligned} +\,&:(R \times S) \times (R \times S) \to (R \times S) \\ &((a, x), (b, y)) \mapsto (a + b, x + y) \end{aligned} \qquad \begin{aligned} \cdot\,&: (R \times S) \times (R \times S) \to (R \times S) \\ &((a, x), (b, y)) \mapsto (a \cdot b, x \cdot y). \end{aligned}$$

Visa att den kartesiska produkten

(i) av två ringar är en ring.

(ii) den kartesiska produkten av två kroppar inte är en kropp.

Övning 3.16. Givet en mängd X och en ring R kan vi definiera mängden av funktioner R^X med ringoperationerna som fås från genom elementivs addition respektive multiplikation

$$\begin{aligned} +\,&:R^X \times R^X \to R^X \\ &(f + g)(x) = f(x) + g(x) \end{aligned} \qquad \begin{aligned} \cdot\,&:R^X \times R^X \to R^X \\ &(f \cdot g)(x) = f(x) \cdot g(x). \end{aligned}$$

Visa att för en mängd X med åtminstone 2 element så gäller

(i) R^X är en ring

(ii) Om K är en kropp så är inte K^X en kropp.

Övning 3.17. Visa att $(\mathbb{Z}/3\mathbb{Z}, +)$ och $(\mathbb{Z}/2\mathbb{Z}, +)$ är isomorfa med delgrupper till S_3. (denna uppgift har alltså inget med ringar att göra).

Övning 3.18. Visa att $(B, \text{xor}, \&)$ är en kommutativ ring med etta. Här är (B, xor) gruppen av Booleans från övning 2.3 och där vi nu definierar

$$\begin{aligned} \&\,&:B \times B \to B \\ &(p, q) \mapsto \text{både } p \text{ och } q \text{ är 'sant'}. \end{aligned}$$

Övning 3.19. Formulera med hjälp av xor och & ett påstående för triplett (p, q, s) i B (från exempel 3.18) som är sant om och endast om exakt ett av p, q, s är sant.

Övning 3.20 ($\star$)**.** Givet en mängd A med åtminstone 2 element, visa att tripletten $(2^A, \Delta, \cap)$ är en kommutativ ring med etta men inte en kropp. Kom ihåg dessa från avsnittet *mängdoperationer* i kapitel 1.

Ge också en konkret tolkning av:

- Identitetselementen 0 och 1
- Inversen med avseende på Δ för en mängd X
- Vad $1 \Delta X$ är för något.

Gör detta antingen genom att verifiera alla ringaxiomen eller visa att funktionen

$$\begin{aligned} \text{supp} : (\mathbb{Z}/2\mathbb{Z})^A &\to (2^A, \Delta, \cap) \\ f &\mapsto \{x \mid f(x) = 1\} \end{aligned}$$

är en bijektiv funktion som uppfyller

$$\text{supp}(f+g) = \text{supp}(f) \,\Delta\, \text{supp}(g) \qquad \text{supp}(f \cdot g) = \text{supp}(f) \cap \text{supp}(g).$$

Övning 3.21 ($\star$)**.** Givet en kropp K, definiera mängden av *affina funktioner* som delmängden till mängden av polynomfunktioner $K[x]$ som består av funktioner på formen

$$A = \{f \mid f(x) = a \cdot x + b \text{ för ett par av element } a \in K \setminus \{0\}, b \in K\}$$

- Bevisa att $(A, \circ)$ är en grupp, där $\circ$ är sammansättning av funktioner.
- Om $K = \mathbb{Z}/3\mathbb{Z}$ visa att A är isomorf med permutationsgruppen S_3. *Notera att elementen i A är bijektioner från {0,1,2} till sig själv*

Övning 3.22 ($\star$)**.** Bevisa att ett heltal är delbart med tre om och endast om dess talsumma är delbar på tre. Exempelvis så är 621 delbart på tre eftersom $6 + 2 + 1 = 9$ är delbart på tre.

4 Ringar fortsättning

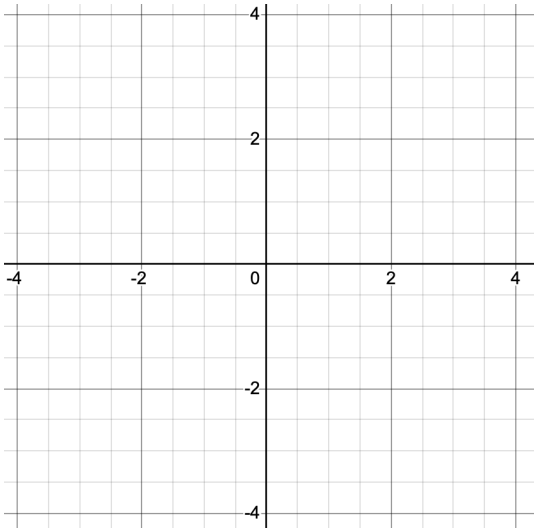
Bland de viktigaste exemplen på ringar har vi ringen av polynom i flera variabler och med koefficienter i reella talen. Vi kommer generalisera detta till att betrakta polynom vars koefficienter ligger i en kropp K istället. Vi kommer att behöva dessa för att definiera elliptiska kurvor inte bara för de reella talen.

4.1 Tallinjen och linjer i talplanet

För de reella talen är polynomringar mängder vars element är funktioner med målmängd $\mathbb{R}$ och där definitionsmängden är tallinjen, talplanet eller mer generellt $\mathbb{R}^n$. Konceptet av en linje och ett plan är inte unikt för de reella talen. Samma koncept går att generalisera för en godtycklig kropp. Vi kommer att börja definiera vad en linje är och sedan jobba med att försöka förstå den.

Definition 4.1.1. Om K är en kropp så kan denna även kallas för en *tallinje*. I denna kontext brukar K^2 kallas för ett *talplan*. Elementen i K kallas *tal* och elementen i K^2 kallas *vektorer*. △

Exempel 4.1.2. Reella talplanet $\mathbb{R}^2$ känner vi igen.

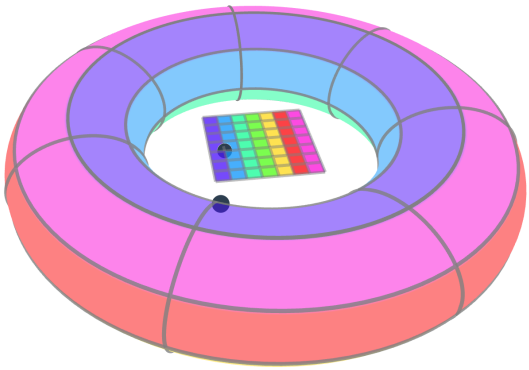


Figur 4.1: Det reella talplanet.

▲

Exempel 4.1.3 (Det modulära talplanet)**.** Låt oss visualiserar talplanet för $\mathbb{Z}/(7)$. Börja med att bläddra tillbaka till figuren 3.1 för att komma ihåg hur vi visualiserade den modulära tallinjen, alltså $\mathbb{Z}/(7)$ självt. Antingen så var det 7 punkter på rad av längd 6 (eftersom vi börjar räkna från 0). Vi kunde också positionera elementen i en cirkel/klocka. Informellt så kan man tänka sig att vi i cirkeln 'limmat' ihop ändarna på intervallet 0 till 6 fast med lite mellanrum mellan ändpunkterna. Talplanet är mängden av ordnade par av punkter i $\mathbb{Z}/(7)$. Vi visualisera detta som en kvadrat med sidlängden 6 och som innehåller $7^2 = 49$ heltalspunkter. Exempelvis ligger punkten $(1, 3)$ där. När vi istället ser den modulära tallinjen som en cirkel/klocka motsvarar detta

för talplanet att man måste limma ihop över med undersidan samt vänster med högersidan. Då får man en sorts doughnut, eller *torus* som matematiker brukar säga. Båda versionerna visualiseras i figuren nedan med en svart punkt som motsvarar elementet $(1, 3)$. Det finns även en länk som låter dig välja andra modulo än 7 och visualisera en punkt som rör sig runt i det modulära talplanet (detta görs genom att variera parametrarna a, b som motsvarar punkten x respektive y koordinat).



Figur 4.2: (Länk till animation)
Alternativt: https://www.math3d.org/k1KUwklvy

. ▲

Definition 4.1.4. En delmängd $L \subset K^2$ kallas för en *linje* om och endast om det finns två tal $k, m \in K$ sådan att

$$L = \{(x, y) \mid y = k \cdot x + m\}$$

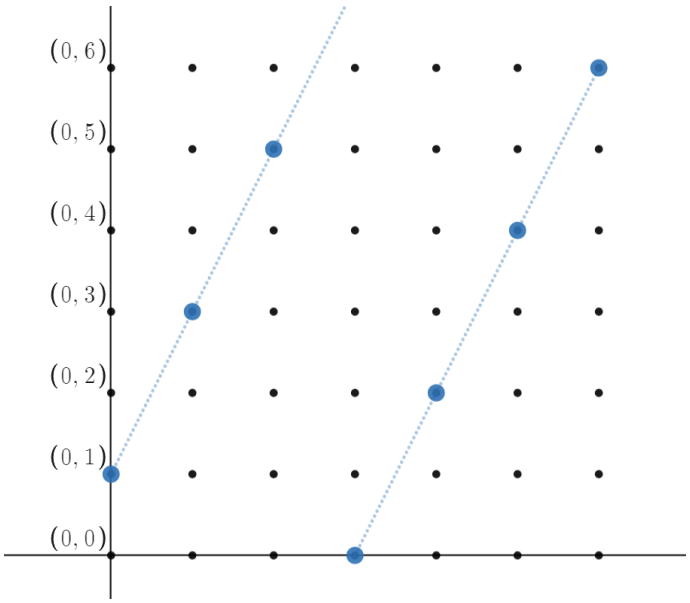
eller

$$L = \{(x, y) \mid x = m\}. \qquad \triangle$$

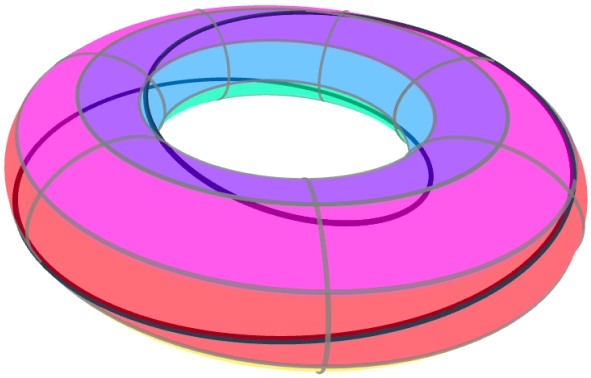
Exempel 4.1.5. Låt oss betrakta linjen $y = 2x + 1$ fast modulo 7. Detta är alltså mängden

$$L = \{(0, 1), (1, 3), (2, 5), (3, 0), (4, 2), (5, 4), (6, 6)\}.$$

Den kan visualiseras på två olika sätt som vi demonstrerar i följande figurer.



Figur 4.3: Linjen $y = 2x + 1$ modulo 7 i rutnätsformat.
(Länk till animation)



Figur 4.4: Linjen $y = 2x + 1$ modulo 7 i torusformat.
(Länk till animation)

I figurerna så ritar vi heldragna linjer, i verkligheten ska vi bara betrakta punkterna med heltalskoordinater såsom markerat i rutnätsvarianten. Lek gärna runt med olika val av räta linjer med olika val av modulo i länkarna i figurerna. Notera att i rutnätsversionen så tillåter vi bara primtal p sådana att $p\%4 = 3$ vilket kan verka konstigt, men vi kommer senare se att dessa primtal kan vara lite enklare att jobba med.

▲

4.2 Polynomringar

Givet en kropp K så kan vi definiera en ny ring $K[x]$ eller $K[x, y]$. Vi kommer behöva denna ring för att definiera elliptiska kurvor över ändliga kroppar.

Definition 4.2.1. Låt K vara en kropp. Vi definierar $K[x]$, polynomringen över K i en variabel, som uttryck på formen

$$\sum_{i=0}^{d} c_i x^i = c_0 + c_1 x + c_2 x_2 + \ldots + c_d x^d$$

för $d \in \mathbb{N}$ och $c_i \in K$. Vi kan definiera addition och multiplikation av uttryck på det vanliga sättet. Vi kallar sådana uttryck för *polynom*. Elementen c_i kallas för *koefficienter* och K kallas för *koefficientkroppen*. Talet d kallas för graden av polynomet. △

Med varje polynom $f(x) = c_0 + c_1 x + ... + c_d x^d$ följer en evalueringsfunktion $f\colon K \to K$ som definieras som

$$f(k) = c_0 + c_1 k + ... + c_d k^d.$$

Exempel 4.2.2. Några exempel på polynom i $\mathbb{R}[x]$ är

$$\begin{aligned} f(x) &= \pi + \sqrt{2}x + 7x^5 \\ g(x) &= 1 - 3x + \frac{4}{9}x^2 \\ h(x) &= \sqrt{3}x + x^3. \end{aligned}$$

Här skriver vi alltså $1 - 3x + \frac{4}{9}x^2$ istället för $1 + (-3)x + \frac{4}{9}x^2$ för $g(x)$. Vi evaluerar polynomen i $x = 2$ och får

$$\begin{aligned} f(2) &= \pi + 2\sqrt{2} + 224 \\ g(2) &= 1 - 6 + \frac{16}{9} \\ h(2) &= 2\sqrt{3} + 8. \end{aligned}$$

▲

Exempel 4.2.3. Polynomet $1 + x + x^2 + x^3$ är ett element i $\mathbb{R}[x]$. ▲

Exempel 4.2.4. Några exempel på element i $(\mathbb{Z}/(5))[x]$ är

$$\begin{aligned} f(x) &= 1 + 3x + 4x^2 \\ g(x) &= 2 + 4x^2 + 3x^5 \\ h(x) &= -x + x^5 \end{aligned}$$

där det är underförstått att alla multiplikationer och additioner görs modulo 5. Notera att eftersom $-1 = 4$ i $\mathbb{Z}/(5)$ (det vill säga att 4 är den additiva inversen till 1) så är det rimligt att skriva

$$g(x) = 2 - x^2 + 3x^5.$$

Vi evaluerar f, g, h i $x = 2$

$$\begin{aligned} f(2) &= (1 + 3 \cdot 2 + 4 \cdot 2^2)\%5 = 23\%5 = 3 \\ g(2) &= (2 - 2^2 + 3 \cdot 32)\%5 = 94\%5 = 4 \\ h(2) &= (-2 + 32)\%5 = 0. \end{aligned}$$

▲

Sats 4.2.5 (Fermats lilla sats)**.** *Givet ett primtal p så gäller det att för alla element $x \in \mathbb{Z}/(p)$.*

$$x^p - x = 0.$$

Bevis. Om $x = 0$ så följer det direkt att $0^{p-1} = 0$. Annars så bevisar vi istället att för varje nollskilt element $a \in \mathbb{Z}/(p)$ så gäller $a^{p-1} = 1$ vilket tydligt implicerar påståendet i satsen. När $a \neq 0$ så är a en generator för $(\mathbb{Z}/(p), +)$ enligt hjälpsats 3.4.5. Alltså har vi att följande mängder är samma (kom ihåg att en mängd inte håller koll på ordningen av sina element),

$$\{1, 2, \ldots, p-1\} = \{a, 2 \cdot a \ldots, (p-1) \cdot a\}.$$

Kom ihåg definitionen av fakultet, $(p-1)! = (p-1)(p-2) \cdot \ldots \cdot 3 \cdot 2 \cdot 1$. Vi har från det vi precis fastslagit att i modulo p så får vi istället

$$\begin{aligned}(p-1)!\%p &= (p-1)(p-2) \cdot \ldots \cdot 3 \cdot 2 \cdot 1 = \\ &= a \cdot (2 \cdot a) \cdot \ldots \cdot ((p-1) \cdot a) = (p-1)! \cdot a^{p-1}.\end{aligned}$$

För att avsluta beviset inser vi att $(p-1)!$ inte är delbart med p och alltså gäller det att $(p-1)!\%p \neq 0$ har en multiplikativ invers eftersom $\mathbb{Z}/(p)$ är en kropp. Vi kan då multiplicera med inversen på båda sidor om likheten

$$(p-1)!\%n = ((p-1)!\%n) \cdot a^{p-1} \quad \overset{\ldots\cdot((p-1)!\%n)^{-1}}{\Longrightarrow} \quad 1 = (a^{p-1}\%n).$$

□

Definition 4.2.6. Vi säger att ett polynom h *delar* ett polynom f, vilket skrivs som $h|f$ om det finns ett polynom g sådant att

$$f = g \cdot h.$$

△

Hjälpsats 4.2.7. *Om $f \in K[x]$ och $f(a) = 0$ så kallar vi a för en* rot *till f. Varje rot a till f har egenskapen att*

$$f = (x - a) \cdot g$$

där graden av g är lägre än graden av f.

Bevis. Vi börjar med att betrakta fallet $a = 0$. Anta att $f(0) = 0$. Det innebär att $c_0 = 0$. Det betyder att

$$f(x) = 0 + c_1 x + c_2 x_2 + \ldots + c_d x^d = x \cdot (c_1 + c_2 x + \ldots + c_d x^{-1}).$$

Där syns det att $g = (c_1 + c_2 x + \ldots + c_d x^{-1})$ har lägre grad. Så fallet $a = 0$ är löst. Betrakta nu ringhomomorfin

$$\begin{aligned}\tau_a \colon K[x] &\to K[x] \\ f(x) &\mapsto f(x - a).\end{aligned}$$

Notera att homomorfin är en isomorfi eftersom den har en invers, nämligen $\tau_a^{-1} = \tau_{(-a)}$. Detta betyder alltså att alla polynom också går att skriva på formen

$$f(x) = \alpha_0 + \alpha_1(x-a) + \alpha_2(x-a)^2 + \ldots + \alpha_d(x-a)^d.$$

På samma sätt är det nu klart att $f(a) = 0$ om och endast om $\alpha_0 = 0$ vilket betyder att

$$f(x) = (x-a)(\alpha_1 + \alpha_2(x-a) + \ldots + \alpha_d(x-a)^{d-1}). \qquad \square$$

Hjälpsats 4.2.8. *Ett polynom $f \in K[x]$ av grad d som har fler än d rötter är konstant 0.*

Bevis. Vi gör ett motsägelsebevis. Säg att

$$f(x) = c_0 + c_1x + ... + c_dx^d$$

har $D > d$ nollställen. Låt a vara ett nollställe. Enligt 4.2.7 kan vi skriva $f(x) = (x-a)g(x)$, där graden av $g(x)$ är som mest $d-1$. Låt $b \neq a$ vara ett nollställe till f. Då är

$$0 = f(b) = (b-a)g(b).$$

Då $a \neq b$ så är $b - a \neq 0$, så vi kan multiplicera båda sidor med $(b-a)^{-1}$ och få $g(b) = 0$. Alltså har g ett nollställe i b. Det innebär att g har åtminstone $D - 1 > d - 1$ nollställen. Vi kan fortsätta med att faktorisera g, tills vi till slut kommer till ett polynom av grad 0 (alltså en konstant) som måste ha åtminstone ett nollställe. Det enda sådana polynomet är 0, men då är det ursprungliga polynomet 0. $\square$

Definition 4.2.9. Låt K vara en kropp. Definera $K[x,y]$, polynomringen över K i två variabler som uttryck på formen

$$\sum_{i,j=0}^{d_1,d_2} c_{i,j}x^iy^j = c_{00} + c_{1,0}x + c_{0,1}y + c_{1,1}xy + \ldots + c_{d_1,d_2}x^{d_1}y^{d_2}$$

för $d_1, d_2 \in \mathbb{N}$. Vi definierar addition och multiplikation på det vanliga sättet. Vi kallar dess element för *polynom*. Elementen $c_{i,j} \in K$ kallas för *koefficienter* och K kallas för *koefficientkroppen*. $\triangle$

Vi definierar informellt graden av ett polynom i flera variabler som högsta möjliga summan av exponenterna av x och y i en term. Exempelvis har polynomet $x^2y^3 + y^4$ grad 5.

Exempel 4.2.10. Ett par exempel på element i $\mathbb{R}[x,y]$ är

$$\begin{aligned} f(x) &= 1 + x - y^2 + x^3 \\ g(x) &= 1 - x^2 - y^2 \\ h(x) &= 2 + x/3 + x^2y^3 \end{aligned}$$

Vi evaluerar f, g, h i $(1, 2)$ alltså $x = 1, y = 2$

$$\begin{aligned} f(1,2) =& 1 + 1 - 2^2 + 1^3 = 3 - 4 = -1. \\ g(1,2) =& 1 - 1^2 - 2^2 = -4. \\ h(1,2) =& 2 + 1/3 + 1^2 \cdot 2^3 = 10 + 1/3 = 31/3. \end{aligned}$$ ▲

Exempel 4.2.11. Ett par exempel på element i $\mathbb{Z}/(5)[x, y]$ är

$$\begin{aligned} f(x) =& 1 + x - y^2 + x^3 \\ g(x) =& 1 - x^2 - y^2 \\ h(x) =& 2 + \frac{1}{3}x + x^2y^3. \end{aligned}$$

Notera att eftersom $-1 = 4$ (det vill säga att 4 är den additiva inversen till 1) så är det rimligt att skriva

$$g(x) = 1 + 4x^2 + 4y^2.$$

På samma sätt vet vi att $3 \cdot 2 = 1$ så vi kan skriva om

$$h(x) = 2 + 2x + x^2y^3.$$

Vi evaluerar f, g, h i $(1, 2)$, alltså $x = 1, y = 2$ och får

$$\begin{aligned} f(1,2) =& (1 + 1 - 2^2 + 1^3)\%5 = (-1)\%5 = 4, \\ g(1,2) =& (1 - 1^2 - 2^2)\%5 = (-4)\%5 = 1, \\ h(1,2) =& (2 + 2 \cdot 1 + 1^2 \cdot 2^3)\%5 = 12\%5 = 2. \end{aligned}$$ ▲

Hjälpsats 4.2.12. *Givet två punkter $P = (x_1, y_1)$ och $Q = (x_2, y_2)$ så finns det en unik linje mellan dem. Om $x_1 = x_2$ så är det linjen $x - x_1 = 0$ och annars är det*

$$y = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1) + y_1.$$

Bevis. Om $x_1 = x_2$ så är det klart att det måste vara den linjen eftersom en linje som inte är lodrät inte har något par av punkter med samma x-värde (eftersom om man vet x så vet man y). I det andra fallet, låt oss betrakta en linje $y = kx + m$ som innehåller P och Q. Då kan vi se att

$$\frac{y_2 - y_1}{x_2 - x_1} = \frac{kx_2 + m - (kx_1 + m)}{x_2 - x_1} = k\frac{x_2 - x_1}{x_2 - x_1} = k.$$

Så det finns ett unikt alternativ för valet av k. På samma sätt ser man att m-värdet måste vara $y_1 = kx_1 + m \implies m = y_1 - kx_1$. □

4.3 Algebraisk geometri

Vi avslutar kapitel 4 med att definiera en kurva i planet, alternativt en *planär* kurva vilket kommer att innefatta de elliptiska kurvorna i nästa kapitel och som är temat för denna kurs.

Definition 4.3.1. En *planär kurva* är mängden av lösningar i K^2 till en ekvation på formen

$$f(x, y) = 0$$

där $f \in K[x, y] \setminus \{0\}$ för en kropp K. △

Exempel 4.3.2. Alla linjer är planära kurvor. Detta eftersom de är lösningsmängden till ekvationer på formen

$$y - kx - m = 0 \quad \text{eller} \quad x - m = 0.$$ ▲

Exempel 4.3.3. Betrakta lösningsmängden till

$$x^2 + 2xy + y^2 - 3x - 3y + 2 = 0.$$

Vi kan faktorisera uttrycket ovan, och få den ekvivalenta ekvationen

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

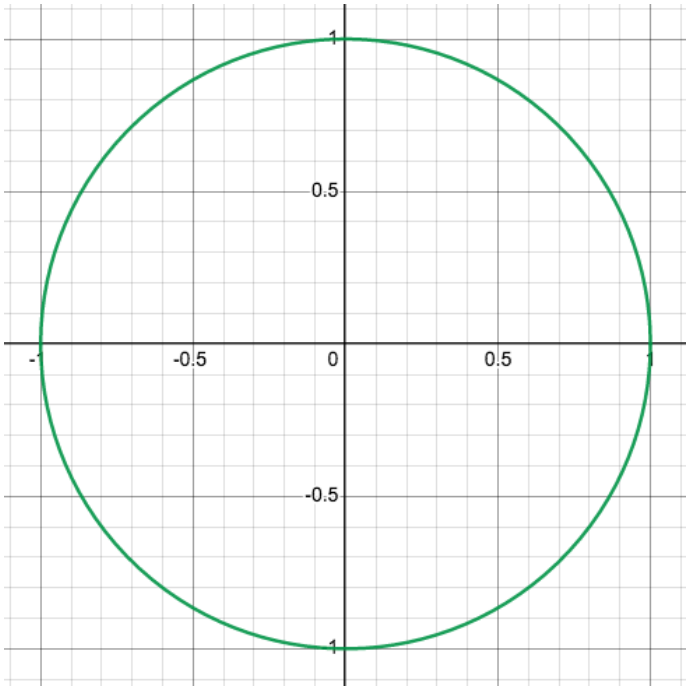
Men uttrycket ovan är noll om och endast om

$$x + y - 1 = 0 \quad \text{eller} \quad x + y - 2 = 0.$$

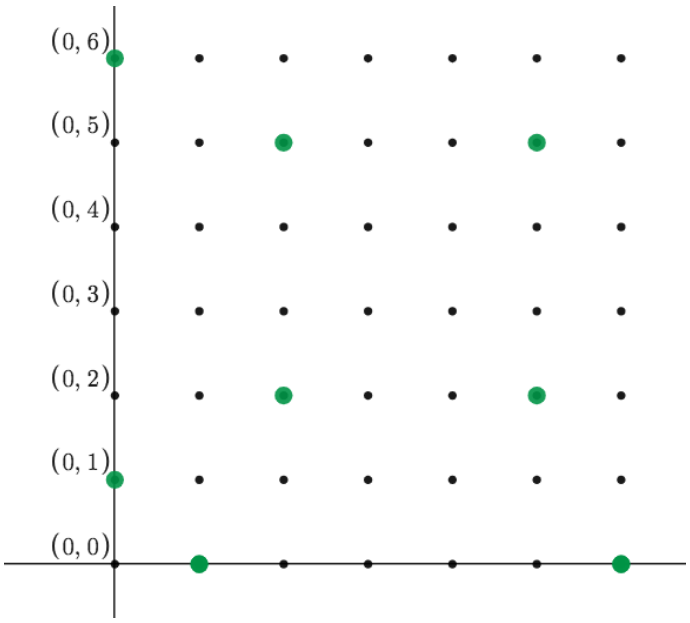
Dessa ekvationer definierar båda linjer, så man kan nu se att denna mängd är unionen av två linjer! ▲

Exempel 4.3.4. Betrakta lösningsmängden till

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



Figur 4.5: Kurvan som motsvarar $x^2 + y^2 - 1$ för $K = \mathbb{R}$.

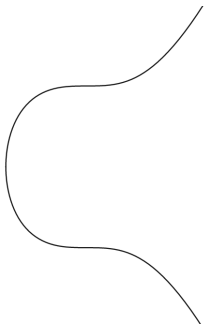


Figur 4.6: Kurvan som motsvarar $x^2 + y^2 - 1$ för $K = \mathbb{Z}/(7)$.

▲

Exempel 4.3.5. Betrakta lösningsmängden till

$$x^3 + 1 - y^2 = 0.$$



Figur 4.7: Kurvan som motsvarar $y^2 = x^3 + 1$ för $K = \mathbb{R}$.

▲

Definition 4.3.6. Givet ett polynom $f \in K[x, y]$

$$f(x, y) = \sum_{i,j=0}^{d_1,d_2} c_{i,j} x^i y^j = c_{0,0} + c_{1,0}x + c_{0,1}y + c_{1,1}xy + \ldots + c_{d_1,d_2} x^{d_1} y^{d_2}.$$

definierar vi dess *partiella derivator* som

$$\frac{\partial f}{\partial x} = \sum_{i,j=0}^{d_1,d_2} i \cdot c_{i,j} x^{i-1} y^j = c_{1,0} + c_{1,1}y + \ldots + d_x c_{d_x,d_y} x^{d_1-1} y^{d_2}.$$

$$\frac{\partial f}{\partial y} = \sum_{i,j=0}^{d_1,d_2} i \cdot j \cdot c_{i,j} x^i y^{j-1} = c_{0,1} + c_{1,1}x + \ldots + d_y c_{d_1,d_2} x^{d_1} y^{d_2-1}$$

△

Exempel 4.3.7. Låt oss derivera polynomet

$$f(x,y) = y^2 - x^3 - 1 = (x^0y^2 + (-1)x^3y^0 - x^0y^0).$$

$$\frac{\partial f}{\partial x}(x,y) = 0 - 3x^2 - 0 = -3x^2.$$
$$\frac{\partial f}{\partial y}(x,y) = 2y - 0 - 0 = 2y.$$

▲

Vi definierar nu tangentlinjen till en kurva. Att förklara vart definitionen kommer från är överkurs och vi nöjer oss med att observera att den stämmer överens med vår intuition i ett exempel.

Definition 4.3.8. *Tangentlinjen* till en planär kurva

$$Y = \{(x,y) \mid f(x,y) = 0\}$$

i punkten $(a,b) \in Y$ ges av ekvationen

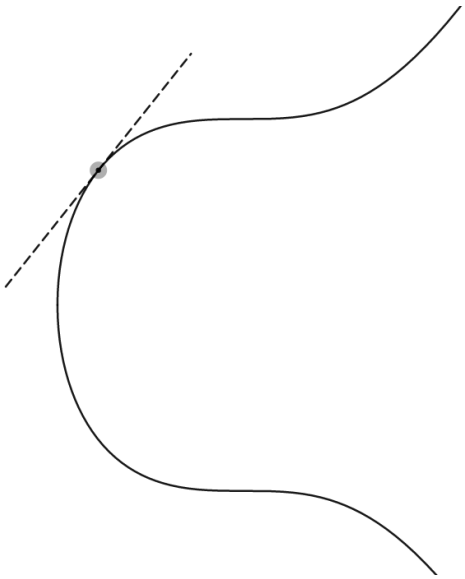
$$\left(\frac{\partial f}{\partial x}(a,b)\right) \cdot (x-a) + \left(\frac{\partial f}{\partial y}(a,b)\right)(y-b) = 0$$

Om båda partiella derivatorna är 0 i någon punkt $(a,b) \in Y$ så existerar inte tangentlinjen där. En punkt där tangentlinjen inte finns kallas för *singularitet*. En kurva med en singularitet säger vi är *singulär*. Om tangentlinjen alltid finns så säger vi att kurvan är *glatt*. △

Exempel 4.3.9. Kurvan $y^2 - x^3 - 1 = 0$ är glatt eftersom en singulär punkt är samma sak som lösningen till ekvationssystemet $f(x,y) = \frac{\partial f}{\partial x}(x,y) = \frac{\partial f}{\partial y}(x,y) = 0$. Alltså,

$$f(x,y) = y^2 - x^3 - 1 = 0$$
$$\frac{\partial f}{\partial x}(x,y) = -3x^2 = 0$$
$$\frac{\partial f}{\partial y}(x,y) = 2y = 0.$$

Genom att titta på de två partiella derivatorna ser vi att den enda möjliga lösningen är origo $x = 0, y = 0$. Denna är dock inte en punkt på kurvan, så kurvan är glatt!

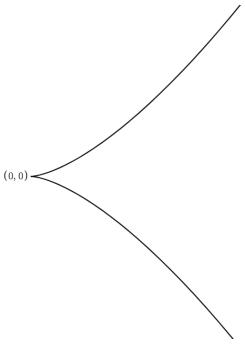


Figur 4.8: Kurvan som motsvarar $y^2 - x^3 - 1 = 0$ för $K = \mathbb{R}$ med en av dess tangentlinjer.

▲

Lek gärna runt med tangentlinjen i olika punkter i animationen i följande länk.(Länk till animation). Det är svårare att förstå intuitionen för denna kurva (den som ges av samma ekvation) i det modulära talplanet. En tangentlinje skär alltid sin kurva i en punkt med multiplicitet åtminstone 2 (se definitionen nedan).

Exempel 4.3.10. Kurvan $y^2 - x^3 = 0$ är singulär eftersom det inte finns någon tangentlinje i punkten $(0, 0)$.



Figur 4.9: Kurvan som motsvarar $y^2 - x^3 = 0$ för $K = \mathbb{R}$.

▲

Definition 4.3.11 (Multiplicitet)**.** Låt C vara en kurva given av ekvationen $f(x, y) = 0$. Låt även L vara en linje som ges av ekvationen

$$y = \lambda x + \nu.$$

Om $f(x, y)$ har grad n kommer $f(x, \lambda x + \nu)$ vara ett polynom i x av grad n. Säg att L skär C i punkten (x_0, y_0). Då säger vi att skärningen är av multiplicitet k om polynomet $f(x, \lambda x + \nu)$ har roten x_0 med multiplicitet k. Eftersom

polynomet har grad k kommer summan av alla skärningspunkters multiplicitet vara som mest graden av f.

△

Exempel 4.3.12. Betrakta kurvan $f(x, y) = x^2 + y^2 - 1 = 0$, alltså en cirkel med radie 1 och center i origo. Linjen $y = 0$ skär kurvan två gånger, då $x = 1$ och då $x = -1$. Polynomet

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

har 2 rötter, båda med multiplicitet ett. Allså har båda skärningspunkter multiplicitet 1.

Linjen $y = 1$ skär kurvan en gång, då $x = 0$. Polynomet

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

har roten $x = 0$ med multiplicitet 2, så linjen skär kurvan en gång med multiplicitet 2.

▲

Övningar

Övning 4.1. Skriv ned följande polynom 'utan minustecken och utan bråkstreck'. Till exempel så skrivs $1 -_3 x +_3 \frac{1}{2}x^2$ skrivs som om $1 +_3 2x + 2x^2$ eftersom $2 +_3 1 = 0$ och $2 \cdot_3 2 = 1$.

(i) $f(x, y) = 1 +_p x -_p x^2$

(ii) $g(x, y) = 1 -_p \frac{1}{2}xy$

för $p = 5$ och $p = 7$ och evaluera dem i $x = 3, y = 2$.

Övning 4.2. Beräkna de partiella derivatorna för polynomen

(i) $1 + 2x + y + 3x^2$

(ii) $2x - 3xy + y^2 + x^2$

Övning 4.3. Skriv ned ekvationen för den unika linjen som genom punkterna $(1, 2)$ och $(5, 7)$ (oavsett kroppen K, det är okej att använda skriva $1/4$ även i modulo p).

Övning 4.4. Bevisa att en linje i talplanet modulo p innehåller exakt p element.

Övning 4.5. Rita ut linjen $y = x/2 + 5$ för kropparna $\mathbb{R}$ och $\mathbb{Z}/(7)$.

Övning 4.6. Rita ut linjen $y = 3x - 4$ för kropparna $\mathbb{R}$ och $\mathbb{Z}/(7)$.

Övning 4.7. Rita ut linjen $\{(x, y) \mid (x, y) = t * (3, 1) + (0, 1)$ för något $t \in K\}$ för kropparna $\mathbb{R}$ och $\mathbb{Z}/(7)$.

Övning 4.8. Skriv ned ekvationer för tangentlinjerna till kurvan $x^2+y^2-25 = 0$ i punkten $(0,5)$ och $(3,4)$.

Övning 4.9. Är kurvan $x^2 + y^2 - 25 = 0$ slät (när $K = \mathbb{R}$)?

Övning 4.10. Bevisa att

$$\begin{aligned}\phi :&\mathbb{Z}/(2) \times \mathbb{Z}/(2)\\ &(a,b) \mapsto (a+b,b)\end{aligned}$$

är en gruppisomorfi. Dra slutsatsen att vilket par av nollskilda element som helst genererar gruppen $\mathbb{Z}/(2) \times \mathbb{Z}/(2)$ med $+$.

Övning 4.11. Är kurvan $y^2 - x^2 - x^3 = 0$ slät (när $K = \mathbb{R}$)?

Övning 4.12 ($\star$). Bevisa att tangetlinjen till en planär kurva i en punkt $(a,b) \in C$ skär kurvan skär den i en dubbelpunkt genom att följa dessa steg

(i) Givet $(a,b) \in K^2$. Bevisa att alla polynom kan skrivas på formen

$$f(x,y) = \sum_{i,j=0}^{d_x,d_y} \alpha_{i,j}(x-a)^i(y-b)^j$$

för några tal $\alpha_{i,j} \in K$.

(ii) Bevisa att $\alpha_{0,0} = f(a,b)$

(iii) Acceptera att $\alpha_{1,0} = \frac{\partial f}{\partial x}(a,b)$ och $\alpha_{0,1} = \frac{\partial f}{\partial y}(a,b)$. (Bevis är överkurs)

(iv) Visa att $f(a+t\frac{\partial f}{\partial y}(a,b), b-t\frac{\partial f}{\partial x}(a,b))$ har en dubbelrot i $t=0$, alltså att den är delbar med t^2.

Övning 4.13. Visa att en ringhomomorfi mellan kommutativa ringar med etta R, S måste skicka 0an i R till 0an i S men ettan i R måste inte skicka 1an i R till 1an i S. Gör detta genom att bevisa att följande funktioner är ringhomomorfier som exemplifierar detta

$$\begin{aligned}0 :&R \to R\\ &x \mapsto 0\end{aligned}$$

där $0 \neq 1 \in R$.

$$\begin{aligned}\phi :&\mathbb{Z}/(6) \to \mathbb{Z}/(6)\\ &a \to 3 \cdot_6 a.\end{aligned}$$

5 Elliptiska kurvor

5.1 Introduktion

I följande kapitel kommer vi att för det mesta jobba över kroppen $\mathbb{R}$, för att lättare visualisera kurvor. Det går dock att göra saker över andra kroppar. För att studera kurvor på ett systematiskt sätt brukar man klassificera kurvor utifrån graden på det polynom som definierar dem. Därmed skulle en kurva på formen

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

där $a, b, c, d, e, f \in \mathbb{R}$ vara en kurva av andra graden. Mycket av en kurvas egenskaper bestäms av dess grad. En linje är en kurva av grad 1. Något som inte är lika uppenbart är att två linjer utgör en kurva av grad två. Om ena linjen beskrivs av ekvationen $ax + by + c = 0$, och den andra av ekvationen $dx + ey + f = 0$, kommer ekvationen

$$(ax + by + c)(dx + ey + f) = 0$$

beskriva en kurva av andra graden som är unionen av de två linjerna. På samma satt utgör 3 linjer en kurva av tredje graden o.s.v.

Vi råkar vara intresserade av kurvor av tredje graden. En generell kurva av tredje graden är en kurva

$$ax^3 + bx^2y + cxy^2 + dy^3 + ex^2 + fxy + gy^2 + hx + iy + j = 0$$

för något val av $a, b, c, d, e, f, g, h, i, j \in \mathbb{R}$.

Vi kommer att behöva en sats om tredje gradens kurvor i framtiden som vi inte kommer att bevisa. Inte för att beviset är för svårt, utan mer för att det innehåller många detaljer.

Sats 5.1.1. *Låt C_1, C_2 och C_3 vara tre kurvor av grad 3 så att C_1 och C_2 skär varandra i 9 punkter. Om C_3 går genom 8 av de här 9 punkterna, går kurvan även igenom den nionde.*

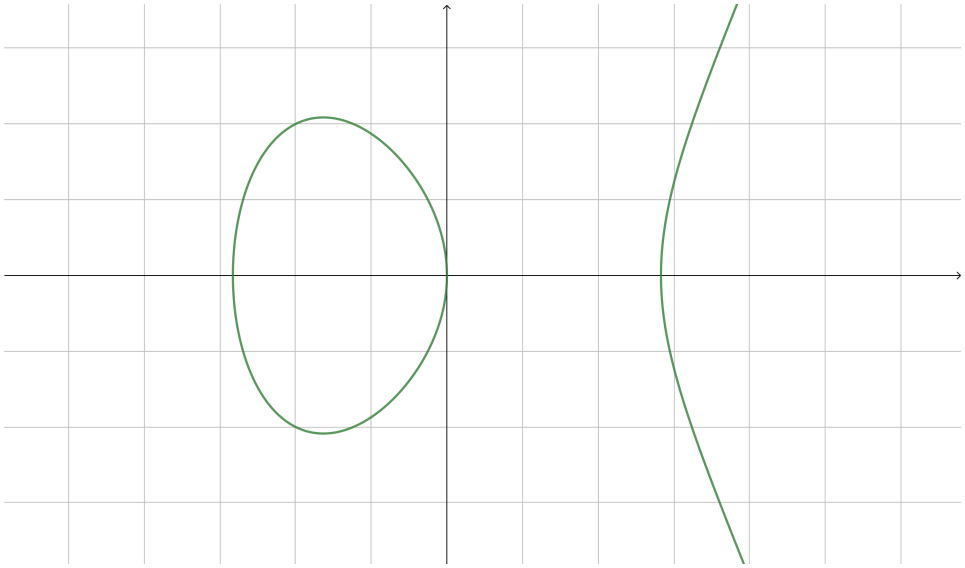
I praktiken kommer vi inte att behöva så många bokstäver för att beskriva tredjegradskurvor. Vi är intresserade av tredjegradskurvor på en särskild form.

Definition 5.1.2 (Elliptisk kurva)**.** En elliptisk kurva är en glatt kurva C som kan beskrivas av en ekvation på formen

$$y^2 = x^3 + ax^2 + bx + c$$

där $a, b, c \in \mathbb{R}$. △

Det här kan verka som en väldigt specifik typ av tredjegradskurva, men det visar sig att de flesta kurvor kan skrivas om på den här formen med några enkla transformationer. Vi kan också förenkla ekvationen vi har ytterligare. Genom



Figur 5.1: En elliptisk kurva med ekvation $y^2 = x^2 - 2x$. Notera att kurvan är symmetrisk kring x-axeln.

att göra variabelbytet $x \to x - a/3$ försvinner termen x^2. Byter vi namn på konstanterna får vi den nya ekvationen

$$y^2 = x^3 + ax + b.$$

Vi kommer ibland använda den enklare formen.

Vi kommer i fortsättningen att lägga till en extra punkt utöver lösningarna (x, y) till $y^2 = x^3 + ax^2 + bx + c$. Den här punkten, kallad punkten vid oändligheten, betecknas $\mathcal{O}$. Visuellt vill vi se det som den punkt som ligger oändligt långt bort i y-axelns riktning, och åt båda håll på samma gång. Om L är en linje säger vi att $\mathcal{O}$ ligger på en linje om linjen är parallell med y-axeln.

När vi nu har en kurva på en så fin form finns det ett väldigt enkelt sätt att kontrollera om kurvan är glatt eller inte. En singularitet till kurvan $y^2 = x^3 + ax + b$ är en punkt på kurvan sådan att

$$0 = \frac{\partial}{\partial y}[y^2 - x^3 - ax - b] = 2y, \quad 0 = \frac{\partial}{\partial x}[y^2 - x^3 - ax - b] = -3x^2 - a.$$

Vi är alltså intresserade av när det finns en punkt $(\alpha, 0)$ på vår kurva så att $\alpha^2 = -a/3$. Eftersom punkten $(\alpha, 0)$ ligger på kurvan uppfyller den

$$0^2 = \alpha^3 + a\alpha + b = -a\alpha/3 + a\alpha + b \iff 2a\alpha = -3b.$$

Vi kvadrerar båda sidor och får

$$(2a\alpha)^2 = (-3b)^2 \iff -\frac{4}{3}a^3 = 9b^2 \iff -4a^3 = 27b^2.$$

Definition 5.1.3. Om kurvan C är definierad av $y^2 = x^3 + ax + b$ låter vi

$$\Delta_C = 4a^3 + 27b^2$$

vara diskriminanten av C. Kurvan C har en singularitet om och endast om $\Delta_C = 0$. △

Exempel 5.1.4. Betrakta kurvorna C_1 och C_2 beskrivna av ekvationerna $y^2 = x^3 - 2x$ och $y^2 = x^3$. Kurvan C_1 har diskriminant $4 \cdot (-2)^2 + 0 = 16$, medan C_2 har diskriminant 0. Vi vet från förra kapitlet att C_2 har en singularitet i $(0, 0)$. ▲

Efter allt jobb är det naturligt att ställa sig frågan, "varför bryr vi oss"? Vad är det som gör tredje gradens kurvor särskilt intressanta, mer än kurvor av andra grader? Det beror på att punkterna i en elliptisk kurva har strukturen av en grupp. Det är en komplicerad gruppstruktur, och inte så enkel som att addera eller multiplicera koordinater.

Låt C vara kurvan $y^2 = x^3 + ax^2 + bx + c$. De allra flesta linjerna skär C i 3 punkter. Alltså är ett naturligt sätt att associera en tredje punkt R till 2 punkter P, Q att betrakta linjen L genom P och Q, och att låta R vara den tredje punkten L skär C i. Vi skriver

$$R = P * Q$$

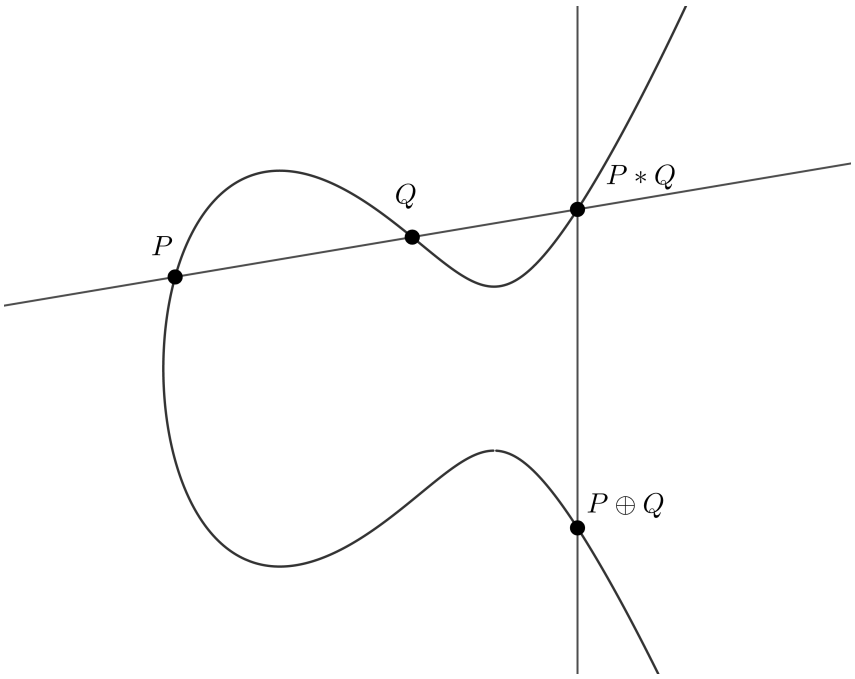
för P, Q, R som i exemplet. Alltså ligger alltid $P, Q, P * Q$ på en linje.

Operationen $*$ definierar inte en grupp, men med en liten förändring kan vi definiera en operation $\oplus$ som ger oss en grupp.

Sats 5.1.5. *Låt C bestå av punkter (x, y) i planet så att $y^2 = x^3 + ax^2 + b$, tillsammans med $\mathcal{O}$, punkten vid oändligheten. Definiera operationen $\oplus \colon C \times C \to C$ genom att låta $P \oplus Q$ vara reflektionen av punkten $P * Q$ i x-axeln, där $P * Q$ är punkten så att P, Q och $P * Q$ ligger på en linje (se figur 5.2).*

Vi går igenom beviset steg för steg. Det finns några detaljer som behöver redas ut i definitionen av addition. Precis som innan säger vi att punkten $\mathcal{O}$ ligger på en linje om linjen är parallell med y-axeln. Om vi därmed tar två punkter P, Q som har samma x-koordinat, kommer $P * Q$ skära C i punkten $\mathcal{O}$.

Givet två punkter P och Q, kommer linjen genom P och Q skära vår kurva genom punkter som sammanlagt har multiplicitet tre. Punkterna P och Q kommer att motsvara två av dessa multipliciteter, och vi låter $P * Q$ vara den tredje multipliciten. Det innebär att det kan hända att tex $P * Q = P$, om linjen skär kurvan i P med multiplicitet 2.



Figur 5.2: Addition av två olika punkter

Med andra ord, vi definierar $P \oplus P$ genom att ta tangentlinjen L till C i P, betrakta den andra skärningspunkten R av tangentlinjen och C, och att sedan reflektera R i x-axeln (se figur 5.3). Vi låter också $\mathcal{O} \oplus \mathcal{O} = \mathcal{O}$.

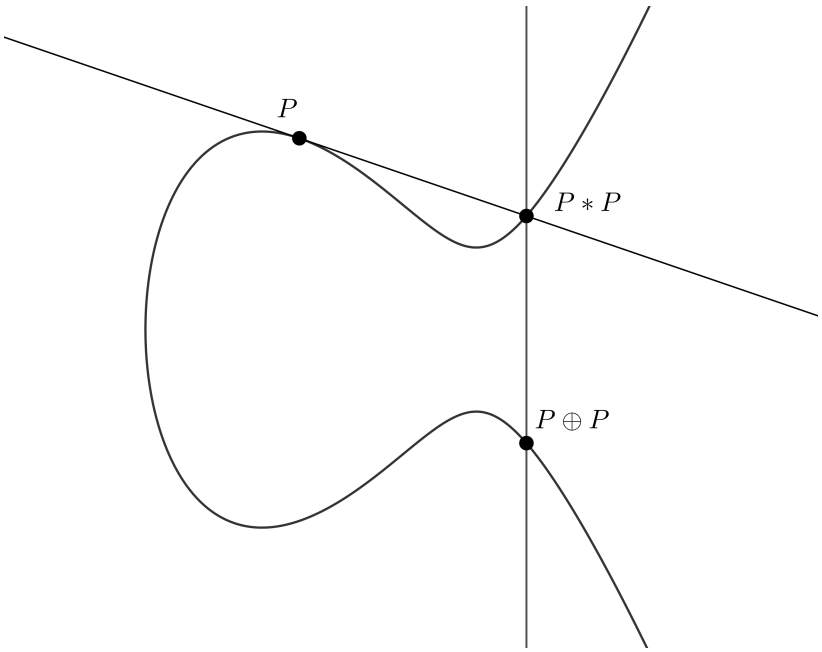
Kommutativitet följer direkt. Identiteten ges av $\mathcal{O}$, då linjen genom $\mathcal{O}$ och en annan godtycklig punkt $P \in Y$ är linjen genom P parallell med y-axeln, så tredje skärningspunkten med C ges av reflektionen av P i x-axeln. Alltså är $\mathcal{O} \oplus P = P \oplus \mathcal{O} = P$.

Inversen till P ges av P_-, reflektionen av P i x-axeln, då linjen genom P och P_- skär C i $\mathcal{O}$, och reflektionen av $\mathcal{O}$ i x-axeln är O. Alltså är $P \oplus P_- = P_- \oplus P = \mathcal{O}$.

Det återstår att visa associativitet. Betrakta figur 5.4. Vi har på bilden ritat tre punkter, P, Q, R. Vi har även två triplar med linjer, tre som är streckade och tre som inte är det.

De streckade linjerna är linjen genom $\mathcal{O}, P * Q, P \oplus Q$, linjen genom $Q * R, Q, R$ och linjen genom $P, Q \oplus R, P * (Q \oplus R)$.

De återstående linjerna är linjen genom $P, Q, P * Q$, linjen genom $Q * R, \mathcal{O}, Q \oplus R$ och linjen genom $P \oplus Q, R, (P \oplus Q) * R$.



Figur 5.3: Addition med samma punkt

Vi kan se de här två samlingar av linjer som två tredjegradskurvor, och vi ser att de möter varandra och C i 8 punkter, nämligen $\mathcal{O}, P, Q, R, P*Q, Q*R, P \oplus Q$ och $Q \oplus R$. Alltså måste enligt sats 5.1.1 de tre tredjegradskurvor alltså mötas i samma nionde punkt. Men de tre streckade linjerna möter C i $(P \oplus Q) * R$ i sin nionde skärningspunkt, och de återstående linjerna möter C i $P * (Q \oplus R)$ is sin nionde skärningspunkt. Alltså måste $(P \oplus Q) * R = P * (Q \oplus R)$.

Det återstår bara att notera att $(P \oplus Q) \oplus R$ är $(P \oplus Q) * R$ reflekterat i x-axeln, och att $P \oplus (Q \oplus R)$ är $P * (Q \oplus R)$ reflekterad i x-axeln. Därmed är

$$(P \oplus Q) \oplus R = P \oplus (Q \oplus R).$$

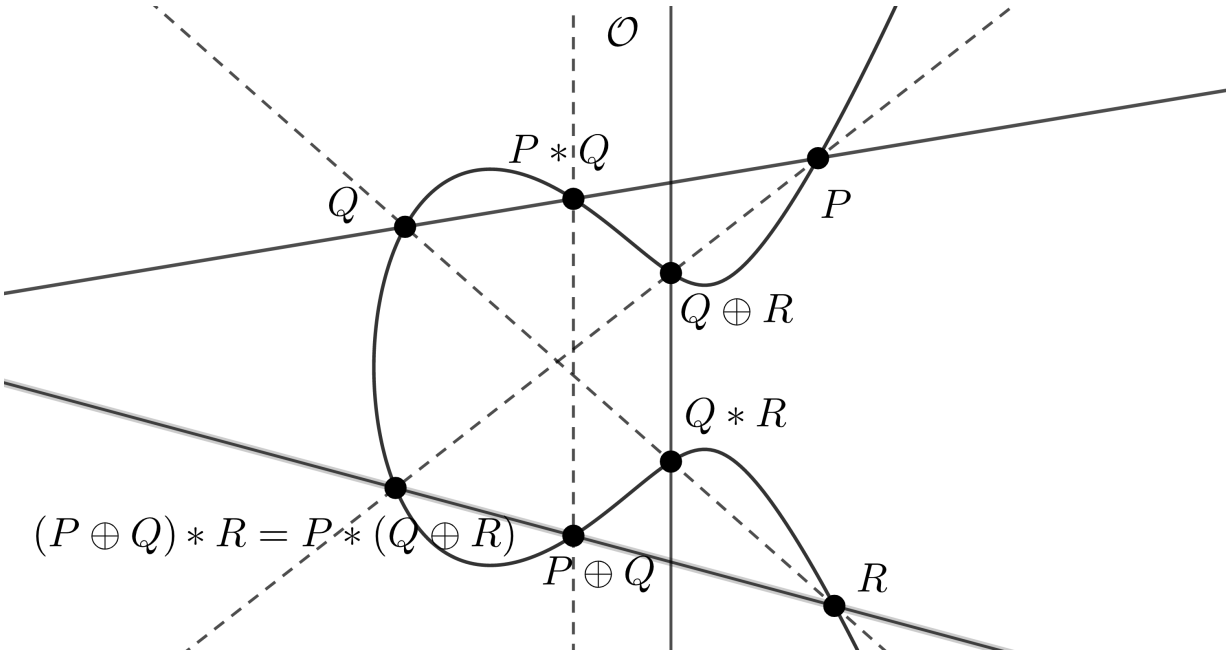
I framtiden kommer vi att använda notationen $P + Q$ för addition istället för $P \oplus Q$. Vi kommer även att skriva $-P$ för inversen till P, och nP för

$$\underbrace{P + P + \cdots + P}_{n \text{ gånger}}.$$

Men kom ihåg att det inte har något att göra med vanlig addition, utan är en ny grupplag.

5.2 Formler för addition

Vi vill nu beskriva vår nya addition på ett rent algebraiskt sätt, vilket kommer att behövas ibland. Låt C vara kurvan definierad av $y^2 = x^3 + ax^2 + bx + c$, och



Figur 5.4: Associativitet hos en elliptisk kurva.

låt $P = (x_1, y_1)$, $Q = (x_2, y_2)$. Reflektion i x-axeln innebär att skicka (x, y) till $(x, -y)$, så $-P = (x, -y)$. Låt $-(P+Q) = (x_3, y_3)$. Vi vet att $P, Q, -(P+Q)$ ligger på en linje. Alltså ligger (x_3, y_3) på linjen

$$y = \lambda x + \nu, \quad \text{där } \lambda = \frac{y_2 - y_1}{x_2 - x_1}, \quad \text{och } \nu = y_1 - \lambda x_1.$$

Men (x_3, y_3) måste även ligga på kurvan C. Att gemensamt ligga på linjen $y = \lambda x + \nu$ och på kurvan C innebär att

$$(\lambda x + \nu)^2 = x^3 + ax^2 + bx + c$$

$$\iff x^3 + (a - \lambda^2)x^2 + (b - 2\lambda\nu)x + (c - \nu^2) = 0.$$

Vi vet att x_1, x_2 och x_3 alla löser ekvationen, och VL är en tredjegradsekvation, så vi kan faktorisera

$$x^3 + (a - \lambda^2)x^2 + (b - 2\lambda\nu)x + (c - \nu^2) = (x - x_1)(x - x_2)(x - x_3).$$

Om vi utvecklar HL och betraktar x^2 termen, får vi

$$a - \lambda^2 = -x_1 - x_2 - x_3$$

$$\iff x_3 = \lambda^2 - a - x_1 - x_2.$$

Vi har också att

$$y_3 = \lambda x_3 + \nu.$$

Nu återstår bara att märka att om $-(P+Q) = (x_3, y_3)$, så är $(P+Q) = (x_3, -y_3)$. Alltså har vi bevisat följande formler.

Hjälpsats 5.2.1. *Låt C vara en elliptisk kurva given av $y^2 = x^3+ax^2+bx+c$. Låt $P = (x_1, y_2)$, $Q = (x_2, y_2)$, och anta $P \neq Q$. Då är*

$$P + Q = (\lambda^2 - a - x_1 - x_2, -\lambda(\lambda^2 - a - x_1 - x_2) - \nu),$$

där

$$\lambda = \frac{y_1 - y_2}{x_1 - x_2}, \quad \nu = y_1 - \lambda x_1.$$

Exempel 5.2.2. Betrakta kurvan $y^2 = x^3 + 17$. Vi har två punkter, $P = (-1, 4)$ och $Q = (2, 5)$. Linjen genom P och Q i det här fallet är

$$y = \frac{1}{3}x + \frac{13}{3}.$$

Alltså är $\lambda = \frac{1}{3}$ och $\nu = \frac{13}{3}$. Om (x_3, y_3) är koordinaterna för $P + Q$ kan vi beräkna

$$x_3 = \lambda^2 - x_1 - x_2 = -\frac{8}{9}.$$

Vi får också

$$y_3 = -\lambda x_3 - \nu = -\frac{109}{27}.$$

Alltså är

$$P + Q = \left(-\frac{8}{9}, -\frac{109}{27}\right).$$

Som vi märker kan komplexiteten av lösningarna gå upp ganska fort.

▲

Tyvärr fungerar inte formeln i det fallet då $P = Q$, och tyvärr är situationen lite värre här. Om $P = Q = (x_1, y_1)$, kommer punkten $P * P = (x_3, -y_3)$ istället ligga på linjen

$$y = \lambda x + \nu, \quad \lambda = \frac{dy}{dx} \mid_{y=y_1} = \frac{f'(x_1)}{2y_1},$$

där $f(x) = x^3 + ax^2 + bx + c$. Förutom det problemet fungerar formeln som innan.

Hjälpsats 5.2.3. *Låt C vara en elliptisk kurva given av $y^2 = x^3+ax^2+bx+c$ och låt $P = (x_1, y_1)$. Då är*

$$2P = (\lambda^2 - a - 2x_1, -\lambda(\lambda^2 - a - 2x_1) - \nu),$$

där

$$\lambda = \frac{f'(x_1)}{2y_1}, \quad \nu = y_1 - \lambda x_1, \quad f(x) = x^3 + ax^2 + bx + c.$$

5.3 Punkter av liten ordning

Givet en elliptisk kurva C på formen

$$y^2 = x^3 + ax + b,$$

vilka punkter $P \in E$ har ordning 2? Att ha ordning 2 innebär att $P+P = \mathcal{O}$. Vi adderar $-P$ till båda sidor och får $P = -P$. Men $-P$ är punkten P reflekterad genom x-axeln, så om $P = -P$ måste P ligga på x-axeln, alltså ha y-koordinat 0. Om $P = (x, 0)$, gäller då alltså

$$0 = x^3 + ax + b.$$

Polynomet $x^3 + ax + b$ kan antingen ha ett eller tre reella nollställen. Eftersom att polynomet är udda kommer det att anta väldigt stora värden för stora x, och väldigt negativa värden för små x. Alltså måste det anta värdet 0 åtminstone en gång. Om det antar värdet 0 en gång och x_0 är det nollstället, kommer vi att kunna faktorisera $x^3 + ax + b$ som $(x - x_0)(x^2 + cx + d)$. Polynomet $x^2 + cx + d$ kommer att antingen ha två reella nollställen eller inga reella nollställen.

Om $x^3 + ax + b$ har ett reellt nollställe kommer kurvan att skära x-axeln 3 gånger, vilket innebär att den kommer att ha två komponenter. Om det enbart finns ett nollställe skär kurvan x-axeln en gång, vilket innebär att kurvan är sammanhängande.

Sammanfattningsvis är punkter med ordning 2 på formen $(0, x_0)$ där x_0 är ett nollställe till polynomet $x^3 + ax + b$, och det finns antingen 1 eller 3 sådana punkter.

Vi betraktar nu istället punkter av ordning 3. En punkt P har ordning 3 om $3P = \mathcal{O}$. Om vi adderar $-P$ till båda sidor får vi $2P = -P$. Punkten $2P$ är punkten man får genom att dra tangentlinjen vid punkten P, betrakta skärningen med C och sedan reflektera punkten i x-axeln. Men $-P$ är P reflekterad i x-axeln, så om $2P = -P$ måste tangentlinjen vid P skära kurvan C i P! Det här fungerar endast om tangentlinjen i P skär C endast en gång med multiplicitet 3. En sådan punkt kallas för en inflektionspunkt.

Vi är intresserade av att hitta en mer algebraisk beskriving av dessa punkter, och för att göra det använder vi de formler vi tagit fram. Låt $P = (x, y)$ vara en punkt av ordning 3. Eftersom $2P = -P$ och $-P = (x, -y)$, kommer x-koordinaten för P vara samma som x-koordinaten för 2P. Men x-koordinaten för $2P$ ges av

$$\lambda^2 - a - 2x$$

där $\lambda = f'(x)/2y$. I vårt fall är $f'(x) = 3x^2 + 2ax + b$. Därmed är

$$\lambda^2 = \frac{(3x^2 + 2ax + b)^2}{4y^2} = \frac{9x^4 + 12ax^3 + (4a^2 + 6b)x^2 + 4abx + b^2}{4x^3 + 4ax^2 + 4bx + 4c}.$$

Med lite extra uträkningar kommer vi till slut fram till

$$\lambda^2 - a - 2x = \frac{x^4 - 2bx^2 - 8cx + b^2 - 4ac}{4x^3 + 4ax^2 + 4bx + 4c}.$$

Den här läskiga formeln har sina egna användningar, så vi skriver ned den.

Hjälpsats 5.3.1 (Explicit dupliceringsformel)**.** *Låt C vara en elliptisk kurva given av ekvationen $y^2 = x^3 + ax^2 + bx + c$. Om $P = (x, y)$ är en punkt på kurvan så är x-koordinaten för $2P$*

$$\frac{x^4 - 2bx^2 - 8cx + b^2 - 4ac}{4x^3 + 4ax^2 + 4bx + 4c}.$$

Om vi nu återgår till punkter av ordning 3, om x-koordinaten för P är samma som x-koordinaten för $2P$ innebär det att

$$x = \frac{x^4 - 2bx^2 - 8cx + b^2 - 4ac}{4x^3 + 4ax^2 + 4bx + 4c}.$$

Om vi förlänger med nämnaren och flyttar över resultatet till vänsterled får vi ekvationen

$$0 = 3x^4 + 4ax^3 + 6bx^2 + 12cx + 4ac - b^2.$$

Alltså är de punkter av ordning 3 de punkter vars x-koordinat är en lösning till ekvationen ovan. För varje sådan x-koordinat finns det två val av y-koordinat, eftersom y^2 då är bestämd. Men polynomet har grad 4, så det finns som mest 8 punkter av ordning 3. Vi sammanfattar resultatet nedan:

Hjälpsats 5.3.2 (Punkter av ordning 2 eller 3)**.** *Låt C vara en elliptisk kurva given av ekvationen $y^2 = x^3 + ax^2 + bx + c$.*

- *De punkter $P \in C$ med ordning 2 är exakt de punkter med y-koordinat 0, och x-koordinat en lösning till ekvationen*

  $$x^3 + ax^2 + bx + c.$$

  Det finns antingen 1 sådan punkt eller 3 sådana punkter.

- *De punkter $P \in C$ med ordning 3 är exakt de punkter där tangentlinjen skär kurvan med multiplicitet 3. Dessa punkter är de som har x-koordinater som är lösningar på ekvationen*

  $$3x^4 + 4ax^3 + 6bx^2 + 12cx + 4ac - b^2.$$

  Det finns som mest 8 sådana punkter.

5.4 Andra kroppar

Säg att vi är intresserade av en elliptisk kurva C över en annan kropp K som inte nödvändigtvis är $\mathbb{R}$. Den ges av en ekvation

$$y^2 = x^3 + ax^2 + bx + c$$

där $a, b, c \in K$. Punkterna på kurvan är alla lösningar i K^2. Gruppstrukturen fungerar på exakt samma sätt i det här fallet, och man kan använda de formler vi nyligen härlett för att definiera addition.

Ibland kan vi ha två kroppar som innehåller varandra som $\mathbb{Q} \subset \mathbb{R}$. Om vi då exempelvis har att göra med kurvan $y^2 = x^3 + 3x + 1$ kan det vara svårt att veta om vi är intresserade av de rationella punkterna eller de reella. Därmed skriver vi $C(\mathbb{Q})$, $C(\mathbb{R})$ för de rationella respektive de reella lösningarna. Vidare skriver vi $C(K)$ för lösningar i K^2.

Mängden $C(K)$ är alltid en grupp, men strukturen beror på vilket K vi väljer. Mycket av det kvarstående materialet i kompendiet går ut på att beskriva strukturen för olika val av K. Då $K = \mathbb{R}$ är gruppen $C(\mathbb{R})$ ganska stor och har lite för enkel struktur för att vara särskilt intressant. Om vi istället begränsar oss till rationella punkter, $C(\mathbb{Q})$, visar det sig att gruppen är mer komplicerad, men också så pass enkel att den går att hantera. Faktum är att den alltid är ändligt genererad i det här fallet.

Kapitel 7 går ut på att bevisa det påståendet, vilket inte kommer att vara särskilt enkelt.

Övningar

Övning 5.1. Låt A vara en abelsk grupp. Definiera funktionen $n : A \to A$, multiplikation med n, som

$$na = \underbrace{a + a + \cdots + a}_{n \text{ gånger}}.$$

Visa att n är en homomorfi.

Övning 5.2. Låt C vara en tredjegradskurva. Låt P, Q vara punkter på C, och låt $P * Q$ vara den punkt så att $P, Q, P * Q$ ligger på en linje.

(i) Förklara varför $P * Q = Q * P$.

(ii) Visa att det inte finns ett element P_0 så att $P_0 * P = P$ för denna lag, alltså visa att det inte finns ett identitetselement.

(iii) Visa att lagen inte heller är associativ, $P * (Q * R) \neq (P * Q) * R$.

(iv) Förklara varför $P * (P * Q) = Q$

Övning 5.3. Låt S vara en mängd och $* : S \times S \to S$ en kompositionsregel med följande egenskaper

- $P * Q = Q * P$ för alla $P, Q \in S$.
- $P * (P * Q) = Q$ för alla $P, Q \in S$.

Välj ett element $\mathcal{O} \in S$ och definiera en ny regel $\oplus$ på S genom

$$P + Q = \mathcal{O} * (P * Q)$$

Visa att

(i) $P + Q = Q + P$ och $\mathcal{O} + P = P$ för alla $P, Q \in S$, alltså att den nya regeln är kommutativ och att $\mathcal{O}$ är ett identitetselement.

(ii) Visa att givet $P, Q \in S$, har ekvationen $X + P = Q$ en unik lösning i X, nämligen $X = P * (Q * \mathcal{O})$. Mer specifikt, visa att om vi definierar $-P$ som $P * (O * O)$ så är $-P$ den unika lösningen på $X + P = \mathcal{O}$.

(iii) Visa att $+$ är associativ, och då även definierar en grupp, om och endast om

$$R * (\mathcal{O} * (P * Q)) = P * (\mathcal{O} * (Q * R)) \text{ för alla } P, Q, R \in S.$$

Övning 5.4. (i) Betrakta kurvan given av ekvationen $u^3 + v^3 = 1$. Visa att om (u, v) är en punkt på kurvan, kommer

$$x = \frac{12}{u + v} \text{ och } y = 36\frac{u - v}{u + v}$$

uppfylla ekvationen $y^2 = x^3 - 432$.

(ii) Visa att om (x, y) ligger på kurvan $y^2 - 432$ och

$$u = \frac{36 + y}{6x}, \quad v = \frac{36 - y}{6x}$$

att (u, v) uppfyller $u^3 + v^3 = 1$.

(iii) Visa att de tidigare två transformationerna är inverser till varandra när de är definierade. För vilka punkter är de väldefinierade?

Övning 5.5. Låt C vara en elliptisk kurva definierad av $y^2 = x^3 + ax^2 + bx + c$.

(i) Vad blir ekvationen om vi gör variabelbytet $x \to x - a/3$?

(ii) Använd föregående exempel för att hitta en diskriminant Δ_C till C som inte är på en förenklad form, alltså ett uttryck i a, b, c som är 0 exakt då C har en singulär punkt.

Övning 5.6. Betrakta punkten P=(3,8) på kurvan

$$y^2 = x^3 - 43x + 166.$$

Beräkna $2P, 4P$ och $8P$ (här är additionen med avseende på den elliptiska kurvan). Jämför P med $8P$.

Övning 5.7. Vilka av följande kurvor har 3 punkter av ordning 2 och vilka har 1?

(i) $y^2 = (x-1)(x-2)(x-3)$

(ii) $y^2 = x^3 + x$

(iii) $y^2 = x^3 - 1$

(iv) $y^2 = x^3 + x - 2$

6 Elliptiska kurvor över ändliga kroppar

De gruppoperationer vi definierat fungerar över vilken kropp som helst. Det är bara lite krångligare att rita ut linjer och elliptiska kurvor. Antag att kurvan Y är på formen

$$y^2 = x^3 + bx + c$$

sådan att

$$\Delta_Y = 16b^3 - 27c^2 \not\equiv 0 \text{ i } \mathbb{Z}/(n).$$

Kom ihåg att gruppoperationen för $P = (x_1, y_1)$ och $Q = (x_2, y_2)$ (ingen av dem är identiteten $\mathcal{O}$) börjar med att studera linjen $y = \lambda x + \nu$ där

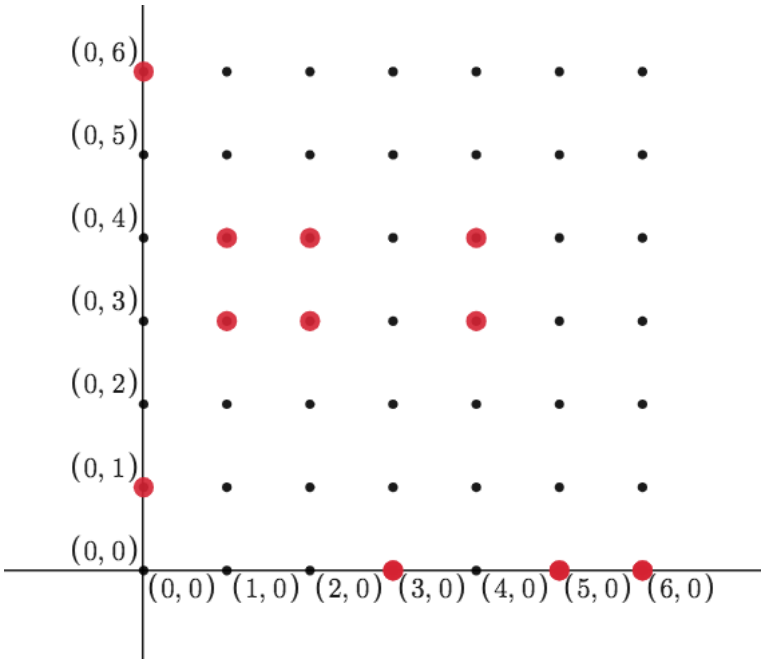
$$\lambda = \begin{cases} \frac{y_2 - y_1}{x_2 - x_1} & \text{om } P \neq Q \\ \frac{3x_1^2 + b}{2y_1} & \text{om } P = Q \end{cases}$$

enligt hjälpsatserna 5.2.1 och 5.2.3. Här är $\nu = y_1 - \lambda x_1$. Detta säger alltså att $P + Q = (x_3, y_3)$ har koordinater

$$x_3 = \lambda^2 - a - x_1 - x_2, \quad y_3 = -\lambda x_3 - \nu.$$

6.1 Hasse–Weils sats

Exempel 6.1.1. Betrakta kurvan Y som ges av $y^2 = x^3 + 1$ över $\mathbb{Z}/(7)$. Diskriminanten är $\Delta_Y = 0 - 27 = (-27)\%7 = 1 \neq 0$ så kurvan är glatt. Den har 12 punkter, varav vi ritar ut 11 element i följande bild



Figur 6.1: Kurvan som motsvarar $y^2 - x^3 - 1 = 0$ för $K = \mathbb{Z}/(7)$.
(Länk till animation)

▲

och det sista elementet är det som ligger borta i oändligheten $\mathcal{O}$. Alla ändliga abelska grupper är ändligt genererade, alltså med hjälp av satsen 3.1.11 måste Y vara isomorf med en produkt av olika $\mathbb{Z}/(n_i)$ sådan att $n_1 \cdot n_2 \cdot \ldots = 12$ i detta fall, eftersom antalet element i en produkt av grupper är produkten av antalet element i respektive grupp. Tolv har primtalsfaktorisering $3 \cdot 2 \cdot 2$. Så vi kan för tillfället inte avgöra om vi jobbar med exempelvis $\mathbb{Z}/(3) \times \mathbb{Z}/(2^2)$ eller $\mathbb{Z}/(3) \times \mathbb{Z}/(2) \times \mathbb{Z}/(2)$ utan att göra ett par ytterligare beräkningar. Notera att $\mathbb{Z}/(12) \cong \mathbb{Z}/(3) \times \mathbb{Z}/(4)$ och att $\mathbb{Z}/(6) \cong \mathbb{Z}/(2) \times \mathbb{Z}(3)$ så vi behöver inte ta hänsyn till de fallen. För att se vilken grupp vi jobbar med så har vi ett knep vi kan använda. Antalet element av ordning 2 (alltså $x + x = 0$) i $\mathbb{Z}/(3) \times \mathbb{Z}/(2) \times \mathbb{Z}/(2)$ är 4, alltså alla element på formen $(0, a, b)$, medan i $\mathbb{Z}/(3) \times \mathbb{Z}/(4)$ har vi att elementen av ordning 2 är bara 2, alltså $(0, 0, 0), (0, 2, 0), (0, 0, 2)$.

Låt oss studera punkten $P = (0, 1)$ och dess multiplar. För $P + P = 2P$ är $\lambda = \frac{0+0}{2} = \frac{0}{2} = 0$. Då får vi att $\nu = 1 - 0 = 1$ och enligt hjälpsats 5.2.3

$$2P = (0^2 - 0 - 0, -0(0^2 - 0 - 0) - \nu) = (0, 0 - 1) = (0, 6).$$

Eftersom vi ser att P och $2P$ ligger på en lodrät linje så är $3P = \mathcal{O}$. Så detta är vår kopia av $\mathbb{Z}/(3)$. Notera att de tre punkterna $(3, 0), (5, 0), (6, 0)$ ligger på x-axeln och är sina egna inverser (som fås genom att byta tecken på y-koordinaten). Alltså har vi att $2P = \mathcal{O}$ för dessa tre, och med identiteten blir det 4 sådana element totalt. Alltså är vår kurva, som grupp isomorf med $\mathbb{Z}/(3) \times \mathbb{Z}/(2) \times \mathbb{Z}/(2)$. Enligt övning 4.10 kan vi låta ett godtyckligt par av punkterna på x-axeln agera som generatorer för $\mathbb{Z}/(2) \times \mathbb{Z}/(2)$. Alltså drar vi slutsatsen att

$$\begin{aligned} \phi \colon \mathbb{Z}/(3) \times \mathbb{Z}/(2) \times \mathbb{Z}/(2) &\to Y \\ (a, b, c) &\to a(0, 1) + b(3, 0) + c(6, 0) \end{aligned}$$

Man kan fråga sig hur många punkter som finns på en elliptisk kurva i allmänhet. Det närmaste vi kommer är Hasse–Weils sats som vi tyvärr inte kan bevisa i denna kurs.

Sats 6.1.2 (Hasse–Weil)**.** *Givet en elliptisk kurva C över heltalen modulo p så är antalet punkter på den kurvan ungefär $p + 1$. Mer exakt så gäller det att*

$$|\#C - p - 1| \leqslant 2\sqrt{p}.$$

Exempel 6.1.3. Vi kan verifiera Hasse–Weil i kurvan från exempel 6.1.1 hade 12 punkter och p var 7. Vi kan då verifiera att

$$12 - 7 - 1 = 4 \leqslant 2\sqrt{7} \sim 5.29.$$

Man kan ju ifrågasätta om 12 och $7 + 1$ ligger nära varandra men ju större p blir desto mindre blir feltermen relativt sett. ▲

Kryptosystem

Under 1970-talet skedde en revolutionerande framsteg inom kryptografi med introduktionen av kryptosystem med offentliga nycklar av Diffie, Hellman,

Merkle, Rivest, Shamir, Adelman och andra. Ett kryptosystem tillåter två parter, oftast kallade Bob och Alice, att utbyta information över en kommunikationskanal som kan avlyssnas men där kommunikationen sker på ett sätt som förhindrar deras eventuella motståndare, Eve, att avgöra informationen i meddelanden som skickats.

Matematiskt sett kan man betrakta ett grundläggande kryptosystem som en injektiv funktion

$$f\colon \{\text{meddelanden}\} \to \{\text{krypterade meddelanden}\}.$$

Mängden av meddelanden och krypterade meddelanden är ofta densamma i denna kurs och den brukar dessutom vara en grupp G där de olika elementen motsvarar olika meddelanden. Till exempel så kan varje element signifiera en bokstav eller en symbol på ditt tangentbord.

Exempel 6.1.4. Det finns 256 olika ASCII tecken, de första 32 är lite obskyra men resten kan du se i tabellen nedan.

ASCII printable characters

32	space	64	@	96	`
33	!	65	A	97	a
34	"	66	B	98	b
35	#	67	C	99	c
36	$	68	D	100	d
37	%	69	E	101	e
38	&	70	F	102	f
39	'	71	G	103	g
40	(	72	H	104	h
41	)	73	I	105	i
42	*	74	J	106	j
43	+	75	K	107	k
44	,	76	L	108	l
45	-	77	M	109	m
46	.	78	N	110	n
47	/	79	O	111	o
48	0	80	P	112	p
49	1	81	Q	113	q
50	2	82	R	114	r
51	3	83	S	115	s
52	4	84	T	116	t
53	5	85	U	117	u
54	6	86	V	118	v
55	7	87	W	119	w
56	8	88	X	120	x
57	9	89	Y	121	y
58	:	90	Z	122	z
59	;	91	[	123	{
60	<	92	\	124	|
61	=	93	]	125	}
62	>	94	^	126	~
63	?	95	_		

Extended ASCII characters

128	Ç	160	á	192	└	224	Ó
129	ü	161	í	193	┴	225	ß
130	é	162	ó	194	┬	226	Ô
131	â	163	ú	195	├	227	Ò
132	ä	164	ñ	196	─	228	õ
133	à	165	Ñ	197	┼	229	Õ
134	å	166	ª	198	ã	230	µ
135	ç	167	º	199	Ã	231	þ
136	ê	168	¿	200	╚	232	Þ
137	ë	169	®	201	╔	233	Ú
138	è	170	¬	202	╩	234	Û
139	ï	171	½	203	╦	235	Ù
140	î	172	¼	204	╠	236	ý
141	ì	173	¡	205	═	237	Ý
142	Ä	174	«	206	╬	238	¯
143	Å	175	»	207	¤	239	´
144	É	176	░	208	ð	240	≡
145	æ	177	▒	209	Ð	241	±
146	Æ	178	▓	210	Ê	242	‗
147	ô	179	│	211	Ë	243	¾
148	ö	180	┤	212	È	244	¶
149	ò	181	Á	213	ı	245	§
150	û	182	Â	214	Í	246	÷
151	ù	183	À	215	Î	247	¸
152	ÿ	184	©	216	Ï	248	°
153	Ö	185	╣	217	┘	249	¨
154	Ü	186	║	218	┌	250	·
155	ø	187	╗	219	█	251	¹
156	£	188	╝	220	▄	252	³
157	Ø	189	¢	221	¦	253	²
158	×	190	¥	222	Ì	254	■
159	ƒ	191	┐	223	▀	255	nbsp

Figur 6.2: Tabell med ASCII tecken

Låt säga att du vill skicka ett krypterat sms, kom ihåg att ett sms har upp till 160 bokstäver. Då kan vi låta mängden av meddelanden och krypterade meddelanden vara $(\mathbb{Z}/(256))^{160}$ där varje element motsvarar en sekvens av ASCII tecken, alltså ett sms. ▲

Låt säga att Bob vill kryptera sitt meddelande m genom att beräkna $f(m)$ och sedan skicka värdet till Alice som kallar det krypterade meddelandet c. Hon vill läsa meddelandet och måste då avkryptera meddelandet genom att beräkna $f^{-1}(c) = f^{-1}(f(m)) = m$.

Definition 6.1.5. När mängden meddelanden och krypterade meddelanden är en grupp G så kan vi definiera ett *Ceasarchiffer* som en funktion på formen

$$f(x) = a * x$$

för något $a \in G$. △

Exempel 6.1.6. Låt mängden av meddelanden vara $(\mathbb{Z}/(256))^{160}$ som vi tidigare identifierat med sms med ASCII tecken. Låt oss använda $a = (4, 4, 4, 4, \ldots, 4)$ för ett Ceasarchiffer exempelvis, då skulle vi kryptera meddelandet *hejsan* (följt av 154 blanksteg) motsvarande sekvensen av siffror

$$(104, 101, 106, 115, 97, 110, 32, 32, 32, \ldots)$$

till att bli

$$(108, 105, 110, 119, 101, 114, 36, 36, 36, \ldots)$$

vilket motsvarar det krypterade meddelandet *linwer* (följt av 154 dollartecken \$). ▲

I klassiska privata nyckelkryptosystem behöver Bob och Alice komma överens om en funktion f, en så kallad *privat nyckel*, innan de kan kommunicera. Exempelvis i ett Ceasarchiffer måste alltså båda parter känna till elementet a. Det är viktigt att funktionen f, den privata nyckeln, är en väl bevarad hemlighet som bara känns till av Bob och Alice, de kan kommunicera säkert så länge den förblir okänd för Eve.

I klassiska privata nyckelkryptosystem kan vem som helst som vet hur man beräknar funktionen f också enkelt beräkna f^{-1}. Till exempel är ett Ceasarchiffer inte så svårt att knäcka eftersom man lätt kan hitta konstanten a och sedan applicera $-a$ på det krypterade meddelandet.

Antag nu att Bob och Alice aldrig har träffats, och deras enda kommunikationsmedel är via e-post eller textmeddelanden, som deras personliga motståndare Eve har resurserna att övervaka. Då kan de inte komma överens om någon privat nyckel. Offentlig-nyckel-kryptografi löser detta problem. I ett kryptosystem med offentliga nycklar kan Alice publicera sin krypteringsnyckel f, och trots att Eve känner till Alices offentliga krypteringsnyckel f, kan de inte beräkna den inversa funktionen f^{-1} som krävs för att avkryptera meddelanden. Vi behöver alltså konstruera en injektiv funktion som är svår att invertera.

Det mest kända systemet för offentlig-nyckel, som du kanske har sett, är RSA-systemet. I detta system är Alices offentliga nyckel ett stort positivt heltal N som är en produkt av två stora primtal p och q, $N = pq$. Såvitt vi vet behöver Eve faktorisera N och hitta p och q för att avkryptera meddelanden i RSA-systemet. Detta anses som svårt.

6.2 Diskreta logaritmproblemen

Vi skiftar vårt fokus mot andra kryptosystem med offentliga nycklar som bygger på gruppteori och elliptiska kurvor. Om G är en given grupp och $a, b \in G$

är element i G, kan vi fråga efter en exponent $m \in \mathbb{N}$ som löser formeln $a^m = b$ i gruppen G, det vill säga

$$\underbrace{a * \cdots * a}_{m \text{ gånger}} = b.$$

Detta kallar vi för ett *gruppteoretiskt logaritmproblem.*

Definition 6.2.1 (Diskreta logaritmproblemet)**.** Betrakta gruppen av de nollskilda talen med multiplikation modulo något primtal p, alltså $\mathbb{Z}/(p) - \{0\}$ som brukar betecknas med $(\mathbb{Z}/(p))^\times$ för något primtal p. Det gruppteoretiska logaritmproblemet för den gruppen kallas för *diskreta logaritmproblemet* (DLP). △

Definition 6.2.2 (Diskreta logaritmproblemet för en elliptisk kurva)**.** Betrakta en elliptisk kurva modulo p, $C(\mathbb{Z}/(p))$. Det gruppteoretiska logaritmproblemet för en sådan grupp kallas för för diskreta logaritmproblemet för en elliptisk kurva eller *elliptic curve discrete logarithm problem* (ECDLP). △

6.3 Kvadratroten modulo p

Vi vet inte hur vi ska lösa de gruppteoretiska logaritmproblemen i allmänhet. I denna kurs nöjer vi oss med att analogt fundera på hur man kan beräkna kvadraträtter modulo p. Med detta menar vi att om $x^2 = a$ modulo p så sägs x vara en kvadratrot till a. Vi låter oss själva skriva $x = \pm\sqrt{a}$ när vi menar detta, vi kommer att förklara mer om detta senare. Låt oss se ett motiverande exempel på hur det kan vara användbart. Låt oss försöka lösa ekvationen

$$x^2 + bx + c = 0 \quad \text{(modulo ett primtal } p > 2\text{)}$$

Vi kan faktiskt använda kvadratkomplettering här. Vi kan skriva om ekvationerna genom att addera och multiplicera saker på båda sidor om likhetstecknet

$$\begin{aligned}
& x^2 + \alpha x + \beta = 0 \\
\iff & \Big(x^2 + 2(2^{-1}\alpha)x + \underbrace{(2^{-1}\alpha)^2\Big) - (2^{-1}\alpha)^2}_{=0} + \beta = 0 \\
\iff & \Big(x + 2^{-1}\alpha\Big)^2 - (2^{-1}\alpha)^2 + \beta = 0 \\
\iff & \Big(x + 2^{-1}\alpha\Big)^2 = (2^{-1}\alpha)^2 - \beta.
\end{aligned}$$

Om vi i detta skede kan hitta en kvadratrot har vi att

$$\begin{aligned}
& x^2 + \alpha x + \beta = 0 \\
\iff & x + 2^{-1}\alpha = \pm\sqrt{(2^{-1}\alpha)^2 - \beta} \\
\iff & x = -2^{-1}\alpha \pm \sqrt{(2^{-2})\alpha^2 - \beta}.
\end{aligned}$$

Vi kan sammanfatta detta i en hjälpsats:

Hjälpsats 6.3.1. *Den välkända* pq-formeln *fungerar över* $\mathbb{Z}/(p)$ *där* $p > 2$ *är ett primtal. Det vill säga*

$$x^2 + \alpha x + \beta = 0$$

har alla sina lösningar på formen

$$x = -\frac{\alpha}{2} \pm \sqrt{\frac{\alpha^2}{4} - \beta}.$$

där $\sqrt{x}$ *är ett tal i* $\mathbb{Z}/(p)$ *så att* $(\sqrt{x})^2 = 1$.

För att hitta kvadraträtter, eller i allmänhet när man vill hitta lösningar till en ekvation $f(x, y) = 0$ modulo p kan man göra trial and error eftersom det bara finns ändligt många alternativ att pröva. Man kan pröva alla p^2 olika värden på (x, y) och se vilka som uppfyller ekvationen. Detta kan dock vara väldigt tidskrävande, så låt oss nu försöka konstruera en mer effektiv lösningsmetod. Låt oss försöka förstå ekvationen för kvadraträtter lite bättre.

Hjälpsats 6.3.2. *I en kropp* K *där* $1 + 1 \neq 0$ *har varje tal* a *som mest två kvadraträtter. Mer precist så menar vi att*

$$x^2 = a$$

har max två olika lösningar och dessa lösningar är varandras additiva inverser. Denna ekvation har exakt en lösning om och endast om $a = 0$.

Bevis. Om man noga följer beviset för satsen 3.4.6 så går det beviset ut på att bevisa att två nollskilda tal alltid har en nollskild produkt modulo p. Alltså får vi då att $x^2 = 0$ betyder att $x = 0$ enligt övningen 3.9. Om $a \neq 0$, så finns antingen ingen lösning eller så finns åtminstone en lösning. Antag att $\sqrt{a}$ är en lösning, det vill säga $\sqrt{a}^2 = a$. Då får vi att

$$(x + \sqrt{a})(x - \sqrt{a}) = x^2 - \sqrt{a}x + \sqrt{a}x - \sqrt{a}^2 = x^2 - a.$$

Det är då tydligt att lösningarna till $x^2 - a = 0$ är exakt $\sqrt{a}$ och $-\sqrt{a}$. Notera att $-\sqrt{a} \neq \sqrt{a}$ eftersom $\sqrt{a} + \sqrt{a} = 2\sqrt{a} \neq 0$ eftersom både 2 och $\sqrt{a}$ inte är noll per antagande och vi får två olika lösningar. □

Från detta kan man exempelvis se att en elliptisk kurva som mest kan ha $2p$ antal punkter modulo p eftersom $y = \pm\sqrt{x^3 + ax + b}$ för varje val av x. Nu vill vi konstruera ett test för om ett tal har en kvadratrot eller inte.

Hjälpsats 6.3.3. *Exakt hälften av alla nollskilda element i* $\mathbb{Z}/(p)$ *har kvadraträtter om* $p > 2$.

Bevis. Enligt hjälpsats 6.3.2 har bilden av funktionen $x \mapsto x^2$ exakt $(p-1)/2$ nollskilda element eftersom det bara är x och $-x$ som har samma kvadrat. □

För de reella talen $\mathbb{R}$ vet vi sedan tidigare att de negativa talen inte har reella kvadraträtter och att de positiva har det. När man räknar modulo p är det svårt att säga vad som är positivt och vad som är negativt, så när exakt vet vi om ett tal har en kvadratrot då?

Hjälpsats 6.3.4. *För varje tal $a \in \mathbb{Z}/(p)$ med $p > 2$ gäller det att*

$$a^{\frac{p-1}{2}} = \begin{cases} 0 \text{ om } a = 0 \\ 1 \text{ om } a \text{ har två olika kvadratrötter} \\ -1 \text{ om } a \text{ inte har några kvadratrötter} . \end{cases}$$

Bevis. Påståendet $a^{\frac{p-1}{2}} = 0 \iff a = 0$ följer från beviset av sats 3.4.6 eftersom nollskilda tal har nollskild produkt och produkt med 0 alltid blir 0 enligt övningen 3.9. Kom ihåg från beviset av satsen 4.2.5 att för $a \neq 0$ gäller det att

$$a^{p-1} = 1.$$

Eftersom p är ett udda primtal så är $p-1$ jämnt. Alltså är $a^{\frac{p-1}{2}}$ ett väldefinierat tal och det är en kvadratrot till 1 eftersom $(a^{\frac{p-1}{2}})^2 = a^{p-1} = 1$. Nu vet vi från hjälpstats 6.3.4 att 1 har exakt två kvadratrötter. Så alltså måste $a^{\frac{p-1}{2}}$ vara ± 1. Notera att om a är en kvadrat $a = b^2$ så får vi att $a^{\frac{p-1}{2}} = b^{p-1} = 1$ enligt vårt bevis av Fermats lilla sats. Alltså har vi enligt hjälpsats 4.2.8 hittat *alla* rötter till $x^{\frac{p-1}{2}} - 1 = 0$, alltså de a som är kvadrater av något annat tal. Eftersom vi (enligt konjugatregeln och Fermats lilla sats) har vi att

$$x(x^{\frac{p-1}{2}} - 1)(x^{\frac{p-1}{2}} + 1) = x(x^{p-1} - 1) = 0,$$

så är alla resterande tal lösningar till $x^{\frac{p-1}{2}} + 1 = 0$. □

Okej, det är ju bra att veta när ett tal har en kvadratrot men hur beräknar vi den kvadratroten modulo p? Det är i allmänhet en svår fråga men det finns ett enkelt recept i fallet då p går att skriva på formen $p = 4k + 3$ för något $k \in \mathbb{N}$.

Hjälpsats 6.3.5. *Om $p = 4k + 3$ för $k \in \mathbb{N}$ är ett primtal och a har en kvadratrot modulo p så kan vi beräkna*

$$\sqrt{a} = a^{k+1}.$$

Bevis. Detta är en rättfram beräkning,

$$a = a \cdot 1 = a \cdot (a^{(p-1)/2}) = (a^{(p+1)/2}) = a^{(4k+4)/2} = (a^{k+1})^2.$$ □

Övningar

Övning 6.1. Rita ut kurvan $y^2 = x^3 + x$ över $\mathbb{Z}/(3)$ och $\mathbb{Z}/(7)$ som en delmängd till $\{0, 1, 2\}^2$ respektive $\{0, 1, \ldots, 6\}^2$. Kontrollera även att Hasse-Weil stämmer.

Övning 6.2. Rita ut kurvan $y^2 = x^3 + x$ över $\mathbb{Z}/(5)$ som en delmängd till $\{0, 1, 2, 3, 4\}^2$ som punkterna $\{(0, 0), (2, 0), (3, 0)\}$ och $\mathcal{O}$. Avgör om kurvan är isomorf med $\mathbb{Z}/(2) \times \mathbb{Z}/(2)$ eller $\mathbb{Z}/(4)$.

Övning 6.3 ($\star$)**.** Rita ut en elliptisk kurva modulo $p > 5$, använd fundamentalsatsen för ändligt genererade abelska grupper och Lagranges sats 3.2.4 för att visa vad för gruppstruktur den har genom att skriva den som en produkt av kopior av heltalen modulo p_i som i exempel 6.1.1.

Övning 6.4. Betrakta den elliptiska kurvan $y^2 = x^3 + x$ över $\mathbb{Z}/(7)$. Addera punkterna $P = (1, 4)$ och $Q = (3, 3)$ på denna kurva.

Övning 6.5. Betrakta den elliptiska kurvan $y^2 = x^3 + 3x$ över $\mathbb{Z}/(7)$. Addera punkterna $P = (1, 2)$ och $Q = (3, 1)$ på denna kurva.

Övning 6.6. Avgör vilka av elementen i $\mathbb{Z}/(17)$ som har (minst) en kvadratrot. Du behöver inte beräkna kvadratroten. Minräknare är tillåtna.

Övning 6.7. Avgör om vilka element i $\mathbb{Z}/(7)$ som har en kvadratrot eller inte. Om kvadraträtter finns, skriv ned dem.

Övning 6.8. Avgör om vilka element i $\mathbb{Z}/(11)$ som har en kvadratrot eller inte. Om kvadraträtter finns, skriv ned dem.

Övning 6.9. Testa dig fram med olika värden på $0 < m < 11$ för att hitta en lösning till instansen $2^m \equiv_{11} 9$ av det diskreta logaritmproblemet.

Övning 6.10. Visa att $m = 4$ är en lösning till $mP = (0, 1)$ av det diskreta logaritmproblemet för den elliptiska kurvan

$$x^3 + 1 \text{ modulo } 11$$

och $P = (2, 3)$

Övning 6.11. Använd *pq* formeln för att lösa ekvationen

$$x^2 + x + 1 = 0 \quad (\text{ modulo } 7)$$

Övning 6.12. Använd *pq* formeln för att lösa ekvationen

$$x^2 - 6x + 4 = 0 \quad (\text{ modulo } 11)$$

Övning 6.13. Ange ett andragradspolynom som har rötterna $5, 11$ modulo 19.

Övning 6.14. Tolka ordet *hemlighet* som ett element i $(\mathbb{Z}/(256))^9$ med hjälp av ASCII. Vad krypteras det till med ett Ceasarchiffer där $a = (5, 5, \ldots, 5)$?

7 Mordells sats

7.1 Introduktion

Som vi nämnde innan har ändligt genererade abelska grupper en väldigt enkel form. Vi upprepar satsen här.

Sats 7.1.1 (Klassifikation av ändligt genererade grupper)**.** *Låt G vara en ändligt genererad abelsk grupp. Då finns heltal $r, d_1, d_2, ...d_k$ så att*

$$G \cong \mathbb{Z}^r \times \mathbb{Z}/(d_1) \times \mathbb{Z}/(d_2) \times ... \times \mathbb{Z}/(d_k).$$

Talet r kallas för rangen av G.

Vi återvänder nu till diskussionen om elliptiska kurvor. Låt E nu vara en elliptisk kurva definierad av ekvationen

$$y^2 = x^3 + ax^2 + bx + c$$

där a, b och c är rationella tal. Det visar sig att gruppen av alla reella punkter av E, $E(\mathbb{R})$, bildar en grupp som är lite för stor för att ha särskilt intressant struktur. Om vi därmed betraktar E:s rationella punkter, $E(\mathbb{Q})$, får vi en mycket mer intressant struktur. Vi kommer att spendera resten av kapitlet med att bevisa följande sats.

Sats 7.1.2 (Mordells sats)**.** *Låt E vara en elliptisk kurva definierad över $\mathbb{Q}$. Då är $E(\mathbb{Q})$ ändligt genererad.*

Kombinerar vi satserna ovan får vi att det för varje elliptisk kurva E finns naturliga tal $r, d_1, ..., d_k$ så att

$$E(\mathbb{Q}) \cong \mathbb{Z}^r \times \mathbb{Z}/(d_1) \times \mathbb{Z}/(d_2) \times ... \times \mathbb{Z}/(d_k).$$

Talet r, alltså antalet kopior av $\mathbb{Z}$ som finns i $E(\mathbb{Q})$, kallas för rangen av E. Det visar sig att en slumpmässigt vald elliptisk kurva har rang ett 50% av tiden och rang noll 50% av tiden. Det finns elliptiska kurvor av högre rang än så, men de är mer sällsynta.

7.2 Bevis av Mordells sats

För att bevisa Mordells sats behöver vi först ett enkelt sätt att mäta ”komplexiteten” av ett rationellt tal.

Definition 7.2.1. Låt $x = a/b$ vara ett rationellt tal på förkortad form, alltså är a och b heltal utan någon gemensam delare. Vi definierar höjden av x, skrivet som $H(x)$, som $\max(|a|, |b|)$. △

Exempel 7.2.2. Vi har att $H(1/3) = 3$, $H(1/36) = 36$. Alltså mäter inte höjden storleken av ett rationellt tal, utan snarare ”komplexiteten”, hur mycket data det tar att representera talet. ▲

Den viktigaste egenskapen av höjdfunktionen är att det bara finns ett ändligt antal rationella tal med höjd mindre än eller lika med ett givet tal.

Definition 7.2.3. Om $P = (x, y)$ är en punkt på E säger vi att höjden av P, skrivet $H(P)$, är $H(x)$. Vi låter den logaritmiska höjden av P, skriven $h(P)$, vara $\log H(P)$. △

För att bevisa Mordells sats kommer vi att behöva tre hjälpsatser. Vi kommer tyvärr inte kunna ge ett helt fullständigt bevis, utan nöjer oss med att bevisa den första hjälpsatsen och ett specialfall av den sista hjälpsatsen. Den andra hjälpsatsen kräver inte några nya eller svåra metoder för att visa, men är väldigt tekniskt. Att vi inte kan ge ett fullständigt bevis av den tredje hjälpsatsen beror dock på att den kräver svårare metoder som är överkurs.

Första hjälpsatsen relaterar addition av två punkter till höjdfunktionen:

Hjälpsats 7.2.4. *Låt P_0 vara en punkt på E. Då finns en konstant C_0 som enbart beror på P_0 så att*

$$h(P + P_0) \leqslant 2h(P) + C_0$$

för alla punkter P på E.

Vi kommer också att behöva ett separat lemma för hur höjd beter sig med avseende på duplicering.

Hjälpsats 7.2.5. *Det finns en konstant C som bara beror på E så att*

$$h(2P) \geqslant 4h(P) - C$$

för alla punkter P på E.

Sist men inte minst kommer vi att behöva ett mycket svårare påstående om gruppen $E(\mathbb{Q})$.

Hjälpsats 7.2.6. *Indexet $[E(\mathbb{Q}) : 2E(\mathbb{Q})]$ är ändligt. Med andra ord, det finns ett ändligt antal punkter $P_0, P_1, ..., P_k$ så att alla punkter i $E(\mathbb{Q})$ kan skrivas på formen $P_i + 2P$ för någon punkt P i $E(\mathbb{Q})$ och $0 \leqslant i \leqslant k$.*

Vi avslutar avsnittet med att visa att Mordells sats gäller om vi får anta de här tre viktiga hjälpsatserna.

Bevis av sats 7.1.2. Låt P vara en punkt i $E(\mathbb{Q})$. Vi använder hjälpsats 7.2.6 och låter $Q_0, Q_1, Q_2, ..., Q_k$ vara punkter motsvarande satsen. Vi vet att det kommer att finnas ett index i_0 och en punkt P_0 i $E(\mathbb{Q})$ så att

$$P = Q_{i_0} + 2P_0.$$

Upprepar vi samma procedur för P_0 får vi ett index i_1 och en punkt P_1 så att

$$P_0 = Q_{i_1} + 2P_1.$$

Vi kan upprepa proceduren några gånger till och kommer fram till en sekvens $P_0, ..., P_m$ och $i_0, ..., i_m$ så att

$$P = Q_{i_0} + 2P_0, \quad P_0 = Q_{i_1} + 2P_1, ... \quad P_{m-1} = Q_{i_m} + 2P_m.$$

Lägger vi ihop ekvationerna $P = Q_{i_0} + 2P_0$ och $P_0 = Q_{i_1} + 2P_1$ får vi

$$P = Q_{i_0} + 2(Q_{i_1} + 2P_1) = Q_{i_0} + 2Q_{i_1} + 4P_1.$$

Fortsätter vi på samma sätt får vi

$$P = Q_{i_0} + 2Q_{i_1} + ... + 2^m Q_{i_m} + 2^{m+1} P_m.$$

Vi ser att P ligger i gruppen genererad av $Q_{i_0}, ..., Q_{i_m}$ och P_m. Om vi kan visa att P_m har en begränsad höjd för tillräckligt stora m kommer gruppen $E(\mathbb{Q})$ genereras av $Q_{i_0}, ..., Q_{i_m}$ och en mängd punkter av begränsad höjd, vilket det endast kan finnas en ändlig mängd av.

Vi vill nu jämföra höjden av två punkter i sekvensen, säg P_j och P_{j+1}, med målet att visa att $h(P_{j+1})$ är mindre än $h(P_j)$. Enligt hjälpsats 7.2.4 finns en konstant C_i för varje punkt Q_i så att $h(P - Q_i) \leqslant 2h(P) + C_i$ för alla punkter P i $E(\mathbb{Q})$. Om vi låter $C' = \max(C_i)_{i=0}^k$ så gäller

$$h(P - Q_i) \leqslant 2h(P) + C'$$

för alla P och i. Låt nu C vara konstanten från hjälpsats 7.2.5. Då gäller

$$4h(P_{j+1}) \leqslant h(2P_{j+1}) + C = h(P_j - Q_{i_{j+1}}) + C \leqslant 2h(P_j) + C + C'.$$

Vi skriver om det som

$$h(P_{j+1}) = \frac{1}{2}h(P_j) + \frac{C + C'}{4}.$$

Om $h(P_j) \geqslant C + C'$, får vi

$$h(P_{j+1}) = \frac{1}{2}h(P_j) + \frac{C + C'}{4} \leqslant \frac{1}{2}h(P_j) + \frac{1}{4}h(P_j) = \frac{3}{4}h(P_j).$$

Om vi upprepat multiplicerar ett tal med $3/4$ går det mot 0, så vi måste antingen kunna hitta m så att $h(P_m) \leqslant C + C'$ eller vara i en situation där

$$h(P_m) \leqslant \frac{3}{4}h(P_{m-1}) \leqslant ... \leqslant \left(\frac{3}{4}\right)^m h(P_0),$$

i vilket fall vi ändå måste ha $h(P_m) \leqslant C + C'$ för tillräckligt stora m.

Slutligen har vi att alla punkter P i $E(\mathbb{Q})$ kan skrivas på formen

$$P = a_0Q_0 + a_1Q_1 + ... + a_kQ_k + 2^mP_m$$

där $h(P_m) \leqslant C + C'$. Då det finns som mest ett ändligt antal punkter med $h(P) \leqslant C + C'$, måste $E(\mathbb{Q})$ vara ändligt genererad.

□

7.3 Addition och höjd

Vi börjar med att bevisa hjälpsats 7.2.4. Innan vi sätter igång med beviset visar vi att de rationella punkterna $P = (x, y)$ på E har en särskild form. Säg att $x = A/M$, och $y = B/N$, där bråken är på förkortad form, så A, N och B, M har inga gemensamma delare. Vi stoppar in våra koordinater i ekvationen som definierar E och får

$$\left(\frac{B}{N}\right)^2 = \left(\frac{A}{M}\right)^3 + a\left(\frac{A}{M}\right)^2 + b\left(\frac{A}{M}\right) + c.$$

Vi multiplicerar båda sidor med M^3N^2 och får

$$B^2M^3 = A^3N^2 + aA^2N^2M + bAN^2M^2 + cN^2M^3.$$

Vi ser att N^2 delar alla termer i högerledet och alltså då även B^2M^3, men B och N har inga gemensamma delare, så N^2 delar M^3.

Vi vill nu också visa motsatsen, att M^3 delar N^2. Vi märker först att M delar A^3N^2, eftersom att alla andra termer i ekvationen är delbara med M. Eftersom A och M inte har några gemensamma delare delar M talet N^2. Men då är samtliga termer utom A^3N^2 delbara med M^2, vilket innebär att även A^3N^2 är det och därmed delar M^2 talet N^2. Vi kan upprepa argumentet igen och får att M^3 delar N^2.

Om N^2 delar M^3 och vice versa måste $N^2 = M^3$. Eftersom M^2 delar N^2 måste M dela N, vilket betyder att $e = N/M$ är ett heltal. Vi får att

$$e^2 = \frac{N^2}{M^2} = \frac{M^3}{M^2} = M, \quad e^3 = e^2 \cdot e = M\frac{N}{M} = N.$$

Alltså är $x = A/e^2$, $y = B/e^3$.

Vi får nu att $H(P) = H(x) = H(A/e^2) = \max(|A|, e^2)$. Det visar sig att $H(P)$ också begränsar y-koordinaten, på det sättet att det finns en konstant K som bara beror på E så att $|B| \leqslant KH(P)^{3/2}$. För att visa det betraktar vi ekvationen som definierar E multiplicerad med e^6

$$B^2 = A^3 + aA^2e^2 + bAe^4 + ce^6.$$

Enligt triangelolikheten gäller

$$|B|^2 \leqslant |A|^3+|aA^2e^2|+|bAe^4|+|ce^6| \leqslant H(P)^3+|a|H(P)^3+|b|H(P)^3+|c|H(P)^3.$$

Alltså gäller påståendet med $K = (1 + |a| + |b| + |c|)^{1/2}$. Nu är vi redo för hjälpsatsen:

Bevis av hjälpsats 7.2.4. Vi kan ignorera ett ändligt antal val av P genom att välja ett litet större värde på C_0. Alltså antar vi att $P \notin \{\mathcal{O}, P_0, -P_0\}$. Låt $P_0 = (x_0, y_0)$ och $P = (x, y)$. Låt nu $P_0 + P = Q = (x_1, y_1)$. Vi har från additionsformeln att

$$x_1 + x_0 + x = \lambda^2 - a, \quad \lambda = \frac{y - y_0}{x - x_0}.$$

Alltså gäller

$$x_1 = \frac{(y - y_0)^2}{(x - x_0)^2} - a - x - x_0 = \frac{(y - y_0)^2 - (x - x_0)^2(x + x_0 + a)}{(x - x_0)^2}.$$

Om vi expanderar uttrycket ovan får vi $y^2 - x^3$ som ledande term, vilken vi kan byta ut mot $ax^2 + bx + c$. Sammanlagt får vi ett uttryck på formen

$$x_1 = \frac{c_1y + c_2x^2 + c_3x + c_4}{c_5x^2 + c_6x + c_7}.$$

Vi kan anta att c_i är heltal eftersom att vi annars kan multiplicera med deras nämnare. Vi stoppar in $x = A/e^2$ och $y = B/e^3$, och får

$$x_1 = \frac{c_1Be + c_2A^2 + c_3Ae^2 + c_4e^4}{c_5A^2 + c_6Ae^2 + c_7e^4}.$$

Här är både täljare och nämnare heltal, även om bråket kanske inte är reducerat. Vi får nu

$$H(x_1) \leqslant \max(|c_1Be + c_2A^2 + c_3Ae^2 + c_4e^4|, |c_5A^2 + c_6Ae^2 + c_7e^4|).$$

Vi har sedan innan att

$$e \leqslant H(P)^{1/2}, \quad A \leqslant H(P), \quad B \leqslant KH(P)^{3/2}.$$

Alltså gäller

$$|c_1Be+c_2A^2+c_3Ae^2+c_4e^4| \leqslant |c_1Be|+|c_2A^2|+|c_3Ae^2|+|c_4e^4| \leqslant (|c_1K|+|c_2|+|c_3|+|c_4|)H(P)^2$$

och på samma sätt

$$|c_5A^2 + c_6Ae^2 + c_7e^4| \leqslant |c_5A^2| + |c_6Ae^2| + |c_7e^4| \leqslant (|c_5| + |c_6| + |c_7|)H(P)^2.$$

Om vi låter $M = \max(|c_5| + |c_6| + |c_7|, |c_1K| + |c_2| + |c_3| + |c_4)$ så gäller

$$H(P + P_0) = H(x_1) \leqslant MH(P)^2.$$

Om vi tar logaritmer av båda sidor av ekvationen får vi den efterfrågade olikheten med $C_0 = \log M$. □

7.4 En mystisk homomorfi

Vi kommer att behöva lite hjälp för bevisa hjälpsats 7.2.6. Därmed introducerar vi en ny elliptisk kurva för att hjälpas åt med att visa satsen. Här kan vi inte utföra beviset för en generell elliptisk kurva utan kommer att behöva anta att $f(x)$ har åtminstone en rationell rot. Om x_0 är det rationella talet så att $f(x_0) = 0$, kan vi skifta koordinatsystemet så att $(x_0, 0)$ är i origo, det kommer inte att påverka gruppstrukturen hos $C(\mathbb{Q})$. Alltså kan vi anta att $f(0) = 0$, eller att $f(x) = x^3 + ax^2 + bx$ för något $a, b \in \mathbb{Q}$.

Den nya elliptiska kurvan som kommer att vara viktig för våra beräkningar, skriven $\overline{C}$, är given av följande ekvation

$$y^2 = x^3 + \bar{a}x^2 + \bar{b}x,$$

där

$$\bar{a} = -2a, \quad \bar{b} = a^2 - 4b.$$

För att få någon sorts förståelse för den här nya elliptiska kurvan, kan det vara bra att märka att den har samma form som vår tidigare elliptiska kurva, alltså att $\bar{f}(0) = 0$. Om vi vidare applicerar konstruktion två gånger i rad, får vi en kurva $\overline{\overline{C}}$, given av ekvationen

$$y^2 = x^3 + \bar{\bar{a}}x^2 + \bar{\bar{b}}x$$

där

$$\bar{\bar{a}} = -2\bar{a} = 4a, \quad \bar{\bar{b}} = 4a^2 - 4(a^2 - 4b) = 16b.$$

Alltså ges $\overline{\overline{C}}$ av ekvationen

$$y^2 = x^3 + 4ax^2 + 16bx.$$

Genom att skala om (x, y) som $(4x, 8y)$ får vi ekvationen för C skalat med en faktor av 64. Alltså är kurvan $\overline{\overline{C}}$ väldigt lik C, och har framför allt samma gruppstruktur.

Vi vill nu definiera en homomorfi $\varphi\colon C(\mathbb{Q}) \to \overline{C}(\mathbb{Q})$. Den kommer vidare ha egenskapen att om vi applicerar konstruktionen igen och får en homomorfi

$$\bar{\varphi}\colon \overline{C} \to \overline{\overline{C}} \cong C(\mathbb{Q})$$

kommer $\varphi \circ \overline{\varphi}$ att vara multiplikation med 2. Vi kommer alltså ha faktoriserat multiplikation med 2 i två mindre homomorfier, vilket är viktigt då vårt huvudintresse är $[C(\mathbb{Q}) : 2C(\mathbb{Q})]$.

Vi definierar $\varphi(x, y) = (\bar{x}, \bar{y})$ på följande sätt,

$$\bar{x} = x + a + \frac{b}{x} = \frac{y^2}{x^2}, \quad \bar{y} = y\frac{x^2 - b}{x^2}.$$

För att φ ska vara en funktion från C till $\overline{C}$, måste $(\bar{x}, \bar{y})$ ligga på $\overline{C}$, vilket kan visas med en beräkning.

Notera att formeln ovan inte fungerar då $(x, y) = (0, 0)$, en punkt vi antagit ligger på kurvan. Vi säger därmed att

$$\varphi(0,0) = \overline{\mathcal{O}}, \quad \varphi(\mathcal{O}) = \overline{\mathcal{O}}.$$

Det är inte uppenbart alls från vår definition att φ är en homomorfi, faktum är att definition är någorlunda omotiverad. Men det enda som återstår för att visa att φ är en homomorfi är direkta beräkningar, och ganska långa sådana. Därmed går vi inte igenom beviset i det här kompendiet.

Sats 7.4.1. *Låt C och $\overline{C}$ vara elliptiska kurvor givna av ekvationer*

$$y^2 = x^3 + ax^2 + bx, \quad y^2 = x^3 + \bar{a}x^2 + \bar{b}x,$$

där

$$\bar{a} = -2a, \quad \bar{b} = a^2 - 4b.$$

(i) *Då finns en homomorfi $\varphi\colon C(\mathbb{Q}) \to \overline{C}(\mathbb{Q})$ given av*

$$\varphi(x,y) = \left(\frac{y^2}{x^2}, y\frac{x^2-b}{x^2}\right) \text{ om } (x,y) \neq \mathcal{O}, (0,0)$$

$$\varphi(0,0) = \overline{\mathcal{O}}, \quad \varphi(\mathcal{O}) = \overline{\mathcal{O}}.$$

(ii) *Om vi upprepar processen får vi en homomorfi*

$$\bar{\varphi}\colon \overline{C}(\mathbb{Q}) \to \overline{\overline{C}}(\mathbb{Q}).$$

Gruppen $\overline{\overline{C}}(\mathbb{Q})$ är isomorf med gruppen $C(\mathbb{Q})$ via isomorfin $(x, y) \to (x/4, y/8)$. Sätter vi ihop de två homomorfierna får vi en homomorfi

$$\psi\colon \overline{C}(\mathbb{Q}) \to C(\mathbb{Q})$$

given av

$$\psi(x,y) = \left(\frac{y^2}{4x^2}, y\frac{x^2-\bar{b}}{8x^2}\right) \text{ om } (x,y) \neq \overline{\mathcal{O}}, (0,0)$$

$$\psi(0,0) = \mathcal{O}, \quad \psi(\overline{\mathcal{O}}) = \mathcal{O}\,.$$

(iii) *Vi har att $\psi \circ \varphi(P) = 2P$.*

7.5 Bevis av sista hjälpsatsen

Med hjälp av den nya homomorfin kan vi äntligen bevisa den sista hjälpsatsen. För att göra notationen enklare, skriver vi Γ för $C(\mathbb{Q})$, $\overline{\Gamma}$ för $\overline{C}(\mathbb{Q})$. Något som kommer att vara viktigt för framtida beräkningar är bilden av φ, alltså $\varphi(\Gamma)$. Vi noterar några viktiga saker om den gruppen.

Hjälpsats 7.5.1. *Låt $P = (x, y) \in \overline{\Gamma}$. Då är $P \in \varphi(\Gamma)$ om och endast om x är kvadraten av ett rationellt tal.*

Bevis. Om

$$P = \varphi(X, Y) = \left(\frac{Y^2}{X^2}, Y\frac{X^2 - b}{X^2} \right)$$

följer det direkt att $x = (Y/X)^2$ är en kvadrat av ett rationellt tal. Säg istället att $x = q^2$ där q är ett rationellt tal. Vi påstår att punkterna

$$Q_1 = (x_1, y_1) = \left(\frac{1}{2}\left(q^2 - a + \frac{y}{q} \right), x_1 q \right),$$

$$Q_2 = (x_2, y_2) = \left(\frac{1}{2}\left(q^2 - a - \frac{y}{q} \right), -x_2 q \right).$$

ligger i Γ och att vidare $\varphi(Q_1) = P$. För att visa det beräknar vi först

$$x_1 x_2 = \frac{1}{4}\left((q^2 - a)^2 - \left(\frac{y}{q} \right)^2 \right) = \frac{1}{4}\left((x - a)^2 - \frac{y^2}{x} \right)$$

$$= \frac{1}{4}\left(\frac{x^3 - 2ax^2 + a^2 x - y^2}{x} \right) = \frac{1}{4}\left(\frac{4b}{x} \right) = b.$$

Här använder vi att $y^2 = x^3 - 2ax^2 + (a^2 - 4b)x$ eftersom (x, y) ligger på $\overline{\Gamma}$.

Att visa att Q_1 ligger på C motsvarar att visa att

$$\left(\frac{y_1}{x_1} \right)^2 = x_1 + a + \frac{b}{x_1}.$$

Vi visade just att $x_1 x_2 = b$, så $x_2 = b/x_1$. Från definitionerna gäller också att $y_1/x_1 = q$. Alltså är ekvationen ovan ekvivalent med

$$q^2 = x_1 + a + x_2.$$

Att visa att den sista likheten stämmer är en enkel beräkning.

Det återstår att visa att $\varphi(Q_1) = (x, y)$, vilket innebär att visa att

$$\frac{y_1^2}{x_1^2} = x, \quad \frac{y_1(x_1^2 - b)}{x_1^2} = y.$$

Den första likheten är väldigt enkel, och vi kan skriva vänsterledet av den andra likheten som

$$\frac{y_1(x_1^2 - b)}{x_1^2} = \frac{x_1 q(x_1^2 - x_1 x_2)}{x_1^2} = q(x_1 - x_2).$$

Men att $q(x_1 - x_2) = y$ är igen en enkel likhet. Alltså är $\varphi(Q_1) = P$.

□

Hjälpsats 7.5.2. *Låt $S = \{p_1, p_2, ..., p_k\}$ vara mängden av primtal som delar b. Om $P = (x, y) \in \Gamma$ så kan vi skriva x på formen*

$$x = \pm p_1^{n_1} \cdot ... \cdot p_k^{n_k} q^2$$

där $n_i \in \{0, 1\}$ och q är ett rationellt tal.

Bevis. Vi vet att (x, y) har formen $x = m/e^2$ och $y = n/e^3$. Om vi substituerar in det i ekvationen för vår elliptiska kurva får vi

$$n^2 = m^3 + am^2e^2 + bme^4 = m(m^2 + ame^2 + be^4).$$

Vi låter $d = \text{sgd}(m, m^2 + ame^2 + be^4)$. Det här talet kommer såklart dela d, och därmed kommer det också dela be^4. Men e/m är förkortat, så d har ingen gemensam delare med e. Alltså delar d talet b. Eftersom produkten av m och $(m^2 + ame^2 + be^4)$ är en kvadrat, kommer de primtal som delar m men inte $(m^2 + ame^2 + be^4)$ att ha en jämn exponent. Därmed har m formen

$$m = \pm(N)^2 p_1^{n_1} \cdot ... \cdot p_k^{n_k}$$

där $n_i \in \{0, 1\}$ och N är ett heltal.

Nu är $x = m/e^2$, så

$$x = \pm p_1^{n_1} \cdot ... \cdot p_k^{n_k} (N/e)^2$$

har den efterfrågade formen. □

Vårt slutmål är att visa att 2Γ har ändligt index i Γ. Det visar sig att det räcker att visa att indexen av $\varphi(\Gamma)$ i $\overline{\Gamma}$ och $\psi(\overline{\Gamma})$ i Γ är ändliga. Då de här olika påståenden är symmetriska räcker det att visa att $(\Gamma : \psi(\overline{\Gamma}))$ är ändligt.

Vårt mål kommer att vara att hitta en homomorfi från Γ till en ändlig grupp, som dessutom har kärna $\varphi(\overline{\Gamma})$. Vi låter $p_1, p_2, ..., p_k$ vara de primtal som delar b. Vi låter A vara gruppen $\{\pm 1\} \times (\mathbb{Z}/(2))^k$. Gruppen $\{\pm 1\}$ är såklart isomorf med $\mathbb{Z}/(2)$, det här är för att göra notationen enklare. Vi definierar nu en funktion $\alpha \colon \Gamma \to A$ genom

$$\alpha(x, y) = (\{\pm 1\}, n_1, n_2, ..., n_k),$$

där $x = \pm p_1^{n_1} \cdot ... \cdot p_k^{n_k} q^2$ som i hjälpsats 7.5.2.

Sats 7.5.3. (i) *Funktionen $\alpha \colon \Gamma \to A$ är en homomorfi.*

(ii) *Kärnan av α är $\psi(\overline{\Gamma})$.*

Bevis. (i) Funktionen α respekterar inverser eftersom

$$\alpha(-P) = \alpha(x, -y) = \alpha(x),$$

och $a = -a$ för alla element $a \in A$. Alltså räcker det att visa att om $P + Q + R = \mathcal{O}$ i $C(\mathbb{Q})$ så är $\alpha(P) + \alpha(Q) + \alpha(R) = 0$. Vi vet att $P + Q + R = \mathcal{O}$ om och endast om P, Q och R ligger på en linje. Om linjen är $y = \lambda x + \nu$ och x-koordinaterna av P, Q, R är x_1, x_2, x_3, har vi sett att x_1, x_2, x_3 är lösningarna till ekvationen

$$x^3 + (a - \lambda^2)x^2 + (b - 2\lambda\nu)x + (c - \nu^2) = 0.$$

Med andra ord,

$$(x - x_1)(x - x_2)(x - x_3) = x^3 + (a - \lambda^2)x^2 + (b - 2\lambda\nu)x + (c - \nu^2).$$

Om vi expanderar vänsterledet och tittar på konstanten på båda sidor, ser vi att

$$x_1x_2x_3 = \nu^2 - c.$$

Men $c = 0$ i vårt fall! Alltså är $x_1x_2x_3 = \nu^2$ och är alltså en kvadrat. Men då är primtalsexponenterna i $x_1x_2x_3$ alla jämna, så

$$\alpha(x_1) + \alpha(x_2) + \alpha(x_3) = \alpha(x_1x_2x_3) = 0.$$

(ii) Om $\alpha(x) = 0$, så är $n_i = 0$ för $1 \leqslant i \leqslant k$, så $x = q^2$, alltså är x en kvadrat. Det här händer precis då $(x, y) \in \psi(\overline{\Gamma})$ enligt hjälpsats 7.5.1.

□

Vi har nu en homomorfi $\psi\colon \Gamma \to A$ med kärna $\overline{\Gamma}$, och A är en ändlig grupp. Därmed vet vi att $(\Gamma : \psi(\overline{\Gamma}))$ är ändlig.

Det enda som återstår är nu att visa att om $(\Gamma : \psi(\overline{\Gamma}))$ och $(\overline{\Gamma} : \varphi(\Gamma))$ är ändliga, är även

$$(\Gamma : \psi(\varphi(\Gamma))) = (\Gamma : 2\Gamma)$$

ändlig.

Sats 7.5.4. *Låt A och B vara abelska grupper, och anta att det finns homomorfier $\varphi\colon A \to B$ och $\psi\colon B \to A$ så att*

$$\varphi \circ \psi(b) = 2b, \quad \psi \circ \varphi(a) = 2a.$$

Anta också att $(A : \psi(B))$ och $(B : \varphi(A))$ är ändliga. Då är även $(A : 2A)$ ändligt.

Bevis. Enligt antaganden finns det ett ändligt antal $b_1, ..., b_n$ så att $a_i + \varphi(A)$ täcker B, och på samma sätt en motsvarande mängd $a_1, ..., a_m$ för A. Vi påstår att mängden

$$a_i + \psi(b_j) + 2A, \quad 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n$$

täcker A. Om $a \in A$, finns det a_i och $b \in B$ så att $a = a_i + \psi(b)$. Men det finns också b_j och $a' \in A$ så att $b = b_j + \varphi(a')$. Lägger vi ihop de här två likheterna får vi

$$a = a_i + \psi(b_j + \varphi(a')) = a_i + \psi(b_j) + \psi \circ \varphi(a') = a_i + \psi(b_j) + 2a'$$

vilket skulle visas. □

Övningar

Övning 7.1. Låt C vara en elliptisk kurva med heltalskoefficienter. En välkänd sats om elliptiska kurvor är Nagell-Lutz sats, den säger att om (x, y) är en rationell punkt på C av ändligt ordning så är x, y heltal. Vidare är y-koordinaten 0 eller delar Δ_C, diskriminanten av vår kurva.

Använd satsen ovan för att hitta alla punkter på av ändlig ordning på kurvorna

(i) $y^2 = x^3 + x$.

(ii) $y^2 = x^3 + x + 1$.

(iii) $y^2 = x^3 + 1$.

Övning 7.2. Visa att mängden punkter med höjd $H(x)$ mindre än N innehåller som mest $2N^2 + N$ element.

Övning 7.3. Betrakta de rationella talen, $\mathbb{Q}$.

(i) Visa att den additiva gruppen $(\mathbb{Q}, +)$ inte är ändligt genererad.

(ii) Är den multiplikativa gruppen $(\mathbb{Q} - \{0\}, *)$ ändligt genererad?

Övning 7.4 ($\star\star$)**.** Betrakta den singulära kurvan C givet av $y^2 = x^3$.

(i) Visa att $(0, 0)$ är C:s enda singularitet.

(ii) Visa att om du betraktar $C' = C - \{(0, 0)\}$ och definierar addition av punkter i C' på det vanliga, att C' har en gruppstruktur.

(iii) Visa att $C(\mathbb{Q})$ inte är ändligt genererad.

8 Elliptiska kurvor över de komplexa talen

8.1 Introduktion

Materialet i det här kapitlet beror mycket på analys, alltså derivator och integraler. På grund av det kommer vi inte att kunna vara lika noggranna med att bevisa alla påståenden, utan vi kommer att vifta lite med händerna vid många tillfällen.

8.2 De komplexa talen

Matematiker har sedan länge märkt att det finns en del problem med de tal vi normalt sätt använder oss av, nämligen kroppen $\mathbb{R}$, de tal som finns på tallinjen. Ett problem med $\mathbb{R}$ som gör kroppen ”ofullständig” är att det finns vanliga polynomekvationer utan en lösning i $\mathbb{R}$. Exempelvis har ekvationen

$$x^2 + 1 = 0$$

ingen reell lösning.

Det visar sig att trots att kroppen $\mathbb{R}$ inte är fullständig på det sätt vi vill, är den inte långt ifrån att vara det. Faktum är att det räcker att lägga till en lösning på ekvationen ovan, så får vi en kropp som är fullständig på precis det sätt vi vill.

Definition 8.2.1. Vi definierar kroppen av komplexa tal, skrivet $\mathbb{C}$, som tal på formen

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

Låt $a + bi, c + di \in \mathbb{C}$. Vi definierar addition genom

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

och multiplikation genom

$$(a + bi) \cdot (c + di) = (ac - bd) + (ad + bc)i. \qquad \triangle$$

Med lite jobb kan vi visa att de komplexa talen bildar en kropp. Vi kan se från definitionen av multiplikation att eftersom $i = 0 + 1 \cdot i$ så gäller

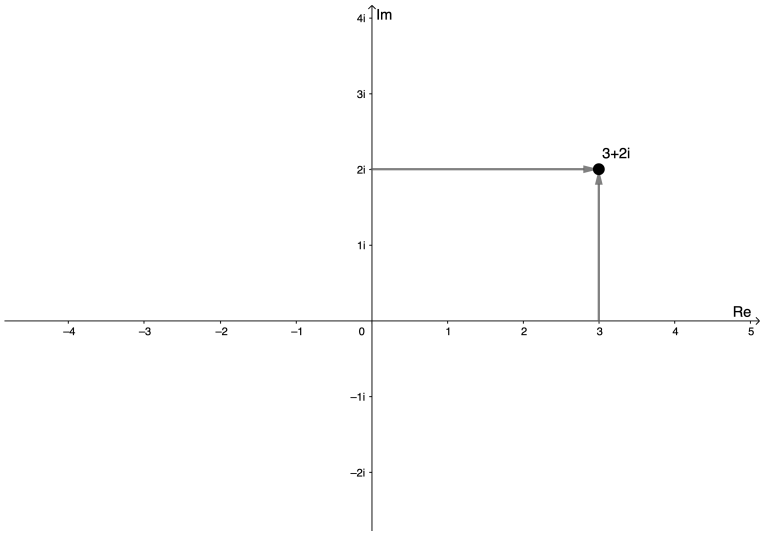
$$i^2 = (0 + 1 \cdot i)(0 + 1 \cdot i) = (0 \cdot 0 - 1 \cdot 1) + (0 \cdot 1 + 1 \cdot 0)i = -1.$$

Alltså uppfyller i ekvationen $x^2 + 1 = 0$. Det visar sig att det räcker att lägga till i för att kroppen $\mathbb{C}$ ska bli ”fullständi”:

Sats 8.2.2 (Algebrans fundamentalsats)**.** *Varje ekvation på formen*

$$a_0 + a_1 x + ... + a_n x^n = 0$$

där $a_i \in \mathbb{C}$ har en lösning i $\mathbb{C}$.



Figur 8.1: Komplexa talplanet, där y-axeln är den imaginära axeln och x-axeln är den reella axeln.

En kropp med den här egenskapen kallas för en algebraiskt sluten kropp. Kroppar med den här egenskapen är ofta mer ”naturliga” när man undersöker lösningar till polynomekvationer, och vissa påståenden blir mycket enklare att formulera.

Om man visualiserar $\mathbb{R}$ som en linje, kan man visualisera $\mathbb{C}$ som ett plan, det komplexa talplanet.

8.3 Weierstrass funktion

Låt E vara en elliptisk kurva på formen

$$y^2 = 4x^3 + g_1 x + g_2.$$

Vi låter

$$E(\mathbb{C}) = \{(x, y) \in \mathbb{C}^2 \mid y^2 = 4x^3 + g_1 x + g_2\} \cup \{\mathcal{O}\},$$

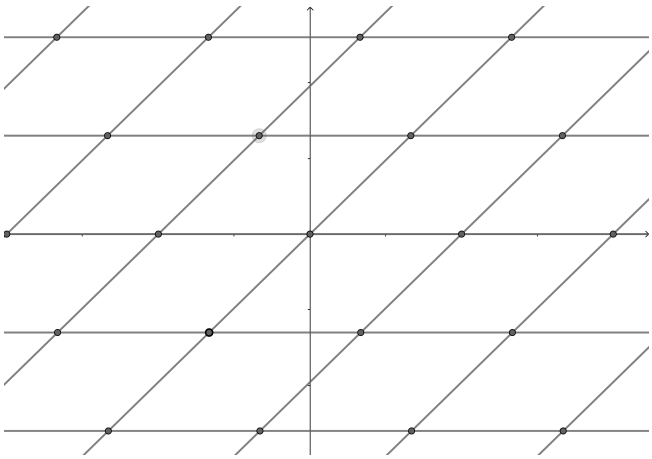
alltså de punkter med båda koordinater komplexa tal som löser ekvationen. Notera att punkterna i $\mathbb{C}$ oftast representeras som ett plan, så en graf till E skulle vara en tvådimensionell figur i ett fyradimensionellt rum. Alltså är det svårt att visualisera de komplexa lösningarna. Det finns dock saker vi kan säga om dem, exempelvis vilken form de har. Först behöver vi lite definitioner.

Definition 8.3.1. Ett gitter är en delgrupp Λ till $(\mathbb{C}, +)$ som är på formen

$$\mathbb{Z}\omega_1 + \mathbb{Z}\omega_2 = \{a\omega_1 + b\omega_2 | a, b \in \mathbb{Z}\}$$

där ω_1 och ω_2 är komplexa tal. Vi kräver ytterligare att punkterna ω_1, ω_2 och 0 inte ligger på en linje.

△



Figur 8.2: Ett exempel på ett gitter i xy-planet med $\omega_1 = 3$ och $\omega_2 = 2 + 2i$.

Det intuitiva sättet att se på ett gitter är som ett oändligt upprepande mönster av punkter, där punkterna utgör korsningar i ett rutnät.

Anledningen till att vi kräver att ω_1, ω_2 och 0 inte ska vara på en linje är för att Λ inte ska ligga på en linje utan istället fylla ut hela planet.

Vi vill nu definiera en komplexvärd funktion $\wp$ som är definierad överallt utom just på ett givet gitter Λ. Här innebär komplexvärd att det är en funktion som man stoppar in komplexa tal i och får ut komplexa tal, alltså en funktion $\wp\colon \mathbb{C} \to \mathbb{C}$. Vi går tillväga på följande sätt:

Definition 8.3.2. Låt Λ vara ett gitter. Då definierar vi Weierstrass funktion för Λ, skriven $\wp_\Lambda$ eller $\wp$, som en funktion $\wp\colon \mathbb{C} \to \mathbb{C}$ given av

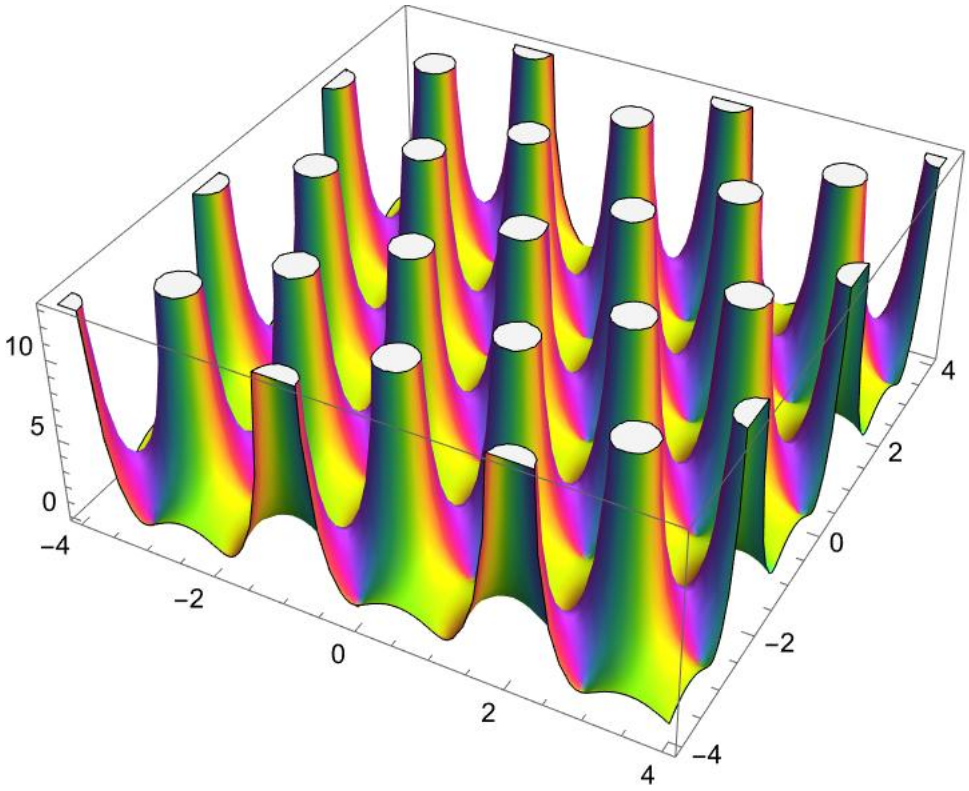
$$\wp(z) = \frac{1}{z^2} + \sum_{\lambda \in \Lambda} \frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2}. \qquad \triangle$$

Här är mängden vi summerar över, Λ, en oändlig mängd, så vi tolkar summan ovan som ett gränsvärde. Man måste såklart egentligen visa att den oändliga serien faktiskt konvergerar. Man kan se från definitionen av $\wp$ att värdet $\wp(z)$ går mot oändligheten då z närmar sig element i λ, de kallas för poler till $\wp$. Funktionen $\wp$ kallas för Weierstrass funktion efter matematikern Karl Weierstrass. Det finns flera funktioner döpta efter Weierstrass, så om det behövs extra förtydligande kallar man ibland funktionen för Weierstrass elliptiska funktion.

En annan sak som följer direkt från definitionen av $\wp$ är att om $\omega \in \Lambda$ så är

$$\wp(z+\omega) = \frac{1}{(z+\omega)^2} + \sum_{\lambda \in \Lambda} \frac{1}{(z+\omega-\lambda)^2} - \frac{1}{\lambda^2} = \wp(z).$$

Funktionen är det man kallar för periodisk med avseende på Λ, den upprepar sig om man förskjuter det komplexa planet med ett tal i Λ.



Figur 8.3: En graf av $\wp(z)$ för något val av Λ. Här representerar höjden på grafen beloppet av $\wp(z)$ och färgen är argumentet.

Det går att definiera derivator för komplexvärda funktioner, och den visar sig ha många av de egenskaper man förväntar sig. Exempelvis är derivatan av x^n funktionen nx^{n-1} om $n \neq -1$. Framför allt är derivatan av $1/z^2$ lika med funktionen $-2/z^3$. Man kan härleda liknande formler för andra derivator. Beräknar vi derivatan av $\wp$ ser vi att

$$\wp'(z) = -\frac{2}{z^3} + \sum_{\lambda \in \Lambda} \frac{2}{(\lambda - z)^3}.$$

Om vi definierar

$$G_n = \sum_{\lambda \in \Lambda} \lambda^{-n}$$

kan man med lite ihärdiga räkningar komma fram till att

$$(\wp'(z))^2 = 4\wp(z)^3 + g_1\wp(z) + g_2$$

där $g_1 = -60G_4$, $g_2 = -140G_6$. Det visar sig alltså att funktionen $(\wp, \wp') \colon \mathbb{C} \to \mathbb{C}^2$ är en funktion vars bild landar precis i lösningarna till en elliptisk kurva.

Säg att $\lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$. Eftersom $\wp$ (och därmed även $\wp'$) uppfylller $\wp(z+\omega_1) = \wp(z)$, $\wp(z + \omega_2) = \wp(z)$, kommer funktionen $\wp$ att upprepa sig om vi går tillräckligt långt i riktningen av ω_1 eller riktningen av ω_2. Alltså räcker det att fokusera på ett parallellogram med hörn i 0, $\omega_1, \omega_2, \omega_1 + \omega_2$.

Eftersom $\wp(z + \omega_1) = \wp(z)$ kommer $\wp$ att anta samma värden på linjen genom $0, \omega_2$ som på linjen genom $\omega_1, \omega_1 + \omega_2$. På samma sätt kommer $\wp$ att anta samma värden på linjen genom $0, \omega_1$ som på linjen genom $\omega_2, \omega_1 + \omega_2$. Alltså kommer $\wp$ att klistra ihop motstående sidor på vårt parallellogram. Samma sak gäller för $\wp'$. Om du föreställer dig att du gör det med en bit papper, får du först en cylinder, och sedan en doughnut. Alltså kommer bilden av vår rektangel under $(\wp, \wp')$ att se ut som en doughnut. Det matematiska namnet för en doughnut är en torus.

Det går att visa att funktionen $\wp \colon \mathbb{C} \to C$ också är surjektiv, den träffar alltså alla punkter på vår elliptiska kurva. Då vi vet att bilden av $\mathbb{C}$ under $\wp$ ser ut som en donut/torus kommer även vår elliptisk kurva $\mathbb{C}$ att se ut som en donut/torus. Tyvärr är det en donut i 4 dimensioner, vilket inte är särskilt lätt att föreställa sig!

Det finns en till egenskap hos Weierstrass funktion som är något svårare att visa.

Sats 8.3.3. *Låt $x, y \in \mathbb{C}$ och $\wp$ vara Weierstrass funktion för något Λ. Låt C vara den elliptiska kurva som funktionen $(\wp, \wp') \colon \mathbb{C} \to \mathbb{C}^2$ landar i. Då gäller*

$$(\wp(x + y), \wp'(x + y)) = (\wp(x), \wp'(x)) + (\wp(y), \wp'(y))$$

där additionen i högerled är addition av punkter på C. Med andra ord, $(\wp, \wp') \colon \mathbb{C} \to C(\mathbb{C})$ är en homomorfi.

Det här visar på att gruppstrukturen av $C(\mathbb{C})$ inte är särskilt komplicerad. För att beskriva den i detalj definierar vi gruppen S som mängden

$$[0,1) = \{x \in \mathbb{R} | 0 \leqslant x < 1\}.$$

Vi ger den additionen

$$a + b = \begin{cases} a + b \text{ om } a + b < 1 \\ a + b - 1 \text{ annars} \end{cases}.$$

Ett sätt att tänka på S är som en cirkel där varje punkt representerar en vinkel, och addition är addition av vinklar. Om man kommer över ett helt varv, börjar man om igen.

Som tidigare kan vi skriva $\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$, och betrakta C som ett parallellogram med hörn i $0, \omega_1, \omega_2$ och $\omega_1 + \omega_2$. Eftersom $(\wp, \wp')$ är en homomorfi, kommer också att addition av punkter i C att representeras av addition av komplexa tal i vårt parallelogram, där vi möjligtvis subtraherar ω_1 eller ω_2 från resultatet. Vi kan definiera en homomorfi $f \colon S \times S \to C(\mathbb{C})$ där vi låter

$$f(a, b) = \wp(a\omega_1 + b\omega_2).$$

Med lite jobb kan man visa att det faktiskt är en isomorfi. Alltså gäller det att

$$C(\mathbb{C}) \cong S \times S.$$

8.4 Punkter av ändlig ordning

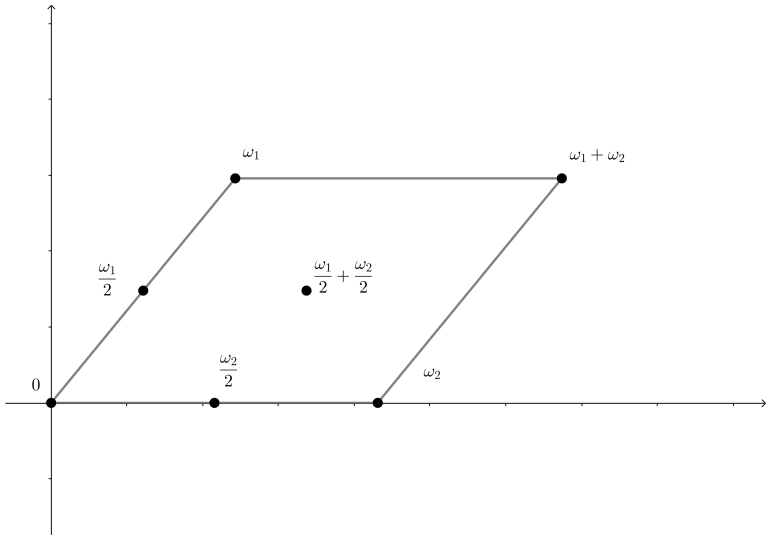
Låt G vara en godtycklig abelsk grupp och $n \in \mathbb{N}$. Mängden punkter $g \in G$ så att $g^n = 1$ kommer att bilda en delgrupp till G. Vi skriver $G[n]$ för den här delgruppen. De kommer att innehålla de punkter i G som har ordning n, men även 1 och de punkter som har en ordning som delar n.

Hur ser $C(\mathbb{C})[n]$ ut då C är en godtycklig elliptisk kurva? Vi besvarar först samma fråga för S. Om vi låter $s \in S$ så är $ns = 0$ om och endast om ns är ett heltal. Alltså kommer de punkter s så att $ns = 0$ att vara $\{0, 1/n, 2/n, ..., (n-1)/n\}$. Den här delgruppen är isomorf med $\mathbb{Z}/(n)$, så $S[n] \cong \mathbb{Z}/(n)$. Eftersom $C(\mathbb{C}) \cong S \times S$, får vi att

$$C(\mathbb{C})[n] \cong S[n] \times S[n] \cong (\mathbb{Z}/(n))^2.$$

Det finns därmed alltid n^2 punkter i $C(\mathbb{C})[n]$, till skillnad från de reella punkterna där vi såg att $C(\mathbb{R})[2]$ kunde innehålla 2 eller 4 punkter.

Det här är ett exempel varför det är mer naturligt att betrakta komplexa lösningar. När vi betraktade lösningar över $\mathbb{R}$ var strukturen av $C(\mathbb{R})[n]$ olika



Figur 8.4: En visuell representation av $C[2] = \{0, \omega_1/2, \omega_2/2, (\omega_1 + \omega_2)/2)$.

beroende på C, men $C(\mathbb{C})[n]$ ser alltid likadan ut.

Om vi återgår till att se $C(\mathbb{C})$ som ett parallellogram med hörn i $0, \omega_1, \omega_2$ och $\omega_1 + \omega_2$, kommer $C(\mathbb{C})[n]$ att utgöras av

$$\frac{i}{n}\omega_1 + \frac{j}{n}\omega_2,$$

för $0 \leqslant i, j < n$.

8.5 Vad har en elliptisk kurva med ellipser att göra?

En av de mest förvirrande delarna av teorin om elliptiska kurvor är just namnet, en elliptisk kurva har väldigt lite med ellipser att göra. I det här avsnittet tänkte jag försöka förklara den här kopplingen.

En ellips är en kurva av andra graden given av en ekvation på formen

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1.$$

Ett väldigt klassiskt problem är att försöka räkna ut omkretsen på en ellips. I det specialfall då $a = b$ blir svaret $2a\pi$. Tyvärr finns det ingen enkel formel för godtyckliga a och b.

En integral som dyker upp vid beräkning av omkretsen till en ellips är

$$E(x) = \int_0^x \frac{1}{\sqrt{(1-t^2)(1-kt^2)}} dt.$$

där $k = 1 - b^2/a^2$. Den här går inte att lösa med välkända funktioner, vilket är anledningen till att vi inte kan få en sluten formel för omkretsen till en ellips.

Det är dock fortfarande en intressant funktion.

Om $k = 0$, så att vi är i cirkelfallet, får vi att

$$E(x) = \int_0^x \frac{1}{\sqrt{(1-t^2)}} dt.$$

Den här integralen går att lösa direkt, och faktum är att

$$E(x) = \sin^{-1}(x).$$

Vi kan dra inspiration från det här specialfallet och försöka betrakta inversen till $E(x)$ istället för $E(x)$, eftersom det kan visa sig vara en mer naturlig funktion. Vi gör därmed definitionen

$$sn(x) = E^{-1}(x).$$

Det visar sig att om vi betraktar sn som en komplexvärd funktion, är den periodisk! Och inte bara periodisk, utan dubbelt periodisk, med två perioder ω_1 och ω_2. Sådana funktioner kallas för elliptiska funktioner av den anledningen. Funktionen $\wp$ är också en elliptisk funktion enligt den definition, och därifrån kommer namnet elliptisk kurva.

Övningar

Övning 8.1. Låt $a + bi \in \mathbb{C}$ så att $a + bi \neq 0$. Vad är $(a + bi)^{-1}$?

Övning 8.2. Betrakta följande komplexvärda funktioner och försök beskriv med ord hur de transformerar det komplexa talplanet.

- $f(z) = z + 3i$.
- $f(z) = -z$.
- $f(z) = iz$.

Övning 8.3. Låt A vara en abelsk grupp. Visa att mängden

$$A[m] = \{P \in A : mP = 0\}$$

är en grupp.

Övning 8.4. Betrakta kurvan C given av ekvationen $y^2 = x^3 + x$. Vad är $C(\mathbb{C})[2]$?

Övning 8.5. Låt Λ vara ett gitter genererat av $\omega_1 = 1 + i$, $\omega_2 = 1 - i$, och låt Λ' vara ett gitter genererat av $\alpha_1 = 2$, $\alpha_2 = 3 - i$. Visa att $\Lambda = \Lambda'$.

Övning 8.6. Visa att $\wp$ är en jämn funktion, alltså att $\wp(z) = \wp(-z)$.

Övning 8.7 ($\star$)**.** Låt $0 < a \leqslant b$, och låt C vara ellipsen

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

(i) Visa att omkretsen av ellipsen ges av integralen

$$4a\int_0^{\pi/2} k^2\sin^2\theta d\theta,$$

där $k=\sqrt{1-\frac{b^2}{a^2}}$. Använd formeln för omkrets av en kurva.

(ii) Gör substitutionen $t=\sin\theta$ och visa att vi istället får integralen

$$4a\int_0^1 \sqrt{\frac{1-k^2t^2}{1-t^2}}dt = 4a\int_0^1 \frac{1-k^2t^2}{\sqrt{(1-t^2)(1-k^2t^2)}}dt.$$

Övning 8.8. Betrakta kurvan C given av ekvationen $s^2=(1-t^2)(1-k^2t^2)$. Visa att transformationen

$$x=\frac{1}{t-1},\quad y=x^2s=\frac{s}{(t-1)^2}$$

ger en transformation mellan C och den elliptiska kurvan

$$y^2=2(k^2-1)x^3+(5k^2-1)x^2+4k^2x+k^2.$$

Lösningar till udda övningsuppgifter

Kapitel 1

Övning 1.1. (i) 0, 1, 2, 3, 4.

(ii) 1, 2, $\{2, 3\}$.

(iii) -3, -2, -1, 0, 1, 2, 3.

(iv) 1, 2.

(v) -3, -2, -1, 0, 1, 2, $\{2, 3\}$.

Övning 1.3. (i) B är en äkta delmängd av A.

(ii) A och B är lika.

(iii) Mängderna är disjunkta eftersom $B = \emptyset$ och alla mängder är disjunkta med den tomma mängden. B är en äkta delmängd till A.

(iv) Mängderna är varken disjunkta, lika eller äkta delmängder av varandra.

(v) Mängderna är disjunkta eftersom $A = \emptyset$.

Övning 1.5. Observera att dessa svar är förslag. Uppgifterna har flera korrekta svar.

(i) $\{n \in \mathbb{Z} \mid n = 2k \text{ för något heltal } k \geqslant 1\}$.

(ii) $\{p/2 \mid p \text{ är ett heltal }\}$.

(iii) $\{r \in \mathbb{R} \mid r \notin \mathbb{Q} \text{ och } |r| < 1\}$.

Övning 1.7. (i) Definitionsmängd: $\{1, 2, 3, \ldots\}$. Målmängd: $\mathbb{N}$.

(ii) Definitionsmängd: $\{T \mid T \text{ är en triangel }\}$. Målmängd: $\mathbb{R}$.

(iii) Definitionsmängd: $\{p(x) \mid p(x) \text{ är ett andragradspolynom}\}$. Målmängd: $\{p(x) \mid p(x) \text{ är ett förstagradspolynom}\}$.

Övning 1.9. (i) Detta är en funktion. Den är definierad för alla värden i definitionsmängden, den ger alltid samma värde för ett givet argument, och alla funktionsvärden ligger i målmängden.

(ii) Detta är inte en funktion, då funktionens värden inte ligger i målmängden ($f(2) = \sqrt{2} \notin \mathbb{Q}$).

(iii) Detta är inte en funktion, eftersom dess värden är slumpmässiga.

(iv) Detta är en funktion. Eftersom ordet Balkong börjar på B, så har funktionen ett definierat värde som ligger i målmängden.

Övning 1.11. (i) Funktionerna är lika. De har samma definitionsmängd, målmängd och $\sqrt{x^2} = |x|$ för alla x.

(ii) Funktionerna är inte lika, då deras definitionsmängder inte är samma.

(iii) Funktionerna är inte lika, då deras målmängder är olika.

Övning 1.13. I exempel 1.5.3 är värdemängden $\{3, 4\}$. Funktionen är inte injektiv och inte surjektiv. I exempel 1.5.4 är värdemängden $\{0, 1\}$ och funktionen är bijektiv, alltså surjektiv och injektiv.

Övning 1.15. Vi gör ett direkt bevis. Om $A \cap B =$ betyder det att inga element i A tas bort i $A \setminus B$. Alltså har vi $A \cap B = \emptyset \implies A \setminus B = A$. På samma sätt har vi att om $A \setminus B = A$ betyder det att det inte finns några gemensamma element i A och B som tas bort i $A \setminus B$, alltså har vi $A \setminus B = A \implies A \cap B = \emptyset$.

Övning 1.17. Vi gör ett direkt bevis. Mängden $A \setminus B$ är mängden av element som ligger i A och inte i B. Alltså är det mängden vi får när vi 'tar bort elementen som ligger i B *från* A'. Alltså tar vi egentligen bara bort elementen som ligger i $A \cap B$ från A. De som inte låg i A från första början struntar vi alltså i.

Övning 1.19. Vi gör ett direkt bevis. Låt $g : X \to Y$ och $f : Y \to Z$ vara bijektioner. Vi verifierar att $f \circ g$ är injektiv,

$$(f \circ g)(x) = (f \circ g)(y) \overset{f \text{ injektiv}}{\implies} g(x) = g(y) \overset{g \text{ injektiv}}{\implies} x = y.$$

På samma sätt ser vi att funktionen är surjektiv eftersom bilden av $f \circ g$ är bilden med f av värdemängden för g. Eftersom g är surjektiv blir bilden för $f \circ g$ densamma som bilden för f och eftersom f antas vara surjektiv har vi att bilden av f är hela Z. Alltså är $f \circ g$ surjektiv.

Övning 1.21. Vi gör ett direkt bevis. Hur många par av element A respektive B finns det? Jo, exakt lika många som det finns sätt att välja ett element ur den ena och sedan den andra vilket är $(\#A) \cdot (\#B)$.

Övning 1.23. Vi löser problemet genom att konstruera den inversa funktionen. Notera att vi har en funktion

$$\begin{aligned} \phi : \{0,1\}^A &\to \{B \mid B \subset A\} \\ f &\mapsto \{a \in A \mid f(a) = 1\}. \end{aligned}$$

Det är nu tydligt att $\phi \circ \chi$ och $\chi \circ \phi$ är varandras inverser.

Övning 1.25. Om n är udda, så finns det ett heltal k så att $n = 2k + 1$. Då gäller att

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

Eftersom $2k^2 + 2k$ är ett heltal, bevisar detta att n^2 är udda.

Övning 1.27. Antag motsatsen, det vill säga att $a < c/2$, $b < c/2$ och $a + b \geqslant c$. Då gäller att

$$a + b < c/2 + b < c/2 + c/2 = c.$$

Alltså gäller $a + b < c$, vilket är en motsägelse.

Övning 1.29. Sätt $a = b = \sqrt{2}$. Om a^b är rationellt så är vi klara, för då är a och b irrationella tal så att a^b är rationellt. Om a^b är irrationellt, så kan vi sätta $c = a^b$ och få att

$$c^b = \sqrt{2}^{\sqrt{2}\cdot\sqrt{2}} = \sqrt{2}^2 = 2,$$

vilket är rationellt.

Kommentar: *Beviset säger oss bara att något av fallen gäller, inte vilket. Man säger att beviset är icke-konstruktivt.*

Övning 1.31. För det första påståendet gör vi ett direkt bevis. Om $r = n$ finns det inget att visa, så antag att $r < n$. Detta är ekvivalent med existensen av ett positivt heltal tal k sådan att $n = r + k$. Då produkten av två positiva heltal är ett positivt heltal kan vi skriva

$$mr < mr + mk = m(r + k) = mn.$$

För det andra påståendet gör vi ett motsägelsebevis. Antag att $r < n$ och att det finns ett positivt heltal k sådant att $r = nk$. Då kan vi med den tidigare uppgiften se att

$$1 \leqslant k \implies n \leqslant nk.$$

I sin tur har vi då att $r < n \leqslant nk = r$. Det är en motsägelse att ett tal är mindre än sig självt.

För det tredje påståendet så gör vi ett direkt bevis.

$$nk - n\ell = n(k - \ell).$$

Kapitel 2

Övning 2.1. i) Nej. Det finns ett identitetselement 0 men däremot finns det inga inverser eftersom summan av två positiva tal aldrig är 0.

ii) Ja, se exempel 2.0.3.

iii) Nej. Problemet är att det saknas inverser, det finns inget heltal k sådant att $2k = 1$ exempelvis.

iv) Nej. Talet 1 fungerar som identitetselement men det finns inget bråktal q sådant $0 \cdot q = 1$, alltså har inte 0 någon multiplikativ invers.

v) Ja.

Övning 2.3. Låt oss verifiera gruppaxiomen för B

(i) (Associativitet) För alla element gäller $p, q, s \in B$ gäller

$$(p \operatorname{xor} q) \operatorname{xor} s = p \operatorname{xor} (q \operatorname{xor} s)$$

eftersom båda sidor är sanna om och endast om exakt ett eller tre utav p, q, s är 'sant'.

(ii) (Identitet) Elementet 'falskt' har egenskapen att för alla $p \in B$ gäller $p \operatorname{xor} \text{falskt} = \text{falskt} \operatorname{xor} p = p$ eftersom alla tre är sanna om och endast om p är sant.

(iii) (Inverser) Vi har att

$$\text{falskt} \operatorname{xor} \text{falskt} = \text{falskt} \qquad \text{sant} \operatorname{xor} \text{sant} = \text{falskt}$$

vilket betyder att båda elementen i B är sin egen invers (kom ihåg att 'falskt' är identiteten).

Till sist har vi en isomorfin från B till $\mathbb{Z}/(2)$ som skickar 'falskt' till 0 och 'sant' till 1. Isomorfin från S_2 till $\mathbb{Z}/(2)$ är den som skickar identiteten till 0 och det andra elementet till 1. Sammansättningen av dessa två ger en isomorfi mellan B och S_2.

Övning 2.5. Ja. Vi Verifierar gruppaxiomen

Associativ Detta ärvs från S_3.

Identitet Identitetsfunktionen uppfyller villkoret för att ligga i K.

Inverser Det är en direkt konsekvens att om f är bijektiv och $f(\{1,2\}) = \{1,2\}$ gäller det att

$$f^{-1}(\{1,2\}) = f^{-1}(f(\{1,2\})) = \mathrm{id}\{1,2\} = \{1,2\}.$$

Detsamma gäller för $\{3,4\}$. Alltså ligger f^{-1} i K om f gör det.

Övning 2.7. (a) Vi använder gruppaxiomet (ii) för att visa att

$$e = e * e$$

och sedan hjälpsats 2.2.1 för att dra slutsatsen att e är sin egen unika invers.

(b) Detta följer direkt från definitionen av att x^{-1} är inversen till x eftersom

$$x^{-1} * x = x * (x^{-1}) = e.$$

och x^{-1} har en unik invers med denna egenskap att $x^{-1} * (x^{-1})^{-1} = (x^{-1})^{-1} * (x^{-1}) = e$

Övning 2.9. Kalla delgrupperna för H och K. Per antagande finns det ett par av element sådan att de ligger den ena delgruppen men inte den andra, det vill säga $h \in H \setminus K$ och $k \in K \setminus H$. Vi gör nu ett motsägelsebevis. Antag att $H \cup K$ är en delgrupp. Notera då att en delgrupp är sluten så vi har $h * k \in H \cup K$, då kan vi anta utan förlust av generalitet att $h * k \in H$. Vi har också att $h^{-1} \in H \subset H \cup K$. Då måste vi även ha att $h^{-1} * (h * k) \in H$ men $h^{-1} * (h * k) = k$ vilket vi antog inte vara ett element i H. Vi har nått en motsägelse. □

Övning 2.11. Eleven kan i detta stadium verifiera gruppaxiomen hur de än vill. Detta är grupptabellen för $\mathbb{Z}/(3\mathbb{Z})$ vilket är ett specialfall av heltalen modulo n som vi bevisar vara en abelsk grupp i nästa kapitel.

Övning 2.13. a) Det är tydligt att identiteten e för $*$ också är det för $*_{op}$. På samma sätt är inversen för varje x densamma för operationen $*$ och $*_{op}$. Det klurigaste att verifiera är alltså associativiteten.

$$a*_{op}(b*_{op}c) = a*_{op}(c*b) = (c*b)*a = \{* \text{ är associativ } \} = c*(b*a) = (b*a)*_{op}c = (a*_{op}b)*_{op}c.$$

b) Man får den ena grupptabellen genom att spegla den andra i diagonalen.

c) Detta är tydligt eftersom definition av att vara en abelsk grupp är ord för ord detsamma som att $a * b = a *_\circ b$.

Övning 2.15. Alla bijektioner från $\{0, 1, \ldots, n\}$ till sig själv går att konstruera genom att första välja vart 1 ska skickas. Detta kan göras på n olika sätt. 2 ska sedan skickas på ett annat element som inte är samma som 1an skickades till, alltså har vi $(n-1)$ alternativ. Genom att repetera detta argument tills vi bestämt ett funktionsvärde för alla element kan vi multiplicera ihop alla möjligheter och då får vi just $n \cdot (n-1) \cdot \ldots$ antal alternativ.

Övning 2.17. Låt e och i vara identitetselementen i H respektive G. Det är tydligt att den konstanta funktionen $\phi(x) = e$ är en grupphomomorfi och att den är identiten i H^G. Definiera funktionen $x \mapsto \phi(x)^{-1}$, detta är den enda möjliga kandidaten för en invers till ϕ med avseende på $*$ (inte $\circ$). Funktionen är inte alltid en grupphomomorfi men den är det om H är abelsk eftersom

$$\phi(x*y)^{-1} = (\phi(x) \star \phi(y))^{-1} \overset{2.2.2}{=} \phi(y)^{-1} \star \phi(x)^{-1} \overset{H \text{ abelsk}}{=} \phi(x)^{-1} * \phi(y)^{-1}.$$

Slutligen för ett par av grupphomomorfier gäller det att $\phi \star \psi$ också är en grupphomomorfi när H är abelsk,

$$(\phi \star \psi)(x*y) = (\phi(x) \star \phi(y)) \star (\psi(x) \star \psi(y)) \overset{H \text{ abelsk}}{=} ((\phi \star \psi)(x)) \star ((\phi \star \psi)(y)).$$

Övning 2.19. (i) Vi kan verifiera att $x^2 \in J$ och $x^3 \in U$.

(ii) Vi verifierar att både J och U är slutna under addition och invertering. Om f, g är jämna ser vi att

$$(f+g)(-x) = f(-x) + g(-x) = f(x) + g(x) = (f+g)(x).$$

Om f, g är udda ser vi att

$$(f+g)(-x) = f(-x) + g(-x) = -f(x) - g(x) = -(f+g)(x).$$

Dessutom så gäller det att om f jämn så är $-f$ jämn, och detsamma gäller för udda polynom.

(iii) Vi verifierar att

$$\phi(f+g) = \left(\frac{(f(x) + f(-x)) + (g(x) + g(-x))}{2}, \frac{(f(x) - f(-x)) + (g(x) - g(-x))}{2} \right)$$
$$= \left(\frac{(f(x) + f(-x))}{2}, \frac{(f(x) - f(-x))}{2} \right) + \left(\frac{(g(x) + g(-x))}{2}, \frac{(g(x) - g(-x))}{2} \right) = \phi(f) + \phi(g).$$

(iv) Inversfunktionen till ϕ är

$$\psi(f, g) = f + g.$$

Vi verifierar att

$$\psi(\phi(f(x)) = \frac{f(x) + f(-x)}{2} + \frac{f(x) - f(-x)}{2} = f(x)$$
$$\phi(\psi(f, g)) = \phi(f + g) = \phi(f) + \phi(g) = (f, 0) + (g, 0) = (f, g).$$

Övning 2.21. Alla axiom följer direkt från att G är en grupp därför axiomen verifieras för alla $x \in X$ separat.

(i) (Associativ) $((f + g) + h)(x) = (f(x) + g(x)) + h(x) = f(x) + (g(x) + h(x)) = (f + (g + h))(x)$.

(ii) (Identitet) Identitselementet är funktionen som alltid ger ut $e \in G$, alltså $e(x) = e$.

(iii) (Inverser) Inversfunktionen definieras som $f^{-1}(x) = (f(x))^{-1}$.

Kapitel 3

Övning 3.1. Kom ihåg att använda hjälpsats 3.4.8 och 2.2.4, vilket betyder att alla beräkningar modulo n kan göras med vanliga heltal och sedan ta %.

Alltså, i modulo 3 får vi

(i) $1 +_3 1 = (1 + 1)\%3 = 2\%3 = 2$

(ii) $1 +_3 1 +_3 1 = (1 + 1 + 1)\%3 = 3\%3 = 0$

(iii) $2 +_3 5 = (2 + 5)\%3 = 7\%3 = 1$

(iv) $2 \cdot_3 5 = (2 \cdot 5)\%3 = 10\%3 = 1$

(v) $0 \cdot_3 4 = (0 \cdot 4)\%3 = 0\%3 = 0$

(vi) $2 \cdot_3 (5 +_3 4) = (2 \cdot (5 + 4))\%3 = 18\%3 = 0$

(vii) $5 \cdot_3 (3 +_3 9) = (5 \cdot (3 + 9))\%3 = 60\%3 = 0$

(viii) $-_3 2 = 1$ eftersom $1 +_3 2 = 3\%3 = 0$ eller alternativt $(-2)\%3 = 1$.

(ix) $8 -_3 2 = (8-2)\%3 = 6\%3 = 0$ eller $8 -_3 2 = 8 +_3 (-_3 2) = 8 +_3 1 = 9\%3 = 0$

I modulo 11 får vi

(i) $1 +_{11} 1 = 2\%11 = 2$

(ii) $1 +_{11} 1 +_{11} 1 = 3\%11 = 3$

(iii) $2 +_{11} 5 = (2 + 5)\%11 = 7\%11 = 7$

(iv) $2 \cdot_{11} 5 = (2 \cdot 5)\%11 = 10\%11 = 10$

(v) $0 \cdot_{11} 4 = (0 \cdot 4)\%11 = 0\%11 = 0$

(vi) $2 \cdot_{11} (5 +_{11} 4)2 = (2 \cdot (5 + 4))\%11 = 18\%11 = 7$

(vii) $-_{11}2 = 9$ eftersom $9 +_{11} 2 = 11\%11 = 0$ eller alternativt $(-2)\%11 = 9$

(viii) $8 -_{11} 2 = (8-2)\%11 = 6\%11 = 6$ eller $8 -_{11} 2 = 8 +_{11} (-_{11}2) = 8 +_{11} 9 = 17\%11 = 6$

Övning 3.3. Tabellen blir

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

Övning 3.5. Mängden av polynom är delring till mängden av funktioner från $\mathbb{R} \to \mathbb{R}$ och är alltså en ring i sig självt.

Övning 3.7. Vi gör ett motsägelsebevis. Antag att $a, b \neq 0$ och $ab = 0$. Då kan vi multiplicera med $b^{-1}a^{-1}$ på båda sidor och få att

$$1 = b^{-1}a^{-1}ab = b^{-1}a^{-1} \cdot 0 \overset{vn3.9}{=} 0.$$

Detta motsäger dock övningen 3.6.

Övning 3.9. Detta är en direkt beräkning,

$$x \cdot (0) = x \cdot (0 + 0) \overset{distributivitet}{=} x \cdot 0 + x \cdot 0.$$

I detta skede vet vi inte vad $x \cdot 0$ är men eftersom R är en ring så finns det en additiv invers $-(x \cdot 0)$ så att när vi adderar den till båda sidor om likheten ovan så får vi

$$0 = x{\cdot}(0){+}(-(x{\cdot}0)) = (x{\cdot}0{+}x{\cdot}0){+}(-(x{\cdot}0)) \overset{+associativ}{=} x{\cdot}0{+}(x{\cdot}0{+}(-(x{\cdot}0)) = x{\cdot}0.$$

Detta avslutar uppgiften.

Övning 3.11. Detta är en direkt beräkning, låt betecka $(-1) \cdot (-1)$ med a. Då har vi att

$$a = (-1) \cdot (-1) = (-1) \cdot (-1 + 0) = (-1) \cdot (-1 + 1 + (-1))$$
$$\overset{\cdot distributiv}{=} (-1) \cdot (-1) + (-1) \cdot 1 + (-1) \cdot (-1) = a + (-1) + a.$$

Då kan vi addera inversen $-a$ på båda sidor och få

$$a - a = a + (-1) + a - a \iff 0 = a + (-1).$$

Adderar vi nu 1 till båda sidor får vi att

$$0 + 1 = a + (-1) + 1 \iff 1 = a + 0 = a.$$

Övning 3.13. Beviset är identiskt med övningen 2.2.8 upprepat för både $+$ och $\cdot$.

Övning 3.15. Den första deluppgiften följer direkt ifrån att varje komponent 'lever sitt egna liv' och alltså ärvs alla nödvändiga egenskaper från de två ringarna man tar produkten av.

Vad går då fel när man tar två kroppar? Vi gör ett motsägelsebevis. Låt F, K vara två kroppar och antag att $F \times K$ är en kropp. Då är $(0,1)$ ett nollskilt element i $F \times K$ och den har en invers (a,b) som uppfyller att

$$(a,b)\cdot(0,1) = (a\cdot 0, b\cdot 1) = (1,1) \implies a\cdot 0 = 1.$$

Detta motsäger övningarna 3.9 och 3.6.

Övning 3.17. Man kan läsa ur grupptabellen för S_3 att $\{\mathrm{id}, \tau, \pi\}$ är isomorf med $\mathbb{Z}/3\mathbb{Z}$ och $\{\mathrm{id}, f\}$ är isomorf med $\mathbb{Z}/2\mathbb{Z}$.

Övning 3.19. Notera att vi ser i övningen 2.3 att $(p \operatorname{xor} q) \operatorname{xor} s$ är sant om och endast om exakt ett eller exakt 3 av p, q, s är sanna. Vi kan nu se att villkoret $p\&q\&s$ är sant om och endast om alla tre är sanna. Då är det bara att ta xor av dessa eftersom den ena är ett specialfall av den andra i någon mening. Vårt svar är alltså:

$$p \operatorname{xor} q \operatorname{xor} s \operatorname{xor} (p\&q\&s).$$

Övning 3.21.

- Vi vet redan att det associativa axiomet är uppfyllt eftersom sammansättning av funktioner alltid gör det. Vi behöver visa att sammansättningen av affina funktioner är affin, identiteten är affin och att inverser finns. Vi kan snabbt se att identiteten är affin på formen $1\cdot x + 0$. En direkt beräkning ger sedan, då vi låter $\phi(x) = ax + b$ och $\psi(x) = cx + d$, att

  $$\psi\circ\phi(x) = c(ax+b) + d = (ca)x + (cb + d).$$

  De nya konstanterna för $\psi\circ\phi$ är alltså ca och $(cb+d)$ som då blir affin. Slutligen ser vi att inversen till $ax + b$ är $f(x) = (a^{-1})x + (-a^{-1}b)$.

- Isomorfin ges av

  $$\begin{array}{lll} x \mapsto \mathrm{id} & x+1 \mapsto \tau & x+2 \mapsto \pi \\ 2x \mapsto f & 2x+2 \mapsto g & 2x+1 \mapsto h \end{array}$$

Kapitel 4

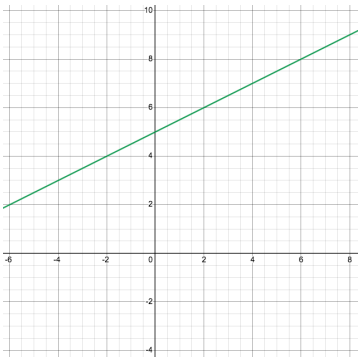
Övning 4.1.

- $f = 1 +_5 x -_5 x^2 = 1 +_5 x +_5 4x^2$,
- $f(3,2) = (1 + 3 + 4\cdot 3^2)\%5 = 40\%5 = 0$.
- $g = 1 -_5 \frac{1}{2}xy = 1 -_5 3xy = 1 +_5 2xy$,
- $g(3,2) = (1 + 2\cdot 3\cdot 2)\%5 = 13\%5 = 3$.

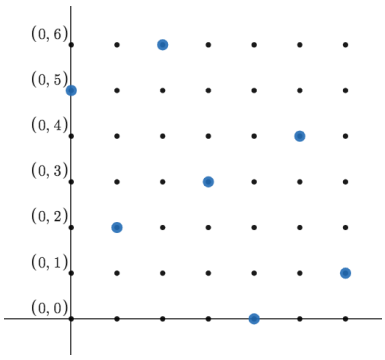
- $f = 1 +_7 x -_7 x^2 = 1 +_7 x +_7 6x^2$,
- $f(3,2) = (1 + 3 + 6 \cdot 3^2)\%7 = 58\%7 = 2$
- $g = 1 -_7 \frac{1}{2}xy = 1 -_7 4xy = 1 +_7 3xy$,
- $g(x,y) = (1 + 3 \cdot 3 \cdot 2)\%7 = 19\%7 = 5$.

Övning 4.3. Vi beräknar enligt hjälpsats 4.2.12 får vi att

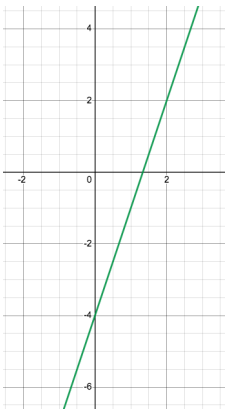
$$y = \frac{7-2}{5-1}(x-1) + 2 = \frac{5}{4}x + (2 - 5/4).$$



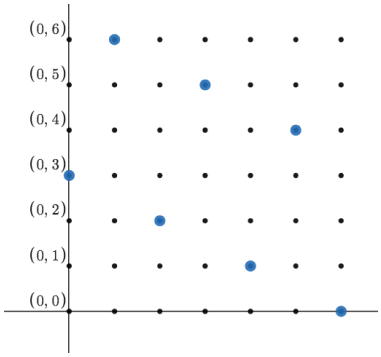
(a) Linjen $y = x/2 + 5$ över $\mathbb{R}$



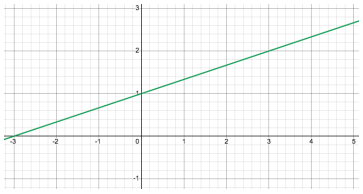
(b) Linjen $y = x/2 + 5$ över $\mathbb{Z}/(7)$.



(a) Linjen $y = 3x - 4$ över $\mathbb{R}$



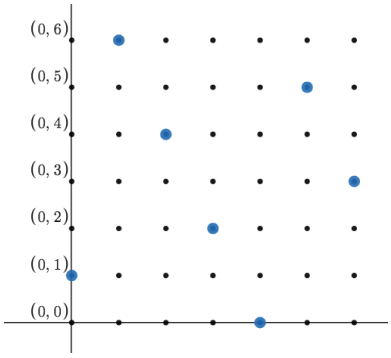
(b) Linjen $y = 3x - 4$ över $\mathbb{Z}/(7)$.



(a) Linjen av punkter på formen

$$t * (3, 1) + (0, 1)$$

över $\mathbb{R}$

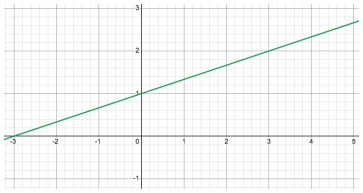


(b) Linjen av punkter på formen

$$t * (3, 1) + (0, 1)$$

över $\mathbb{Z}/(7)$.

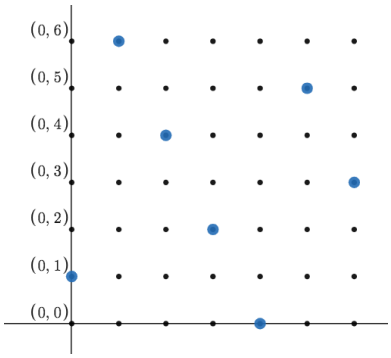
Övning 4.5.



(a) Linjen av punkter på formen

$$t * (3, 1) + (0, 1)$$

över $\mathbb{R}$



(b) Linjen av punkter på formen

$$t * (3, 1) + (0, 1)$$

över $\mathbb{Z}/(7)$.

Övning 4.7.

Övning 4.9. Ja, det finns ingen lösning till systemet $x^2+y^2-1=0$ samtidigt som $2x = 2y = 0$.

Övning 4.11. Nej,

$$\frac{\partial f}{\partial x} = -2x - 3x^2 \frac{\partial f}{\partial x} = 2y.$$

I origo, $x = 0, y = 0$ är f och dess partiella derivator noll. Alltså är detta en singulär punkt på kurvan, där det saknas tangentlinje.

Övning 4.13. Den konstanta nollfunktionen kan snabbt verifieras vara en ringhomomorfi med hjälp av övning 3.9 och vi vet att $0 \neq 1$ från övningen 3.6.

Vi verifierar nu att funktionen ϕ är en ring-homomorfi genom att använda att $3 \cdot_6 3 = 3$,

$$\phi(a +_6 b) = 3 \cdot_6 (a +_6 b) \overset{distributivitet}{=} (3 \cdot_6 a) +_6 (3 \cdot_6 b) = \phi(a) +_6 \phi(b).$$
$$\phi(a \cdot_6 b) = 3 \cdot_6 (a \cdot_6 b) = (3 \cdot_6 3) \cdot_6 (a \cdot_6 b) = (3 \cdot_6 a) \cdot_6 (3 \cdot_6 b) = \phi(a) \cdot_6 \phi(b).$$

Anmärk att en ringhomomorfi är per definition en grupphomomorfi mellan $(R,+)$ och $(S,+)$. Vi kan nu avsluta lösningen med att verifiera $\phi(1) = 3 \neq 1$ och att $\phi(0) = 3 \cdot_6 0 = 0$.

Kapitel 5

Övning 5.1. Vi har att

$$n(a+b) = (a+b) + (a+b) +_{n \text{ ggr}} +(a+b)$$

$$= a + a + ...{}_{n \text{ ggr}} + a + b + b + ...{}_{n \text{ ggr}} + b = na + nb.$$

Övning 5.3. (i) Vi har att $P + Q = \mathcal{O} * (P * Q) = \mathcal{O} * (Q * P) = Q + P$. Vidare gäller $\mathcal{O} + P = \mathcal{O} * (\mathcal{O} * (P)) = P$.

(ii) Låt $X + P = Q$. Vi kan skriva om likheten som $\mathcal{O} * (X * P) = Q$. Om vi utför $\mathcal{O} *$ till båda sidor får vi

$$\mathcal{O} * Q = \mathcal{O} * (\mathcal{O} * (X * P)) = X * P.$$

Vi utför $* P$ till båda sidor och får

$$P * (\mathcal{O} * Q) = (\mathcal{O} * Q) * P = (X * P) * P = X.$$

Därmed har X de efterfrågade utseendet.

(iii) Ekvationen $P + (Q + R) = (P + Q) + R$ är ekvivalent med associativitet. Vi kan skriva om likheten som

$$\mathcal{O} * (P * (\mathcal{O} * (Q * R))) = \mathcal{O} * ((\mathcal{O} * (P * Q)) * R).$$

Om vi utför $\mathcal{O} *$ på båda sidor försvinner $\mathcal{O} *$ faktorn och vi får likheten

$$P * (\mathcal{O} * (Q * R) = R * (\mathcal{O} * (P * Q).$$

Övning 5.5. (i) Vi får

$$y^2 = (x - a/3)^3 + a(x - a/3)^2 + b(x - a/3) + c$$

$$= x^3 + (a^2/3 - ab/3 + b)x + 2a^3/27 + c.$$

(ii) Då polynomet ovan är på den förenklade formen, alltså att termen framför x^2 är 0, kan vi använda formeln för diskriminanter och får

$$4(a^2/3 - ab/3 + b)^3 + 27(2a^3/27 + c)^2$$

$$= \frac{1}{27}(-4a^3c + a^2b^2 + 18abc - 4b^3 - 27c^2).$$

Övning 5.7. (i) 3 punkter, $(1,0), (2,0), (3,0)$.

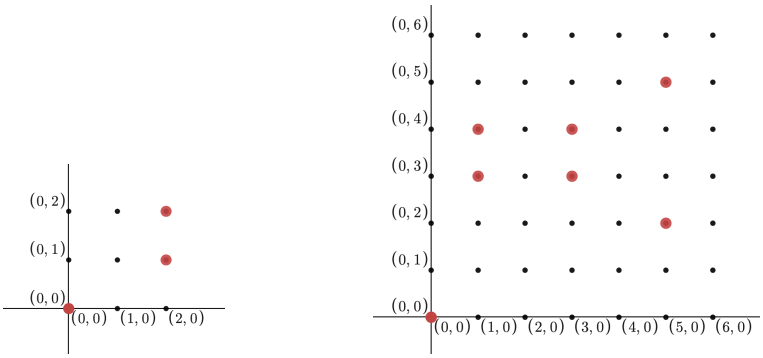
(ii) 1 punkt, $(0,0)$.

(iii) 1 punkt, $(1,0)$.

(iv) 1 punkt, $(1,0)$.

Kapitel 6

Övning 6.1. Vi ritar upp kurvan



(a) Kurvan $y^2 = x^3 + x$ modulo 3 **(b)** Kurvan $y^2 = x^3 + x$ modulo 7

Hasse-Weil säger att

$$1 = |3 - 3 - 1| = |\#C(\mathbb{Z}/(3)) - 3 - 1| \leqslant 2\sqrt{3} \sim 3.46$$

och

$$1 = |7 - 7 - 1| = |\#C(\mathbb{Z}/(7)) - 7 - 1| \leqslant 2\sqrt{7} \sim 5.29$$

vilket stämmer.

Övning 6.3. Det beror på kurvan du väljer.

Övning 6.5. Här i detta facit utgår vi ifrån att alla operationer gör modulo 7. Formeln för addition på en elliptisk kurva säger att

$$\lambda = \frac{1-2}{3-1} = \frac{-1}{2} \text{ eftersom } P \neq Q$$

modulo 7 så är $-1/2$ alltså den additiva inversen till den multiplikativa inversen till 2. Vi ser att $2^{-1} = 4$ och då blir $-1/2 = -4 = 3$. Alltså har vi att $\lambda = 3$. Då säger formeln för addition ytterligare att

$$\begin{aligned} x_3 &= \lambda^2 - x_1 - x_2 = 3^2 - 1 - 3 = 2 - 4 = 5 \\ y_3 &= -\lambda x_3 - (y_1 - \lambda x_1) = -3 \cdot 5 - (2 - 3) = -1 - (-1) = 0 \end{aligned}$$

. Vi kan då konstatera att

$$P + Q = (5, 0)$$

vilket är svaret på frågan.

Övning 6.7. Vi gör en tabell där vi beräknar $a^{\frac{7-1}{2}} = a^3$ modulo 7 för alla element. Hjälpsats 6.3.4 säger att vi letar efter tal sådana att $a^3 = 1$. Sedan använder vi 6.3.5 för att beräkna kvadratroten som $a^{\frac{7-3}{4}+1} = a^2$.

(i) $0 = \sqrt{0}$

(ii) $1^3 = 1 \qquad \sqrt{1} = \pm 1$

(iii) $2^3 = 1 \qquad \sqrt{2} = \pm 2^2 = \pm 4$

(iv) $3^3 = 6 = -1$

(v) $4^3 = 1 \qquad \sqrt{4} = \pm 4^2 = \pm 2$

(vi) $5^3 = -1$

(vii) $6^3 = -1$

Övning 6.9. Vi vill alltså lösa ekvationen $2^m = 9 = -2$ modulo 11. Vi gör en tabell där vi beräknar $a^m = a^5$

(i) $2^0 = 1$

(ii) $2^1 = 2$

(iii) $2^2 = 4$

(iv) $2^3 = 8$

(v) $2^4 = 5$

(vi) $2^5 = 10 = -1$

(vii) $2^6 = -2 = 9$

(viii) $2^7 = -4$

(ix) $2^8 = -8 = 3$

(x) $2^9 = 6$

(xi) $2^10 = 1$

Svar: $m = 6$.

Övning 6.11. Vi använder pq formeln med koefficienterna $\alpha = 1, \beta = 1$ och får

$$x = -\frac{1}{2} \pm \sqrt{\frac{1^2}{4} - 1} = -2^{-1} \pm \sqrt{4^{-1} - 1} = -4 \pm \sqrt{2 - 1} = 3 \pm 1$$

alltså är lösningarna $x = 2$ och $x = 4$.

Övning 6.13. Vid närmare eftertanke ser man att

$$(x - 5)(x - 11)$$

är ett andragradspolynom som har just 5 och 11 som rötter. Vi skriver ut det och beräknar modulo 19 att,

$$(x - 5)(x - 11) = (x + 14)(x + 8) = x^2 + (14 + 8)x + 14 \cdot 8 = x^2 + 3x + 17.$$

Svaret är alltså: $x^2 + 3x + 17$. Om du har angett ett annat polynom som svar bör det vara en multipel av detta polynom.

Kapitel 7

Övning 7.1. (i) Diskriminanten i det här fallet är 4, så möjliga y är $0, \pm 1, \pm 2, \pm 4$. För varje val av y finns en begränsad mängd val av x, eftersom att högerledet annars blir för litet/stort. De enda heltalslösningarna vi hittar är $(0, 0)$, som har ordning 2.

(ii) Diskriminanten är $4 + 27 = 31$. Alltså är möjligheterna för y, 0 eller $\pm 1, \pm 31$. Ingen av dessa värden på y ger en lösning i x, så det finns ingen lösning i det här fallet.

(iii) Diskriminanten är 27, så möjligheter för y är $0, \pm 3, \pm 27$. Vi hittar lösningar $(-1, 0)$ och $(2, \pm 3)$. Punkten $(-1, 0)$ har ordning 2. Om $P = (2, 3)$ har är $2P = (0, 1)$ och om $P = (2, -3)$ är $2P = (0, -1)$. Alltså har punkterna $(0, \pm 1)$ och $(2, \pm 3)$ ändlig ordning.

Övning 7.3. (i) Summan av en ändligt mängd bråk $a_1/b_1, a_2/b_2, ..., a_k/b_k$, kommer aldrig att innehålla primtal i sina bråk som inte delar $b_1 b_2 ... b_k$. Det finns oändligt många primtal, så ett ändligt antal bråk kommer aldrig att generera gruppen.

(ii) Nej. Precis som i föregående del kommer produkten av ett ändligt antal bråk bara innehålla primtal i sina nämnare som var och en av de individuella bråken har i sina nämnare.

Kapitel 8

Övning 8.1. Jag påstår att

$$(a + bi)^{-1} = \frac{a}{a^2 + b^2} - i\frac{b}{a^2 + b^2}.$$

Det följer från multiplikationen

$$(a + bi)\left(\frac{a}{a^2 + b^2} - i\frac{b}{a^2 + b^2}\right) = \frac{a^2 + b^2}{a^2 + b^2} + i\frac{ab - ab}{a^2 + b^2} = 1.$$

Övning 8.3. Om $x, y \in A[m]$ så gäller

$$m(x + y) = mx + my = 0 + 0 = 0.$$

Om vidare $x \in A$ så gäller

$$m(-x) = -mx = -0 = 0.$$

Övning 8.5. Vi har att $\omega_2 = \alpha_2 - \alpha_1, \omega_1 = 2 \cdot \alpha_1 - \alpha_2$. Alltså är

$$\omega_1, \omega_2 \in \mathbb{Z}\alpha_1 + \mathbb{Z}\alpha_2.$$

På ett liknande sätt får vi $\alpha_1 = \omega_1 + \omega_2)$ och $\alpha_2 = \omega_1 + 2\omega_2$, så att

$$\alpha_1, \alpha 2 \in \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2.$$

Sammanlagt får vi

$$\mathbb{Z}\alpha_1 + \mathbb{Z}\alpha_2 \subset \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2 \subset \mathbb{Z}\alpha_1 + \mathbb{Z}\alpha_2.$$

Därmed måste likhet gälla överallt.

Övning 8.7. (i) Vi kan parametrisera vår ellips med uttrycket $r(\theta) = (a\cos(\theta), b\sin(\theta))$, $0 \leqslant \theta \leqslant 2\pi$. Enligt formeln för båglängd har vi att båglängden av en fjärdedel av ellipsen ges av

$$\int_0^{\pi/2} \sqrt{\left(\frac{d}{d\theta}[a\cos(\theta)]\right)^2 + \left(\frac{d}{d\theta}[b\sin(\theta)]\right)^2}\, d\theta = \int_0^{\pi/2} \sqrt{a^2\sin^2(\theta) + b^2\cos(\theta)} d\theta$$

$$= \int_0^{\pi/2} a\sqrt{1 - \left(1 - \frac{b^2}{a^2}\right)\sin^2(\theta)} d\theta = a\int_0^{\pi/2} \sqrt{1 - k^2\sin^2(\theta)} d\theta.$$

Hela omkretsen är 4 gånger så stor.

(ii) Om vi gör substitutionen får vi nya intervall $0 < t < 1$, och $d\theta = 1/\sqrt{1-t^2}dt$. Sammalagt får vi

$$4a\int_0^{\pi/2} \sqrt{1 - k^2\sin^2(\theta)} d\theta = 4a\int_0^1 \sqrt{1-k^2t^2}\frac{dt}{\sqrt{1-t^2}}$$

$$= 4a\int_0^1 \frac{\sqrt{1-k^2t^2}}{\sqrt{1-t^2}} dt.$$

Referenser och förslag till vidare läsning

Joseph H. Silverman, John T. Tate: *Rational Points on Elliptic Curves*
Springer International Publishing Switzerland 1992, 2015

Nigel Smart: *Cryptography, an introduction*
Mcgraw Hill-College, 2004 (Gratis)

J. W. S. Cassels: *Lectures on Elliptic Curves*
Cambridge University Press, 1991