KTH Engineering Sciences

KTHs Matematiska Cirkel

Diofantiska ekvationer

Dan Petersen
Kathrin Vorwerk

Institutionen för matematik, 2011
Finansierat av Marianne och Marcus Wallenbergs Stiftelse

Innehåll

Några ord på vägen

Detta kompendium är skrivet för att användas som litteratur till KTHs Matematiska Cirkel under läsåret 2011–2012 och består av åtta avsnitt. Kompendiet är inte tänkt att läsas enbart på egen hand, utan ska ses som ett skriftligt komplement till undervisningen på de åtta träffarna. En bra idé kan vara att försöka läsa varje kapitel själv innan varje föreläsning, så att man redan innan vet vad målet med föreläsningen är och vad som kan visa sig vara svårt.

Som den mesta matematik 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. Dessa är dels ordnade efter ungefärlig svårighetsgrad: övningar kan ha en ($\star$), två ($\star\star$) eller tre ($\star\star\star$) stjärnor. Dessutom har de udda övningarna facit längst bak i kompendiet och syftet med dessa är att eleverna ska kunna räkna dem och på egen hand kontrollera att de förstått materialet. De 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.

Vi vill dock betona att få av uppgifterna är helt enkla. Detta betyder dels att läsaren inte bör titta i facit efter några få minuter, utan att först prata med kompisar om uppgiften, kanske lägga den åt sidan ett tag och tänka på annat, och sedan försöka lite till. Dessutom innebär det att få av eleverna kommer att kunna klara samtliga uppgifter, så ett krav på att eleven ska ha löst alla uppgifter bör inte ingå i examinationen. Dock rekommenderar vi starkt att alla elever åtminstone tittar på och försöker sig på alla övningar.

De flesta övningar kommer att ha många olika möjliga lösningar och det som står i facit bör endast ses som ett förslag.

KTHs Matematiska Cirkel finns också på Facebook. Om ni har funderingar om materialet eller har kört fast med en övning får ni gärna fråga där.

KTHs Matematiska Cirkel finansieras av Marianne och Marcus Wallenbergs Stiftelse. Vi tackar Dan Laksov och Roy Skjelnes, Institutionen för Matematik vid KTH, Alan Sola vid Institut Mittag-Leffler, och Toomas Liiv för givande kommentarer om denna skrift.

Några ord om Cirkeln

KTHs Matematiska Cirkel, i dagligt tal benämnd Cirkeln, startade 1999. Dess ambition 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 skall särskilt stimulera elevernas matematikintresse och inspirera dem till fortsatta naturvetenskapliga och matematiska studier. Lärarna på Cirkeln kan vid behov ge eleverna förslag på ämnen till projektarbeten vid gymnasiet eller förslag till annan förkovran inom matematik.

Till varje kurs skrivs ett kompendium som distribueras gratis till eleverna. Detta material, föreläsningsschema och övriga uppgifter om KTHs Matematiska Cirkel finns tillgängligt på

`http://www.math.kth.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. Detta, och att flera ämnen är på universitetsnivå, gör att lärarna och eleverna kan uppleva programmet som tungt, och alltför långt över gymnasienivån. Meningen är emellertid inte att lärarna och eleverna skall behärska ämnet fullt ut och att lära in det på samma sätt som gymnasiekurserna. Det viktigaste är att eleverna kommer i kontakt med teoretisk matematik och får en inblick i *matematikens väsen*. Vår förhoppning är att lärarna med denna utgångspunkt skall ha lättare att upplysa intresserade elever om KTHs Matematiska Cirkel och övertyga skolledarna om vikten av att låta både elever och lärare delta i programmet.

Några ord om betygssättning

Ett speciellt problem tidigare år har varit betygssättningen. Detta borde emellertid bara vara ett problem om lärarna använder sig av samma standard som de gör när de sätter betyg på ordinarie gymnasiekurser. Om utgångspunkten istället är att eleverna skall få insikt i matematiken genom att gå på föreläsningarna och att eleven gör sitt bästa för att förstå materialet och lösa uppgifterna, blir betygsättningen lättare. Självklart betyder det mycket vad eleverna har lärt av materialet i kursen, men lärarna kan bara förvänta sig att ett fåtal elever behärskar ämnet fullt ut. I det perspektivet blir det lätt att använda de officiella kriterierna:

Godkänd: Eleven har viss insikt i de moment som ingår i kursen och kan på ett godtagbart sätt redovisa valda delar av kursen såväl muntligt som skriftligt. Detta kan ske genom att eleven håller föredrag inför klassen, redovisar eller lämnar en rapport till sin matematiklärare.

Väl godkänd: Eleven har god insikt i flera moment från kursen. Eleven kan redovisa dessa moment både skriftligt och muntligt och dessutom uppvisa lösningar på problem som givits på kursen. Detta kan ske genom att eleven håller föredrag inför klassen, redovisar eller lämnar en rapport till sin matematiklärare.

Mycket väl godkänd: Eleven har mycket god insikt i flera moment av kursen och lämnar skriftliga redovisningar av flera delar av kursen eller lämnar lösningar på problem som givits på kursen. Detta kan ske genom att eleven håller föredrag inför klassen, redovisar eller lämnar en rapport till sin matematiklärare.

Det är också till exempel möjligt att skolorna samarbetar, så att elever från en skola redovisar eller lämnar rapport för en lärare i en annan skola.

Författarna, september 2011

1 Grundläggande begrepp och bevisföring

I det här kapitlet kommer vi att ge en introduktion till matematisk bevisföring. Som fallstudie kommer vi att studera *division med rest*. Innan dess kommer vi dock att introducera lite terminologi. I matematiken använder man ofta *mängder* och *funktioner* som ett bekvämt språk för att beskriva saker och ting, och detta kommer vi också att göra i detta kompendium. Vi ger därför en introduktion till denna terminologi.

1.1 Mängder

Låt oss titta på ett av de mest grundläggande begreppen i matematiken, nämligen mängder. En mängd är en samling objekt, som till exempel tal, och dessa objekt kallar vi för *element* i mängden. Det enklaste sättet att beskriva en mängd är att räkna upp dess element. Ett sådant exempel är

$$A = \{1, 3, a, 7\}.$$

Detta betyder att A är en mängd som innehåller elementen $1, 3, a$ och 7. Vi bryr oss inte om i vilken ordning eller hur många gånger elementen räknas upp och därmed gäller till exempel

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

En mängd kan också ha oändligt många element, och då går det inte att räkna upp alla element. Ett exempel på en oändlig mängd är

$$\{1, 2, 3, 4, \dots\}.$$

De tre punkterna betyder här att *alla* positiva heltal ingår i mängden.

Exempel 1.1.1. Mängden som består av alla udda heltal mellan 0 och 10 kan också skrivas som

$$\{1, 3, 5, 7, 9\}. \qquad \blacktriangle$$

Om A är en mängd och x är ett element i mängden A så skriver vi $x \in A$ och säger att x *tillhör* A. Exempelvis gäller $b \in \{a, b, 10, 3\}$. Att ett element x inte tillhör mängden A skrivs $x \notin A$. Den *tomma mängden* innehåller ingenting och betecknas $\varnothing$.

Definition 1.1.2. Låt A och B vara mängder. Om alla element i mängden A också är element i mängden B så sägs A vara en *delmängd* till B. Detta betecknas $A \subseteq B$.

Exempel 1.1.3. Mängden $\{1, a\}$ är en delmängd till $\{1, 3, a\}$, eftersom alla element i $\{1, a\}$ finns i mängden $\{1, 3, a\}$. Vi skriver $\{1, a\} \subseteq \{1, 3, a\}$. $\blacktriangle$

Ett användbart sätt att beskriva en mängd är som en delmängd av en annan mängd. Det finns ett speciellt skrivsätt för detta, nämligen

$$\{x \in D \ : \ \text{villkor på } x\}.$$

Med detta menar man delmängden bestående av de element i D som uppfyller de givna villkoren. Som exempel kan vi definiera

$$B = \{n \in \{1, 2, 3, \ldots\} \ : \ n \text{ är udda}, \}$$

och

$$C = \{y \in \{1, 2, 3, 4\} \ : \ y > 2\}.$$

Mängden B är delmängden av de positiva heltalen som består av alla udda positiva heltal, medan C är delmängden av $\{1, 2, 3, 4\}$ bestående av element större än 2. Alltså har vi

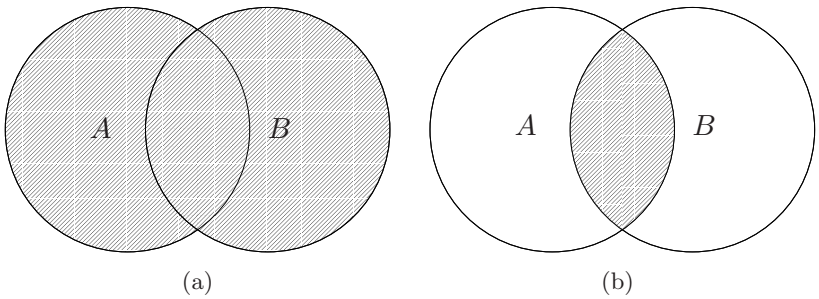
$$B = \{1, 3, 5, 7, 9, 11, \ldots\} \quad \text{och} \quad C = \{3, 4\}.$$

Exempel 1.1.4. Låt $A = \{4, 5, 8, 4711, 12, 18\}$ och $B = \{x \in A \ : \ x > 10\}$. Då är $B = \{12, 18, 4711\}$ medan $\{x \in A \ : \ x < 3\} = \varnothing$. Vidare har vi att $4 \in A$ men $4 \notin B$. ▲

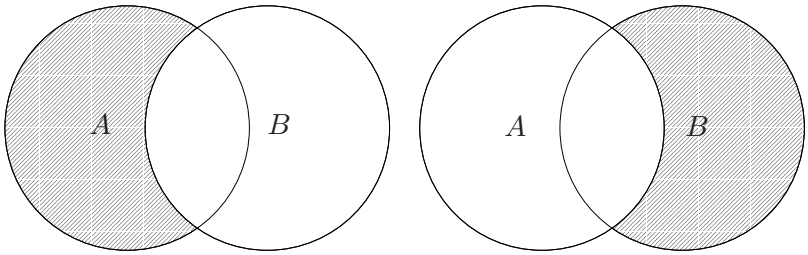
Definition 1.1.5. Antag att A och B är mängder. *Unionen* av A och B består av de element som ligger i någon av mängderna och betecknas $A \cup B$. *Snittet* av A och B består av de element som ligger i båda mängderna och betecknas $A \cap B$. *Differensen* av A och B består av alla element som ligger i A men inte ligger i B, och betecknas $A \setminus B$.

Exempel 1.1.6. Låt $A = \{1, 3, 5, 6\}$ och $B = \{5, 8, 3, 4711\}$. Då har vi $A \cup B = \{1, 3, 5, 6, 8, 4711\}$, $A \cap B = \{3, 5\}$, $A \setminus B = \{1, 6\}$ och $B \setminus A = \{8, 4711\}$. Till skillnad från unionen och snittet är differensen av två mängder inte symmetrisk i A och B. ▲

Ett användbart sätt att åskådliggöra union, snitt och differens är med hjälp av så kallade *Venndiagram*, som visas i Figur 1.1 – 1.2.



Figur 1.1: Venndiagram som åskådliggör mängderna (a) $A \cup B$ och (b) $A \cap B$.



Figur 1.2: Venndiagram som åskådliggör (a) $A \setminus B$ och (b) $B \setminus A$.

Det är dags att titta på några viktiga talmängder. Den mängd vi använder för att räkna föremål är de *naturliga talen* $\{0, 1, 2, 3, \dots\}$. Denna mängd betecknas $\mathbb{N}$. Tar vi med negativa tal får vi heltalen

$$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}.$$

Beteckningen kommer från tyskans *Zahl* som betyder tal. Mängden av alla kvoter av två heltal p/q där $q \neq 0$ innehåller t.ex. $2/3, -7/243$ och $25/1$. Vi kallar mängden de *rationella talen* och betecknar den med $\mathbb{Q}$. Slutligen betecknar vi med $\mathbb{R}$ de *reella talen*, det vill säga alla tal på tallinjen, exempelvis $0, -1, 3/2, -527/3, \sqrt{2}$ och π. Notera att

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}.$$

Låt oss i förbifarten anmärka att i detta kompendium kommer ett tal n att kallas: *positivt* om $n > 0$; *negativt* om $n < 0$; *icke-positivt* om $n \leqslant 0$ och *icke-negativt* om $n \geqslant 0$.

Exempel 1.1.7. Vi har att $\mathbb{N} = \{n \in \mathbb{Z} \,:\, n \text{ är icke-negativt}\}$. ▲

Exempel 1.1.8. Mängden $\{n \in \mathbb{Z} \,:\, n = 2 \cdot k \text{ för något } k \in \mathbb{Z}\}$ är mängden av alla jämna heltal. Denna mängd kan också skrivas som $\{2 \cdot k \,:\, k \in \mathbb{Z}\}$, eller som $\{\dots, -4, -2, 0, 2, 4, \dots\}$. ▲

Exempel 1.1.9. Låt oss påpeka att en mängd även kan ha andra mängder bland dess element. Exempelvis kan vi låta

$$A = \{2, 3, \{-1, 1\}, 4\},$$

och vi har att $\{-1, 1\} \in A$, det vill säga mängden $\{-1, 1\}$ är ett element i mängden A. ▲

1.2 Funktioner

Innan vi gör en allmän definition av vad en funktion är kan det vara på sin plats att titta på något välbekant, nämligen en formel som $f(x) = x^2 + 1$.

Detta är ett exempel på en funktion. Formeln säger att om vi tar ett tal $x \in \mathbb{R}$ så får vi ett nytt tal $f(x) \in \mathbb{R}$ genom att göra beräkningen $x^2 + 1$; till exempel får vi $f(2) = 2^2 + 1 = 5$. Vi säger att f är en funktion från de reella talen till de reella talen, eftersom både det vi stoppar in, x, och det vi får ut, $f(x)$, är reella tal. Vi brukar beteckna detta med $f\colon \mathbb{R} \to \mathbb{R}$.

Definition 1.2.1. Låt X och Y vara mängder. En *funktion* $f\colon X \to Y$ är ett sätt att till varje element $a \in X$ tilldela ett välbestämt element $b \in Y$. Vi skriver $f(a) = b$. Vi säger att a *avbildas* på b och att b är *bilden* av a.

Anmärkning 1.2.2. Ofta säger man att f är en funktion från X till Y istället för att använda beteckningen $f\colon X \to Y$. Ett vanligt alternativ till ordet funktion är *avbildning*.

Exempel 1.2.3. Betrakta mängderna $A = \{1, 2, 3\}$ och $B = \{1, 2, \ldots, 100\}$. Ett exempel på funktion en $f\colon A \to B$ ges av $f(n) = 2n$ för $n \in A$. Vi har alltså att $f(1) = 2$, $f(2) = 4$ och $f(3) = 6$. Per definition måste vi ha $f(x) \in B$ för alla $x \in A$, och detta gäller ju här eftersom

$$f(1) = 1 \in B, \quad f(2) = 4 \in B, \quad \text{och} \quad f(3) = 6 \in B.$$

I detta exempel definieras funktionen f av formeln $f(n) = 2n$, men det är inte alls nödvändigt att det finns en formel som beskriver hur funktionen verkar. Om vi som här har en funktion från den *ändliga* mängden $A = \{1, 2, 3\}$ kan man till exempel definera funktionen med hjälp av en tabell:

n	$f(n)$
1	2
2	4
3	6

▲

Exempel 1.2.4. Låt $h\colon \mathbb{R} \to \mathbb{R}$ vara den funktion som definieras av formeln $h(x) = 3/2 \cdot x^2 - x^3$. Vi har exempelvis att

$$h(1) = \frac{3}{2} \cdot 1^2 - 1^3 = \frac{1}{2}, \quad \text{och} \quad h(-2) = \frac{3}{2} \cdot (-2)^2 - (-2)^3 = 14. \qquad ▲$$

Definition 1.2.5. En funktion $f\colon X \to Y$ säges vara *injektiv* om följande är sant: om $f(x) = f(y)$ för $x, y \in X$ så gäller att $x = y$.

Uttryckt i ord säger den här definitionen att funktionen aldrig avbildar två olika element i X på samma element i Y.

Definition 1.2.6. En funktion $f\colon X \to Y$ säges vara *surjektiv* om följande är sant: för varje $y \in Y$ existerar ett $x \in X$ sådant att $f(x) = y$.

Varje element i Y är alltså bilden av något x under funktionen f om funktionen är surjektiv.

En funktion kan vara surjektiv utan att vara injektiv, och tvärtom.

Exempel 1.2.7. Låt $\mathbb{R}_+$ beteckna de icke-negativa reella talen. Betrakta funktionen $f\colon \mathbb{R} \to \mathbb{R}_+$ som definieras av $f(x) = x^2$. Då är f surjektiv, men inte injektiv — till exempel har vi $f(-2) = f(2) = 4$.

Ett exempel på en funktion som är injektiv men inte surjektiv ges av funktionen i 1.2.3. Det finns till exempel inget $n \in \{1, 2, 3\}$ sådant att $f(n) = 3$. ▲

Definition 1.2.8. En funktion $f\colon X \to Y$ som är både surjektiv och injektiv säges vara *bijektiv*, eller en *bijektion*.

En bijektion $f\colon X \to Y$ har en så kallad *invers*. Detta är en avbildning som brukar betecknas f^{-1} och som låter oss gå tillbaka från bilden av f till den ursprungliga mängden X.

Definition 1.2.9. Låt $f\colon X \to Y$ vara en bijektion. *Inversen* till f är avbildningen $f^{-1}\colon Y \to X$ som ges av $f^{-1}(y) = x$, där x är det entydiga element i X som uppfyller $f(x) = y$.

Vi ser här att både injektivitet och surjektivitet är viktigt. Om f inte är injektiv kan det finnas många $x \in X$ med $f(x) = y$. Om f inte är surjektiv kan det vara så att det inte finns något x med $f(x) = y$.

Exempel 1.2.10. Betrakta funktionen $f\colon \mathbb{R} \to \mathbb{R}$ som ges av $f(x) = x^3$. Denna funktion är injektiv och surjektiv, och därmed en bijektion.

Inversen till f ges av funktionen $f^{-1}\colon \mathbb{R} \to \mathbb{R}$ som definieras av $f^{-1}(y) = y^{1/3}$. ▲

Exempel 1.2.11. Både definitionsmängden och bildmängden måste beaktas när vi undersöker om en funktion är en bijektion.

Funktionen $f\colon \mathbb{R}_+ \to \mathbb{R}_+$ med $f(x) = x^2$ är en bijektion, med invers $f^{-1}(y) = \sqrt{y}$. Som vi såg tidigare är detta påstående är falskt om vi betraktar f definierad på hela $\mathbb{R}$. ▲

1.3 Matematisk bevisföring

Denna kurs kommer i huvudsak att handla om *bevis* av matematiska påståenden; varje föreläsning kommer att innehålla flera bevis, och en majoritet av övningsuppgifterna går ut på att bevisa någonting. Detta innebär antagligen en omställning från tidigare kurser i matematik. Men vad är då ett bevis egentligen? Här är en möjlig definition.

Definition 1.3.1. Ett *bevis* av ett *påstående* är en *logisk slutledning* som leder från en överenskommen uppsättning av *antaganden* fram till påståendet.

Det förekommer flera viktiga ord i föregående definition. Låt oss diskutera dem ett i taget.

Definition 1.3.2. Ett *påstående* är en logisk utsaga som antingen är sann eller falsk.

Exempel 1.3.3. Här är några exempel på påståenden:

(i) $2A + 5B > -C^2$.

(ii) $X \subseteq (Y \cap Z)$.

(iii) Alla jämna tal är delbara med 3.

(iv) Det finns oändligt många primtal.

Av dessa vet vi inte om de första två är falska eller sanna, eftersom vi inte vet vad A, B, C respektive X, Y, Z betyder. Det tredje påståendet är falskt: ett motexempel ges av det jämna talet 2 som ej är delbart med 3. Det fjärde påståendet är sant och kommer att bevisas i detta kompendium. ▲

Exempel 1.3.4. Här är också några exempel på saker som *inte* är påståenden.

(i) $x^2 + 6x + 5$

(ii) Mängden av alla jämna tal. ▲

Påståenden kan kombineras på många olika sätt, som påminner om de sätt vi kan skapa nya mängder av gamla genom operationerna $\cap$, $\cup$ och $\setminus$. Till exempel kan vi sätta två påståenden bredvid varandra och skriva ordet "och" emellan, och vi får ett nytt påstående. Ett annat ord man kan sätta mellan två påståenden är "eller". En annan sak man kan göra är att skriva "Det är inte sant att..." före ett påstående, och detta ger också ett nytt påstående.

Men viktigast av alla sätt att skapa nya påståenden ur gamla är kanske följande.

Definition 1.3.5. Låt P och Q vara två påståenden, till exempel några av de som stod i vår lista. Med $P \implies Q$ menar vi följande påstående: "om påståendet P är sant, är även påståendet Q sant." I ord säger vi att P *implicerar* Q. Om $P \implies Q$ och $Q \implies P$ så skriver vi att $P \iff Q$. I ord säger vi att P gäller *om och endast om* Q gäller.

För varje par av påståenden P och Q får vi alltså ett nytt påstående, $P \implies Q$. Sanningshalten av $P \implies Q$ kan utläsas ur Tabell 1:

Att $P \implies Q$ alltid är sant om P är falskt kan verka ointuitivt till en början. Ett motiverande exempel för denna princip kan vara följande mening som man kan få höra på en biograf: "Om du har en mobiltelefon med dig, är den avstängd?" Om man inte har sin mobiltelefon med sig skall man alltid svara "Ja", oavsett om man har stängt av den eller inte.

P	Q	$P \implies Q$
sant	sant	sant
sant	falskt	falskt
falskt	sant	sant
falskt	falskt	sant

Tabell 1: Hur $P \implies Q$ beror på P och Q.

Exempel 1.3.6. Det gäller att

$$5a + b = 0 \implies 5a = -b.$$

Här gäller även den omvända implikationen, så vi hade kunnat skriva $\iff$ i stället för $\implies$. Vi har också att

$$5a = -b \implies 5ac = -bc,$$

men här är omvändningen inte nödvändigtvis sann. För att gå från det vänstra påståendet till det högra måste vi nämligen dela med c, vilket vi inte vet är tillåtet om vi inte vet att $c \neq 0$. Vi har dock att

$$5a = -b \iff 5ac = -bc \text{ och } c \neq 0.$$ ▲

Exempel 1.3.7. Påståendet

(Det finns oändligt många primtal) $\implies$ (Alla jämna tal är delbara med 3)

är falskt, eftersom det första påstående är sant medan det andra är falskt. Dock är påståendet

(Alla jämna tal är delbara med 3) $\implies$ (Det finns oändligt många primtal)

lustigt nog sant enligt vår definition av $\implies$. ▲

Exempel 1.3.8. För varje påstående P gäller att $P \implies P$, oavsett om P är sant eller inte. ▲

Definition 1.3.9. En *logisk slutledning* är en sekvens av påståenden

$$P_1, P_2, \ldots, P_n$$

med egenskapen att $P_i \implies P_{i+1}$ för alla i.

Definition 1.3.10. Ett *antagande* är ett påstående som vi förutsätter är sant. Ibland kallas dessa synonymt för *axiom* eller *postulat*.

Vi vet nu alltså vad ett bevis av ett påstående Q är: det är en kedja av mindre, enklare påståenden som låter oss dra slutsatsen att Q är sant, endast utgående

ifrån en mindre uppsättning antaganden som vi har bestämt oss för att starta med.

När vi skriver ett bevis brukar vi dock inte bara skriva en lång följd av påståenden med $\implies$ mellan – i stället brukar man försöka uttrycka beviset i vanliga ord och meningar. I stället för symbolen $\implies$ används konstruktioner som "vilket innebär att..." eller "eftersom... så..." eller "från vilket vi drar slutledningen att...", och så vidare.

Speciellt värt att nämna är begreppet *motsägelsebevis*. Detta är en speciell bevisteknik där man i stället för att visa att ett påstående P är sant, så bevisar man att det *inte kan vara falskt*. Med detta menar vi att man börjar med antagandet att P inte gäller, och försöker att härleda ett påstående som man vet inte stämmer, som att $0 = 1$. Enligt Tabell 1 så kan bara ett falskt påstående implicera ett falskt påstående, så vårt antagande att P inte gällde måste ha varit falskt.

Om denna förklaring känns abstrakt, blir det förhoppningsvis mer konkret i det bevis som kommer i slutet av detta kapitel, där ett motsägelsebevis används.

I detta kompendium kommer vi att förutsätta att läsaren känner till följande:

(i) De olika sorternas tal: heltal, rationella, reella.

(ii) Hur man jämför tal med varandra: relationerna $\leqslant$, $\geqslant$ och $=$ samt olika varianter såsom $<, >$ och $\neq$.

(iii) Operationerna addition, subtraktion, multiplikation och division, och deras grundläggande räkneregler, såsom att $a+b = b+a$ eller att $0 \cdot a = 0$ för alla a.

I så stor utsträckning vi bara kan kommer vi att försöka påpeka om vi i ett bevis använder oss av ett antagande som inte står med på denna lista. Det här är inte så lätt som det låter: ofta smyger det sig in ett antagande i ett bevis man inte har tänkt på att man använder, eller så tar man något för givet som egentligen inte är uppenbart.

Vår lista på antaganden är inte så precist formulerad: vi skriver bara "grundläggande räkneregler", men räknar inte upp alla dessa. Vi ber om läsarens överseende.

1.4 Ett bevis

För att inte denna första föreläsning endast skall bli till torrsim, kommer vi nu att försöka bevisa ett påstående. Det kommer till och med att vara ett påstående som kommer att användas många gånger i detta kompendium, nämligen att man kan utföra *division med rest* mellan heltal.

Ämnet för detta kompendium är nämligen heltalsekvationer, och när vi delar två heltal med varandra behöver ju inte resultatet bli ett heltal. Om vi endast

är intresserade av heltal vill man kanske inte lämna heltalens värld alltför ofta, och då kan det vara bekvämt att arbeta med *heltalsdivision*, där vi i stället får en kvot och en *rest* vid division. Till exempel kan man skriva att

$$\frac{11}{5} = 2 + \frac{1}{5},$$

så att kvoten av 11 vid division med 5 är 2 och resten är 1. Detta påstående kan i sin tur formuleras om som

$$11 = 2 \cdot 5 + 1,$$

vilket ger ett helt ekvivalent påstående som endast innehåller heltal.

Vi vill visa att detta alltid går att göra. Detta är dock inte ett helt enkelt bevis, även om påståendet som skall bevisas kan verka elementärt! Beviset kan därför behöva läsas upprepade gånger innan läsaren lyckats smälta det. Läsaren uppmuntras att leta efter var varje antagande används, att försöka ändra i bevismetoderna och se om något går fel, att försöka konstruera alternativa bevis av samma resultat, och så vidare.

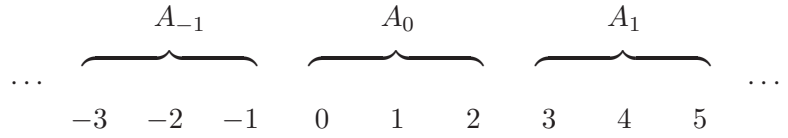
Sats 1.4.1 (Divisionssatsen)**.** *Låt a, b vara heltal med $b > 0$. Då finns heltal q och r, där $0 \leqslant r < b$, sådana att*

$$a = b \cdot q + r.$$

Vi kallar q för kvot *och r för* rest*.*

Anmärkning 1.4.2. Talen q och r är faktiskt unikt bestämda av villkoren i satsen. Detta kommer läsaren att visa i en övning i slutet av kapitlet.

En idé till ett bevis av Sats 1.4.1 skulle kunna vara att fixera talet b och betrakta mängderna $A_n = \{b \cdot n + m \ : \ 0 \leqslant m < b\}$, där $n \in \mathbb{Z}$. För exempelvis $b = 3$ ser mängderna ut så här:



Bilden antyder att för varje $a \in \mathbb{Z}$ finns ett unikt n sådant att $a \in A_n$. Och enligt definitionen av A_n finns ett unikt m med $0 \leqslant m < b$ sådant att $a = b \cdot n + m$. Om man låter $q = n$ och $r = m$ är man därmed klar.

Har vi nu bevisat satsen? **Nej**, eftersom vi inte har *bevisat* att för varje $a \in \mathbb{Z}$ finns ett unikt n med $a \in A_n$. Vi har gjort det troligt att påståendet stämmer, men vi har inte bevisat det.

Det är faktiskt så att de antaganden vi har formulerat ovan inte riktigt räcker till för att bevisa satsen. (Läsaren uppmuntras dock att försöka!) Vi kommer att behöva något mer för att kunna bevisa påståendet. Ett antagande som är tillräckligt är följande:

Princip 1.4.3 (Minimumprincipen)**.** *Låt $X \subseteq \mathbb{N}$ vara en delmängd av de naturliga talen. Om X är icke-tom, så innehåller X ett minsta element.*

Minimumprincipen är ekvivalent med det påstående som brukar kallas *induktionsprincipen*, om någon läsare skulle känna till denna. Vi kommer att använda denna princip på flera ställen i detta kompendium.

Idén för hur vi ska använda minimumprincipen för att visa divisionssatsen är följande. För varje värde på talet q får man ur ekvationen $a = bq + r$ ett unikt värde på r. Vi vill hitta ett värde på r som uppfyller $0 \leqslant r < b$, och naturligt är kanske då att välja r till det minsta möjliga talet som fortfarande uppfyller $0 \leqslant r$. Minimumprincipen är vad som garanterar att ett sådant minsta möjliga tal existerar.

Bevis av Sats 1.4.1. Definiera mängden

$$X = \{a - nb \ : \ n \in \mathbb{Z}\} \cap \mathbb{N}.$$

Mängden X består alltså av alla icke-negativa heltal som kan skrivas som $a-nb$ för något värde på n.

Vi hävdar först att $X \neq \varnothing$. Vi delar upp i två fall:

(i) Om $a \geqslant 0$, så kan vi välja $n = 0$. Vi finner att $a - nb = a \geqslant 0$, så att $a \in X$.

(ii) Om $a < 0$, så kan vi välja $n = a$. Vi finner att $a - nb = a - ab = a(1 - b)$. Men $a < 0$ och $1 - b \leqslant 0$, så produkten av a och $1 - b$ är icke-negativ. Alltså kommer $a(1 - b) \in X$.

Eftersom vi visat att X inte är tom, och vi per definition har att $X \subseteq \mathbb{N}$, så ger Princip 1.4.3 att X har ett minsta element. Kalla detta minsta element för r, och låt q vara motsvarande värde på n, så att $a = qb + r$. Vi hävdar nu att

$$0 \leqslant r < b.$$

Ty antag motsatsen, att $r \geqslant b$. I så fall är

$$a - (q + 1)b = a - qb - b = r - b \geqslant 0,$$

så att $a - (q + 1)b \in X$. Men vi har också att

$$a - (q + 1)b < a - qb = r,$$

vilket säger emot att r skulle vara det minsta elementet i X. Beviset är klart. □

Övningar

Övning 1.1 ($\star$)**.** Låt $A = \{1, 2, 3, 4, \ldots\}$, $B = \{1, 3, 5, 7, \ldots\}$, $C = \{2, 4, 6, 8, \ldots\}$ och $D = \{1, 4, 19, 36, 101\}$. Bestäm mängderna

(i) $B \cup C$,

(ii) $B \cap C$,

(iii) $D \cap C$,

(iv) $\{x \in D \ : \ x \in B\}$,

(v) $\{x \in A \ : \ x = y + 1 \text{ för något } y \in D\}$,

(vi) $\{x + 1 \ : \ x \in D\}$.

Övning 1.2 ($\star$)**.** Ge ett exempel på en funktion från mängden $\{1, 2, 3, 4\}$ till mängden $\{A, B, C\}$. Hur många olika funktioner $f : \{1, 2, 3, 4\} \to \{A, B, C\}$ finns det?

Övning 1.3 ($\star$)**.** Avgör vilka av följande utsagor som är påståenden enligt vår definition av ett påstående. Vilka av dessa är sanna, vilka är falska, och vilka behöver vi mer information för att avgöra?

(i) Mängden av de naturliga talen.

(ii) x är ett positivt heltal.

(iii) Talet x är jämnt.

(iv) Varje mängd innehåller minst ett element.

(v) $x = 5$.

(vi) x är lösningen till ekvationen $3x + 5 = 11$.

Övning 1.4 ($\star$)**.** Använd påståenden från föregående övning och bilda olika sammansatta påståenden på formen $P \implies Q$. Hitta minst två sådana påståenden som är sanna respektive falska.

Övning 1.5 ($\star$)**.** Utför division med rest för följande exempel: 27/6, 142/5 och 1429/3.

Övning 1.6 ($\star\star$)**.** Låt $\mathbb{N} = \{0, 1, 2, \ldots\}$ och $B_n = \{1, 2, \ldots, n\}$ för $n = 1, 2, 3, \ldots$. Visa att $\mathbb{N} \setminus \{0\} = B_1 \cup B_2 \cup B_3 \cup \cdots$.

Övning 1.7 ($\star\star$)**.** (i) Ge ett exempel på en funktion $f : \{1, 2, 3\} \to \{A, B, C\}$ som är både injektiv och surjektiv.

(ii) Ge ett exempel på en funktion $f : \{1, 2, 3\} \to \{A, B, C, D\}$ som är injektiv men inte surjektiv.

(iii) Ge ett exempel på en funktion $f : \{1, 2, 3, 4\} \to \{A, B, C\}$ som är surjektiv men inte injektiv.

(iv) Ge ett exempel på en funktion $f : \{1, 2, 3\} \to \{A, B, C\}$ som är varken injektiv eller surjektiv.

Övning 1.8 ($\star\star$)**.** I Princip 1.4.3 antas det att mängden X är en delmängd av $\mathbb{N}$, det vill säga, endast innehåller positiva heltal.

(i) Hitta en delmängd av $\mathbb{Z}$ som saknar minsta element.

(ii) Hitta en delmängd av de positiva reella talen som saknar minsta element.

Poängen med denna uppgift är att visa att hypotesen $X \subseteq \mathbb{N}$ inte kan försvagas i Princip 1.4.3.

Övning 1.9 ($\star\star$)**.** Låt r vara ett reellt tal. Om det gäller att

$$r = \lfloor r \rfloor + \overline{r}$$

där $\lfloor r \rfloor$ är ett heltal och $\overline{r}$ är ett reellt tal som uppfyller $0 \leqslant \overline{r} < 1$, så säger vi att $\lfloor r \rfloor$ är r *avrundat nedåt till närmaste heltal*. Visa att talet q från divisionssatsen är a/b avrundat nedåt till närmaste heltal.

Övning 1.10 ($\star\star$)**.** Visa att heltalen q och r från Sats 1.4.1 är unika.

Ledning: Antag att q, r, q', r' är olika lösningar. Visa att $b(q - q') = r' - r$ och att $-b < r' - r < b$. Visa att enda heltalslösningen till detta är $q = q'$, $r = r'$.

Övning 1.11 ($\star\star\star$)**.** Visa med hjälp av Princip 1.4.3 att en icke-tom nedåt begränsad delmängd av $\mathbb{Z}$ har ett minsta element och att en icke-tom uppåt begränsad delmängd av $\mathbb{Z}$ har ett största element.

(En mängd X är *nedåt begränsad* respektive *uppåt begränsad* om det finns ett tal $z \in \mathbb{Z}$ sådant att $x \geqslant z$ respektive $x \leqslant z$ gäller för alla $x \in X$.)

Övning 1.12 ($\star\star\star$)**.** I denna övning ska du titta på andra sätt att dividera med rest än det vanliga.

(i) Låt a och b vara heltal med $b > 0$. Visa att det finns unika heltal q och r, där $-b < r \leqslant 0$, sådana att $a = b \cdot q + r$.

Ledning: Använd Divisionssatsen för $-a$.

(ii) Låt a och b vara heltal med $b < 0$. Visa att det finns unika heltal q och r, där $-b < r \leqslant 0$, sådana att $a = b \cdot q + r$.

Ledning: Använd Divisionssatsen för $-a$ och $-b$.

2 Den diofantiska ekvationen $aX + bY = c$

I detta avsnitt kommer vi att påbörja vår studie av diofantiska ekvationer, det vill säga att hitta heltalslösningar till ekvationer med heltalskoefficienter. Mer specifikt kommer vi i detta kapitel att studera ekvationen $aX + bY = c$, där a, b och c är fixerade heltal.

Först kommer vi att diskutera delbarhet och största gemensamma delare. Med hjälp av de satser vi bevisar om den största gemensamma delaren till två tal kan vi sedan ge en fullständig beskrivning av *när* denna ekvation har heltalslösningar, och dessutom ger vi i dessa fall en fullständig beskrivning av *alla* lösningar till ekvationen.

För att kunna använda denna beskrivning av lösningarna till ekvationen krävs att man kan beräkna den största gemensamma delaren av talen a och b. Vi avslutar detta kapitel med en beskrivning av *Euklides algoritm*, som kan användas för att snabbt beräkna den största gemensamma delaren även av ganska stora tal.

2.1 Diofantiska ekvationer

Ämnet för denna kurs är *diofantiska ekvationer*, så låt oss inleda med en definition av detta. Namnet kommer av den grekiske matematikern Diofantos (c:a 250 e.Kr.) vars mest kända verk är de tretton böckerna som utgör *Arithmetika*, som till stor del handlar om just diofantiska ekvationer.

Definition 2.1.1. En *diofantisk ekvation* är en ekvation som uppfyller att:

(i) det finns en eller flera okända $X_1, \ldots, X_n$ som alltid antas vara heltal;

(ii) alla koefficienter som ingår i ekvationen är heltal;

(iii) de enda räknesätten som ingår i ekvationen är addition, subtraktion och multiplikation.

Om antalet okända är litet kallar vi ofta variablerna för X och Y, eller X, Y, Z, eller något liknande.

Diofantiska ekvationer är därmed en gren av talteorin, det vill säga, studiet av heltalen.

Exempel 2.1.2. Om vi studerar ekvationen

$$5X^2 - 7Y^4 = 2$$

och endast är intresserade av heltalslösningar för X och Y, så säger vi att detta är en diofantisk ekvation. Läsaren invänder kanske att i denna ekvation ingick ett räknesätt till utöver addition, subtraktion och multiplikation,

nämligen exponentiering. Dock är detta inte ett problem, eftersom ekvationen kan skrivas om som

$$5 \cdot X \cdot X - 7 \cdot Y \cdot Y \cdot Y \cdot Y = 2. \qquad \blacktriangle$$

Exempel 2.1.3. Följande ekvationer bör inte kallas diofantiska ekvationer.

(i) $5X^2 - \frac{1}{7} \cdot Y^4 = 2$ är inte en diofantisk ekvation, eftersom $\frac{1}{7}$ ej är ett heltal.

(ii) $3^X - Y^2 = 5$ är inte heller en diofantisk ekvation, eftersom det i termen 3^X ingår ett räknesätt som varken är addition, subtraktion eller multiplikation. $\blacktriangle$

När vi studerar diofantiska ekvationer kommer vi oftast inte att uttryckligen säga att vi enbart är intresserade av heltalslösningar till ekvationen. Detta tillåts vara underförstått.

De viktigaste frågor man kan ställa sig om en diofantisk ekvation är följande: Har ekvationen några lösningar? Om ja, har den ändligt många eller oändligt många lösningar? Om antalet är ändligt, hur många finns det? Går det att skriva ned alla lösningar? Om antalet är oändligt, går det att ge en beskrivning av hur alla lösningar ser ut?

Innan vi kan börja lösa ekvationer kommer vi att behöva lite teoretisk bakgrund.

2.2 Delbarhet

Av stor vikt inom talteorin är egenskapen hos ett tal att dela andra tal. Till exempel vet vi att $12 = 4 \cdot 3$, och vi säger därför att både 3 och 4 delar talet 12. Den allmänna definitionen är inte mycket annorlunda.

Definition 2.2.1. Låt a och b vara heltal. Om det finns ett heltal q sådant att $a = b \cdot q$ så säger vi att b *delar* a eller att a är *delbart* med b, och skriver

$$b \mid a.$$

Om b inte delar a så skriver vi $b \nmid a$.

Exempel 2.2.2. Vi har att

$$3 \mid 24 \quad \text{eftersom} \quad 24 = 8 \cdot 3.$$

Men delbarhet fungerar också för negativa tal. Det gäller att

$$-3 \mid 24 \quad \text{eftersom} \quad 24 = (-8) \cdot (-3).$$

och också att

$$3 \mid -24 \quad \text{samt} \quad -3 \mid -24.$$

Däremot gäller inte $5 \mid 24$. Detta skriver vi alltså som $5 \nmid 24$. $\blacktriangle$

Anmärkning 2.2.3. Enligt definitionen av delbarhet delar alla heltal talet 0. Att $0 = a \cdot 0$ innebär ju att $a \mid 0$ för alla $a \in \mathbb{Z}$.

Hjälpsats 2.2.4. *Antag att $d \mid a$ och $d \mid b$. Då gäller att $d \mid ax + by$ för alla heltal x och y.*

Bevis. Låt $a = qd$ och $b = pd$. Då är $ax + by = qdx + pdy = (qx + py)d$, så d delar $ax + by$. □

Anmärkning 2.2.5. Speciellt har vi i beviset för föregående hjälpsats visat följande:

(i) Antag att $d \mid a$. Då gäller även att $d \mid ka$ för varje $k \in \mathbb{Z}$.

(ii) Antag att $d \mid a$ och $d \mid b$. Då gäller även att $d \mid (a + b)$.

2.3 Gemensamma delare

Vi ska nu göra en precis definition av vad vi menar med den *största gemensamma delaren* till två tal. Betrakta som exempel talen 8 och 12. Talet 8 har följande positiva delare:

$$1, \quad 2, \quad 4, \quad 8,$$

och talet 12 har följande positiva delare.

$$1, \quad 2, \quad 3, \quad 4, \quad 6, \quad 12.$$

Det största talet som är en delare till både 8 och 12 är alltså 4. Vi säger att 4 är den största gemensamma delaren till 8 och 12, och skriver $4 = \text{sgd}(8, 12)$. Låt oss nu göra en allmän definition av detta.

Definition 2.3.1. Låt n vara ett heltal. Betrakta mängden

$$D(n) = \{a \in \mathbb{Z} \ : \ a > 0, \ a \mid n\}.$$

Denna mängd kallar vi *delarmängden till n.*

Mängden $D(n)$ innehåller alltså alla positiva delare till n. Vi har exempelvis att $D(12) = \{1, 2, 3, 4, 6, 12\}$.

Definition 2.3.2. Låt a, b vara heltal, där inte både a och b är 0. Elementen i $D(a) \cap D(b)$ kallas *gemensamma delare* till a och b. Det största talet i mängden $D(a) \cap D(b)$ kallar vi för *den största gemensamma delaren till a och b.* Vi betecknar detta tal med $\text{sgd}(a, b)$.

På engelska kallas den för *greatest common divisor*, och den gängse beteckningen är $\gcd(a, b)$.

Anmärkning 2.3.3. När man ger en definition som denna måste man fundera på om mängden $D(a) \cap D(b)$ verkligen innehåller ett största element, för alla val av a och b, så att definitionen har en innebörd. Det följer från Övning 1.11 att två saker skulle kunna gå fel: antingen att mängden är tom, eller att mängden innehåller obegränsat stora element. Det första är omöjligt eftersom $1 \in D(a)$ för varje a, och därmed är även $1 \in D(a) \cap D(b)$. Det andra är också omöjligt eftersom inte både a och b kan vara noll, och om $a \neq 0$ så kommer alla delare till a att vara mindre än eller lika med a. Därmed innehåller inte $D(a)$, och därmed inte heller $D(a) \cap D(b)$, obegränsat stora element.

Definition 2.3.4. Låt a, b vara heltal, inte båda 0. Om $\text{sgd}(a, b) = 1$ så säger vi att a och b är *relativt prima*.

Exempel 2.3.5. Betrakta talen $9 = 3 \cdot 3$, $10 = 2 \cdot 5$, och $12 = 2 \cdot 2 \cdot 3$. Klart är att

$$D(9) = \{1, 3, 9\}, \quad D(10) = \{1, 2, 5, 10\}, \quad D(12) = \{1, 2, 3, 4, 6, 12\}.$$

Alltså gäller

$$\text{sgd}(9, 10) = 1, \quad \text{sgd}(9, 12) = 3, \quad \text{sgd}(10, 12) = 2.$$

Detta betyder att talen 9 och 10 är relativt prima, medan varken 9 och 12 eller 10 och 12 är relativt prima. ▲

Exempel 2.3.6. Det kan vara intressant att fundera på vad största gemensamma delaren till ett positivt tal och 0 är. Detta exempel kommer att bli viktigt senare, när vi beskriver Euklides algoritm. Låt $a > 0$. Enligt definitionen är det faktiskt så att

$$D(0) = \{1, 2, 3, \ldots\},$$

och därmed får vi

$$D(a) \cap D(0) = D(a).$$

Observera nu att det största talet i $D(a)$ är a, och därför är

$$\text{sgd}(a, 0) = a.$$

En konsekvens av detta är att det enda positiva tal som är relativt primt med 0 är talet 1. ▲

En av de viktigaste satserna om största gemensamma delare är följande.

Sats 2.3.7. *Låt a och b vara heltal, inte bägge noll. Det finns heltal x och y sådana att*

$$ax + by = \text{sgd}(a, b).$$

Vi visar denna sats i flera steg, och på vägen kommer vi se att vi får ett ännu mer precist påstående. Låt $\langle a, b\rangle$ beteckna mängden $\{ax + by \ : \ x, y \in \mathbb{Z}\}$, det vill säga mängden av *alla* element som kan skrivas på formen $ax + by$. Det kommer att visa sig att det minsta positiva elementet i $\langle a, b\rangle$ är precis $\text{sgd}(a, b)$. Notera speciellt att enligt minimumprincipen, Princip 1.4.3, så måste det existera ett minsta positivt element i $\langle a, b\rangle$. (Vi lämnar åt läsaren att kontrollera att $\langle a, b\rangle$ innehåller minst ett positivt element.)

Först visar vi dock följande:

Sats 2.3.8. *Låt a och b vara heltal, inte båda noll. Låt d vara det minsta positiva elementet i $\langle a, b\rangle$. Då gäller att*

$$\langle a, b\rangle = \{dq \ : \ q \in \mathbb{Z}\}.$$

Bevis. Antag att $c \in \langle a, b\rangle$. Vi kan enligt Divisionssatsen 1.4.1 skriva

$$c = qd + r$$

med $0 \leqslant r < d$. Men $c = ax + by$, $d = ax' + by'$, så vi har att

$$r = c - qd = a(x - qx') + b(y - qy').$$

Speciellt är $r \in \langle a, b\rangle$. Men $0 \leqslant r < d$ och d är det minsta positiva elementet i $\langle a, b\rangle$, så $r = 0$, och $c = dq$. Detta visar inklusionen

$$\langle a, b\rangle \subseteq \{dq \ : \ q \in \mathbb{Z}\}.$$

För omvändningen, tag ett tal $c \in \{dq \ : \ q \in \mathbb{Z}\}$. Då kan vi skriva $c = dq$ för något q. Men vi vet att $d = ax + by$ för några tal x och y, vilket ger att $c = axq + byq \in \langle a, b\rangle$. Detta visar att de bägge mängderna är lika. □

Sats 2.3.9. *Låt a och b vara heltal, inte båda noll. Det minsta positiva elementet d i $\langle a, b\rangle$ är den största gemensamma delaren till a och b.*

Bevis. Eftersom $a \in \langle a, b\rangle$ och $b \in \langle a, b\rangle$, så visar föregående sats att $d \mid a$ och $d \mid b$, så att d är en gemensam delare.

Låt nu c vara en gemensam delare. Om c delar a och b delar c även $ax + by$ för alla x och y enligt Hjälpsats 2.2.4, så c delar varje element i $\langle a, b\rangle$. Speciellt delar c talet d, så vi kan skriva $d = cq$ för ett heltal $q \geqslant 1$, och därmed är $d \geqslant c$. Så d är större än eller lika med varje annan gemensam delare, och därmed den största gemensamma delaren. □

Följdsats 2.3.10. *Varje gemensam delare till a och b är delbar med den största gemensamma delaren.*

Bevis. Vi såg i beviset av föregående sats att om c är en gemensam delare, så måste $c \mid d$. □

Anledningen att detta är förvånande är att vi definierade den största gemensamma delaren som den gemensamma delare som är större än alla andra delare. Detta visar att den till och med är *delbar* med alla andra delare.

Följdsats 2.3.11. *Låt a och b vara heltal. Ett heltal c kan skrivas på formen $ax + by$ för några heltal x och y om och endast om* $\text{sgd}(a, b) \mid c$.

Bevis. Detta följer av definitionen av $\langle a, b\rangle$ samt Sats 2.3.8 och Sats 2.3.9. □

2.4 Ekvationen $aX + bY = c$

Låt oss nu studera den diofantiska ekvationen $aX+bY = c$, som är den kanske enklaste av alla diofantiska ekvationer. Vi antar alltså att a, b och c är fixerade heltal, och vi vill beskriva alla heltalslösningar X och Y till ekvationen.

Låt oss börja med det enklaste fallet att $c = 0$.

Sats 2.4.1. *Låt a och b vara heltal, inte bägge noll. Alla lösningar till den diofantiska ekvationen*

$$aX + bY = 0$$

ges av

$$X = \frac{b}{\text{sgd}(a, b)} \cdot N \qquad \textit{och} \qquad Y = -\frac{a}{\text{sgd}(a, b)} \cdot N$$

för $N \in \mathbb{Z}$.

Bevis. Vi ber läsaren att kontrollera att om X och Y ges av formeln ovan, så kommer ekvationen $aX + bY = 0$ att vara uppfylld. Vi vill därför nu endast visa att det inte finns några andra lösningar än dessa.

Låt (X, Y) vara en lösning. Enligt Sats 2.3.7 finns heltal x och y sådana att $ax + by = \text{sgd}(a, b)$. Vi multiplicerar denna ekvation med X och finner att $aXx + bXy = X \cdot \text{sgd}(a, b)$. Samtidigt är $aX = -bY$, så vi finner att

$$-bYx + bXy = X \cdot \text{sgd}(a, b),$$

vilket vi skriver om till

$$\frac{b}{\text{sgd}(a, b)} \cdot (-Yx + Xy) = X.$$

Alltså måste X vara delbart med $b/\text{sgd}(a, b)$ och vi får att $X = \frac{b}{\text{sgd}(a,b)} \cdot N$ för något heltal N. Vi sätter in detta i $aX + bY = 0$ och finner att

$$\frac{ab}{\text{sgd}(a, b)} \cdot N + bY = 0$$

och därmed att $Y = -\frac{a}{\text{sgd}(a,b)} \cdot N$. □

Det allmänna fallet ger inga ytterligare svårigheter.

Sats 2.4.2. *Låt a och b vara heltal, inte bägge noll, och låt c vara ett heltal. Den diofantiska ekvationen*

$$aX + bY = c$$

är lösbar om och endast om $\text{sgd}(a, b) \mid c$. *Antag att detta gäller, och låt* (X_0, Y_0) *vara en lösning. Då ges resterande lösningar av formeln*

$$X = X_0 + \frac{b}{\text{sgd}(a,b)} \cdot N \qquad \textit{och} \qquad Y = Y_0 - \frac{a}{\text{sgd}(a,b)} \cdot N$$

för $N \in \mathbb{Z}$.

Bevis. Följdsats 2.3.11 säger att ekvationen är lösbar om och endast om $\text{sgd}(a, b) \mid c$.

Antag att (X, Y) är någon lösning. Då gäller att $(X - X_0, Y - Y_0)$ uppfyller ekvationen

$$a(X - X_0) + b(Y - Y_0) = 0,$$

eftersom

$$a(X - X_0) + b(Y - Y_0) = aX + bY - aX_0 - bY_0 = c - c = 0.$$

Alltså vet vi att

$$X - X_0 = \frac{b}{\text{sgd}(a,b)} \cdot N \qquad \text{och} \qquad Y - Y_0 = -\frac{a}{\text{sgd}(a,b)} \cdot N$$

enligt föregående sats, vilket är ekvivalent med formeln som skulle visas. □

Exempel 2.4.3. Ekvationen

$$3X - 7Y = 5$$

har lösningar, eftersom $\text{sgd}(3, -7) = 1 \mid 5$. För att beskriva alla lösningar måste vi känna till någon lösning till ekvationen. Med lite letande hittar man kanske lösningen $X_0 = 4$ och $Y_0 = 1$. Nu kan den allmänna lösningen skrivas ned som

$$X = 4 - 7N \qquad \text{och} \qquad Y = 1 - 3N \qquad \text{för } N \in \mathbb{Z}$$

enligt Sats 2.4.2. För säkerhets skull kan vi kontrollera att $N = -1$ ger oss $(X, Y) = (11, 4)$ och $N = 1$ ger oss $(X, Y) = (-3, -2)$. Bägge av dessa är lösningar. ▲

Exempel 2.4.4. Antag att a och b är relativt prima. I så fall gäller att $\text{sgd}(a, b) \mid c$ för alla värden på c, så att ekvationen $aX + bY = c$ alltid har oändligt många lösningar, en för varje värde på talet N. ▲

Anmärkning 2.4.5. (För de läsare som redan sett differentialekvationer.) Metoden vi använder i de föregående två satserna är mycket lik det som i differentialekvationer kallas för att hitta den *homogena lösningen* och sedan hitta en *partikulärlösning*. När vi löser ekvationen med högerledet lika med

noll så gör vi något som motsvarar att hitta den homogena lösningen. Sedan antar vi att vi har *någon* lösning (X_0, Y_0), vilken som helst, och denna är motsvarigheten till partikulärlösningen. Slutligen konstateras att alla lösningar fås genom att addera en homogen lösning till partikulärlösningen.

Anledningen att samma metod fungerar i bägge situationerna är just att summan av två lösningar till den homogena ekvationen återigen är en lösning. Man säger att ekvationen är *linjär*.

2.5 Euklides algoritm

En naturlig fråga efter att ha läst Sats 2.3.7 är hur man faktiskt beräknar den största gemensamma delaren av två givna heltal, och hur man sedan hittar tal x och y som löser ekvationen $ax + by = \text{sgd}(a, b)$. Redan i Euklides *Elementa* (300-talet f.Kr.) beskrivs följande metod som brukar kallas för Euklides algoritm. Grundbulten i algoritmen är följande hjälpsats.

Sats 2.5.1. *Låt a, b vara heltal där $b \neq 0$. Antag att heltalen q, r uppfyller*

$$a = b \cdot q + r.$$

Då gäller

$$\text{sgd}(a, b) = \text{sgd}(b, r).$$

Beviset lämnas som en övning åt läsaren.

Denna observation kan användas för att räkna ut största gemensamma delaren till två tal. Vi illustrerar nu uträkningsmetoden med ett exempel.

Exempel 2.5.2 (Euklides algoritm)**.** Låt oss bestämma $\text{sgd}(6\,396, 525)$. Genom att dividera 6 396 med 525 finner man att

$$6\,396 = 525 \cdot 12 + 96,$$

varur det följer att

$$\text{sgd}(6\,396, 525) = \text{sgd}(525, 96).$$

Vidare har vi att

$$525 = 96 \cdot 5 + 45,$$

och därmed

$$\text{sgd}(525, 96) = \text{sgd}(96, 45).$$

Fortsätter man på detta sätt så ser det ut så här:

$$\begin{array}{rl|rl} 6\,396 &= 525 \cdot 12 + 96 & \text{sgd}(6\,396, 525) &= \text{sgd}(525, 96) \\ 525 &= 96 \cdot 5 + 45 & &= \text{sgd}(96, 45) \\ 96 &= 45 \cdot 2 + 6 & &= \text{sgd}(45, 6) \\ 45 &= 6 \cdot 7 + 3 & &= \text{sgd}(6, 3) \\ 6 &= 3 \cdot 2 + 0 & &= \text{sgd}(3, 0) = 3. \end{array}$$

Vi ser alltså att $\text{sgd}(6\,396, 525) = 3$. Algoritmen går alltså ut på att utföra heltalsdivision ett antal gånger tills den går jämnt ut (det vill säga att resten blir 0), och sedan använda det faktum att $\text{sgd}(n, 0) = n$ för alla $n > 0$, vilket diskuterades i Exempel 2.3.6. ▲

I det allmänna fallet börjar man med att utan inskränkning anta att $b > 0$. Sedan används Sats 1.4.1 för att utföra heltalsdivision om och om igen med a/b som första steg tills resttermen blir 0. Man får:

$$\begin{aligned}
a &= b \cdot q_1 + r_1 && \text{där} \quad 0 < r_1 < b \\
b &= r_1 \cdot q_2 + r_2 && \text{där} \quad 0 < r_2 < r_1 \\
r_1 &= r_2 \cdot q_3 + r_3 && \text{där} \quad 0 < r_3 < r_2 \\
r_2 &= r_3 \cdot q_4 + r_4 && \text{där} \quad 0 < r_4 < r_3 \\
&\vdots \\
r_{n-2} &= r_{n-1} \cdot q_n + r_n && \text{där} \quad 0 < r_n < r_{n-1} \\
r_{n-1} &= r_n \cdot q_{n+1} + r_{n+1} && \text{där} \quad r_{n+1} = 0
\end{aligned} \tag{2.1}$$

Att denna process verkligen tar slut, och att resten på den sista raden blir 0 inser man eftersom

$$b > r_1 > r_2 > \cdots > r_n > 0.$$

En följd av positiva heltal som blir mindre och mindre hela tiden måste så småningom nå 0. (Detta påstående är ekvivalent med minimumprincipen.) Enligt Sats 2.5.1 vet vi att

$$\begin{aligned}
\text{sgd}(a, b) &= \text{sgd}(b, r_1) \\
&= \text{sgd}(r_1, r_2) \\
&= \text{sgd}(r_2, r_3) \\
&\vdots \\
&= \text{sgd}(r_{n-1}, r_n) \\
&= \text{sgd}(r_n, 0) \\
&= r_n,
\end{aligned}$$

det vill säga att $r_n = \text{sgd}(a, b)$.

Skriv nu (2.1) baklänges på formen

$$\begin{aligned}
r_n &= r_{n-2} - q_n r_{n-1} \\
r_{n-1} &= r_{n-3} - q_{n-1} r_{n-2} \\
&\vdots \\
r_1 &= a - q_1 b
\end{aligned}$$

och sätt in ekvationerna i varandra för att få formen

$$r_n = x \cdot a + y \cdot b$$

för två heltal x och y.

Exempel 2.5.3. Talen 4 712 och 585 är relativt prima. Det visas på följande sätt:

$$\begin{array}{rl|l} 4\,712 = 585 \cdot 8 & + 32 & \mathrm{sgd}(4\,712, 585) = \mathrm{sgd}(585, 32) \\ 585 = 32 \cdot 18 & + 9 & \qquad\qquad\quad = \mathrm{sgd}(32, 9) \\ 32 = 9 \cdot 3 & + 5 & \qquad\qquad\quad = \mathrm{sgd}(9, 5) \\ 9 = 5 \cdot 1 & + 4 & \qquad\qquad\quad = \mathrm{sgd}(5, 4) \\ 5 = 4 \cdot 1 & + 1 & \qquad\qquad\quad = \mathrm{sgd}(4, 1) \\ 4 = 1 \cdot 4 & + 0 & \qquad\qquad\quad = \mathrm{sgd}(1, 0) = 1. \end{array} \tag{2.2}$$

▲

Exempel 2.5.4. Föregående exempel visar att $\mathrm{sgd}(4\,712, 585) = 1$, men uträkningarna kan också användas för att hitta de tal x, y som enligt Sats 2.3.7 finns och uppfyller

$$1 = 4\,712 \cdot x + 585 \cdot y.$$

Använd (2.2) baklänges och få

$$\begin{array}{rll} 1 & = 5 - 4 \cdot 1 & = 5 - (9 - 5 \cdot 1) \cdot 1 \\ & = -9 + 5 \cdot 2 & = -9 + (32 - 9 \cdot 3) \cdot 2 \\ & = 32 \cdot 2 - 9 \cdot 7 & = 32 \cdot 2 - (585 - 32 \cdot 18) \cdot 7 \\ & = -585 \cdot 7 + 32 \cdot 128 & = -585 \cdot 7 + (4\,712 - 585 \cdot 8) \cdot 128 \\ & = 4\,712 \cdot 128 - 585 \cdot 1\,031. & \end{array}$$

Alltså gäller

$$x = 128 \quad \text{och} \quad y = -1\,031.$$

▲

Övningar

Övning 2.1 ($\star$)**.** Visa att om $a \mid b$ och $b \mid c$, så gäller att $a \mid c$.

Övning 2.2 ($\star$)**.** Låt a vara ett naturligt tal. Antag att talet a har sifferföljden $d_n, d_{n-1}, \ldots, d_1, d_0$. Till exempel är då dess första siffra d_n och dess sista siffra d_0.

(i) Övertyga dig själv om att $a = 10^n d_n + 10^{n-1} d_{n-1} + \ldots + 10 d_1 + d_0$ gäller. Vi tar detta som *definition* av vad vi menar med att siffrorna $d_n, \ldots, d_0$ är decimalutvecklingen av heltalet a.

(ii) Visa att sista siffran är resten vid division av a med 10 som i Divisionssatsen.

(iii) Visa att a är delbart med 2 om och endast om dess sista siffra är jämn.

(iv) Visa att a är delbart med 5 om och endast om dess sista siffra är 0 eller 5.

Övning 2.3 ($\star$)**.** Använd Euklides algoritm för att bestämma:

(i) $\text{sgd}(15, 27)$.

(ii) $\text{sgd}(615, 135)$.

(iii) $\text{sgd}(269, 196)$.

(iv) $\text{sgd}(8\,860, 1\,075)$.

Övning 2.4 ($\star$)**.** Talen 139 och 117 är relativt prima. Enligt Sats 2.3.7 finns heltal x och y sådana att $1 = x \cdot 139 + y \cdot 117$. Använd tekniken i Exempel 2.5.3 för att bestämma x och y.

Övning 2.5 ($\star$)**.** Bestäm alla heltalslösningar till den diofantiska ekvationen $ax + by = c$ för:

(i) $a = 15, b = 27, c = 7$.

(ii) $a = 615, b = 135, c = 15$.

(iii) $a = 269, b = 196, c = 5$.

Ledning: Använd Följdsats 2.3.11 för att avgöra om det finns lösningar. Om det finns, använd samma metod som i Exempel 2.5.3 för att beräkna en lösning till ekvationen. Använd sedan Sats 2.4.2 för att beskriva alla lösningar.

Övning 2.6 ($\star$)**.** Definiera den största gemensamma delaren av n heltal $a_1, a_2, \ldots, a_n$, inte alla noll, med hjälp av mängderna $D(a_1), D(a_2), \ldots, D(a_n)$. Vi skriver

$$\text{sgd}(a_1, a_2, \ldots, a_n).$$

Övning 2.7 ($\star$)**.** Visa att om $ax + by = 1$ för några heltal x och y, så är a och b relativt prima varandra.

Övning 2.8 ($\star\star$)**.** Låt $a > 1$ vara ett heltal. Bestäm

$$\text{sgd}(a, a + 1) \text{ och } \text{sgd}(a, a + 2).$$

Ledning: $\text{sgd}(a, a + 2)$ kommer att bero på om a är jämnt eller udda.

Övning 2.9 ($\star\star$)**.** Låt a och b vara heltal, inte bägge noll. Visa att

$$\frac{a}{\text{sgd}(a, b)} \text{ och } \frac{b}{\text{sgd}(a, b)}$$

är relativt prima.

Övning 2.10 ($\star\star$)**.** Låt a, b vara heltal där $b \neq 0$. Antag att heltalen q, r uppfyller

$$a = b \cdot q + r.$$

Då gäller

$$\text{sgd}(a, b) = \text{sgd}(b, r).$$

Övning 2.11 ($\star\star\star$)**.** Låt a och b vara relativt prima heltal. Antag att $a \mid bc$. I så fall gäller $a \mid c$.

Ledning: Använd Sats 2.3.7 och samma "trick" som används i beviset till Sats 2.4.1 – att multiplicera ekvationen med rätt konstant.

Övning 2.12 ($\star\star\star$)**.** Låt a och b vara relativt prima heltal. Antag att heltalet c uppfyller att

$$a \mid c \quad \text{och} \quad b \mid c.$$

Visa att $ab \mid c$.

Övning 2.13 ($\star\star\star$)**.** Låt a och b vara nollskilda heltal. Ett heltal M kallas en *gemensam multipel* till a och b om $a \mid M$ och $b \mid M$. Låt oss för enkelhets skull anta att a och b är positiva. Visa att:

(i) Det existerar en minsta positiv gemensam multipel till a och b. Vi kallar detta tal för den *minsta gemensamma multipeln* till a och b och skriver $\text{mgm}(a, b)$.

(ii) Varje gemensam multipel till a och b är delbar med $\text{mgm}(a, b)$.

Ledning: Titta på beviset till Sats 2.3.8.

(iii) Det gäller att

$$\text{mgm}(a, b) = \frac{ab}{\text{sgd}(a, b)}.$$

Ledning: Visa att $\frac{ab}{\text{sgd}(a,b)}$ är en gemensam multipel till a och b. Enligt (i) är det då en multipel av $\text{mgm}(a, b)$. Del (i) och övning 2.9 kan användas för att visa att de måste vara lika.

3 Modulär aritmetik

I detta kapitel introducerar vi det användbara begreppet *moduloräkning*, som kan ses som ett systematiskt sätt att räkna med resten vid heltalsdivision. Ett sätt att tänka på räkning modulo n är att man ersätter varje heltal m med det tal man får som rest vid heltalsdivision av m med n. Det visar sig att man kan definiera addition, subtraktion och multiplikation av dessa rester med varandra på ett meningsfullt sätt.

Sedan visar vi med exempel hur moduloräkning kan användas för att studera diofantiska ekvationer. Kapitlet avslutas med en diskussion om under vilka omständigheter man även kan använda det fjärde räknesättet division i moduloräkning.

3.1 Moduloräkning

Låt oss nu introducera moduloräkning. Detta är ett annorlunda sätt att räkna med heltal på, men egentligen bygger det på de vanliga räknesätten. Allt utgår från följande definition:

Definition 3.1.1. Låt $n \geqslant 1$ vara ett heltal. Vi säger att heltalen a och b är *kongruenta modulo* n, och skriver

$$a \equiv b \pmod{n},$$

om $n \mid (a - b)$, det vill säga att $a - b = qn$ för något heltal q.

Exempel 3.1.2. Låt $n = 12$. Exempelvis har vi att

$$16 \equiv 4 \pmod{12}, \quad -27 \equiv -3 \pmod{12} \quad \text{och} \quad 22 \equiv -2 \pmod{12},$$

eftersom 12 är en delare till alla tre talen $16 - 4 = 12$, $(-27) - (-3) = -24$ och $22 - (-2) = 24$. ▲

När man räknar modulo ett tal, exempelvis 12, kan man se det som att man formellt identifierar alla tal som är kongruenta med varandra. Enligt exemplet ovan är 16 och 4 ”samma tal” modulo 12, liksom 22 och -2. Därför verkar det inte alldeles orimligt, så länge man räknar modulo 12, att $16 + 22$ borde bli ”samma tal” som $4 - 2$. Mer korrekt kan vi skriva detta som

$$16 + 22 \equiv 4 - 2 \pmod{12}.$$

Denna kongruens gäller eftersom

$$16 + 22 - (4 - 2) = 36 = 12 \cdot 3.$$

Detta beteende är ingen speciellt för $n = 12$, utan gäller i allmänhet, vilket visas nedan.

Sats 3.1.3. *Låt n vara ett positivt heltal. Antag att heltalen $a, \tilde{a}$ samt $b, \tilde{b}$ uppfyller*

$$a \equiv \tilde{a} \pmod{n} \quad \textit{och} \quad b \equiv \tilde{b} \pmod{n}.$$

Då gäller

$$a + b \equiv \tilde{a} + \tilde{b} \pmod{n}$$

samt

$$a \cdot b \equiv \tilde{a} \cdot \tilde{b} \pmod{n}.$$

Bevis. Per definition vet vi att $n \mid (a - \tilde{a})$ och $n \mid (b - \tilde{b})$. Det betyder att det finns heltal x och y sådana att

$$a - \tilde{a} = nx \quad \text{och} \quad b - \tilde{b} = ny.$$

Nu följer

$$\begin{aligned}(a + b) - (\tilde{a} + \tilde{b}) &= (a - \tilde{a}) + (b - \tilde{b}) \\ &= nx + ny \\ &= n \cdot (x + y).\end{aligned}$$

Alltså gäller $n \mid (a + b) - (\tilde{a} + \tilde{b})$, vilket betyder att

$$a + b \equiv \tilde{a} + \tilde{b} \pmod{n}.$$

Vidare,

$$\begin{aligned}ab - \tilde{a}\tilde{b} &= ab - a\tilde{b} + a\tilde{b} - \tilde{a}\tilde{b} \\ &= a \cdot (b - \tilde{b}) + (a - \tilde{a}) \cdot b \\ &= a \cdot ny + nx \cdot b \\ &= n \cdot (ya + xb),\end{aligned}$$

och därmed $n \mid (ab - \tilde{a}\tilde{b})$, det vill säga

$$ab \equiv \tilde{a}\tilde{b} \pmod{n}.$$ □

Exempel 3.1.4. För att förstå hur kraftfullt föregående resultat är, kan det vara instruktivt att titta på ett exempel. Antag att vi vill veta om

$$67 \cdot 56^5 + 789 \cdot 345 \equiv 2 \pmod{5}.$$

Det naiva sättet att avgöra detta skulle vara att räkna ut värdet på vänsterledet (förslagsvis med en miniräknare), och se om svaret minus 2 är delbart med 5. Men föregående sats säger att kongruensen inte förändras om vi ersätter vart och ett av talen 67, 56, 789 och 345 med tal som de är kongruenta med modulo 5. Eftersom

$$67 = 65 + 2 \equiv 2 \pmod{5} \qquad \text{och} \qquad 56 = 55 + 1 \equiv 1 \pmod{5},$$

så finner vi att $67 \cdot 56^5 \equiv 2 \cdot 1^5 \equiv 2 \pmod 5$ enligt föregående sats. På samma sätt har vi att

$$789 \equiv 4 \pmod 5 \qquad \text{och} \qquad 345 \equiv 0 \pmod 5,$$

så att $789 \cdot 345 \equiv 4 \cdot 0 \equiv 0 \pmod 5$. Alltså är

$$67 \cdot 56^5 + 789 \cdot 345 \equiv 2 + 0 \equiv 2 \pmod 5,$$

så kongruensen är bevisad, helt utan miniräknare. ▲

Exempel 3.1.5. Varje heltal är udda eller jämnt. Detta påstående kan tolkas som en utsaga om hur man räknar modulo 2. Ett tal n är jämnt precis om $2 \mid n$, det vill säga om

$$n \equiv 0 \pmod 2,$$

och n är udda precis om $n - 1$ är jämnt, det vill säga

$$n \equiv 1 \pmod 2.$$

Alltså är varje heltal kongruent med 0 eller 1 modulo 2. Med denna tolkning av udda och jämna tal kan vi från Sats 3.1.3 avläsa följande välbekanta påståenden:

(i) summan av två tal är jämnt om bägge talen är jämna eller bägge är udda;

(ii) summan av två tal är udda om det ena talet är jämnt och det andra är udda;

(iii) produkten av två tal är jämnt om något av talen är jämnt, annars är produkten udda. ▲

En fördel med att räkna modulo ett heltal n är att man kan räkna som om det endast existerade ändligt många olika heltal. Låt oss precisera vad vi menar med detta påstående.

Definition 3.1.6. Låt n vara ett positivt heltal. Vi definierar $\mathbb{Z}_n$ som mängden av heltal $\{0, 1, \ldots, n-1\}$.

Sats 3.1.7. *Låt $n \geqslant 1$ vara ett heltal. För varje heltal m finns ett unikt element $r \in \mathbb{Z}_n$ sådant att*

$$m \equiv r \pmod n.$$

Vi säger att r är resten *av m modulo n, och skriver $r = \overline{m}$ om n är underförstått från sammanhanget.*

Bevis. Följande påståenden är ekvivalenta:

$$\begin{aligned} m \equiv r \pmod{n} \quad &\iff \\ n \text{ delar } (m-r) \quad &\iff \\ \text{det finns ett heltal } q \text{ sådant att } m - r = qn \quad &\iff \\ \text{det finns ett heltal } q \text{ sådant att } m = qn + r. & \end{aligned}$$

Men enligt divisionsalgoritmen (Sats 1.4.1) finns det för alla heltal m två heltal q och r med $0 \leqslant r < n$ (det vill säga, med $r \in \mathbb{Z}_n$) sådana att

$$m = qn + r,$$

och enligt Övning 1.10 är q och r unikt bestämda. Tydligen är därför r det unika talet i $\mathbb{Z}_n$ som uppfyller $m \equiv r \pmod{n}$. □

Anmärkning 3.1.8. Detta förklarar också varför vi kallar r för resten modulo n: talet r är precis det som traditionellt kallas resten vid division av m med n. När vi talar om värdet av något tal m modulo n kan vi därför alltid anta att vi menar något av talen i mängden $\mathbb{Z}_n$.

Exempel 3.1.9. Låt $a = 228 \cdot 115$. Låt $r = \overline{a}$ vara resten av a modulo 8. Vi ska bestämma r.

Börja med att utföra heltalsdivision med 8 av talen 228 och 115:

$$228 = 8 \cdot 28 + 4 \quad \text{och} \quad 115 = 8 \cdot 14 + 3.$$

Detta betyder att $228 - 4 = 8 \cdot 28$ och $115 - 3 = 8 \cdot 14$, det vill säga att $8 \mid (228 - 4)$ och $8 \mid (115 - 3)$. Alltså gäller

$$228 \equiv 4 \pmod{8} \quad \text{och} \quad 115 \equiv 3 \pmod{8}.$$

Enligt satsen ovan får vi

$$a = 228 \cdot 115 \equiv 4 \cdot 3 = 12 \pmod{8}.$$

Men eftersom det är så att $12 \equiv 4 \pmod{8}$ så följer

$$a \equiv 4 \pmod{8}$$

och därmed $\overline{a} = r = 4$. ▲

Poängen med ovanstående exempel är att det är mycket enklare att heltalsdividera talen 228 och 115 med 8 än vad det är att utföra heltalsdivisionen med deras produkt $a = 228 \cdot 115 = 26\,200$. Vi ska förtydliga detta med ytterligare ett exempel.

Exempel 3.1.10. Låt $a = 3^{100}$, och låt $r = \overline{a}$ vara resten av a modulo 6. Vi ska bestämma r.

Observera att

$$3^3 = 27 = 6 \cdot 4 + 3 \equiv 3 \pmod 6.$$

och därmed

$$3^9 = (3^3)^3 = 3^3 \equiv 3 \pmod 6.$$

Alltså gäller

$$a = 3^{100} = 3 \cdot 3^{99} = 3 \cdot (3^9)^{11} \equiv 3 \cdot 3^{11} = 3^{12} \pmod 6.$$

Vidare följer

$$a \equiv 3^{12} = (3^3)^4 \equiv 3^4 = 3 \cdot 3^3 \equiv 3 \cdot 3 = 9 \equiv 3 \pmod 6,$$

vilket betyder att vi kan skriva

$$a = 6 \cdot q + 3,$$

för något heltal q. Alltså får vi att $r = 3$. ▲

I detta exempel får vi en väldigt stor vinst. Talet $a = 3^{100}$ är enormt stort — om man skriver ut det så består det av 48 siffror, vilket antagligen är alldeles för stort för de flesta miniräknare. Så utan moduloräkning blir det mycket svårt att bestämma r.

3.2 Moduloräkning och diofantiska ekvationer

Moduloräkning är en kraftfull teknik för att studera diofantiska ekvationer, dels i allmänhet för att hitta villkor på eventuella lösningar, eller i synnerhet för att visa att en ekvation helt saknar lösningar.

Anledningen att moduloräkning är så användbart är att vi i vår definition av en diofantisk ekvation krävde att det enda som skall ingå i ekvationen är heltal, addition, subtraktion och multiplikation — se Definition 2.1.1 — och dessa räkneoperationer är precis de som går att reducera modulo n för något n, se Sats 3.1.3 och Övning 3.5.

Principen illustrerar vi lättast med ett exempel.

Exempel 3.2.1. Antag att vi vill visa att den diofantiska ekvationen

$$11x^4 - 10xy^3 + 2z^4 - 19 = 0$$

saknar lösningar. Det finns ingen allmän metod för att besvara denna sorts fråga, men något som fungerar ibland (om man har tur) är att prova att reducera ekvationen modulo n för olika värden på n. Om det finns en lösning (x, y, z) i heltalen så kommer det också att finnas en lösning modulo n för varje n (eller hur?). Alltså räcker det att hitta ett enda heltal n för vilket det inte finns någon lösning för att bevisa att ingen heltalslösning kan existera. Och för varje n finns det bara ändligt många värden att pröva, eftersom vi bara behöver låta x, y och z vara varje möjligt tal i $\mathbb{Z}_n$.

Låt oss reducera ekvationen ovan modulo n för de första värdena på n:

(i) Modulo 2 finner vi ekvationen

$$x^4 + 1 \equiv 0 \pmod 2,$$

vilket t.ex. har lösningen $x = 1$. Vi prövar nästa tal.

(ii) Modulo 3 finner vi

$$2x^4 - xy^3 + 2z^4 - 1 \equiv 0 \pmod 3,$$

vilket t.ex. har lösningen $(1, 0, 1)$.

(iii) Modulo 4 finner vi

$$3x^4 + 2xy^3 + 2z^2 + 1 \equiv 0 \pmod 4.$$

Denna har t.ex. lösningen $(1, 0, 0)$.

(iv) Modulo 5 finner vi slutligen

$$x^4 + 2z^4 + 1 \equiv 0 \pmod 5.$$

Det finns 5 möjliga värden för x respektive z i $\mathbb{Z}_5$, så det finns sammanlagt 25 kombinationer. Provar man att sätta in var och en av dessa 25 möjligheter i ekvationen, så finner man att ingen av dem ger en lösning. Voilà! Vi har alltså visat att ekvationen inte kan ha några heltalslösningar. ▲

Anmärkning 3.2.2. I slutet av föregående exempel visade det sig att man behövde prova alla 25 kombinationer av x och z för att visa att det saknades lösningar. Man kan dock göra en viss besparing av arbete genom att vara systematisk, och först prova vilka värden som egentligen kan antas av x^4 (och därmed också vilka värden som kan antas av z^4). Man finner nämligen att det endast finns två möjligheter, vilket kommer att visas i Övning 3.2: antingen är $x^4 = 0$ i $\mathbb{Z}_5$, eller så är $x^4 = 1$ i $\mathbb{Z}_5$. Det räcker alltså att prova med $x^4 \in \{0, 1\}$ och $z^4 \in \{0, 1\}$. Resultatet kan sammanfattas i följande tabell:

$x^4 \pmod 5$	$z^4 \pmod 5$	$x^4 + 2z^4 + 1 \pmod 5$
0	0	1
1	0	2
0	1	3
1	1	4

Som synes är det omöjligt att $x^4 + 2z^4 + 1 \equiv 0 \pmod 5$.

Att det fanns så få olika möjliga värden på x^4 är ingen slump, som vi ska få se i nästa kapitel.

Exempel 3.2.3. Här är ett exempel från "verkligheten" om hur viktigt moduloräkning är för att studera diofantiska ekvationer.

D. J. Lewis, i *Diophantine equations: p-adic methods*, Studies In Number Theory, (1969), påstår på sidan 26, "The equation $x^3 + 117y^3 = 5$ is known to have at most 18 integral solutions[1] but the exact number is unknown". Inget

[1] integral solutions = heltalslösningar

bevis eller referens ges.

R. Finkelstein och H. London bevisade sedan i artikeln *On D. J. Lewis's equation* $x^3+117y^3 = 5$, Canadian Mathematical Bulletin 14 (1971) att ekvationen saknar lösningar. Deras bevis är inte jättesvårt men använder en del algebraisk talteori och satser om strukturen för så kallade kubiska utvidgningar av de rationella talen.

Men efter detta påpekar Valeriu Şt. Udrescu i en kort artikel *On D. J. Lewis's equation* $x^3 + 117y^3 = 5$, Revue Roumaine de Mathématiques Pures et Appliquées 18 (1973) att om man reducerar ekvationen modulo 9 så får man kvar $x^3 \equiv 5 \pmod 9$, som man direkt kan kontrollera inte har någon lösning:

$x \pmod 9$	$x^3 \pmod 9$
0	0
1	1
2	8
3	0
4	1
5	8
6	0
7	1
8	8

Hade Finkelstein eller London provat med lite moduloräkning hade de alltså antagligen kunnat ge ett mycket kort och enkelt bevis att det inte finns några lösningar! Sannolikt är att Lewis råkat ut för ett slarvfel i sin artikel och att han menade någon annan ekvation. Det är dock okänt vilken. ▲

Anmärkning 3.2.4. När man konstruerar tabeller som föregående kan det ofta löna sig att använda lite "räknetricks" för att slippa multiplicera stora tal med varandra. Till exempel vet vi att $(-x)^3 = -(x^3)$ för alla x. Detta gör att vi endast behöver räkna ut $x^3 \pmod 9$ för $x = 0, 1, 2, 3, 4$. För resterande tal använder vi att $5^3 \equiv -4^3$, $6^3 \equiv -3^3$, $7^3 \equiv -2^3$ och $8^3 \equiv -1^3$.

3.3 Enheter och division modulo n

Definition 3.3.1. Ett element $a \in \mathbb{Z}_n$ sägs vara en *enhet* modulo n om det finns ett tal $x \in \mathbb{Z}$ sådant att $xa \equiv 1 \pmod n$. Vi kallar x för en (multiplikativ) *invers* till a modulo n.

Anmärkning 3.3.2. Ett sätt att tänka på definitionen av enhet är följande. Enligt Sats 3.1.3 kan vi tänka på räkneoperationerna addition och multiplikation inte bara i $\mathbb{Z}$, $\mathbb{Q}$ eller $\mathbb{R}$, utan även i $\mathbb{Z}_n$ genom att hela tiden räkna modulo n. Till exempel kan vi säga att talen 5 och 7 har summan 3 och produkten 8 i $\mathbb{Z}_9$. Dessutom kan vi tydligen tala om subtraktion i $\mathbb{Z}_n$. Till exempel kan vi säga att 5 minus 7 blir 7 i $\mathbb{Z}_9$. Så vi har tre av fyra räknesätt. Men hur är det med division? Vad ska vi mena med 5 dividerat med 7?

Relationen mellan enheter och division är följande. En möjlig definition av division som gäller i, säg, de reella talen, är att division med x är detsamma som multiplikation med x^{-1}, där $x^{-1} = 1/x$ är ett tal som uppfyller $x^{-1} \cdot x = 1$. Men jämför detta med definitionen av en enhet!

Alltså kan man tänka att i $\mathbb{Z}_n$ har vi dels räknesätten addition, subtraktion och multiplikation, men utöver dessa har vi också en partiellt definierad divisionsoperation: vi kan utföra divisionen b/a genom att multiplicera b med a^{-1}, *men detta är bara möjligt om talet a är en enhet.* Man kan jämföra detta med situationen i de reella talen, där vi kan utföra divisionen a/b endast om talet b är nollskilt. Alltså finns det även i $\mathbb{R}$ vissa tal som ej går att dividera med.

Idéerna i denna lite informella beskrivning av räkneoperationer modulo n kan preciseras och formaliseras ytterligare, exempelvis genom att införa begreppet *kommutativ ring*. Vi kommer dock inte att göra detta i det här kompendiet.

Hjälpsats 3.3.3. *Produkten av två enheter är en enhet.*

Bevis. Antag att a och b är enheter. Då finns x och y med $ax \equiv by \equiv 1 \pmod{n}$. Men då är $(ab)(xy) = (ax)(by) \equiv 1 \pmod{n}$, så att ab är en enhet. □

Följande sats beskriver exakt vilka tal som är enheter i $\mathbb{Z}_n$.

Sats 3.3.4. *Låt $n \geqslant 1$ vara ett heltal, och låt a vara ett heltal. Följande påståenden är ekvivalenta:*

(i) $\text{sgd}(a, n) = 1$.

(ii) a är en enhet modulo n.

(iii) Om heltalen b, c uppfyller $ab \equiv ac \pmod{n}$ så måste $b \equiv c \pmod{n}$.

En vanlig metod när man vill bevisa att flera påståenden är ekvivalenta är att ge en ”cirkulär följd” av implikationer: i detta fall kommer vi till exempel att visa att (i) $\implies$ (ii), (ii) $\implies$ (iii) samt att (iii) $\implies$ (i). Ur detta följer att alla påståenden är parvis ekvivalenta.

Bevis. Vi visar först att (i) $\implies$ (ii). När a och n är relativt prima finns enligt Sats 2.3.7 heltal x, y sådana att

$$1 = xa + yn.$$

Detta betyder att $1 - xa = n \cdot y$, det vill säga att $n \mid (1 - xa)$, och därmed gäller $xa \equiv 1 \pmod{n}$, så att a är en enhet.

Vi visar sedan att (ii) $\implies$ (iii). Antag att heltalen b, c uppfyller

$$ab \equiv ac \pmod{n}, \tag{3.1}$$

och att $xa \equiv 1 \pmod{n}$. Multiplicera ekvation (3.1) med x, och vi finner att

$$xab \equiv xac \pmod{n}$$

vilket i sin tur ger att $b \equiv c \pmod{n}$ eftersom $xa \equiv 1 \pmod{n}$.

Slutligen visas att (iii) $\implies$ (i). Antag att d delar både a och n. Då gäller att

$$a \cdot \frac{n}{d} = \frac{a}{d} \cdot n \equiv 0 \pmod{n}.$$

Samtidigt gäller

$$a \cdot 0 \equiv 0 \pmod{n}.$$

Men nu har vi visat att $a \cdot (n/d) \equiv a \cdot 0 \pmod{n}$, så enligt vårt antagande måste det nu gälla att $n/d \equiv 0 \pmod{n}$. Med andra ord är n/d delbart med n. Detta är bara möjligt om $d = 1$, så a och n är relativt prima. □

Implikationen (ii) $\implies$ (iii) visar på det fruktbara i att tänka på enheter som ”tal man kan dividera med”. Ty vi är givna ekvationen

$$ab \equiv ac \pmod{n}$$

och vill dra slutledningen att

$$b \equiv c \pmod{n}.$$

Det naturliga sättet att göra detta är att ”dividera med talet a”, och antagandet att a är en enhet är vad som låter oss göra detta eftersom vi kan multiplicera med inversen a^{-1}.

Övningar

Övning 3.1 ($\star$)**.** Låt a vara ett heltal. Om a är jämnt är $a^2 \equiv 0 \pmod{4}$ och om a är udda så är $a^2 \equiv 1 \pmod{4}$.

Övning 3.2 ($\star$)**.** Låt a vara ett heltal. Om $5 \mid a$ så är $a^4 \equiv 0 \pmod{5}$ och om $5 \nmid a$ så är $a^4 \equiv 1 \pmod{5}$. Detta visar att $a^4 \in \{0, 1\}$ i $\mathbb{Z}_5$.

Övning 3.3 ($\star$)**.** Visa följande utan att utföra ”den långa” multiplikationen:

(i) $1234567 \cdot 90123 \equiv 1 \pmod{10}$.

(ii) $2468 \cdot 13579 \equiv -3 \pmod{25}$.

Övning 3.4 ($\star$)**.** Beräkna $2^{99} \pmod{5}$, $3^{99} \pmod{7}$ och $4^{99} \pmod{9}$.

Övning 3.5 ($\star$)**.** Låt $n \geqslant 1$ vara ett heltal. Antag att $a \equiv b \pmod{n}$. Visa att $-a \equiv -b \pmod{n}$.

Övning 3.6 ($\star$)**.** Vilka tal är enheter in $\mathbb{Z}_5$, $\mathbb{Z}_8$ och $\mathbb{Z}_{10}$? För varje enhet x, beräkna x^{-1}.

Övning 3.7 ($\star\star$)**.** Låt x ha decimalutvecklingen $d_n, \ldots, d_0$ som i Övning 2.2.

(i) Visa att $x \equiv d_n + d_{n-1} + \ldots + d_1 + d_0 \pmod 9$. Dra slutsatsen att ett tal är delbart med 9 om och endast om summan av dess siffror är delbar med 9.

(ii) Visa att $x \equiv d_0 - d_1 + \ldots + (-1)^n d_n \pmod{11}$. Dra slutsatsen att ett tal är delbart med 11 om och endast om dess alternerande summa av siffror är delbar med 11.

Övning 3.8 ($\star\star$)**.** För heltal x, y gäller följande: Om $xy = 0$ och $x \neq 0$, så är $y = 0$.

(i) Visa att detta gäller modulo 3 genom att testa alla möjliga fall.

(ii) Visa att detta inte gäller modulo 4 och inte heller modulo 6.

Övning 3.9 ($\star\star$)**.** Visa att följande diofantiska ekvationer inte är lösbara:

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

(ii) $7x^3 + 2 = y^3$.

Ledning: Reducera ekvationerna modulo rätt tal.

Övning 3.10 ($\star\star$)**.** I Exempel 3.2.3 visade det sig att ekvationen

$$(x+3)^3 \equiv x^3 \pmod 9$$

gäller för varje heltal x. Detta kan nämligen läsas av ur tabellen som ger värdet på x^3 för alla $x \in \mathbb{Z}_9$. Ge ett "direkt" bevis av detta påstående, det vill säga, ett bevis som inte går ut på att prova alla olika värden av x i $\mathbb{Z}_9$.

Övning 3.11 ($\star\star\star$)**.** Visa att om a är en enhet modulo n och $b = a^{-1}$, så är även b en enhet modulo n. Vad är b^{-1}?

4 Primtal

I detta avsnitt kommer vi att använda de satser som vi har bevisat i de föregående två kapitlen för att ge tre korta bevis av viktiga resultat. Vi visar först att varje positivt heltal kan skrivas på ett unikt sätt som en produkt av primtal. Sedan bevisar vi att det finns oändligt många primtal.

Kapitlet avslutas med det som brukar kallas för Eulers sats, som handlar om hur många gånger man måste multiplicera en enhet med sig själv för att få resultatet 1.

4.1 Grundläggande definitioner

Definition 4.1.1. Ett heltal $p > 1$ sägs vara ett *primtal* om de enda positiva heltal som delar p är talen 1 och p. Ett heltal som inte är ett primtal kallas *sammansatt*.

Det kan vara värt att poängtera att *alla* heltal delas av 1 och sig självt. Om n är ett heltal har vi ju att $n = n \cdot 1$, vilket enligt definitionen av delare betyder både att $n \mid n$ och att $1 \mid n$. Men primtalen är alltså de enda heltalen, större än 1, som inte delas av något annat positivt heltal än just dessa två.

Exempel 4.1.2. De första tio primtalen är:

$$2, \quad 3, \quad 5, \quad 7, \quad 11, \quad 13, \quad 17, \quad 19, \quad 23, \quad 29. \qquad \blacktriangle$$

Om ett tal n kan *faktoriseras*, det vill säga skrivas som en produkt av andra positiva heltal än 1 och n, är det inte ett primtal. Exempelvis är inte 12 ett primtal eftersom $12 = 3 \cdot 4$, vilket bland annat betyder att talen 3 och 4 delar 12. Tal som kan faktoriseras är alltså sammansatta.

Exempel 4.1.3. Följande tal är *inte* primtal, eftersom de kan faktoriseras:

$$6 = 2 \cdot 3, \quad 36 = 4 \cdot 9, \quad 64 = 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2. \qquad \blacktriangle$$

Hjälpsats 4.1.4. *Om p och q är olika primtal, så är p och q relativt prima.*

Bevis. Antag att ett tal n delar p och q. De enda tal som delar p är 1 och p, och de enda som delar q är 1 och q. Eftersom $p \neq q$ är den enda möjligheten att $n = 1$. □

4.2 Primtalsfaktorisering

Låt oss nu titta på hur man kan faktorisera sammansatta tal. Betrakta talet 60. Det finns ett antal olika sätt att faktorisera det på, exempelvis:

$$60 = 6 \cdot 10, \quad 60 = 3 \cdot 20, \quad 60 = 3 \cdot 4 \cdot 5, \quad 60 = 2 \cdot 2 \cdot 3 \cdot 5.$$

Observera särskilt den sista faktoriseringen. Den består bara av primtal. Försök nu hitta en faktorisering av 60 som består av andra primtal än dessa. Försök dock inte för länge, för det går inte. Detta är inget speciellt för talet 60, utan en egenskap som alla heltal har.

Det visar sig nämligen att alla naturliga tal inte bara kan skrivas som en produkt av primtal, utan dessutom att det bara finns ett enda sätt att göra detta på. Visserligen är det så att

$$60 = 2 \cdot 5 \cdot 3 \cdot 2 = 5 \cdot 3 \cdot 2 \cdot 2,$$

vilket betyder att man kan ordna om primtalsfaktorerna i talet 60 för att få faktoriseringar som ser olika ut. Men det väsentliga är att det alltid är samma primtalsfaktorer som förekommer (i fallet med 60 är primtalsfaktorerna 2, 3 och 5), och antalet gånger varje primtal förekommer är också detsamma (för talet 60 förekommer 2 två gånger, medan 3 och 5 förekommer en gång vardera). Låt oss nu bevisa detta.

Sats 4.2.1. *Varje heltal $m > 1$ har* minst *en faktorisering i primtalsfaktorer.*

Bevis. Antag motsatsen. I så fall är mängden av heltal större än 1 som saknar primtalsfaktorisering icke-tom, så den har ett minsta element M enligt minimumprincipen (1.4.3).

Det kan inte vara så att M är ett primtal, för då är det redan faktoriserat i primtal ($M = p$, där p är ett primtal). Alltså är M inte ett primtal, vilket betyder att det finns en delare a till M som inte är 1 eller M. Enligt definitionen av delare finns också b sådant att $M = ab$. Uppenbarligen kan inte heller b vara 1 eller M. Vi kan utan inskränkning anta att $a, b > 0$. Alltså gäller $2 \leqslant a < M$ och $2 \leqslant b < M$. Men eftersom M är det minsta tal $\geqslant 2$ som inte har en primtalsfaktorisering, så måste både a och b ha primtalsfaktoriseringar. Vi kan alltså skriva

$$a = q_1 q_2 \cdots q_k \quad \text{och} \quad b = r_1 r_2 \cdots r_l,$$

där $q_1, q_2, \ldots, q_k, r_1, r_2, \ldots, r_l$ är primtal. Men nu gäller ju

$$M = ab = q_1 q_2 \cdots q_k r_1 r_2 \cdots r_l,$$

vilket betyder att även M har en primtalsfaktorisering. Detta är en motsägelse, vilket betyder att det inte kan finnas tal $m \geqslant 2$ som saknar primtalsfaktorisering. □

Sats 4.2.2. *Varje heltal $m > 1$ har* högst *en faktorisering i primtalsfaktorer.*

Bevis. Antag motsatsen. Då finns det ett minsta heltal M med egenskapen att M har flera olika primtalsfaktoriseringar enligt minimumprincipen (1.4.3).

Låt

$$M = p_1 p_2 \cdots p_k = q_1 q_2 \cdots q_l$$

vara två distinkta faktoriseringar av M.

Antag att $p_1 \neq q_i$ för alla $1 \leqslant i \leqslant l$. Enligt Hjälpsats 4.1.4 är då p_1 och q_i relativt prima för varje i, så enligt Sats 3.3.4 är varje q_i en enhet modulo p_1. Enligt Hjälpsats 3.3.3 är en produkt av enheter igen en enhet, så att

$$M = q_1 q_2 \cdots q_l$$

är en enhet modulo p_1. Men detta är omöjligt eftersom $M \equiv 0 \pmod{p_1}$, och 0 kan inte vara en enhet.

Det följer att vi har $p_1 = q_i$ för något $1 \leqslant i \leqslant l$. Genom att dividera bägge sidor med p_1, finner vi att

$$p_2 p_3 \cdots p_k = q_1 q_2 \cdots q_{i-1} q_{i+1} \cdots q_l.$$

Men talet

$$p_2 p_3 \cdots p_k = \frac{M}{p_1}$$

är mindre än M, vilket betyder att det har högst en primtalsfaktorisering. Alltså är de två faktoriseringarna $p_2 p_3 \cdots p_k$ och $q_1 q_2 \cdots q_{i-1} q_{i+1} \cdots q_l$ av M/p_1 lika med varandra upp till ordningen av faktorerna. Multiplicerar vi med $p_1 = q_i$ finner vi därför också att de två faktoriseringarna

$$M = p_1 p_2 \cdots p_k = q_1 q_2 \cdots q_l$$

måste vara lika med varandra, vilket ger en motsägelse. □

Tillsammans har vi visat följande påstående:

Sats 4.2.3 (Aritmetikens fundamentalsats)**.** *Varje heltal $m > 1$ har en unik faktorisering i primtal.*

Det vi gjorde i detta beviset är en mycket vanlig strategi när man skall bevisa att det finns ett unikt objekt med en viss egenskap. Man kan nämligen ofta dela in ett sådant bevis i två delar: först ett bevis att det finns minst ett objekt med denna egenskap (existens), och sedan ett bevis att det finns högst ett sådant objekt (unicitet).

4.3 Existens av primtal

En sak som återstår att fråga sig om primtal är hur många det finns. Detta besvaras av nästa sats, vars bevis var känt redan av Euklides på 300-talet f.Kr. Beviset vi ger är ett av världshistoriens mest kända och eleganta.

Sats 4.3.1. *Det finns oändligt många primtal.*

Bevis. Antag att det bara finns ändligt många primtal $p_1, p_2, \ldots, p_n$. Sätt

$$m = 1 + p_1 p_2 \cdots p_n.$$

Enligt aritmetikens fundamentalsats finns primtal $q_1, q_2, \ldots, q_k$ sådana att

$$m = q_1 q_2 \cdots q_k.$$

Eftersom $p_1, p_2, \ldots, p_n$ enligt antagandet är alla primtal som finns, så måste q_1 vara ett av dessa. Vi har alltså $q_1 = p_j$ där $1 \leqslant j \leqslant n$. Speciellt är både m och $m - 1 = p_1 p_2 \cdots p_n$ delbara med q_1. Men enligt Sats 2.2.4 är då även deras differens det, så vi har att

$$q_1 \mid (m - (m - 1)) = 1.$$

Men talet 1 är inte delbart med något primtal, vilket är en motsägelse. Alltså måste det finnas oändligt många primtal. □

4.4 Eulers sats

I förra kapitlet märkte vi i ett exempel att om $5 \nmid x$, så gäller att $x^4 \equiv 1 \pmod 5$. I detta delavsnitt kommer vi att se en förklaring till detta fenomen.

Definition 4.4.1. Låt $n \geqslant 1$ vara ett heltal. Vi låter $\phi(n)$ vara antalet heltal i $\{1, ..., n\}$ som är relativt prima med n. Funktionen $\phi(n)$ kallas *Eulers ϕ-funktion.*

Anmärkning 4.4.2. Den grekiska bokstaven ϕ uttalas *fi.*

Exempel 4.4.3. Här följer de första värdena på $\phi(n)$.

n	1	2	3	4	5	6	7	8	9
$\phi(n)$	1	1	2	2	4	2	6	4	6

▲

Anmärkning 4.4.4. Enligt Sats 3.3.4 är mängden av tal i $\mathbb{Z}_n$ som är relativt prima med n precis samma som mängden av enheter i $\mathbb{Z}_n$. Speciellt ser man därför att antalet enheter i $\mathbb{Z}_n$ ges av $\phi(n)$.

Sats 4.4.5. *(Eulers sats.) Om a är en enhet i $\mathbb{Z}_n$, så gäller att $a^{\phi(n)} \equiv 1 \pmod n$.*

Bevis. Låt mängden av enheter i $\mathbb{Z}_n$ vara

$$X = \{b_1, \ldots, b_{\phi(n)}\}.$$

Vi hävdar nu att mängden

$$Y = \{\overline{ab_1}, \overline{ab_2}, \ldots, \overline{ab_{\phi(n)}}\}$$

innehåller samma element av $\mathbb{Z}_n$, det vill säga, att $Y = X$. Enligt Hjälpsats 3.3.3 är varje element av Y en enhet, så $Y \subseteq X$. Enligt Sats 3.3.4 (iii) så är alla elementen ab_i parvis olika modulo n, så att Y och X har samma antal element.

Alltså är $Y = X$. Det följer att produkten av elementen i Y är densamma som produkten av elementen i X, så att

$$b_1 b_2 \cdots b_{\phi(n)} \equiv a b_1 a b_2 \cdots a b_{\phi(n)} \equiv a^{\phi(n)} b_1 \cdots b_{\phi(n)} \pmod{n}.$$

Alltså är $b_1 \cdots b_{\phi(n)} a^{\phi(n)} \equiv b_1 \cdots b_{\phi(n)} \cdot 1 \pmod{n}$, vilket ger att $a^{\phi(n)} \equiv 1 \pmod{n}$ enligt Sats 3.3.4 (iii) eftersom $b_1 \cdots b_{\phi(n)}$ är en enhet enligt Hjälpsats 3.3.3. □

Exempel 4.4.6. Vi har $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$. Av dessa är alla utom 0 relativt prima med 5, så $\phi(5) = 4$. Det följer att för $0 \neq x \in \mathbb{Z}_5$ gäller $x^4 \equiv 1 \pmod{5}$, vilket förklarar "sammanträffandet" vi såg i Anmärkning 3.2.2. ▲

Anmärkning 4.4.7. Eulers sats är hörnstenen i krypteringsmetoden som kallas för RSA, som med största sannolikhet används i till exempel din bankdosa. Materialet vi har gått igenom hittills i kompendiet är tillräckligt för att förstå hur RSA-algoritmen fungerar. Läs om den på till exempel Wikipedia, där den finns pedagogiskt beskriven! En trevlig populärmatematisk bok om kryptering är *Kodboken* av Simon Singh.

Övningar

Övning 4.1 (⋆)**.** Primtalsfaktorisera talen

$$12, \quad 26, \quad 55, \quad 98, \quad 150, \quad 210, \quad 315, \quad 455.$$

Ledning: Prova att successivt dela talet med primtalen $2, 3, 5, 7, 11, \ldots$. Om ett primtal p delar talet a, så är det ett av primtalfaktorerna, och resten av faktoriseringen hittar man genom att faktorisera kvoten a/p på samma sätt. Exempelvis ser vi att $105/3 = 35$. Så för att hitta faktoriseringen av 105 återstår nu att faktorisera 35. Men $35/5 = 7$ så $105 = 3 \cdot 5 \cdot 7$.

An annan användbar observation är att för att faktorisera ett tal n så räcker att dela med primtal som är mindre än $\sqrt{n}$ — varför gäller detta?

Övning 4.2 (⋆)**.** Låt a vara ett heltal och p vara ett primtal. Visa att antingen gäller $\mathrm{sgd}(a, p) = 1$ eller så gäller $p \mid a$.

Övning 4.3 (⋆)**.** I Övning 2.11 visades följande resultat: Låt a och b vara relativt prima heltal. Antag att $a \mid bc$. I så fall gäller $a \mid c$. Ge ett enklare bevis av detta resultat genom att primtalsfaktorisera a, b och c och använda aritmetikens fundamentalsats.

Övning 4.4 (⋆)**.** Visa följande variant av föregående övning: om p är ett primtal och $p \mid ab$, så måste $p \mid a$ eller $p \mid b$ att gälla. Vad händer om p inte är ett primtal?

Övning 4.5 (⋆)**.** Beräkna $2^{99} \pmod{5}$, $3^{99} \pmod{7}$ och $4^{99} \pmod{9}$ med hjälp av Fermats lilla sats.

Övning 4.6 ($\star\star$)**.** Låt a och b vara icke-negativa heltal. Binomialkoefficienten $\binom{a}{b}$ är definierad genom

$$\binom{a}{b} = \frac{a!}{b!(a-b)!}$$

där $a! = a \cdot (a-1) \cdot \ldots \cdot 2 \cdot 1$ om $a \geqslant 1$ och $0! = 1$. (Det visar sig att detta blir alltid ett heltal!)

(i) Kontrollera att $5 | \binom{5}{k}$ för $k = 1, 2, 3, 4$ och att $7 | \binom{7}{k}$ för $k = 1, 2, \ldots, 6$.

(ii) För vilka k gäller det att $4 | \binom{4}{k}$ och för vilka k gäller det att $6 | \binom{6}{k}$?

(iii) Visa att $p | \binom{p}{k}$ för $k = 1, \ldots, p-1$ om p är ett primtal.

Övning 4.7 ($\star\star$)**.** Låt a, b och c vara positiva heltal. Antag att $a^2 = bc$ och $\text{sgd}(b, c) = 1$. Visa att b och c själva är kvadrater, alltså att $b = s^2$ och $c = t^2$ där $s > 0$ och $t > 0$ är heltal. Visa också att $\text{sgd}(s, t) = 1$.

Övning 4.8 ($\star\star$)**.** Visa Fermats lilla sats i två olika versioner:

(i) Visa att $\phi(p) = p - 1$ om och endast om talet p är ett primtal.

(ii) Visa Fermats lilla sats: $a^{p-1} \equiv 1 \pmod{p}$ om p är ett primtal och $p \nmid a$.

Ledning: Detta är ett specialfall av Eulers sats.

(iii) Visa följande variant av Fermats lilla sats: $a^p \equiv a \pmod{p}$ om p är ett primtal och $a \in \mathbb{Z}$.

Övning 4.9 ($\star\star$)**.** Låt p vara ett primtal, och $n \neq 0$ ett heltal. Definiera $\text{ord}_p(n)$ som det unika heltal $a \geqslant 0$ för vilket $p^a \mid n$ men $p^{a+1} \nmid n$.

(i) Förklara varför $\text{ord}_p(n) > 0$ bara kan gälla för ändligt många primtal p.

(ii) Låt $n, m > 0$. Visa att $n = m$ om och endast om $\text{ord}_p(n) = \text{ord}_p(m)$ för alla primtal p.

(iii) Visa att $\text{ord}_p(nm) = \text{ord}_p(n) + \text{ord}_p(m)$ för heltal n och m.

(iv) Visa att $n \mid m$ om och endast om $\text{ord}_p(n) \leqslant \text{ord}_p(m)$ för alla primtal p.

Övning 4.10 ($\star\star\star$)**.** Använd notationen och resultaten från föregående övning.

(i) Visa att $\text{ord}_p(\text{sgd}(n, m)) = \min(\text{ord}_p(n), \text{ord}_p(m))$.

(ii) Visa att $\text{ord}_p(\text{mgm}(n, m)) = \max(\text{ord}_p(n), \text{ord}_p(m))$.

(iii) Använd dessa observationer för att ge enklare bevis av resultaten i Övning 2.13.

Övning 4.11 ($\star\star\star$)**.** En övning om enheter och deras inverser:

(i) Låt p vara ett primtal. Visa att ekvationen $x = x^{-1}$ bara har lösningarna 1 och $p-1$ i $\mathbb{Z}_p$.

Ledning: Visa att ekvationen $x^2 - 1 \equiv 0$ måste gälla och skriv $x^2 - 1$ som produkt av två faktorer.

(ii) Visa att detta inte gäller i $\mathbb{Z}_8$ och $\mathbb{Z}_{15}$.

Övning 4.12 ($\star\star\star$)**.** Låt p vara ett primtal. Visa att $(p-1)! \equiv -1 \pmod{p}$.

Ledning: Använd föregående uppgift. Dela upp produkten $(p-1)!$ i par av enheter och deras inverser, vad blir produkten?

5 Pythagoreiska tripplar I

Den enda diofantiska ekvation som vi har studerat i detalj var en förstagradsekvation i två variabler, nämligen $aX + bY = c$. I detta kapitel kommer vi att börja titta på ekvationer av andra graden. Mer specifikt kommer vi att titta på en av de mest klassiska och välstuderade sådana, nämligen ekvationen

$$X^2 + Y^2 = Z^2. \tag{5.1}$$

Som tidigare intresserar vi oss bara för heltalslösningar X, Y och Z.

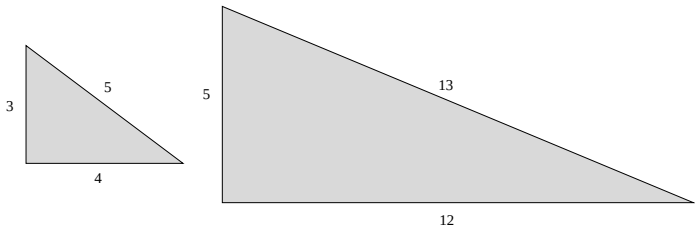
Vi ska hitta en parametrisering av lösningarna till ekvation (5.1), det vill säga, en formel som ger alla lösningar. Det visar sig att det räcker att bestämma en speciell klass av heltalslösningar som vi kallar *primitiva* för att kunna beskriva alla heltalslösningar.

5.1 Pythagoreiska tripplar

Definition 5.1.1. En trippel (X, Y, Z) av heltal kallas för en *pythagoreisk trippel* om $X^2 + Y^2 = Z^2$.

Exempel 5.1.2. En pythagoreisk trippel är $(3, 4, 5)$. Genom att byta tecken på X, Y eller Z i denna trippel fås nya pythagoreiska tripplar, till exempel $(-3, 4, 5)$, $(3, -4, -5)$ och $(-3, -4, -5)$. ▲

Vi kan tolka de pythagoreiska tripplar där X, Y och Z är positiva geometriskt: Enligt *Pythagoras sats* svarar en positiv heltalslösning till (5.1) mot sidlängderna av en rätvinklig triangel. (Eller hur?) Denna geometriska tolkning förklarar namnet *pythagoreisk trippel*.



Figur 5.1: Två rätvinkliga trianglar med sidlängder

Exempel 5.1.3. I Figur 5.1 ser vi två rätvinkliga trianglar. Sidlängderna är 3, 4 och 5 för den ena triangeln samt 5, 12 och 13 för den andra triangeln. Mycket riktigt gäller att $3^2 + 4^2 = 5^2$ samt

$$5^2 + 12^2 = 25 + 144 = 169 = 13^2.$$ ▲

Exempel 5.1.4. Det finns två oändliga familjer av uppenbara pythagoreiska tripplar, nämligen $(X, 0, X)$ och $(0, Y, Y)$ där X och Y är godtyckliga heltal. Dessa lösningar kan tolkas geometriskt som en triangel där en sida har krympt ihop så att dess längd är noll. ▲

Definition 5.1.5. En pythagoreisk trippel är *trivial* om det gäller att $X = Y = Z = 0$. Annars säger vi att trippeln är *icketrivial*.

Vi kommer framför allt att vara intresserade av icketriviala lösningar.

Anmärkning 5.1.6. Förutom att byta tecken kan man även skapa nya lösningar från gamla genom "skalning". Antag att (X, Y, Z) är en pythagoreisk trippel och låt $k \in \mathbb{Z}$ vara ett heltal. Vi gör följande beräkning:

$$(kX)^2 + (kY)^2 = k^2X^2 + k^2Y^2 = k^2(X^2 + Y^2) = k^2Z^2 = (kZ)^2$$

Detta visar att även (kX, kY, kZ) bildar en pythagoreisk trippel.

Antag å andra sidan att (X, Y, Z) är en pythagoreisk trippel och att $k \in \mathbb{Z}$ är ett heltal som delar X, Y och Z. Då blir $\left(\frac{X}{k}, \frac{Y}{k}, \frac{Z}{k}\right)$ en trippel av heltal, och vi kan göra en liknande beräkning som ovan (gör det!) som visar att $\left(\frac{X}{k}, \frac{Y}{k}, \frac{Z}{k}\right)$ också är en pythagoreisk trippel.

Exempel 5.1.7. Vi har redan sett den pythagoreiska trippeln $(3, 4, 5)$. Därmed blir också $(6, 8, 10)$, $(30, 40, 50)$ och allmänt $(3k, 4k, 5k)$ en pythagoreisk trippel om $k \in \mathbb{Z}$ är ett heltal. ▲

Anmärkning 5.1.6 leder oss till följande definition:

Definition 5.1.8. Vi säger att en pythagoreisk trippel (X, Y, Z) är *primitiv* om $Z > 0$ och den största gemensamma delaren till X, Y och Z är 1, dvs $\text{sgd}(X, Y, Z) = 1$.

Hjälpsats 5.1.9. *Varje icketrivial pythagoreisk trippel (X, Y, Z) kan skrivas unikt som (kX_0, kY_0, kZ_0) där $k \in \mathbb{Z}$ och (X_0, Y_0, Z_0) är en primitiv pythagoreisk trippel.*

Bevis. Låt k' vara den största gemensamma delaren till X, Y och Z. (Denna är väldefinierad eftersom det inte gäller att $X = Y = Z = 0$.) Om vi ska ha att $(X, Y, Z) = (kX_0, kY_0, kZ_0)$ och att $\text{sgd}(X_0, Y_0, Z_0) = 1$, så måste $k = \pm k'$. Villkoret att $Z_0 > 0$ bestämmer sedan tecknet på k, så att k är unikt bestämt utifrån (X, Y, Z). □

Exempel 5.1.10. De primitiva pythagoreiska tripplar med $X = 0$ eller $Y = 0$ är

$$(1, 0, 1), \qquad (-1, 0, 1), \qquad (0, 1, 1) \qquad \text{och} \qquad (0, -1, 1). \qquad ▲$$

5.2 Parametrisering av primitiva pythagoreiska tripplar

Vi börjar med tre hjälpsatser som leder oss till den fullständiga beskrivningen av alla primitiva pythagoreiska tripplar i Sats 5.2.4 och Sats 5.2.6.

Hjälpsats 5.2.1. *Om (X, Y, Z) är en primitiv pythagoreisk trippel, så är*

$$\text{sgd}(X, Y) = \text{sgd}(Y, Z) = \text{sgd}(Z, X) = 1.$$

Bevis. Antag motsatsen, t.ex. att $\text{sgd}(X, Y) \neq 1$. Då finns ett primtal p som delar X och Y. Detta primtal måste då även dela X^2 och Y^2, och därav också X^2+Y^2, och därför har vi $p \mid Z^2$. Men enligt Övning 4.4 gäller då $p \mid Z$. Alltså är p en gemensam faktor till både X, Y och Z, vilket säger emot att trippeln var primitiv. På samma sätt visas att $\text{sgd}(Y, Z) = \text{sgd}(Z, X) = 1$. □

Hjälpsats 5.2.2. *Låt (X, Y, Z) vara en primitiv pythagoreisk trippel. Då är Z udda, och det ena av talen X och Y är udda och det andra är jämnt.*

Bevis. Om både X och Y är jämna tal, så är $\text{sgd}(X, Y)$ åtminstone 2 i motsats till Hjälpsats 5.2.1. Alltså kan högst ett av talen X och Y vara jämnt.

Antag att både X och Y är udda. Vi vet att ett udda tal a uppfyller $a^2 \equiv 1 \pmod 4$ enligt Övning 3.1. Vi får att

$$X^2 + Y^2 \equiv 1 + 1 \equiv 2 \equiv Z^2 \pmod 4.$$

Men inget tal Z uppfyller $Z^2 \equiv 2 \pmod 4$, igen enligt Övning 3.1. Alltså kan inte både X och Y vara udda, så exakt ett av talen är jämnt. Då blir $Z^2 = X^2 + Y^2$ udda och därmed även Z. □

I fortsättningen ska vi anta att X är udda och Y är jämnt. Detta är ingen inskränkning eftersom vi annars kan byta namn på båda.

Hjälpsats 5.2.3. *Låt (X, Y, Z) vara en primitiv pythagoreisk trippel med X udda. Då är* $\text{sgd}(Z + Y, Z - Y) = 1$.

Bevis. Låt $d = \text{sgd}(Z+Y, Z-Y)$. Vi finner att d delar $(Z+Y)+(Z-Y) = 2Z$ och att d delar $(Z+Y)-(Z-Y) = 2Y$. Eftersom $\text{sgd}(Z, Y) = 1$ enligt Hjälpsats 5.2.1, kan då bara $d = 1$ och $d = 2$ gälla.

Men d delar också $(Z + Y)(Z - Y) = Z^2 - Y^2 = X^2$ och vi antog att X är udda. Då är $d = 2$ omöjligt och vi får att $d = 1$. □

Vi har nu samlat allt vi behöver för att kunna visa de två viktigaste satserna i detta kapitel.

Sats 5.2.4. *Låt (X, Y, Z) vara en primitiv pythagoreisk trippel med X udda. Då finns det unika udda heltal s och t med $t > 0$ och* $\text{sgd}(s, t) = 1$*, sådana att*

$$X = st, \qquad Y = \frac{s^2 - t^2}{2} \qquad Z = \frac{s^2 + t^2}{2}. \tag{5.2}$$

Bevis. Notera först att

$$X^2 = Z^2 - Y^2 = (Z + Y)(Z - Y).$$

Per definition är $Z > 0$, och uppenbarligen är antingen Y eller $-Y$ icke-negativt. Alltså är minst en faktor i $(Z+Y)(Z-Y)$ positiv. Men deras produkt är X^2, som måste vara positivt, så bägge faktorerna är positiva.

Enligt Lemma 5.2.3 gäller $\text{sgd}(X+Y, X-Y) = 1$. Vi använder nu resultatet av övningsuppgift 4.7 med $a = X$, $b = Z + Y$ och $c = Z - Y$ och får att det finns heltal s och t sådana att

$$s^2 = Z + Y \qquad \text{och} \qquad t^2 = Z - Y \tag{5.3}$$

och $\text{sgd}(s, t) = 1$. Talen s och t är unikt bestämda upp till tecken.

Då blir

$$(st)^2 = X^2,$$

och genom att ta kvadratroten finner vi att

$$st = \pm X.$$

Det följer att s och t båda är udda, eftersom X är udda. Dessutom ser vi att s och t blir unikt bestämda om vi kräver att $t > 0$: ty s och t var tidigare bestämda endast upp till tecken, att kräva att $t > 0$ fixerar tecknet till t, och tecknet till s bestäms sedan av villkoret att $st = X$ ska gälla.

Genom att lösa ut Z och Y ur ekvation (5.3) finner vi slutligen också att

$$Z = \frac{s^2 + t^2}{2} \qquad \text{och} \qquad Y = \frac{s^2 - t^2}{2}.$$

Kontrollera detta! □

Exempel 5.2.5. Vi såg tidigare att det fanns fyra primitiva pythagoreiska tripplar där någon av X och Y var noll, nämligen $(1, 0, 1)$, $(-1, 0, 1)$, $(0, 1, 1)$ och $(0, -1, 1)$. Av dessa har endast de första två udda värde på X. Enligt föregående sats finns därför udda och relativt prima heltal s och t, $t > 0$, sådana att

$$st = 1, \qquad \frac{s^2 - t^2}{2} = 0, \qquad \frac{s^2 + t^2}{2} = 1,$$

respektive

$$st = -1, \qquad \frac{s^2 - t^2}{2} = 0, \qquad \frac{s^2 + t^2}{2} = 1.$$

Om $s^2 - t^2 = 0$ måste $s = \pm t$, och om s och t ska vara relativt prima så måste då s och t vara ± 1. Men $t > 0$ så vi får att $t = 1$, och därför bestäms s ur ekvationerna $st = 1$ respektive $st = -1$. Alltså svarar dessa två primitiva pythagoreiska tripplarna mot

$$(s, t) = (1, 1) \qquad \text{respektive} \qquad (s, t) = (-1, 1). \qquad \blacktriangle$$

Sats 5.2.6. *Låt s och t vara udda och relativt prima heltal. Då är (X, Y, Z) definierad genom*

$$X = st, \qquad Y = \frac{s^2 - t^2}{2} \qquad Z = \frac{s^2 + t^2}{2}$$

en primitiv pythagoreisk trippel.

Bevis. Notera först att eftersom s och t är udda är även s^2 och t^2 udda, vilket ger att $s^2 - t^2$ och $s^2 + t^2$ är jämna. Alltså går divisionen med 2 jämnt ut, så att Y och Z blir heltal. Vi lämnar som övning åt läsaren i slutet av kapitlet att kontrollera att (X, Y, Z) definierade enligt denna formel blir en pythagoreisk trippel. Det återstår att visa att trippeln (X, Y, Z) är primitiv. Det är tydligt att $Z > 0$, så vi behöver endast visa att (X, Y, Z) saknar gemensamma delare. Antag motsatsen, att det finns ett primtal p som delar Y och Z. Då delar p även $Z + Y = s^2$ och $Z - Y = t^2$. Enligt Övning 4.4 gäller då att $p \mid s$ och $p \mid t$. Detta säger emot att s och t är relativt prima. Därmed är $\operatorname{sgd}(X, Y, Z) = 1$ och trippeln (X, Y, Z) är primitiv. □

Sats 5.2.4 och Sats 5.2.6 ger en fullständig beskrivning av alla primitiva pythagoreiska tripplar med X udda. Därmed fås indirekt en beskrivning av alla pythagoreiska tripplar: de resterande fås genom att byta plats på X och Y och att multiplicera trippeln med ett godtyckligt heltal k.

Övningar

Övning 5.1 ($\star$)**.** Visa att om s och t är godtyckliga reella tal, och

$$X = st, \qquad Y = \frac{s^2 - t^2}{2} \qquad Z = \frac{s^2 + t^2}{2},$$

så gäller ekvationen $X^2 + Y^2 = Z^2$.

Övning 5.2 ($\star$)**.** (i) Bestäm alla primitiva pythagoreiska tripplar med $Z \leqslant 30$.

(ii) Bestäm alla pythagoreiska tripplar med $-30 \leqslant Z \leqslant 30$.

Övning 5.3 ($\star\star$)**.** Låt (X, Y, Z) vara en primitiv pythagoreisk trippel med X udda.

(i) Visa att $(Z - X)/2$ är ett heltal.

(ii) Visa att $(Z - X)/2$ är en kvadrat av ett heltal.

Övning 5.4 ($\star\star$)**.** (i) Visa att i en pythagoreisk trippel (X, Y, Z) är alltid X eller Y delbart med 3.

(ii) Visa att i en pythagoreisk trippel (X, Y, Z) är alltid X, Y eller Z delbart med 5.

Övning 5.5 ($\star\star$)**.** Antag att (X, Y, Z) är en pythagoreisk trippel med X udda. Visa att Y är delbart med 4 genom att undersöka ekvationen $X^2 + Y^2 = Z^2$ modulo 8.

Övning 5.6 ($\star\star$)**.** De primitiva tripplarna $(3, 4, 5)$ och $(5, 12, 13)$ uppfyller bägge att $Z = Y + 1$.

(i) Bestäm två andra primitiva tripplar som har denna egenskap.

(ii) Bestäm en primitiv pythagoreisk trippel (X, Y, Z) med denna egenskap som dessutom uppfyller $Z > 100$.

(iii) Bestäm en formel som beskriver alla primitiva pythagoreiska tripplar (X, Y, Z) som uppfyller $Z = Y + 1$. (Tips: Vad kan du säga om s och/eller t från Sats 5.2.4 i detta fall?)

Övning 5.7 ($\star\star\star$)**.** De primitiva tripplarna $(3, 4, 5)$, $(15, 8, 17)$ och $(35, 12, 37)$ uppfyller alla att $Z = X + 2$.

(i) Bestäm tre andra primitiva tripplar som har denna egenskap.

(ii) Bestäm en primitiv pythagoreisk trippel (X, Y, Z) med denna egenskap som dessutom uppfyller $Z > 1000$.

(iii) Bestäm en formel som beskriver alla primitiva pythagoreiska tripplar (X, Y, Z) som uppfyller $Z = X + 2$.

6 Pythagoreiska tripplar II

I detta kapitel ska vi härleda en liten variation av formeln för pythagoreiska tripplar som vi såg i förra kapitlet. Om metoden i föregående kapitel var mer algebraisk, är denna mer geometrisk. En fördel är att det vi gör i detta kapitel kan generaliseras till mer allmänna ekvationer än $X^2 + Y^2 = Z^2$. Detta, i sin tur, kommer att bli ämnet för nästa kapitel.

6.1 Pythagoreiska tripplar och enhetscirkeln

Följande sats ger ytterligare en anledning till varför det är naturligare att betrakta primitiva pythagoreiska tripplar än allmänna pythagoreiska tripplar. Vi rekommenderar läsaren att bläddra tillbaka till definitionen av en bijektion i Definition 1.2.8, då begreppet spelar en central roll i detta kapitel.

Sats 6.1.1. *Det finns en bijektion mellan mängden av primitiva pythagoreiska tripplar och mängden av rationella lösningar till*

$$x^2 + y^2 = 1. \tag{6.1}$$

Bevis. Antag att (X, Y, Z) är en primitiv pythagoreisk trippel. Eftersom trippeln är primitiv måste speciellt $Z \neq 0$, så vi kan definiera

$$x = \frac{X}{Z} \qquad \text{och} \qquad y = \frac{Y}{Z}.$$

Dessa är rationella tal och uppfyller $x^2 + y^2 = 1$ ty

$$x^2 + y^2 = \frac{X^2}{Z^2} + \frac{Y^2}{Z^2} = \frac{X^2 + Y^2}{Z^2} = \frac{Z^2}{Z^2} = 1.$$

Antag nu omvänt att (x, y) är en rationell lösning till ekvation (6.1). Vi kan då skriva på ett unikt sätt

$$x = \frac{a}{c} \qquad \text{och} \qquad y = \frac{b}{d}$$

för heltal a, b, c och d med $\text{sgd}(a, c) = 1$ och $\text{sgd}(b, d) = 1$, och där c och d är positiva. Vi påstår att då måste $c = d$, vilket visas i övningsuppgift 6.1. Vi finner att

$$\frac{a^2}{c^2} + \frac{b^2}{d^2} = \frac{a^2 + b^2}{c^2} = 1,$$

så att

$$a^2 + b^2 = c^2.$$

Vi ser att (a, b, c) är en pythagoreisk trippel, som dessutom är primitiv eftersom $\text{sgd}(a, c) = 1$, $\text{sgd}(b, c) = 1$ och $c > 0$.

Det återstår att visa att dessa två konstruktioner är varandras inverser vilket vi lämnar åt läsaren i Övning 6.7. □

Detta ger en ny geometrisk tolkning av sökandet efter pythagoreiska tripplar: Ekvationen

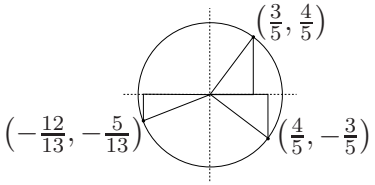
$$x^2 + y^2 = 1$$

beskriver *enhetscirkeln* i planet, det vill säga cirkeln med centrum i origo och radie 1. Hjälpsats 6.1.1 säger att vi för att hitta primitiva pythagoreiska tripplar istället kan leta efter punkter på enhetscirkeln vars koordinater är *rationella tal*.

Definition 6.1.2. Med en *rationell punkt* i planet menar vi en punkt vars x- och y-koordinater är rationella tal.

Exempel 6.1.3. De positiva pythagoreiska tripplarna $(3, 4, 5)$ och $(5, 12, 13)$ som vi sett tidigare svarar mot de rationella punkterna $\left(\frac{3}{5}, \frac{4}{5}\right)$ respektive $\left(\frac{5}{13}, \frac{12}{13}\right)$. Till den pythagoreiska trippeln $(4, -3, 5)$ hör punkten $\left(\frac{4}{5}, -\frac{3}{5}\right)$. Alla tre punkter har rationella koordinater och vi ser att $x^2 + y^2 = 1$ gäller för dem.

Givet punkten $(\frac{4}{5}, -\frac{3}{5})$ som uppfyller $x^2 + y^2 = 1$ konstruerar vi som i beviset ovan trippeln $(4, -3, 5)$ vilken är en primitiv pythagoreisk trippel. Vi ser att vi får tillbaka samma trippel som vi började med.



Figur 6.1: Omskalade rätvinkliga trianglar

▲

6.2 Rationella punkter på enhetscirkeln

För reella tal $c, d \in \mathbb{R}$ kan vi betrakta funktionen f som är given genom

$$y = f(x) = cx + d \tag{6.2}$$

Funktionsgrafen för f är mängden av punkterna $(x, f(x))$ för alla reella tal $x \in \mathbb{R}$ och beskriver en *linje i planet*. Vi kan beskriva varje linje i planet som inte är vertikal på detta sätt. Vertikala linjer ges i stället av ekvationer på formen $x = d$ för $d \in \mathbb{R}$.

Definition 6.2.1. För en linje i planet som är given genom ekvation (6.2), kallas det reella talet c *lutningen* av linjen. Vi säger att lutningen av en vertikal linje är ∞.

Exempel 6.2.2. Lutningen av en horisontell linje är 0. Linjen $y = x$ har lutning 1. ▲

För de elever som är bekanta med derivering nämner vi att lutningen av en linje som inte är vertikal även kan definieras som derivatan av funktionen f. Det gäller att

$$f'(x) = c \qquad \text{för alla } x \in \mathbb{R}.$$

Att lutningen av en vertikal linje ska vara ∞ stämmer överens med vår intuition: Om vi låter lutningen av en linje bli större och större, så kommer den vara närmare och närmare en vertikal linje.

Hjälpsats 6.2.3. *Låt L vara en linje i planet som inte är vertikal. Välj två olika punkter (x_0, y_0) och (x_1, y_1) på L. Då är lutningen av linjen L lika med talet*

$$\frac{y_0 - y_1}{x_0 - x_1}. \tag{6.3}$$

Beviset kommer att ges som övning i slutet av kapitlet.

Anmärkning 6.2.4. Vi skulle kunna ta bort antagandet att linjen L inte är vertikal om vi i stället kom överens om att formellt tolka resultatet av division med noll som ∞. Vi kommer dock att undvika denna sortens "räkning" med ∞ i detta kompendium.

Vi observerar att vi kan beräkna lutningen enligt formel (6.3) för två godtyckliga punkter på linjen. Speciellt beror inte kvoten i formeln på vilka punkter man väljer.

Vi kommer att vara intresserade av linjer som har en rationell lutning och konstaterar följande.

Hjälpsats 6.2.5. *Låt P och Q vara rationella punkter i planet med olika x-koordinater. Då har den unika linjen genom P och Q en rationell lutning.*

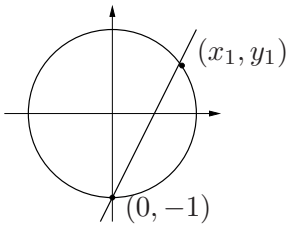
Bevis. Om $P = (x_0, y_0)$ och $Q = (x_1, y_1)$ så kommer lutningen enligt Hjälpsats 6.2.3 att ges av

$$\frac{y_0 - y_1}{x_0 - x_1}.$$

Vi ser att detta blir ett rationellt tal eftersom alla koordinater x_0, x_1, y_0 och y_1 är rationella och dessutom är både differensen och kvoten av två rationella tal igen ett rationellt tal (utom när man dividerar med 0 vilket inte kan hända här). □

Vi ska nu beskriva alla rationella punkter på enhetscirkeln genom att för varje sådan punkt dra en linje genom punkten själv och punkten $(0, -1)$. Det visar sig att man får alla rationella punkter genom att dra linjer med rationella lutningar genom punkten $(0, -1)$. Ett undantag uppstår endast för punkten $(0, 1)$ där man behöver dra en vertikal linje.

Sats 6.2.6. *Låt P vara punkten $(0, -1)$ på enhetscirkeln. Varje linje genom P vars lutning är ett nollskilt rationellt tal möter cirkeln i P och exakt en annan rationell punkt.*



Figur 6.2: Enhetscirkeln med linje genom $(0, -1)$ och andra skärningspunkten

Bevis. Låt $c \in \mathbb{Q} \setminus \{0\}$ vara lutningen. Linjen som går genom punkten $(0, -1)$ med lutning c har ekvationen

$$y = cx - 1.$$

Sätter vi in detta i ekvationen

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

finner vi att

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

Detta är en andragradsekvation i x, vars lösningar svarar mot x-koordinaterna till skärningspunkterna mellan linjen och cirkeln. Vi förväntar oss inte bara från ekvationen utan även från bilden ovan att hitta två lösningar x_0 och x_1 där ena lösningen borde bli 0, svarande mot x-koordinaten till skärningspunkten $(0, -1)$. Multiplicerar vi ut parentesen finner vi i att

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

vilket är ekvivalent med

$$x\left((c^2 + 1)x - 2c\right) = 0.$$

Detta ger oss den förväntade första lösningen $x_0 = 0$ som borde tillhöra skärningspunkten $(0, -1)$. Vi beräknar också att $y_0 = cx_0 - 1 = -1$. Den andra lösningen uppfyller ekvationen

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

så vi får att

$$x_1 = \frac{2c}{c^2 + 1}.$$

Insättning i linjens ekvation $y = cx - 1$ ger att

$$y_1 = \frac{c^2 - 1}{c^2 + 1}.$$

Om c är rationellt är alltså koordinaterna x_1 och y_1 till den andra skärningspunkten rationella, vilket skulle bevisas. □

Exempel 6.2.7. Betrakta linjen $y = f(x) = \frac{1}{2}x - 1$ som går igenom punkten $(0, -1)$. För att hitta den andra skärningspunkten med cirkeln sätter vi in linjens ekvation i cirkelns ekvation. Vi får

$$x^2 + \left(\frac{1}{2}x - 1\right)^2 = 1$$

det vill säga att

$$x\left(\frac{5}{4}x - 1\right) = 0.$$

Lösningarna är $x = 0$ och $x = \frac{4}{5}$. Vi beräkner y-koordinaten för $x = \frac{4}{5}$ med linjens ekvation och får

$$y = \frac{1}{2} \cdot \frac{4}{5} - 1 = -\frac{3}{5}$$

Den andra skärningspunkten är alltså den rationella punkten $(\frac{4}{5}, -\frac{3}{5})$ i överensstämmelse med de formler som vi hittade i beviset ovan. (Kontrollera detta!) ▲

Sats 6.2.8. *Alla rationella punkter på enhetscirkeln utom* $(0, 1)$ *kan skrivas på ett unikt sätt som*

$$\left(\frac{2c}{c^2+1}, \frac{c^2-1}{c^2+1}\right) \tag{6.4}$$

där $c \in \mathbb{Q}$*, och för varje* $c \in \mathbb{Q}$ *ger formeln ovan en rationell punkt på enhetscirkeln.*

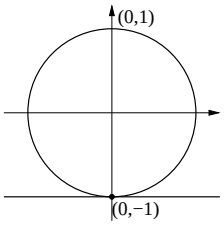
Bevis. Vi har redan sett i slutet av beviset till föregående sats att om $c \in \mathbb{Q} \setminus \{0\}$ får vi en rationell punkt på enhetscirkeln av formeln, och att två olika lutningar c och c' alltid ger två olika punkter. Om $c = 0$ ger formeln punkten $(0, -1)$ som vi också vet är en rationell punkt.

Det återstår att visa att varje rationell punkt utom $(0, 1)$ på cirkeln kan skrivas på detta sätt. Punkten $(0, -1)$ ges av formeln ovan med $c = 0$. För varje annan rationell punkt P på cirkeln så finns en unik linje genom P och $(0, -1)$ och denna linje har en lutning som är ett nollskilt rationellt tal enligt hjälpsats 6.2.5. Enligt beräkningarna i Sats 6.2.6 kan P därför skrivas på den givna formen. □

Anmärkning 6.2.9. Det kanske inte är särskilt förvånande att vi får punkten $(0, -1)$ när vi sätter in $c = 0$ i formel (6.4). Om vi låter lutningen av linjen genom $(0, -1)$ komma närmare och närmare noll, så kommer den andra skärningspunkten av linjen med enhetscirkeln närma sig $(0, -1)$. För lutningen $c = 0$ blir linjen en tangent till cirkeln och vi kan tänka på punkten $(0, -1)$ som en dubbel skärningspunkt.

Anmärkning 6.2.10. För de läsare som vet vad ett gränsvärde är, noterar vi att

$$\lim_{c \to \pm\infty} \left(\frac{2c}{c^2+1}, \frac{c^2-1}{c^2+1}\right) = (0, 1)$$



Figur 6.3: Enhetscirkeln med tangentlinje i punkten $(0, -1)$

Om linjen genom $(0, -1)$ får obegränsat stor lutning, så kommer den andra skärningspunkten att röra sig obegränsat nära punkten $(0, 1)$.

Vi skulle kunna säga att lutningen ∞ också är en rationell lutning. Då blir den vertikala linjen genom $(0, -1)$ tillåten som har punkten $(0, 1)$ som andra skärningspunkt med cirkeln. Vi skulle alltså inte behöva specialbehandla punkten $(0, 1)$ i föregående sats.

Detta stämmer i princip, men man måste vara noggrann med hur man får räkna med ∞ i sådana här sammanhang och vi har därför undvikit det.

6.3 Primitiva pythagoreiska tripplar igen

När vi nu har en enkel beskrivning av alla rationella punkterna på enhetscirkeln vill vi gärna återfå vår formel för primitiva pythagoreiska tripplar. En liten skillnad är att vi här inte längre kommer att anta att X är udda.

Beviset till nästa sats är ganska tungt, så vi uppmanar läsaren att försöka få ett överblick över vad som bevisas egentligen först och sedan sätta sig in i alla detaljer.

Sats 6.3.1. *Låt s och t vara relativt prima heltal. Om s och t är udda, så kommer*

$$\left(st, \frac{s^2 - t^2}{2}, \frac{s^2 + t^2}{2}\right) \tag{6.5}$$

att vara en primitiv pythagoreisk trippel. Annars är

$$(2st, s^2 - t^2, s^2 + t^2) \tag{6.6}$$

en primitiv pythagoreisk trippel. Varje primitiv pythagoreisk trippel kan skrivas unikt på en av de två formerna med $t > 0$ eller $t = 0$ och $s > 0$.

Bevis. Låt $c = \frac{s}{t}$ med s och t relativt prima och $t > 0$. Då svarar c mot den rationella punkten

$$\left(\frac{2c}{c^2 + 1}, \frac{c^2 - 1}{c^2 + 1}\right) = \left(\frac{2st}{s^2 + t^2}, \frac{s^2 - t^2}{s^2 + t^2}\right)$$

på enhetscirkeln, där vi i högerledet har förlängt med t^2. Som i Sats 6.1.1, finner vi den pythagoreiska trippeln

$$\left(2st, s^2 - t^2, s^2 + t^2\right)$$

Men eftersom bråken

$$\frac{2st}{s^2 + t^2} \qquad \text{och} \qquad \frac{s^2 - t^2}{s^2 + t^2}$$

inte behöver vara skrivna på reducerad form, så behöver inte trippeln vara primitiv. Vi tar en paus i beviset för att undersöka när trippeln är primitiv och vilka gemensamma delare vi kan få om den inte är det.

Hjälpsats 6.3.2. *Låt s och t vara relativt prima heltal. Då är den största gemensamma delaren av trippeln $(2st, s^2 - t^2, s^2 + t^2)$ antingen 1 eller 2. Det senare fallet inträffar om och endast om både s och t är udda.*

Bevis av hjälpsats 6.3.2. Antag att något primtal p delar alla tre talen i trippeln

$$(2st, s^2 - t^2, s^2 + t^2).$$

Då måste p också dela

$$(s^2 - t^2) + (s^2 + t^2) = 2s^2$$

och

$$(s^2 + t^2) - (s^2 - t^2) = 2t^2.$$

Eftersom s och t är relativt prima, så är största gemensamma delaren till $2s^2$ och $2t^2$ lika med 2. Alltså måste $p = 2$ gälla. Då är $s^2 - t^2$ och $s^2 + t^2$ jämna tal. Det kan inte gälla att både s och t är jämna eftersom de är relativt prima. Alltså är s och t udda.

Detta visar att om s eller t inte är udda, så är den största gemensamma delaren lika med 1.

Omvänt ser vi att om s och t är udda så är alla tre talen

$$(2st, s^2 - t^2, s^2 + t^2)$$

jämna och den största gemensamma delaren blir 2. □

Fortsättning av beviset av sats 6.3.1. Enligt hjälpsats 6.3.2 så är

$$\left(2st, s^2 - t^2, s^2 + t^2\right)$$

en primitiv pythagoreisk trippel om inte båda s och t är udda. I fall att både s och t är udda, delar vi med den största gemensamma delaren 2 och får att

$$\left(st, \frac{s^2 - t^2}{2}, \frac{s^2 + t^2}{2}\right)$$

är en primitiv pythagoreisk trippel.

Vi har därmed visat att när vi väljer $c \in \mathbb{Q}$ godtyckligt får vi en primitiv pythagoreisk trippel enligt formel (6.5) eller formel (6.6), och vi har sett tidigare att alla primitiva pythagoreiska tripplar utom $(0, 1, 1)$ (som svarar mot punkten $(0, 1)$ på enhetscirkeln) uppstår på detta sätt.

Å andra sidan har vi kvar fallet $t = 0$, i vilket fall s/t inte är ett rationellt tal och vilket motsvarar den vertikala linjen i vår tidigare formulering. Man kontrollerar att $s = 1$, $t = 0$ ger precis den återstående trippeln $(0, 1, 1)$. □

Exempel 6.3.3. Låt oss beräkna några primitiva pythagoreiska tripplar enligt formlerna (6.5) och (6.6).

(i) Låt $s = 2$ och $t = 3$. Lutning av linjen genom $(0, -1)$ är $c = 2/3$ och vi får den rationella punkten $\left(\frac{12}{13}, -\frac{5}{13}\right)$ på enhetscirkeln vilken motsvarat den primitiva pythagoreiska trippeln $(12, -5, 13)$ i enlighet med formel (6.5).

(ii) Låt $s = -2$ och $t = 3$. På samma sätt får vi $c = -2/3$, punkten $\left(-\frac{12}{13}, -\frac{5}{13}\right)$ på enhetscirkeln och den primitiva pythagoreiska trippeln $(-12, -5, 13)$.

(iii) Låt $s = 1$ och $t = 3$. Vi beräknar $c = 1/3$, punkten $\left(\frac{6}{10}, -\frac{8}{10}\right) = \left(\frac{3}{5}, -\frac{4}{5}\right)$ och får trippeln $(3, -4, 5)$ vilket överensstämmer med formel (6.6).

(iv) För $s = 3$ och $t = 1$ får vi $c = 3$, punkten $\left(\frac{3}{5}, \frac{4}{5}\right)$ och vår välkända trippel $(3, 4, 5)$. ▲

Anmärkning 6.3.4. Låt oss jämföra denna formel med den vi härledde i föregående kapitel. Vi antog i förra kapitlet att både s och t var udda, och vi fick en formel som parametriserade tripplar (X, Y, Z) med X udda. Denna formel är exakt samma som den övre av de två formler vi hittade i detta kapitel. I detta kapitel antog vi inte att X var udda, så tydligen parametriserar den undre formeln (när någon av s och t är jämnt) de primitiva tripplar för vilka X är jämnt och Y är udda. Man kan alltså se det som att vi har upptäckt två olika formler, en som ger alla tripplar med Y jämnt och en som ger alla tripplar med X jämnt.

Övningar

Övning 6.1 ($\star$)**.** Antag att (x, y) är en rationell lösning till ekvation (6.1). Skriv $x = a/c$ och $y = b/d$ med heltal a, b, c och d där $\text{sgd}(a, c) = 1$, $\text{sgd}(b, d) = 1$ och där c och d är positiva.

(i) Visa ekvationen $a^2d^2 + b^2c^2 = c^2d^2$.

(ii) Visa att c^2 delar a^2d^2 och att c delar d.

(iii) Visa att $d = c$.

Övning 6.2 ($\star$)**.** Bestäm skärningspunkterna av linjen $y = -\frac{2}{3}x - 1$ med enhetscirkeln.

Övning 6.3 ($\star$)**.** Bestäm ekvationen för linjen som går genom punkterna $(0, -1)$ och $(\frac{5}{13}, -\frac{12}{13})$.

Övning 6.4 ($\star$)**.** Visa att kvoten

$$\frac{y_0 - y_1}{x_0 - x_1}$$

som förekom i Hjälpsats 6.2.3 är lika med lutningen av linjen. Speciellt beror kvoten inte på valet av distinkta punkter (x_0, y_0) och (x_1, y_1) på linjen.

Övning 6.5 ($\star\star$)**.** I Övning 5.5 visades att i en pythagoreisk trippel med X udda så måste Y vara delbart med 4. Ge ett alternativt bevis av detta påstående som använder formeln 6.6 för primitiva pythagoreiska tripplar som härleddes i detta kapitel.

Övning 6.6 ($\star\star$)**.** Tag en primitiv trippel (X, Y, Z), och antag X udda. Vi vet att det finns två udda tal s och t sådana att

$$(X, Y, Z) = \left(st, \frac{s^2 - t^2}{2}, \frac{s^2 + t^2}{2} \right).$$

Å andra sidan kan vi byta X och Y mot varandra, och nu vet vi också att det finns två tal u och v, där det ena är jämnt och det andra är udda, sådana att

$$(Y, X, Z) = (2uv, u^2 - v^2, u^2 + v^2),$$

eftersom detta är allmänna formeln för pythagoreiska tripplar där det *första* talet är jämnt.

Vad är relationen mellan (u, v) och (s, t)?

Övning 6.7 ($\star\star$)**.** I Sats 6.1.1 ges två konstruktioner. Den ena avbildar en primitiv pythagoreisk trippel (X, Y, Z) på en punkt (x, y) på enhetscirkeln genom formeln

$$x = \frac{X}{Z} \qquad \text{och} \qquad y = \frac{Y}{Z}.$$

och den andra avbildar en punkt (x, y) på enhetscirkeln med

$$x = \frac{a}{c} \qquad \text{och} \qquad y = \frac{b}{d}$$

på den primitiva pythagoreiska trippeln (a, b, c).

Visa att dessa två konstruktioner är varandras inverser. Med andra ord: visa att om man utför båda dessa konstruktioner efter varandra, så kommer man tillbaka till vad man började med.

Övning 6.8 ($\star\star\star$)**.** Punkten $(1, 1)$ ligger på cirkeln $X^2 + Y^2 = 2$.

(i) Bestäm ekvationerna till alla linjer genom punkten $(1, 1)$ som har rationell eller oändlig lutning.

Ledning: En linje ges av en ekvation på formen $Y = aX + b$ eller $X = c$. Bestäm villkor på parametrarna a, b och c för att linjen ska gå igenom punkten $(1, 1)$.

(ii) Beräkna i vilka punkter linjerna från (i) skär cirkeln $X^2 + Y^2 = 2$.

(iii) Bestäm en allmän formel för alla rationella punkter på denna cirkel.

(iv) Vi säger att en heltalslösning till den diofantiska ekvationen $X^2 + Y^2 = 2Z^2$ är *primitiv* om $\text{sgd}(X, Y, Z) = 1$ och $Z > 0$. Bestäm alla primitiva lösningar till ekvationen.

Övning 6.9 ($\star\star\star$)**.** (i) Finns det två olika primitiva pythagoreiska tripplar (X, Y, Z) som har samma värde för Z? För att undvika triviala exempel kräver vi att $X > 0$, $Y > 0$ och att X är jämnt.

Ledning: Använd parameterframställningen för primitiva pythagoreiska tripplar från Sats 6.3.1. Du behöver hitta två olika par (s, t) av relativt prima heltal som har samma värde på $s^2 + t^2$. Det finns sådana med $1 \leqslant s, t, s', t' \leqslant 9$. Hitta dem genom att undersöka alla möjliga summor $s^2 + t^2$ och se vilka som förekommer flera gånger.

(ii) Finns det tre olika primitiva pythagoreiska tripplar med denna egenskap?

Ledning: Du kan hitta tre tripplar med samma värde på Z genom att skriva 1105 som summa av två relativt prima kvadrater på tre olika sätt och använda dessa värden för s och t i formeln från Sats 6.3.1. Talet 1105 är det minsta talet som har denna egenskap. Man kan med fördel använda sig av en *dator*.

Detta är en svår uppgift. I allmänhet gäller att man för varje naturligt tal m kan hitta m olika primitiva tripplar med samma värde för Z, men beviset för detta går långt utanför denna uppgift.

7 Kägelsnitt och rationella punkter

I de föregående två kapitlen har vi diskuterat lösningar till den diofantiska ekvationen

$$X^2 + Y^2 = Z^2.$$

I detta kapitel kommer vi att vidga perspektivet och studera en allmän diofantisk ekvation av grad två i tre variabler:

$$aX^2 + bXY + cY^2 + dXZ + eYZ + fZ^2 = 0, \tag{7.1}$$

där a, b, c, d, e, f är heltal.

Innan vi börjar måste vi dock varna (7.1) för att detta kapitel innehåller färre bevis än tidigare. På flera ställen kommer vi att använda oss av påståenden som vi av platsbrist inte kan ge ett rigoröst bevis för. Vi hoppas att läsaren ändå kan finna nytta i denna framställning.

Vi kommer att se att mycket av det vi gjorde i tidigare kapitel kommer att fungera även i detta kapitel:

- Precis som tidigare kan man få nya lösningar från gamla genom *skalning*: om (X, Y, Z) är en lösning och k ett heltal, så är även (kX, kY, kZ) en lösning. Vi kan på ett liknande sätt tala om *primitiva* och *icke-primitiva* lösningar, och det räcker att förstå sig på de primitiva lösningarna.

- Om vi råkar ha en lösning (X, Y, Z) med $Z \neq 0$, så kan vi titta på den rationella punkten $(x, y) = \left(\frac{X}{Z}, \frac{Y}{Z}\right)$ i planet: denna kommer att vara en lösning till ekvationen

  $$ax^2 + bxy + cy^2 + dx + ey + f = 0.$$

  Vi får som tidigare ett samband mellan rationella lösningar till en ekvation i två variabler, och heltalslösningar till vår ursprungliga ekvation upp till skalning.

- För de allra flesta värden på a, b, c, d, e, f så kommer samma trick som vi introducerade i föregående kapitel att fungera: givet någon rationell punkt p på kurvan

  $$ax^2 + bxy + cy^2 + dx + ey + f = 0,$$

  så kommer nästan varje linje genom p med rationell lutning att skära kurvan i en ny punkt som också blir rationell.

Dock får vi också några extra svårigheter i detta mer allmänna fall. Redan i översiktsskissen som vi beskrev ovan syns detta. För att kunna dela med Z måste vi begränsa oss till lösningar med $Z \neq 0$, vilket inte behövdes tidigare. (Varför behövde vi inte anta att $Z \neq 0$ när vi tittade på nollskilda lösningar till $X^2 + Y^2 = Z^2$?)

Dessutom skriver vi att för "de flesta" värden på a, b, c, d, e, f så kommer tangentmetoden att fungera, men det kommer inte att vara sant för alla möjliga värden, vilket ger en extra komplikation. För att metoden skall fungera måste vi anta att kurvan är vad som kallas ett icke-degenererat kägelsnitt.

Slutligen skriver vi att nästan varje linje genom p med rationell lutning skär kurvan i en ny punkt. Detta märkte vi visserligen redan i föregående kapitel, eftersom tangentlinjen inte kommer att skära kurvan i en ny punkt, men för ett allmänt kägelsnitt kan det finnas fler linjer än bara tangenten som inte möter kurvan i några fler punkter.

7.1 Kägelsnitt

När vi studerade ekvationen $X^2 + Y^2 = Z^2$, fann vi det bekvämt att i stället titta på ekvationen $x^2 + y^2 = 1$ i planet, som ju beskriver en cirkel. I detta kapitel kommer vi också att titta på andragradsekvationer i planet. Det kommer att vara bekvämt att använda lite klassisk terminologi rörande just andragradskurvor i planet, som vi nu kommer att sammanfatta. Vi ger dock inga bevis i detta delavsnitt.

Definition 7.1.1. Ett *kägelsnitt* i planet är en delmängd av $\mathbb{R}^2$ på formen

$$\{(x, y) \ : \ ax^2 + bxy + cy^2 + dx + ey + f = 0\},$$

där a, b, c, d, e, f är reella tal, varav minst ett är nollskilt.

Det finns en klassificering av kägelsnitt, som ser ut som följer. Geometriskt kan ett kägelsnitt se ut på följande sätt. (Med detta menar vi att varje kägelsnitt kan fås att ha någon av dessa former efter att ha roterats och flyttats i sidled.)

(i) En *ellips*. Ett standardexempel är lösningarna till ekvationen

$$\frac{x^2}{s^2} + \frac{y^2}{t^2} = 1$$

där talen s och t är konstanter. Ellipsens centrum har placerats i origo, och s och t ger längderna på ellipsens axlar. Om $s = t$ är axlarna lika långa, och vi har en cirkel med radie s.

(ii) En *parabel*, kanske mer känt helt enkelt som en andragradskurva. Typexemplet ser ut som

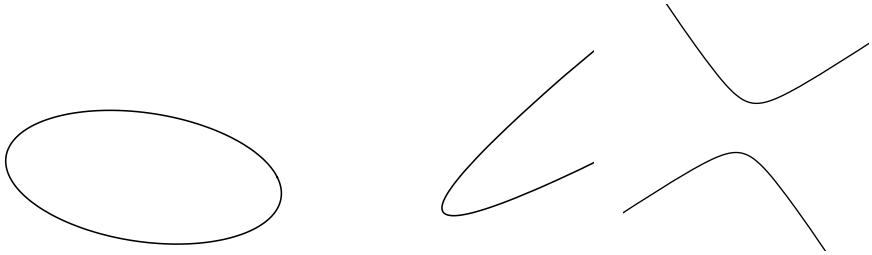
$$y = x^2 + sx + t,$$

där s och t igen är konstanter.

(iii) En *hyperbel*. Ekvationen för en hyperbel liknar mycket ekvationen för en ellips:

$$\frac{x^2}{s^2} - \frac{y^2}{t^2} = 1.$$

Dock är dess utseende ganska annorlunda, som ses i Figur 7.1 nedan.



(a) (b) (c)

Figur 7.1: Exempel på en (a) ellips, (b) parabel och (c) hyperbel i planet.

Ett sätt att beskriva skillnaden mellan dessa tre sorterna av kägelsnitt är följande: en ellips är sammanhängande och begränsad (det vill säga får plats inuti en ändligt stor rektangel); en parabel är sammanhängande och obegränsad; en hyperbel har två sammanhängande komponenter och är obegränsad.

Det finns även några mindre intressanta möjligheter för ett kägelsnitt:

(iv) Man kan få två linjer som möts i en punkt, eller två parallella linjer. Exempel på ekvationer för dessa är

$$xy = 0 \qquad \text{respektive} \qquad x(x-1) = 0.$$

(v) Man kan få bara en linje. Ett exempel på detta är ekvationen

$$x^2 = 0.$$

(vi) Man kan få bara en punkt. Till exempel beskriver ekvationen

$$x^2 + y^2 = 0$$

en cirkel med radie 0, d.v.s. en punkt.

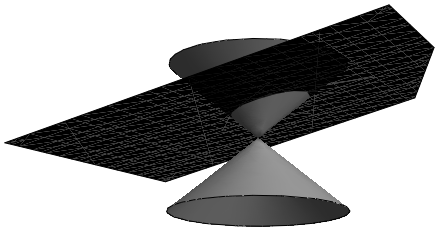
(vii) Man kan få den tomma mängden, det vill säga, en ekvation helt utan lösningar. Ett exempel är ekvationen

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

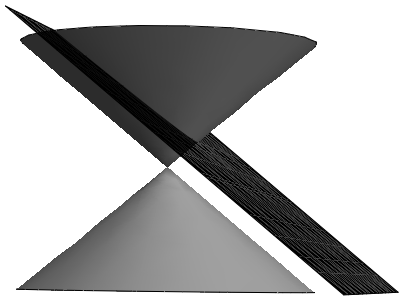
(Varför saknar denna lösningar?)

Anmärkning 7.1.2. Namnet kägelsnitt kommer av att om man tar en ”kägla” (det vill säga en kon) och ”snittar” med ett plan, så kommer resultatet alltid att bli ett kägelsnitt i planet. Se Figur 7.2–7.4.

Definition 7.1.3. Ett kägelsnitt som faller under kategorierna (i)-(iii) ovan kallas *icke-degenererat*. Ett kägelsnitt som faller under kategorierna (iv)-(vii) kallas *degenererat*.



Figur 7.2: Ellipsen som en skärning mellan en kägla och ett plan.



Figur 7.3: Parabeln som en skärning mellan en kägla och ett plan.

Anmärkning 7.1.4. Figurerna 7.2–7.4 visar exempel på snitt mellan käglor och plan. Om planet inte passerar genom spetsen på käglan, så kommer kägelsnittet alltid att vara icke-degenererat, men om planet passerar genom spetsen blir snittet alltid degenererat.

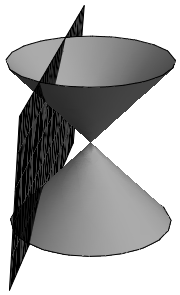
Anmärkning 7.1.5. Som sagt har vi inte gett några bevis för någon del av denna klassificering av kägelsnitt: vi hoppas dock att läsaren kan finna resultatet någorlunda intuitivt trovärdigt, och att detta avsnitt kan ge en viss geometrisk intuition för hur andragradskurvor i planet ser ut.

7.2 Diofantiska andragradsekvationer i tre variabler

Låt a, b, c, d, e och f vara heltal, inte alla noll. Vi kommer att vara intresserade av lösningar till den diofantiska ekvationen

$$aX^2 + bXY + cY^2 + dXZ + eYZ + fZ^2 = 0. \tag{7.2}$$

Precis som tidigare gäller det att om (X, Y, Z) är en lösning, så är även (kX, kY, kZ) det för alla värden på k.



Figur 7.4: Hyperbeln som en skärning mellan en kägla och ett plan.

Vi kommer också att titta på det motsvarande kägelsnittet

$$ax^2 + bxy + cy^2 + dx + ey + f = 0. \tag{7.3}$$

Vi säger att detta kägelsnitt är *associerat* till den diofantiska ekvationen.

Detta leder till att tala om primitiva och icke-primitiva lösningar:

Definition 7.2.1. En trippel (X, Y, Z) av heltal, inte alla noll, är *primitiv* om $\text{sgd}(X, Y, Z) = 1$ och ett av följande villkor är uppfyllt:

(i) $Z > 0$;

(ii) $Z = 0$ och $Y > 0$;

(iii) $Z = Y = 0$ och $X > 0$.

Annars kallas lösningen *icke-primitiv*.

Hjälpsats 7.2.2. *Varje trippel av heltal (X, Y, Z), där inte alla tre talen är noll, kan skrivas unikt som (kX_0, kY_0, kZ_0) där $k \in \mathbb{Z}$ och (X_0, Y_0, Z_0) är primitiv.*

Beviset liknar beviset av Hjälpsats 5.1.9 och lämnas åt läsaren.

Sats 7.2.3. *Det finns en funktion F som avbildar heltalslösningar (X, Y, Z) till ekvation (7.2) med $Z \neq 0$ till rationella punkter på det associerade kägelsnittet (7.3). Funktionen avbildar en trippel (X, Y, Z) på den rationella lösningen*

$$F(X, Y, Z) = (X/Z, Y/Z).$$

Funktionen F ger en bijektion mellan mängden av primitiva lösningar (X, Y, Z) till ekvation (7.2) med $Z > 0$ och mängden av rationella punkter (x, y) på det associerade kägelsnittet (7.3).

Bevis. Om (X, Y, Z) är en lösning till ekvation (7.2), så kontrollerar man enkelt att $(x, y) = (X/Z, Y/Z)$ löser ekvationen (7.3).

Antag att

$$x = \frac{a}{b} \qquad y = \frac{c}{d}$$

är en rationell punkt på kägelsnittet. Då kommer

$$(ad, bc, bd)$$

att vara en heltalslösning till (7.2), och

$$F(ad, bc, bd) = \left(\frac{a}{b}, \frac{c}{d}\right). \qquad \square$$

Dock behöver denna lösning inte nödvändigtvis vara primitiv! (Till skillnad från föregående kapitel, behöver det inte längre vara sant att $b = d$.) Dock kan trippeln (ad, bc, bd) skrivas unikt på formen (kX_0, kY_0, kZ_0) där $k \in \mathbb{Z}$ och (X_0, Y_0, Z_0) är primitiv enligt Hjälpsats 7.2.2. Vi har att $F(X_0, Y_0, Z_0) = (x, y)$ och att ingen annan primitiv trippel avbildas på (x, y) under F, så F är en bijektion.

Anmärkning 7.2.4. Om kägelsnittet är degenererat, är det inte så svårt att hitta de rationella lösningarna. Ty i detta fall är antingen kägelsnittet tomt (i vilket fall det inte finns några lösningar), eller så består det av en punkt (i vilket fall det finns exakt en lösning), eller så består det av en eller två linjer. Att hitta rationella punkter på en linje blir en övning i slutet av detta kapitel. Vi kan alltså fokusera våra ansträngningar på att hitta de rationella punkterna på ett icke-degenerat kägelsnitt: en ellips, parabel eller hyperbel.

Vi sammanfattar detta delavsnitt på följande sätt: för varje rationell punkt på kägelsnittet får vi oändligt många heltalslösningar som alla är multiplar av samma primitiva lösning, och varje heltalslösning med $z \neq 0$ fås på detta sätt.

7.3 Linjer och skärningar

I föregående kapitel såg vi att om man tar en rationell punkt på cirkeln och drar en linje genom den med rationell eller oändlig lutning, så finner man att skärningen med cirkeln ger en ny rationell punkt. Vi kommer nu att studera motsvarande konstruktion för ett allmänt icke-degenererat kägelsnitt. Följande hjälpsats kommer att spela en central roll:

Hjälpsats 7.3.1. *Låt*

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

vara en andragradsekvation med rationella koefficienter, d.v.s. $p, q \in \mathbb{Q}$. Antag att ekvationen har två lösningar eller en dubbelrot, och att den ena lösningen är ett rationellt tal. Då är även den andra lösningen det.

Bevis. Vi förutsätter känt att om ekvationen har två lösningar eller en dubbelrot, så kan den *faktoriseras* som

$$x^2 + px + q = (x - x_1)(x - x_2),$$

där x_1 och x_2 är de två lösningarna. Multiplicerar vi ut högerledet av denna ekvation finner vi att

$$x^2 + px + q = x^2 - (x_1 + x_2)x + x_1x_2,$$

så speciellt måste vi ha att $p = -x_1 - x_2$ och $q = x_1x_2$. Antag nu att x_1 är ett rationellt tal. I så fall ser vi ur ekvationen

$$p = -x_1 - x_2$$

och antagandet att p är ett rationellt tal att även x_2 är rationellt. □

Hjälpsats 7.3.2. *Låt L vara en linje i planet. Om linjen inte är vertikal, så kan man lösa ut y som en funktion av x ur linjens ekvation, och om linjen inte är horisontell, så kan man lösa ut x som en funktion av y. Om ekvationen för L har rationella koefficienter, så kommer ekvationerna för x som funktion av y och vice versa bägge ha rationella koefficienter.*

Bevis. Lämnas åt läsaren. □

Vi kan nu dra följande slutsats:

Sats 7.3.3. *Antag att vi har ett icke-degenerat kägelsnitt i planet med rationella koefficienter, och låt P vara en rationell punkt som ligger på kägelsnittet. Om en linje genom P med rationell eller oändlig lutning möter kägelsnittet i en punkt Q, är även Q en rationell punkt.*

Bevis. Antag först att linjen är varken vertikal eller horisontell. Då har vi ekvationen

$$y - y_0 = c(x - x_0),$$

där (x_0, y_0) är en godtycklig punkt på linjen och c är linjens lutning. Vi kan både skriva y som en funktion av x med rationella koefficienter och skriva x som en funktion av y med rationella koefficienter enligt föregående hjälpsats.

Om vi nu substituerar $y = y_0 + c(x - x_0)$ i ekvationen

$$ax^2 + bxy + cy^2 + dx + ey + f = 0,$$

så fås en andragradsekvation som endast beror på x. Denna kommer alltid att ha minst en rot, som svarar mot x-koordinaten till punkten P. Vi har antagit att linjen möter kägelsnittet i en annan punkt Q, och denna punkt har en annan x-koordinat eftersom linjen inte var vertikal. Alltså finns det minst två rötter.

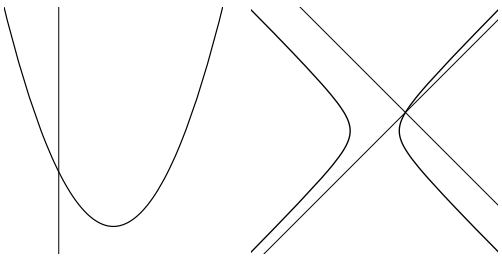
Men det kan inte heller finnas fler än två rötter: den enda andragradsekvation som har fler än två rötter är nollpolynomet, vilken har oändligt många rötter. Detta skulle svara mot att hela linjen ligger på kägelsnittet. Men en rät linje kan endast ligga på ett kägelsnitt om kägelsnittet är degenerat – och vi har antagit att det är icke-degenererat.

Alltså har vi en andragradsekvation med rationella koefficienter, med två skilda rötter, varav den ena roten är rationell. Det följer att även den andra roten är rationell, enligt Hjälpsats 7.3.1.

På samma sätt visas att den andra roten har rationell y-koordinat. Det följer att Q är en rationell punkt, om linjen varken är horisontell eller vertikal.

Antag nu slutligen att linjen är horisontell. I detta fall kan inte x-koordinaten lösas ut som en funktion av y, så vi kan inte med metoden ovan visa att y-koordinaten till Q är rationell. Dock gäller i stället att P och Q har *samma* y-koordinat! (Det är ju det det betyder att linjen är horisontell.) Alltså följer igen att Q måste vara en rationell punkt, men av en annan, enklare, anledning. På samma sätt fungerar det om linjen är vertikal. □

Anmärkning 7.3.4. I föregående kapitel studerade vi en cirkel i planet, och då såg vi att om man drar en linje genom en punkt P på cirkeln så finns endast två fall som kan uppstå: antingen är linjen en tangent, eller så möter den cirkeln i en unik andra punkt. Detta gäller också för allmänna ellipser. Dock är detta påstående falskt för parabler och hyperbler, som följande figurer visar!



Figur 7.5: Linjer som möter (a) parabeln och (b) hyperbeln i exakt en punkt

Geometriskt ser vi att på en parabel finns det alltid exakt en linje som passerar genom en punkt och som varken är en tangent eller möter parabeln i en andra punkt. På en hyperbel finns det exakt två sådana linjer genom varje punkt.

En linje som endast möter kägelsnittet i en punkt (som inte är en tangentlinje) svarar algebraiskt mot att följande händer: när man löser ut y som en funktion av x längs linjen och sätter in i ekvationen för kägelsnittet, så råkar högstagradstermerna ta ut varandra och man är endast kvar med ett förstagradspolynom. Det är alltså detta som inte kan hända för till exempel cirkeln $x^2 + y^2 = 1$. Kontrollera detta!

7.4 Ett exempel

Låt oss i detta delavsnitt studera den diofantiska ekvationen

$$X^2 - 5Y^2 + YZ = 7Z^2. \tag{7.4}$$

Exemplet är ganska långt och kan behöva läsas flera gånger. Metoderna i detta kapitel låter oss hitta lösningar med $Z \neq 0$, så låt oss enbart betrakta dessa. (En övning i slutet av kapitlet är att visa att lösningar saknas om $Z = 0$.) Enligt Sats 7.2.3 är detta samma sak som att hitta rationella punkter på kurvan

$$x^2 - 5y^2 + y = 7. \tag{7.5}$$

För att kunna starta vår metod att rita linjer och skärningar, behöver vi en rationell lösning som startpunkt att dra linjer igenom. Det finns tyvärr ingen allmän metod för att göra detta – ibland kan det vara så att det inte finns någon rationell punkt – men i detta fall har vi kanske tur nog att upptäcka lösningen $x = 5$, $y = 2$, till exempel genom att prova olika små heltalsvärden på y.

Exempel 7.4.1. Vi har nu kägelsnittet $x^2 - 5y^2 + y = 7$ och den rationella punkten $(5, 2)$. Låt oss nu använda denna rationella punkt för att hitta en formel för alla andra rationella punkter. Vi vill dra linjer genom punkten. En linje med lutning c genom punkten $(5, 2)$ kan skrivas som

$$y - 2 = c(x - 5),$$

eller

$$y = c(x - 5) + 2.$$

Vi substituerar detta i ekvationen, och finner att

$$x^2 - 5(c(x - 5) + 2)^2 + c(x - 5) + 2 = 7,$$

det vill säga ett andragradspolynom i x. För de flesta värden på c kommer denna ekvation att ha två lösningar, varav den ena är $x = 5$ och svarar mot punkten vi drog linjen genom. Om vi multiplicerar ut resultatet lite finner vi att

$$x^2 - 25 - 5c^2(x - 5)^2 - 19c(x - 5) = 0,$$

vilket vi skriver om till

$$(x - 5)(x + 5 - 5c^2(x - 5) - 19c) = 0.$$

Faktorn $x - 5$ svarar mot den rationella punkten $(5, 2)$ som vi redan känner till och vi är intresserade av den andra skärningspunkten. Alltså sätter vi

$$x + 5 - 5c^2(x - 5) - 19c = 0,$$

vilket kan skrivas om till

$$x(1 - 5c^2) = -25c^2 + 19c - 5.$$

Om

$$1 - 5c^2 = 0,$$

det vill säga $c = \pm\sqrt{1/5}$, så ser vi att högerledet ovan är nollskilt och vänsterledet är noll. Alltså finns inga lösningar för dessa värden på c, vilket svarar mot att linjen inte möter kurvan i någon mer punkt. Detta beror på att kurvan är en hyperbel, och vi har sett tidigare att för en hyperbel kommer det alltid att finnas två linjer som varken är tangenter eller skär kurvan i ytterligare en punkt. Dock är detta inte ett problem för oss! Vi är nämligen bara intresserade av de värden som antas när c väljs till ett *rationellt* tal, men $\pm\sqrt{1/5}$ är irrationellt (se Övning 7.2).

För alla andra värden på c finner vi punkten

$$x = \frac{-25c^2 + 19c - 5}{1 - 5c^2},$$

på kägelsnittet, och de rationella punkterna är precis de som fås när vi väljer talet c rationellt. Insättning av detta i $y - 2 = c(x - 5)$ger att

$$y - 2 = c \cdot \frac{-25c^2 + 19c - 5 - 5(1 - 5c^2)}{1 - 5c^2} = c \cdot \frac{19c - 10}{1 - 5c^2} = \frac{19c^2 - 10c}{1 - 5c^2},$$

så att

$$y = \frac{9c^2 - 10c + 2}{1 - 5c^2}.$$

Vi finner alltså att alla rationella punkter ges av formeln

$$(x, y) = \left(\frac{-25c^2 + 19c - 5}{1 - 5c^2}, \frac{9c^2 - 10c + 2}{1 - 5c^2} \right)$$

där c är ett rationellt tal. ▲

Exempel 7.4.2. Låt oss nu se hur en formel för alla rationella punkter på kägelsnittet (7.5), som den vi hittade i föregående exempel, låter oss hitta lösningar till den ursprungliga diofantiska ekvationen (7.4). Till exempel kan vi välja $c = 1$, vilket ger att

$$x = \frac{11}{4}.$$

och

$$y = -\frac{1}{4},$$

vilket ger oss den primitiva lösningen

$$(X, Y, Z) = (11, -1, 4)$$

till den ursprungliga ekvationen

$$X^2 - 5Y^2 + YZ = 7Z^2.$$

Vi kan även kontrollera att vänsterledet blir

$$121 - 5 - 4 = 112,$$

och högerledet blir

$$7 \cdot 16 = 112.$$

I allmänhet ser vi att formeln

$$(-25c^2 + 19c - 5, 9c^2 - 10c + 2, 1 - 5c^2),$$

som vi får genom att förlänga med $1 - 5c^2$, alltid ger *rationella* lösningar till ekvation (7.4). För att hitta heltalslösningar skriver vi $c = s/t$ och förlänger med t^2: vi finner då formeln

$$(X, Y, Z) = (-25s^2 + 19st - 5t^2, 9s^2 - 10st + 2t^2, t^2 - 5s^2).$$

Denna formel ger alltså en heltalslösning till (7.4) för alla heltalsvärden på s och t. Enligt Hjälpsats 7.2.2 är varje heltalslösning en multipel av en unik primitiv lösning, så vi kan därmed säga att varje heltalsvärde på s och t ger upphov till någon primitiv lösning, även om lösningen vi får genom att sätta in s och t inte skulle vara primitiv.

Enligt Sats 7.2.3 ger alla olika värden på s, t med $\text{sgd}(s, t) = 1$ och $t > 0$ upphov till olika primitiva lösningar, och alla primitiva lösningar med $Z \neq 0$ uppstår på detta sätt. Eftersom det saknas lösningar med $Z = 0$ har vi därmed funnit en implicit beskrivning av alla lösningar till den diofantiska ekvationen (7.4). ▲

Anmärkning 7.4.3. Låt oss påpeka en liten detalj i föregående exempel. Vi fann först att det inte fanns några heltalslösningar till den ursprungliga diofantiska ekvationen med $Z = 0$. Vi fann senare att de två linjerna genom den rationella punkten $(5, 2)$ som varken skar kurvan i en ytterligare punkt eller tangerade kurvan bägge hade *irrationell* lutning.

Detta är ingen slump utan något som alltid kommer att inträffa: det är nämligen så att de linjer med rationell lutning som inte skär kurvan i en ytterligare punkt alltid svarar mot heltalslösningar med $Z = 0$ och som man ”missat” när man begränsat sig till $Z \neq 0$. Se också diskussionen i nästa delavsnitt.

7.5 En utblick mot projektiv geometri

För en djupare studie av det vi diskuterat i detta kapitel är det givande att i stället för vanlig geometri arbeta med så kallad *projektiv geometri*. En grundlig introduktion till projektiv geometri skulle kräva ett helt kompendium, men vi kan inte motstå att ge några korta indikationer av hur denna teori ser ut. Materialet i denna utblick är svårt, men vi hoppas att det kan verka spännande snarare än avskräckande.

I projektiv geometri arbetar man med det *projektiva planet*, i stället för det vanliga planet. I xy-planet har man som bekant två koordinater. I det projektiva planet har man i stället tre koordinater, som brukar skrivas med kolon mellan sig:

$$(X : Y : Z).$$

Dock har vi en speciell räkneregel: om man multiplicerar alla tre koordinaterna med en nollskild konstant, så beskriver de fortfarande samma punkt. Alltså är t.ex.

$$(3 : 4 : 5) \qquad \text{och} \qquad (30 : 40 : 50)$$

samma punkt i det projektiva planet! Dessa två punkter svarar också mot samma primitiva pythagoreiska trippel, och detta är ingen slump. Vi kräver dock att minst en av de tre koordinaterna är nollskild: det finns ingen punkt i projektiva planet som ges av $(0 : 0 : 0)$.

Antag att en punkt

$$(X : Y : Z)$$

har nollskild Z-koordinat. Då kommer värdet på kvoterna

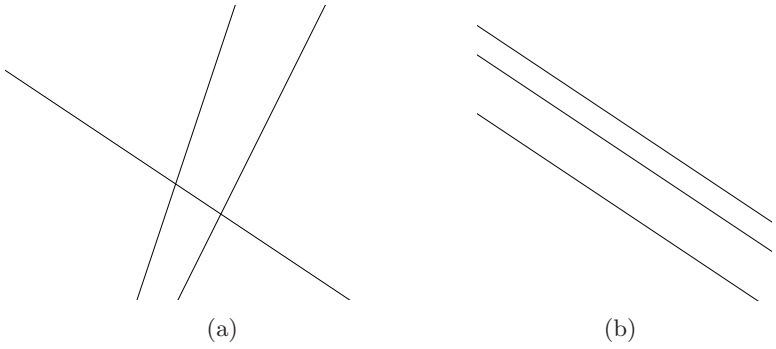
$$\frac{X}{Z} \qquad \text{och} \qquad \frac{Y}{Z}$$

inte att förändras om vi multiplicerar både X, Y och Z med en nollskild konstant. Alltså kan vi tänka på dessa två bråk som vanliga koordinater: det bekanta xy-planet är på ett naturligt sätt en delmängd av det projektiva planet, nämligen delmängden med nollskild Z-koordinat. Med andra ord har vi en funktion från xy-planet till det projektiva planet, som avbildar (x, y) på $(x : y : 1)$. Denna identifierar xy-planet med en delmängd av det projektiva planet.

Att vi fick titta enbart på lösningar med $Z \neq 0$ i detta kapitel beror ur detta perspektiv på att vi inte arbetade projektivt: i det projektiva planet finns *alla* primitiva lösningar representerade med en unik punkt.

Man kan alltså tänka på det projektiva planet som en ”förstoring” av det vanliga planet. Hur ska man då tänka på de kvarvarande punkterna, de med $Z = 0$? Det är ofta naturligt att tänka på dessa som ”punkter i oändligheten”, och att det finns en punkt i oändligheten för varje möjlig ”riktning” man kan ha ut mot oändligheten. I figurerna nedan så kommer linjerna i Figur 7.6(b) att mötas i en unik punkt i oändligheten eftersom de har samma riktning, medan linjerna i Figur 7.6(a) inte möts i oändligheten. (Notera dock att om man går ut mot oändligheten längs en given linje, så kommer man till samma punkt oavsett vilken riktning man går längs linjen. Alltså är det inte riktigt sant att det finns en punkt i oändligheten för varje möjlig riktning, eftersom två rakt motsatta riktningar gör att man hamnar i samma punkt.)

Faktiskt gäller följande: *I det projektiva planet kommer varje par av distinkta linjer att mötas i en unik punkt.* I Figur **??** ser vi att de som inte möts i det vanliga planet, det vill säga är parallella, är precis de som i stället möts i oändligheten.

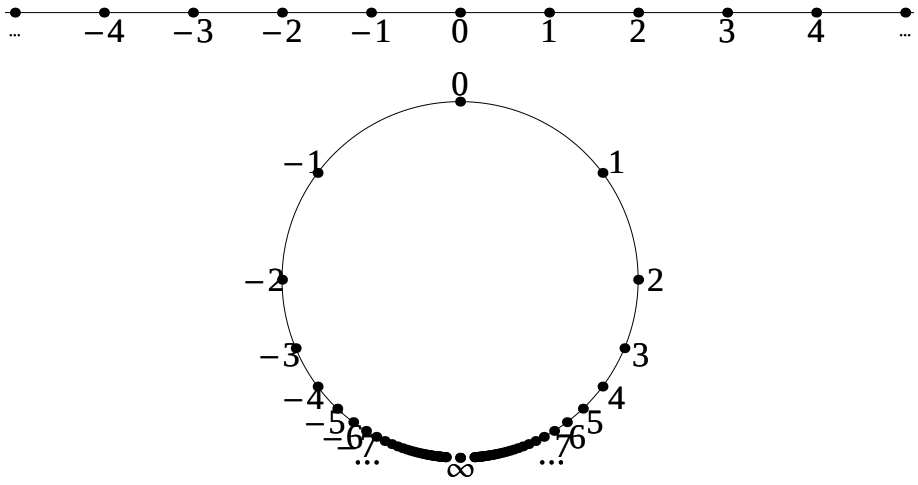


Figur 7.6: Linjer som möts (a) i planet och (b) i oändligheten

Eftersom båda ekvationerna $x = 0$ och $y = 0$ beskriver linjer, så drar vi också slutsatsen att mängden av punkter i oändligheten – som ju ges av $Z = 0$ – också är en linje. Vi kallar denna för linjen i oändligheten. Notera att påståendet i föregående paragraf också gäller för denna linje!

Linjerna i det projektiva planet är dock inte riktigt vanliga linjer. Snarare är de *projektiva linjer*. På samma sätt som det projektiva planet ser ut som ett plan med extra punkter i oändligheten, så kommer en projektiv linje att se ut som en linje med punkter i oändligheten.

Men på en linje finns bara en enda riktning ut mot oändligheten, så en projektiv linje är bara en linje med en extra punkt tillagd som man kan komma till genom att gå oändligt långt i någon av riktningarna. Alltså ser en projektiv linje ut ungefär som en cirkel! Vi ser detta i följande figur:



Figur 7.7: Ovan: Den reella linjen med heltalen markerade. Nedan: den projektiva linjen med heltalen markerade. Den projektiva linjen är ritad som en cirkel för att punkten i oändligheten ska få plats på papperet, men man bör tänka på den som oändligt lång, vilket också antyds av avstånden mellan heltalen.

Detta ger oss ett nytt perspektiv på parabler och hyperbler. En parabel är en kurva som går ut mot oändligheten, möter linjen i oändligheten i en unik punkt, och kommer tillbaka igen. I Figur 7.5(a) möter parabeln linjen i oändligheten i punkten som svarar mot vertikal riktning.

En hyperbel är en kurva som går ut mot oändligheten, möter linjen i oändligheten i en punkt, kommer tillbaks från andra hållet och går ut mot oändligheten en andra gång. Detta åskådliggörs i Figur 7.5(b).

Alltså ser både ellipser, parabler och hyperbler ut ungefär som cirklar ur detta projektiva perspektiv. Den enda skillnaden mellan dem är att en ellips inte möter linjen i oändligheten i någon punkt, en parabel har linjen i oändligheten som tangentlinje, medan en hyperbel möter linjen i oändligheten i två punkter! Klassificeringen av kägelsnitt blir därför mycket mindre komplicerad om man bara arbetar i det projektiva planet.

Konstruktionen med linjer genom en punkt blir också mer lättbegriplig ur ett projektivt perspektiv. Ty vi parametriserade linjerna genom en punkt med hjälp av dess lutning, som kunde vara vilket tal som helst i $\mathbb{R} \cup \{\infty\}$. Men $\mathbb{R} \cup \{\infty\}$ har vi nu sett är en projektiv linje, så det är faktiskt en projektiv linje som parametriserar alla linjer som passerar genom en given punkt.

I det vanliga planet fanns det för hyperbler och parabler vissa linjer som varken tangerade kägelsnittet eller mötte kägelsnittet någon mer punkt. Denna defekt uppstår inte i det projektiva planet: här ser vi i stället att detta är de linjer som möter kägelsnittet i en punkt i oändligheten. I figurerna 7.5(a) och 7.5(b) ser vi också att de linjer som inte möter kurvan i någon mer punkt har samma riktning som kurvan ut mot oändligheten, så att de därför möts i en unik punkt i oändligheten. Detta förklarar varför det alltid finns en sådan linje på en parabel (en parabel har en punkt i oändligheten), medan det finns två sådana linjer för en hyperbel (som har två punkter i oändligheten).

Ett sätt att se på metoden att dra linjer genom en given punkt är att det finns en projektiv linje som parametriserar linjer genom den givna punkten på kägelsnittet, att var och en av dessa linjer ger oss en punkt på kägelsnittet (den andra skärningspunkten), och omvänt: varje punkt på kägelsnittet ger en unik linje genom den givna punkten.

Detta kan användas för att identifiera punkterna på kägelsnittet med punkterna på en projektiv linje, vilket förklarar varför både ellipser, parabler och hyperbler ser ut som cirklar när man tänker på dem projektivt: det är helt enkelt för att en projektiv linje ser ut som en cirkel!

Metoden i detta och föregående kapitel kan nu sammanfattas på följande eleganta sätt: *om man genom att dra linjer på detta sätt identifierar punkterna på kägelsnittet med punkterna på en projektiv linje, och den punkt man startar med har rationella koordinater, så kommer de rationella punkterna på kägelsnittet att identifieras bijektivt med de rationella punkterna på den projektiva linjen.*

Övningar

Övning 7.1 ($\star$)**.** Låt f vara ett moniskt andragradspolynom med heltalskoefficienter. Antag att f har två distinkta reella rötter eller en reell dubbelrot samt att den ena roten är ett heltal. Visa att den andra roten är ett heltal.

Ledning: Ett andragradspolynom har formen $f(x) = ax^2 + bx + c$. Polynomet kallas *moniskt* om $a = 1$ och har *heltalskoefficienter* om a, b och c är heltal. Om f har rötterna x_1 och x_2, så kan vi skriva $f(x) = a(x - x_1)(x - x_2)$. Multiplicera ut och jämför koefficienter!

Övning 7.2 ($\star$)**.** Om man sätter $Z = 0$ i ekvation (7.4), får man den diofantiska ekvationen

$$X^2 = 5Y^2.$$

Visa att denna ekvation saknar heltalslösningar utöver den triviala $X = 0, Y = 0$. *Ledning:* Hur många gånger kan primtalet 5 förekomma i en primtalsfaktorisering av höger- och vänsterledet?

Förklara varför detta är ekvivalent med att $\sqrt{5}$ inte är ett rationellt tal.

Övning 7.3 ($\star$)**.** Enligt Hjälpsats 5.2.1 räcker det att kontrollera att två av talen i en pythagoreisk trippel är relativt prima för att veta att alla tre talen är relativt prima. Gäller detta även för heltalslösningar till en allmän andragradsekvation i tre variabler på formen i ekvation (7.1)?

Övning 7.4 ($\star$)**.** Ett exempel på ett kägelsnitt i planet är en parabel på formen $y = f(x) = ax^2 + bx + c$, där a, b och c är rationella koefficienter. Visa att de rationella punkterna på parabeln kan parametriseras som $(x, f(x))$ för $x \in \mathbb{Q}$.

Övning 7.5 ($\star\star$)**.** (i) Låt $ax + by = c$ vara en linje i planet, där a, b, c är rationella tal och a och b inte bägge är noll. Beskriv hur man parametriserar alla rationella punkter på linjen.

Ledning: Använd x eller y som parameter.

(ii) Beskriv sammanhanget mellan rationella punkter på linjen $ax + by = c$ i planet och primitiva heltalslösningar till den diofantiska ekvationen $aX + bY - cZ = 0$ där a, b och c är heltal.

Ledning: Visa ett påstående motsvarande Sats 7.2.3 för linjära ekvationer. Använd samma avbildning $x = X/Z$ och $y = Y/Z$.

Övning 7.6 ($\star\star$)**.** Bestäm alla lösningar (X, Y, Z) till den diofantiska ekvationen

$$6X + 10Y - 15Z = 0.$$

på följande sätt:

(i) Visa att $5 \mid X$, $3 \mid Y$ och $2 \mid Z$ i varje lösning.

(ii) Sätt $X = 5X'$, $Y = 3Y'$ och $Z = 2Z'$ och sätt in detta i ekvationen. Vilken ny ekvation får du?

(iii) Visa att man får alla lösningar om man sätter $X = 5(b-a)$, $Y = 3a$ och $Z = 2b$ för heltal a och b.

Övning 7.7 ($\star\star$)**.** Ett plan i rummet ges av en ekvation på formen

$$aX + bY + cZ = d.$$

Ge exempel på ekvationer för olika plan i rummet sådana att skärningen med käglan som beskrivs av ekvationen $X^2 + Y^2 = Z^2$ blir

(i) en ellips.

(ii) en parabel.

(iii) en hyperbel.

(iv) två linjer som skär varandra.

(v) en linje.

(vi) en punkt.

Ledning: Titta på Figur 7.2–7.4 och försök att "gissa" ekvationer för plan liknande de i figurerna.

Övning 7.8 ($\star\star$)**.** Vilka kägelsnitt enligt klassificeringen kan man inte få som skärningen av ett plan i rummet och käglan som beskrivs av ekvationen $X^2 + Y^2 - Z^2 = 0$? Ett informellt geometriskt argument räcker.

Övning 7.9 ($\star\star\star$)**.** Visa resultatet av övning 7.4 igen med hjälp av projektiv geometri:

(i) Bestäm den diofantiska ekvationen som är associerad till parabeln $y = f(x) = ax^2 + bx + c$ genom att sätta $y = Y/Z$ och $x = X/Z$ och multiplicera ekvationen med Z^2.

(ii) Visa att punkten $(0 : 1 : 0)$ uppfyller denna ekvation. Punkten $(0 : 1 : 0)$ motsvarar en punkt i oändligheten i planet.

(iii) Vilka diofantiska ekvationer svarar mot projektiva linjer som passerar genom punkten $(0 : 1 : 0)$?

Ledning: En projektiv linje ges av en ekvation på formen $rX + sY + tZ = 0$ där r, s och t är reella tal. Denna linje är en diofantisk ekvation, och ger en linje med rationell lutning, om r, s och t är heltal.

(iv) Bestäm de linjer i planet som är associerade till linjerna från (iii) genom att sätta $x = X/Z$ och $y = Y/Z$. Hur ser de ut och i vilka punkter skär de parabeln?

8 Ett specialfall av Fermats sista sats

Antag att vi vill visa ett påstående P för alla positiva heltal. Låt S vara mängden av alla positiva heltal för vilka påståendet P *inte* stämmer. Antag motsatsen, att S är icke-tom. Enligt minimumprincipen (1.4.3) som vi sett många gånger i detta häfte har då S ett minsta element. För att hitta en motsägelse räcker det därför att visa följande: *För varje $k \in S$ finns något $l \in S$ med $l < k$.* Detta ger en motsägelse om vi låter k vara det minsta elementet i S.

I princip är denna metod med "minimalt motexempel" långt ifrån något nytt: vi har redan använt flera olika varianter av den i detta kompendium. Dock är den speciellt tillämpbar i teorin för diofantiska ekvationer. Pierre de Fermat (1601-1665), en av den moderna talteorins grundare, använde denna metod flitigt under hela sin karriär. I detta sammanhang brukar metoden kallas för *infinite descent*, vilket kanske kan översättas till "oändlig nedgång". Anledningen till namnet är att om man visat att det för varje lösning x_i finns en mindre lösning x_{i+1}, så kan man betrakta en oändligt lång avtagande sekvens

$$x_1 > x_2 > x_3 > \cdots$$

som blir en oändlig avtagande följd av positiva heltal, vilket är omöjligt. Jämför också med Övning 8.6.

I detta kapitel kommer vi att visa att två stycken diofantiska ekvationer saknar lösningar med hjälp av denna metod. Den första är lite enklare och illustrerar förhoppningsvis idén. Den andra är lite mer invecklad och kräver att vi använder en stor mängd resultat som visats tidigare i häftet och i övningar. Till exempel behöver vi använda formeln för pythagoreiska tripplar två gånger! Vad vi får ut i slutändan är fallet $n = 4$ av den så kallade *Fermats sista sats*.

Vi avslutar kompendiet med en historisk utblick om just Fermats sista sats.

8.1 Ett exempel

Vi ska använda reduktion modulo 3 och metoden med minsta motexempel för att visa att den diofantiska ekvationen $a^2 + b^2 = 3(c^2 + d^2)$ inte har några heltalslösningar.

Hjälpsats 8.1.1. *Låt a vara ett heltal. Då är $a^2 \equiv 0 \pmod 3$ eller $a^2 \equiv 1 \pmod 3$.*

Bevis. Lättast är kanske att ställa upp en tabell. Det är lätt att kontrollera att följande gäller:

$a \pmod 3$	$a^2 \pmod 3$
0	0
1	1
2	1

Beviset är klart. (Jämför också med resultatet i Övning 3.1.) □

Hjälpsats 8.1.2. *Om* $a^2 + b^2 \equiv 0 \pmod 3$ *för heltal* a *och* b*, så gäller* $3 \mid a$ *och* $3 \mid b$*.*

Bevis. Med Hjälpsats 8.1.1 och tabellen

$a^2 \pmod 3$	$b^2 \pmod 3$	$a^2 + b^2 \pmod 3$
0	0	0
1	0	1
0	1	1
1	1	2

inser vi att $a^2 + b^2 \equiv 0 \pmod 3$ endast kan gälla om

$$a^2 \equiv b^2 \equiv 0 \pmod 3,$$

Vi har alltså att $3 \mid a^2$, så enligt Övning 4.4 gäller också $3 \mid a$, och på samma sätt att $3 \mid b$. □

Här kommer det egentliga exemplet:

Sats 8.1.3. *Det finns inga heltalslösningar till den diofantiska ekvationen*

$$a^2 + b^2 = 3(c^2 + d^2).$$

Bevis. Antag att (a, b, c, d) är en lösning till ekvationen som minimerar summan $a^2 + b^2$. Om vi reducerar ekvationen modulo 3, finner vi att

$$a^2 + b^2 \equiv 0 \pmod 3.$$

och enligt Hjälpsats 8.1.2 gäller $3 \mid a$ och $3 \mid b$. Vi kan nu skriva $a = 3a'$, $b = 3b'$, och finner att

$$(3a')^2 + (3b')^2 = 3(c^2 + d^2)$$

vilket vid division med 3 ger

$$3(a'^2 + b'^2) = c^2 + d^2.$$

Men detta ger oss en ny lösning: från lösningen

$$(a, b, c, d)$$

har vi hittat lösningen

$$(c, d, a', b') = \left(c, d, \frac{a}{3}, \frac{b}{3}\right).$$

Denna lösning är mindre än den vi antog var minimal, eftersom vi vet att $c^2 + d^2 = \frac{1}{3}(a^2 + b^2) < a^2 + b^2$. □

8.2 Ekvationen $x^4 + y^4 = z^2$

Sats 8.2.1. *Det finns inga heltalslösningar till den diofantiska ekvationen*

$$x^4 + y^4 = z^2. \tag{8.1}$$

med $x, y \neq 0$.

Anmärkning 8.2.2. Om vi tillåter att $x = 0$ eller $y = 0$ finner vi oändligt många lösningar som dock är enkla att beskriva och därmed inte är så intressanta. För $x = 0$, får vi lösningarna $y = c$ och $z = c^2$ för varje heltal c. Likadant finns det en lösning $x = c$, $y = 0$ och $z = c^2$ för varje heltal c.

Bevis. Antag att (x, y, z) är en lösning till ekvation (8.1) som minimerar z^2. En första observation är att vi kan skriva

$$(x^2)^2 + (y^2)^2 = z^2,$$

så att (x^2, y^2, z) är en pythagoreisk trippel.

Fall 1: Trippeln (x^2, y^2, z) är inte primitiv. Antag att det finns ett tal $k > 1$ som delar både x^2, y^2 och z. Då måste det också existera ett primtal p som delar alla tre. Enligt Övning 4.4 kommer då också $p \mid x$ och $p \mid y$, så $p^4 \mid x^4$ och $p^4 \mid y^4$, vilket slutligen ger att

$$p^4 \mid (x^4 + y^4) = z^2.$$

Men då måste $p^2 \mid z$, och om vi definierar

$$x' = \frac{x}{p} \qquad y' = \frac{y}{p} \qquad z' = \frac{z}{p^2}$$

så kommer (x', y', z') att vara en mindre lösning till (8.1), som vi antog var minimal.

Fall 2: Trippeln (x^2, y^2, z) är primitiv. Vi kan anta att x^2 är udda och y^2 jämnt, jämför Hjälpsats 5.2.2. Enligt Sats 6.3.1 finns relativt prima heltal u och v med $v > 0$ sådana att

$$x^2 = u^2 - v^2,$$

$$y^2 = 2uv \tag{8.2}$$

och

$$z = u^2 + v^2.$$

Nu har vi hittat ännu en pythagoreisk trippel, vilket vi ser från ekvationen

$$x^2 + v^2 = u^2.$$

Eftersom u och v är relativt prima, blir även denna trippel primitiv. Och eftersom x är udda och (x, v, u) är en pythagoreisk trippel, måste enligt Hjälpsats 5.2.2 talet v vara jämnt. Vi kan alltså skriva

$$x = s^2 - t^2,$$

$$v = 2st \tag{8.3}$$

och

$$u = s^2 + t^2 \tag{8.4}$$

för relativt prima heltal s och t med $t > 0$. Eftersom y och v är jämna tal, kan vi dividera ekvation (8.2) med 4 vilket ger

$$\left(\frac{y}{2}\right)^2 = u\left(\frac{v}{2}\right).$$

Vi vet att $y/2$, u och $v/2$ är heltal. Dessutom är u och v, och därmed också u och $v/2$, relativt prima med varandra. Från ekvation (8.2) ses att eftersom $v > 0$ och $y^2 > 0$, måste också $u > 0$. Alltså ger Övning 4.7 att u och $v/2$ båda är jämna kvadrater. Låt därför $v/2 = c^2$, där c är ett heltal. Ekvation (8.3) blir

$$c^2 = st,$$

och s och t är relativt prima. Eftersom $t > 0$ och $c^2 > 0$, måste också $s > 0$. Så s och t är jämna kvadrater enligt Övning 4.7. För att sammanfatta har vi funnit att både $\sqrt{s}$, $\sqrt{t}$ och $\sqrt{u}$ är heltal. Men nu säger ekvation (8.4) att

$$(\sqrt{s})^4 + (\sqrt{t})^4 = (\sqrt{u})^2$$

vilket är en mindre lösning än lösningen (x, y, z) som vi antog var minimal, eftersom $\sqrt{u}^2 = u < u^2 + v^2 = z^2$. □

8.3 Fermats sista sats

Låt oss göra några historiska anmärkningar rörande ekvationen vi precis har studerat. Vi har just visat att ekvationen $x^4 + y^4 = z^2$ saknar lösningar, utom de ”triviala” lösningarna där $x = 0$ eller $y = 0$. Speciellt ser vi också att ekvationen

$$x^4 + y^4 = z^4$$

inte kan ha några icke-triviala lösningar, eftersom vi från en sådan lösning direkt skulle hitta en lösning till $x^4 + y^4 = z^2$ genom att kvadrera z. Denna ekvation är ett av de första fallen av vad som kallas *Fermats sista sats*. Satsen lyder som följer:

Sats 8.3.1. *Låt $n \geqslant 3$ vara ett heltal. Då har ekvationen*

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

inga icke-triviala heltalslösningar, det vill säga inga lösningar med både x, y och z nollskilda.

Anmärkning 8.3.2. Som vi tydligt har sett i detta kompendium finns det gott om lösningar då $n = 2$.

Satsen är döpt efter den franske matematikern Pierre de Fermat (c:a 1601-1665), som först formulerade den. Mer precist skrev han ned den som en anteckning i marginalen till sin egen kopia av Diofantos bok *Arithmetika*, en bok om det som idag kallas diofantiska ekvationer. Kapitlet i Arithmetika handlade om just pythagoreiska tripplar, eller med dåtidens terminologi, om hur man kan dela upp en kvadrat i två kvadrater. Fermat skrev:

> *Cubum autem in duos cubos, aut quadratoquadratum in duos quadratoquadratos, et generaliter nullam in infinitum ultra quadratum potestatem in duos ejusdem nominis fas est dividere: cuius rei demonstrationem mirabilem sane detexi. Hanc marginis exiguitas non caperet.*

Eller, på svenska: Att dela en kub i två kuber, eller att dela en bikvadrat[2] i två bikvadrater, eller i allmänhet att dela vilken potens som helst i två potenser av samma grad högre än den andra är omöjligt: jag har ett i sanning underbart bevis för detta påstående. Marginalen är alltför trång för att rymma detsamma.

Notera alltså att Fermat var verksam tiden före matematiker börjat uttrycka sig systematiskt i formler och ekvationer!

Något bevis gav Fermat hur som helst aldrig någon annanstans heller, och det är lätt att föreställa sig att han ursprungligen trodde sig ha ett bevis men hittade ett fel i det. Dock gav Fermat ett bevis för fallet $n = 4$, och det är väsentligen detta bevis som vi har återgett i detta kompendium.

Fermat lämnade efter sig stor samling påståenden som han hävdade sig ha visat, utan något bevis nedskrivet. Efter hans död blev alla påståenden antingen bevisade (eller, i vissa fall, motbevisade) av andra matematiker, vilket förklarar adjektivet ”sista” i namnet på satsen.

Fermats sista sats blev slutligen bevisad 1995 av Andrew Wiles med hjälp av Richard Taylor, med ett bevis som använde några av de allra mest kraftfulla och abstrakta metoder vi känner till i talteorin. Mer specifikt använde beviset modulära former, elliptiska kurvor över $\mathbb{Q}$, och det som kallas modularitet för elliptiska kurvor. (Man behöver inte förstå dessa ord för att begripa historien.)

På slutet av 50-talet formulerade den japanska matematikern Yutaka Taniyama (med viss hjälp av Goro Shimura) en storslagen förmodan, då väldigt oväntad, att alla elliptiska kurvor över $\mathbb{Q}$ är modulära. Denna förmodan baserades på en mycket liten mängd indicier, framför allt att varje exempel på elliptiska kurvor över $\mathbb{Q}$ som de kunde räkna på visade sig vara modulär. Dock hade ingen någon aning hur man skulle visa ett sådant påstående; det var tydligt att helt nya idéer skulle krävas.

I början av 80-talet visade Gerhard Frey att man, givet ett motexempel mot Fermats sista sats, kan konstruera en elliptisk kurva (som idag kallas Frey-kurvan) över $\mathbb{Q}$ som har flera märkliga egenskaper, och föreslog att denna

[2] Bikvadrat = fjärdepotens

kurva inte borde kunna vara modulär. Jean-Pierre Serre kom nära att visa detta, och det sista och svåraste steget visades av Ken Ribet 1986. Nu visste man alltså att Taniyama-Shimuras förmodan implicerar Fermats sista sats!

Vid denna tidpunkt är Andrew Wiles en av världens främsta experter på modularitet för elliptiska kurvor, och Ribets resultat får honom att bestämma sig för att lösa Taniyama-Shimuras förmodan. Han isolerar sig helt och arbetar i hemlighet med problemet i sju år. 1993 tillkännager Wiles för en överrumplad matematikvärld att han kan bevisa påståendet för elliptiska kurvor som dessutom är semistabila, vilket är tillräckligt eftersom Freykurvan (som alltså inte existerar) kan visas vara semistabil. Argumentet är över 100 sidor långt och anses vara ett av de mest invecklade och tekniska matematiska bevis som någonsin skrivits ned. En lucka upptäcks dock i beviset, och Wiles bestämmer sig för att ta in hjälp utifrån. Tillsammans med Richard Taylor lagar han luckan i argumentet 1995, och Fermats sista sats är bevisad. År 2001 bevisas den fullständiga versionen av Taniyama-Shimura av Christophe Breuil, Brian Conrad, Fred Diamond, och Richard Taylor med hjälp av en modifikation av Wiles metoder.

En trevlig populärmatematisk bok om Fermats sista sats är *Fermats gåta* av Simon Singh, som bland annat återger historien oven med fler detaljer och mer dramatik.

Övningar

Övning 8.1 ($\star$)**.** Visa att $a^2 + b^2 = 7(c^2 + d^2)$ saknar icketriviala lösningar genom att reducera modulo 7.

Övning 8.2 ($\star$)**.** I *The Simpsons*-avsnittet "Treehouse of Horror VI" skymtar ekvationen

$$1782^{12} + 1841^{12} = 1922^{12}$$

förbi i bakgrunden. Stämmer den? Beräkna ett ungefärligt värde på tolfte roten av $1782^{12} + 1841^{12}$ med hjälp av dator.

Övning 8.3 ($\star\star$)**.** Använd metoden med minimalt motexempel för att visa följande: om p är ett primtal, finns ingen heltalslösning till ekvationen

$$x^3 + py^3 + p^2z^3 = 0.$$

Övning 8.4 ($\star\star$)**.** Använd det faktum att $x^4 + y^4 = z^2$ saknar positiva heltalslösningar för att visa att ekvationen $u^2 = v^4 + 1$ saknar *rationella* lösningar utom $(u, v) = (\pm 1, 0)$.

Ledning: Visa att om $u = \frac{a}{b}$ och $v = \frac{c}{d}$ är en lösning, så gäller $b = d^2$.

Övning 8.5 ($\star\star\star$)**.** Visa att Fermats sista sats för $n = 4$ och för udda primtalsvärden på n tillsammans implicerar Fermats sista sats för *varje* heltal större än eller lika med 3.

Övning 8.6 ($\star\star\star$)**.** Låt P vara påståendet "varje icke-tom delmängd av $\mathbb{N}$ har ett minsta element", och låt Q vara påståendet "det finns inga oändliga sekvenser

$$x_1 > x_2 > x_3 > \cdots$$

av positiva heltal". Visa att

$$P \iff Q.$$

Lösningar till udda övningsuppgifter

Övning 1.1.

(i) $B \cup C = A$.

(ii) $B \cap C = \emptyset$.

(iii) $D \cap C = \{4, 36\}$.

(iv) $\{x \in D \ : \ x \in B\} = D \cap B = \{1, 19, 101\}$.

(v) $\{x \in A \ : \ x = y + 1 \text{ för något } y \in D\} = \{2, 5, 20, 37, 102\}$.

(vi) $\{x + 1 \ : \ x \in D\} = \{2, 5, 20, 37, 102\}$.

Övning 1.3. Alla utom "Mängden av de naturliga talen" är påståenden. Det enda påståendet för vilket vi kan avgöra om det är sant eller falskt är "Varje mängd innehåller minst ett element", och detta påstående är falskt eftersom den tomma mängden inte innehåller något element.

Övning 1.5. Man finner att $27 = 4{\cdot}6{+}3$, $142 = 28{\cdot}5{+}2$, och $1429 = 476{\cdot}3{+}1$.

Övning 1.7. Ett exempel ges av följande funktioner.

(i) $f(1) = A, f(2) = B, f(3) = C$.

(ii) $f(1) = A, f(2) = B, f(3) = C$.

(iii) $f(1) = A, f(2) = B, f(3) = C, f(4) = A$.

(iv) $f(1) = A, f(2) = A, f(3) = B$.

Övning 1.9. Låt a och b vara givna, där $b > 0$. Vi vet enligt divisionssatsen att det går att skriva

$$a = qb + r$$

med $0 \leqslant r < b$. Dividerar vi denna ekvation med b finner vi att

$$\frac{a}{b} = q + \frac{r}{b}.$$

I denna ekvation är q ett heltal och det gäller att $0 \leqslant (r/b) < 1$ eftersom $0 \leqslant r < b$. Alltså är q lika med a/b avrundat nedåt till närmaste heltal.

(*Anmärkning*: Att göra denna räkning baklänges skulle kunna användas som ett alternativt bevis av divisionssatsen, om man förutsätter känt att varje reellt tal på ett unikt sätt kan avrundas nedåt till närmaste heltal. Notera att för att bevisa detta påstående behöver man dock använda någonting liknande minimumprincipen, för hur vet vi annars att det finns ett största heltal bland alla heltal mindre än r?)

Övning 1.11. Låt X vara en icke-tom nedåt begränsad delmängd av $\mathbb{Z}$. Då finns det $z \in \mathbb{Z}$ sådant att $x > z$ för alla $x \in X$. Sätt

$$Y = \{x - z : x \in X\}$$

Då är Y en icke-tom mängd av positiva heltal och enligt Princip 1.4.3 har Y ett minsta element y_0. Men då är $x_0 = y_0 + z \in X$ ett minsta element i X: För varje $x \in X$ gäller att $x - z \in Y$ och därmed $x - z \geqslant y_0 = x_0 - z$, alltså $x \geqslant x_0$.

Låt i stället X vara en icke-tom uppåt begränsad delmängd av $\mathbb{Z}$. Då finns det $z \in \mathbb{Z}$ sådant att $x < z$ för alla $x \in X$. Sätt nu

$$Y = \{z - x : x \in X\}$$

Då är Y en icke-tom mängd av positiva heltal och har enligt Princip 1.4.3 ett minsta element y_0. Vi påstår att $x_0 = z - y_0$ är ett största element i X: För varje $x \in X$ gäller att $z - x \in Y$ och därmed $z - x \geqslant y_0 = z - x_0$. Alltså är $x_0 \geqslant x$.

Övning 2.1. Antag att $a \mid b$ och att $b \mid c$. Då finns heltal m och n sådana att $b = an$ och $c = bm$. Det följer att $c = anm$ och därmed gäller att $a \mid c$.

Övning 2.3. (i) Vi byter till $a = 27$ och $b = 15$ då $27 > 15$ och får:

$$\begin{aligned} 27 &= 15 \cdot 1 + 12 \\ 15 &= 12 \cdot 1 + 3 \\ 12 &= 3 \cdot 4 + 0 \end{aligned}$$

Därmed är $\text{sgd}(27, 15) = 3$.

(ii) Vi får $\text{sgd}(615, 135) = 15$ med följande beräkning:

$$\begin{aligned} 615 &= 135 \cdot 4 + 75 \\ 135 &= 75 \cdot 1 + 60 \\ 75 &= 60 \cdot 1 + 15 \\ 60 &= 15 \cdot 4 + 0 \end{aligned}$$

(iii) Vi får $\text{sgd}(269, 196) = 1$ från:

$$\begin{aligned} 269 &= 196 \cdot 1 + 73 \\ 196 &= 73 \cdot 2 + 50 \\ 73 &= 50 \cdot 1 + 23 \\ 50 &= 23 \cdot 2 + 4 \\ 23 &= 4 \cdot 5 + 3 \\ 4 &= 3 \cdot 1 + 1 \\ 3 &= 1 \cdot 3 + 0 \end{aligned}$$

(iv) Vi får $\text{sgd}(8\,860, 1\,075) = 5$ från:

$$\begin{array}{rl|rl} 8\,860 &= 1\,075 \cdot 8 + 260 & \text{sgd}(8\,860, 1\,075) &= \text{sgd}(1\,075, 260) \\ 1\,075 &= 260 \cdot 4 + 35 & &= \text{sgd}(260, 35) \\ 260 &= 35 \cdot 7 + 15 & &= \text{sgd}(35, 15) \\ 35 &= 15 \cdot 2 + 5 & &= \text{sgd}(15, 5) \\ 15 &= 5 \cdot 3 + 0 & &= \text{sgd}(5, 0) = 5. \end{array}$$

Övning 2.5. (i) Eftersom $\text{sgd}(15, 27) = 3$ inte delar 7, har ekvationen $15x + 27y = 7$ ingen lösning.

(ii) Vi vet att $\text{sgd}(615, 135) = 15$, alltså finns det lösningar. Vi beräknar en lösning med Euklides algoritm använd baklänges:

$$\begin{array}{rll} 15 & = 75 - 60 \cdot 1 & = 75 - (135 - 75 \cdot 1) \cdot 1 \\ & = -135 + 75 \cdot 2 & = -135 + (615 - 135 \cdot 4) \cdot 2 \\ & = 615 \cdot 2 - 135 \cdot 9 & \end{array}$$

Alltså gäller $15 = 2 \cdot 615 - 9 \cdot 135$ och vi får en lösning med $x_0 = 2$ och $y_0 = -9$. Enligt Sats 2.4.2 har då alla lösningar formen

$$x = 2 + 9N \qquad \text{och} \qquad y = -9 - 41N.$$

(iii) Med samma metod får vi att $1 = -51 \cdot 269 + 70 \cdot 196$ och därmed har alla lösningar formen

$$x = -51 + 196N \qquad \text{och} \qquad y = 70 - 269N.$$

Övning 2.7. Antag att $d \mid x$ och $d \mid y$. Då gäller enligt Hjälpsats 2.2.4 också att $d \mid (ax + by) = 1$. Men det enda positiva tal som delar 1 är 1 självt, så den största gemensamma delaren är 1.

Övning 2.9. Antag att $d > 0$ är en gemensam delare till $a/\text{sgd}(a, b)$ och $b/\text{sgd}(a, b)$. Då finns det heltal m och n sådana att

$$\frac{a}{\text{sgd}(a, b)} = d \cdot n \qquad \text{och} \qquad \frac{b}{\text{sgd}(a, b)} = d \cdot m$$

Det följer att

$$a = \text{sgd}(a, b) \cdot dn \qquad \text{och} \qquad b = \text{sgd}(a, b) \cdot dm$$

Alltså är $\text{sgd}(a, b) \cdot d$ en positiv gemensam delare av a och b. Eftersom $\text{sgd}(a, b)$ är den största gemensamma delaren av a och b är detta endast möjligt om $d = 1$.

Övning 2.11. Enligt Sats 2.3.7 finns det heltal x och y sådana att $ax + by = 1$. Vi multiplicerar ekvationen med c och finner att $acx + bcy = c$. Men $a \mid acx$ och $a \mid bcy$, alltså gäller att $a \mid acx + bcy = c$.

Övning 2.13. (i) Mängden av gemensamma multiplar till a och b är en delmängd av de positiva heltalen. Dessutom innehåller den tydligen elementet ab, och är speciellt icke-tom. Så den har ett minsta element enligt Princip 1.4.3.

(ii) Låt m vara en gemensam multipel till a och b. Enligt Divisionssatsen kan vi skriva $m = q \cdot \text{mgm}(a,b) + r$ där $0 \leqslant r < \text{mgm}(a,b)$. Nu är både m och $\text{mgm}(a,b)$ delbara med både a och b och därmed är även $r = m - q \cdot \text{mgm}(a,b)$ delbart med både a och b. Alltså är r en gemensam multipel till a och b. Men $r < \text{mgm}(a,b)$ och $\text{mgm}(a,b)$ är den minsta gemensamma multipeln. Då måste $r = 0$ gälla och vi får att $m = q \cdot \text{mgm}(a,b)$ är en multipel av $\text{mgm}(a,b)$.

(iii) Eftersom

$$\frac{ab}{\text{sgd}(a,b)} = a \cdot \frac{b}{\text{sgd}(a,b)} = b \cdot \frac{a}{\text{sgd}(a,b)}$$

och både $\frac{b}{\text{sgd}(a,b)}$ och $\frac{a}{\text{sgd}(a,b)}$ är heltal, så är $\frac{ab}{\text{sgd}(a,b)}$ en gemensam multipel till a och b. Alltså gäller enligt (ii) att

$$\frac{ab}{\text{sgd}(a,b)} = q \cdot \text{mgm}(a,b)$$

för något heltal q.

Vi vet att $\text{mgm}(a,b) = kb$ för något heltal k och vi finner att

$$\frac{a}{\text{sgd}(a,b)} = qk$$

vilket visar att $q \mid \frac{a}{\text{sgd}(a,b)}$. På samma sätt visar man att $q \mid \frac{b}{\text{sgd}(a,b)}$. Men enligt Övning 2.9 är $\frac{a}{\text{sgd}(a,b)}$ och $\frac{b}{\text{sgd}(a,b)}$ relativt prima och då blir $q = 1$. Detta visar att

$$\frac{ab}{\text{sgd}(a,b)} = \text{mgm}(a,b).$$

Övning 3.1. Vi ställer upp följande tabell:

$a \pmod 4$	$a \pmod 2$	$a^2 \pmod 4$
0	0	0
1	1	1
2	0	0
3	1	1

Övning 3.3. (i) $1234567 \cdot 90123 \equiv 7 \cdot 3 = 21 \equiv 1 \pmod{10}$.

(ii) $2468 \cdot 13579 \equiv 18 \cdot 4 = 72 \equiv -3 \pmod{25}$.

Övning 3.5. Per definition gäller $n \mid (a-b)$, det vill säga det finns ett heltal q sådant att $a - b = n \cdot q$. Nu följer

$$(-a) - (-b) = -(a-b) = -(n \cdot q) = n \cdot (-q),$$

vilket betyder att $n \mid ((-a) - (-b))$, och därmed $-a \equiv -b \pmod{n}$.

Övning 3.7. (i) Enligt Sats 3.1.3 förändras inte värdet av uttrycket om vi reducerar varje förekomst av talet 10 modulo 9. Vi finner därmed att

$$\begin{aligned} x &= 10^n d_n + 10^{n-1} d_{n-1} + \ldots + 10 d_1 + d_0 \\ &\equiv 1^n d_n + 1^{n-1} d_{n-1} + \ldots + 1 \cdot d_1 + d_0 \pmod 9 \\ &= d_n + d_{n-1} + \ldots + d_1 + d_0 \pmod 9 \end{aligned}$$

Vi ser att högertermen är lika med summan av siffrorna och att den blir delbar med 9 precis när x är delbar med 9.

(ii) Denna uppgift är snarlik den föregående:

$$\begin{aligned} x &= 10^n d_n + 10^{n-1} d_{n-1} + \ldots + 10 d_1 + d_0 \\ &\equiv (-1)^n d_n + (-1)^{n-1} d_{n-1} + \ldots + (-1)^1 \cdot d_1 + (-1)^0 \cdot d_0 \pmod{11} \\ &= (-1)^n d_n + (-1)^{n-1} d_{n-1} + \ldots - d_1 + d_0 \pmod{11} \end{aligned}$$

Vi ser att högertermen är lika med alternerande summan av siffrorna och att den blir delbar med 11 precis när x är delbar med 11.

Övning 3.9. (i) Vi reducerar $3x^2 + 2 = y^2$ modulo 3 och får ekvationen $2 \equiv y^2 \pmod 3$. Men vi har redan sett att y^2 endast kan anta värdena 0 och 1 modulo 3.

(ii) Vi reducerar $7x^3 + 2 = y^3$ modulo 7 och får ekvationen $2 \equiv y^3 \pmod 7$. Men man kontrollerar lätt att y^3 endast antar värdena $0, 1$ och 6 modulo 7.

Övning 3.11. Om a är en enhet och $b = a^{-1}$, så gäller att $ab \equiv 1 \pmod n$ alltså även $ba \equiv 1 \pmod n$. Då är b en enhet och $b^{-1} = a$.

Övning 4.1. Vi har att

$$\begin{aligned} 12 &= 2 \cdot 2 \cdot 3, \\ 26 &= 2 \cdot 13, \\ 55 &= 5 \cdot 11, \\ 98 &= 2 \cdot 7 \cdot 7, \\ 150 &= 2 \cdot 3 \cdot 5 \cdot 5, \\ 210 &= 2 \cdot 3 \cdot 5 \cdot 7, \\ 315 &= 3 \cdot 3 \cdot 5 \cdot 7, \\ 455 &= 5 \cdot 7 \cdot 13. \end{aligned}$$

Övning 4.3. Vi vet att det finns ett tal q sådant att

$$aq = bc.$$

Vi kan primtalsfaktorisera denna ekvation på två sätt: dels genom att skriva en faktorisering av a och en av q efter varandra, och dels genom att skriva en

faktorisering av b och en av c efter varandra. Enligt aritmetikens fundamentalsats är dessa lika, så samma primtal ingår på bägge sidor. Eftersom a och b är relativt prima, kan omöjligt samma primtal ingå i faktoriseringen av a och faktoriseringen av b. Alltså kommer alla primtal som ingår i faktoriseringen av a att ingå i faktoriseringen av c med minst samma multiplicitet. Speciellt måste då $a \mid c$.

Övning 4.5. Enligt Fermats lilla sats är $2^4 \equiv 1 \pmod 5$ och därmed

$$2^{99} = 2^3 \cdot (2^4)^2 4 \equiv 2^3 \cdot 1^2 4 = 8 \equiv 3 \pmod 5.$$

P samma sätt gäller det att $3^6 \equiv 1 \pmod 7$ och vi finner att

$$3^{99} = 3^3 \cdot (3^6)^1 6 \equiv 3^3 \cdot 1^1 6 = 27 \equiv 6 \pmod 7.$$

Slutligen får vi att $4^8 \equiv 1 \pmod 9$, alltså

$$4^{99} = 4^3 \cdot (4^8)^1 2 \equiv 64 \cdot 1^1 2 \equiv 1 \pmod 9.$$

Övning 4.7. Om primfaktoriseringen av a ges av $a = p_1^{s_1} p_2^{s_2} \cdots p_k^{s_k}$ där $p_1, p_2, \ldots, p_k$ är olika primtal, så blir primfaktoriseringen av a^2 lika med

$$a^2 = p_1^{2s_1} p_2^{2s_2} \cdots p_k^{2s_k}$$

Eftersom $\operatorname{sgd}(b, c) = 1$, kan varje primtal p_i för $i = 1, \ldots, k$ bara förekomma i primfaktoriseringen det ena av talen b och c, men inte bägge. Och eftersom $a^2 = bc$ måste alla primtal som delar b eller c förekomma i primfaktoriseringen av a^2. Genom att eventuellt sortera om primtalen $p_1, p_2, \ldots, p_k$ kan vi anta att $p_1, \ldots, p_l$ delar b och att $p_{l+1}, \ldots, p_k$ delar c för något $0 \leqslant l \leqslant k$. Vi får då att primfaktoriseringen av b respektive c måste ha formen

$$b = p_1^{2s_1} p_2^{2s_2} \cdots p_l^{2s_l} \qquad \text{och} \qquad c = p_{l+1}^{2s_{l+1}} \cdots p_k^{2s_k}$$

vilket visar att de själva är kvadrater. Om vi sätter

$$s = p_1^{s_1} p_2^{s_2} \cdots p_l^{s_l} \qquad \text{och} \qquad t = p_{l+1}^{s_{l+1}} \cdots p_k^{s_k}$$

så får vi att $b = s^2$ och $c = t^2$. Dessutom är $\operatorname{sgd}(s, t) = 1$ då alla primtal $p_1, \ldots, p_k$ var olika.

Övning 4.9. (i) Enligt aritmetikens fundamentalsats så förekommer bara ändligt många primtal i primfaktoriseringen av n. Det är precis dessa primtal p som uppfyller att $\operatorname{ord}_p(n) > 0$.

(ii) Enligt aritmetikens fundamentalsats kan vi skriva $a = p_1 p_2 \cdots p_k$ som produkt av primtal. Då är $\operatorname{ord}_p(a)$ antalet gånger p förekommer bland primtalen $p_1, \ldots, p_k$. Så om $\operatorname{ord}_p(a) = \operatorname{ord}_p(b)$ för alla primtal p så måste talen a och b ha samma primtalsfaktorisering. Därför är också $a = b$.

(iii) Låt $s = \mathrm{ord}_p(n)$ och $t = \mathrm{ord}_p(m)$. Då är $n = p^s \cdot n'$ och $m = p^t \cdot m'$ där $p \nmid n'$ och $p \nmid n'$. Vi ser att $nm = (p^s \cdot n') \cdots (p^t \cdot m') = p^{s+t} \cdot n'm'$. Alltså gäller $p^{s+t} \mid nm$. Om nu $p^{s+t+1} \mid nm$, så skulle $p \mid n'm'$ gälla vilket inte kan stämma eftersom p varken delar n' eller m'. Därmed är $\mathrm{ord}_p(nm) = s + t = \mathrm{ord}_p(n) + \mathrm{ord}_p(m)$.

(iv) Om $n \mid m$, så är $m = n \cdot c$ för något heltal c. Det följer att $\mathrm{ord}_p(m) = \mathrm{ord}_p(n) + \mathrm{ord}_p(c) \geqslant \mathrm{ord}_p(n)$ eftersom $\mathrm{ord}_p(c) \geqslant 0$ för alla primtal p.

Antag nu att $\mathrm{ord}_p(n) \leqslant \mathrm{ord}_p(m)$ för alla primtal p. Vi vet från (i) att $\mathrm{ord}_p(m) > 0$ bara för ändligt många primtal $p_1, p_2, \ldots, p_k$. Vi har att $\mathrm{ord}_{p_i}(m) - \mathrm{ord}_{p_i}(n) \geqslant 0$ för $i = 1, \ldots, k$. Sätt

$$c = p_1^{\mathrm{ord}_{p_1}(m) - \mathrm{ord}_{p_1}(n)} p_2^{\mathrm{ord}_{p_2}(m) - \mathrm{ord}_{p_2}(n)} \cdots p_k^{\mathrm{ord}_{p_k}(m) - \mathrm{ord}_{p_k}(n)}.$$

Då är c ett heltal och vi ser att $m = \pm nc$ enligt (ii), eftersom bägge sidor har samma $\mathrm{ord}_p(-)$ för varje primtal p. Alltså gäller $n \mid m$.

Övning 4.11. (i) Vi multiplicerar med x och får att $x^2 \equiv 1 \pmod{p}$ eftersom $x^{-1} = x$. Alltså uppfyller varje lösning x ekvationen $x^2 - 1 \equiv 0 \pmod{p}$. Vi faktoriserar till

$$0 \equiv x^2 - 1 \equiv (x+1)(x-1) \pmod{p}.$$

Detta betyder att

$$p \mid (x-1)(x+1)$$

och eftersom p är ett primtal, så måste $p \mid x - 1$ eller $p \mid x + 1$ gälla enligt Övning 4.4. Därmed är $x = 1$ och $x = -1$ de enda lösningarna i $\mathbb{Z}_p$.

(ii) I $\mathbb{Z}_8$ och $\mathbb{Z}_{15}$ följer från $(x+1)(x-1) \equiv 0 \pmod{n}$ inte att $n \mid x+1$ eller $n \mid x - 1$. T.ex. gäller det att $2 \cdot 4 = 0$ i $\mathbb{Z}_8$ och därmed har ekvationen även lösningen $x = 3$. I $\mathbb{Z}_{15}$ gäller det att $3 \cdot 5 = 0$ och ekvationen har lösningen $x = 4$. Observera att $x = 1$ och $x = -1$ fortfarande är lösningar även i $\mathbb{Z}_8$ och $\mathbb{Z}_{15}$.

Övning 5.1. Vi beräknar

$$X^2 + Y^2 = (st)^2 + \left(\frac{s^2 - t^2}{2}\right)^2 = s^2t^2 + \frac{s^4 - 2s^2t^2 + t^4}{4} =$$
$$= \frac{s^4 + 2s^2t^2 + t^4}{4} = \left(\frac{s^2 + t^2}{2}\right)^2 = Z^2.$$

Övning 5.3. (i) Vi har visat i Hjälpsats 5.2.2 att om X är udda, så är Y jämnt och Z udda. Då är $Z - X$ jämnt och därmed blir $\frac{Z-X}{2}$ ett heltal.

(ii) Vi finner med hjälp av formeln för primitiva tripplar att

$$\frac{Z - X}{2} = \frac{\frac{s^2+t^2}{2} - st}{2} = \frac{s^2 + t^2 - 2st}{4} = \frac{(s-t)^2}{4} = \left(\frac{s-t}{2}\right)^2.$$

Eftersom både s och t är udda, så är $s - t$ jämnt och $\frac{s-t}{2}$ ett heltal.

Övning 5.5. Vi undersöker vilka rester som förekommer som $Y^2 \pmod 8$ med följande tabell:

$Y \pmod 8$	$Y^2 \pmod 8$
0	0
1	1
2	4
3	1
4	0
5	1
6	4
7	1

Vi ser att om X är udda, så gäller $X^2 \equiv 1 \pmod 8$. Vi har sett att Y måste vara jämnt och att $X^2 + Y^2$ därmed lämnar rest $1, 2$ eller 5 modulo 8. Men Z^2 lämnar rest $0, 1$ eller 4 modulo 8. Så det måste gälla att $X^2 + Y^2 \equiv Z^2 \equiv 1 \pmod 8$. Det följer att $Y^2 \equiv 0 \pmod 8$ vilket är fallet endast om $Y \equiv 0 \pmod 8$ eller $Y \equiv 4 \pmod 8$ enligt tabellen ovan. Då är Y delbart med 4 vilket skulle bevisas.

Övning 5.7. (i) Tag till exempel $(-1, 0, 1)$, $(35, 12, 37)$, $(63, 16, 65)$.

(ii) Tag till exempel $(1023, 64, 1025)$.

(iii) Formeln för primitiva tripplar ger att $Z = X + 2$ kan skrivas på formen

$$\frac{s^2 + t^2}{2} = st + 2$$

eller ekvivalent med detta $(s - t)^2 = 4$, eller $s - t = \pm 2$. Vi får alltså alla primitiva tripplar med $Z = X + 2$ genom att välja ett godtyckligt tal $t > 0$ udda och sedan låta $s = t + 2$ eller $s = t - 2$.

Övning 6.1. (i) Vi sätter in $x = a/c$ och $y = b/d$ i ekvationen $x^2 + y^2 = 1$ och får

$$\left(\frac{a}{c}\right)^2 + \left(\frac{b}{d}\right)^2 = 1$$

Vi multiplicerar ekvationen med $c^2 d^2$ och får

$$a^2 d^2 + b^2 c^2 = c^2 d^2$$

(ii) Ekvationen ovan kan skrivas som $a^2 d^2 = c^2 \left(d^2 - b^2\right)$. Alltså gäller att $c^2 \mid a^2 d^2$. Men $\operatorname{sgd}(a, c) = 1$ och då måste c^2 vara en delare till d^2 vilket medför att c delar d.

(iii) På samma sätt kan vi skriva $b^2 c^2 = d^2 \left(c^2 - a^2\right)$ och det följar att d^2 delar c^2 och därmed att d delar c. Om $c \mid d$ och $d \mid c$, så måste $c = d$ gälla.

Övning 6.3. Vi beräknar linjens lutning som

$$c = \frac{y_1 - y_0}{x_1 - x_0} = \frac{-\frac{12}{13} - (-1)}{\frac{5}{13} - 0} = \frac{1}{5}.$$

Därmed har linjen ekvationen $y = \frac{1}{5}x + d$ för något reellt tal d. Vi sätter in punkten $(0, -1)$ och får att $d = -1$. Linjens ekvation blir $y = \frac{1}{5}x - 1$.

Övning 6.5. Notera först att det räcker att bevisa påståendet för primitiva pythagoreiska tripplar, eftersom vi kan göra trippeln primitiv genom att dela med något tal k, och om Y/k är delbart med 4 är definitivt även Y det. Eftersom X är udda är Y jämnt, och enligt ekvation 6.6 kan vi skriva

$$Y = 2st$$

för relativt prima heltal s och t, varav ett är jämnt. Men då är speciellt st jämnt, så $st = 2q$ för något q, och

$$Y = 4q,$$

vilket skulle bevisas.

Övning 6.7. Låt (X, Y, Z) vara en primitiv pythagoreisk trippel. Den första konstruktionen avbildar trippeln på punkten $(\frac{X}{Z}, \frac{Y}{Z})$ på enhetscirkeln. Vi ser att $a = X$, $b = Y$ och $c = d = Z$. Dessutom är $\text{sgd}(a, c) = \text{sgd}(b, d) = 1$ och c samt d är positiva eftersom (X, Y, Z) är primitiv. Den andra konstruktionen avbildar punkten $(\frac{X}{Z}, \frac{Y}{Z})$ alltså på trippeln $(a, b, c) = (X, Y, Z)$ vilket är samma trippel som vi började med.

Låt på samma sätt $\left(\frac{a}{c}, \frac{b}{c}\right)$ vara en rationell punkt på enhetscirkeln med $\text{sgd}(a, c) = \text{sgd}(b, c) = 1$ och $c > 0$. Punkten avbildas på trippeln (a, b, c) av den andra konstruktionen och den första konstruktionen avbildar sedan trippeln (a, b, c) på punkten $\left(\frac{a}{c}, \frac{b}{c}\right)$, samma punkt som vi började med.

Övning 6.9. (i) Vi hittar att $65 = 8^2 + 1^2 = 7^2 + 4^2$. För $s = 8$ och $t = 1$ får vi trippeln $(X, Y, Z) = (16, 63, 65)$ och för $s = 7$ och $t = 4$ får vi $(X, Y, Z) = (56, 33, 65)$.

(ii) Vi hittar (efter en del sökande) att

$$1105 = 33^2 + 4^2 = 31^2 + 12^2 = 24^2 + 23^2$$

vilket ger oss de primitiva pythagoreiska tripplarna $(264, 1073, 1105)$, $(744, 817, 1105)$ och $(1104, 47, 1105)$.

Övning 7.1. Enligt antagande har f formen

$$f(x) = x^2 + bx + c$$

där b och c är heltal. Samtidigt är

$$f(x) = (x - x_1)(x - x_2)$$

där x_1 och x_2 är de två distinkta rötterna om $x_1 \neq x_2$ eller där $x_1 = x_2$ är dubbelroten. Vi multiplicerar ut och får att

$$f(x) = x^2 - (x_1 + x_2)x + x_1x_2$$

Det följer att $b = -x_1 - x_2$ och eftersom både b och x_1 är ett heltal enligt våra antaganden, är även $x_1 = -b - x_2$ ett heltal.

Övning 7.3. Nej. Betrakta till exempel följande andragradsekvation i tre variabler: $X^2 + Y^2 - 18Z^2 = 0$. Den har heltalslösningen $(3, 3, 1)$ som uppfyller att $\text{sgd}(X, Y) = 3$ men $\text{sgd}(X, Z) = \text{sgd}(Y, Z) = 1$.

Övning 7.5. (i) Antag att $b \neq 0$. I så fall kan vi skriva

$$y = \frac{c - ax}{b},$$

och speciellt är y ett rationellt tal så fort x är det. Alltså kan alla rationella punkter skrivas som

$$\left(x, \frac{c - ax}{b}\right).$$

Om vi har $a \neq 0$ fås på samma sätt att

$$\left(\frac{c - by}{a}, y\right)$$

är en parametrisering av alla rationella punkter.

(ii) Vi ska visa följande påstående:

Det finns en bijektion mellan alla rationella punkter på linjen $ax+by = c$ och alla primitiva heltalslösningar till ekvationen $aX + bY - cZ = 0$ där $Z > 0$.

Givet en heltalslösning till ekvationen $aX + bY - cZ = 0$ där $Z \neq 0$, kan vi sätta $x = \frac{X}{Z}$ och $y = \frac{Y}{Z}$. Insättning visar att (x, y) uppfyller ekvationen $ax + by = c$.

Givet en rationell punkt (x, y) på linjen $ax + by = c$ där $x = \frac{p}{q}$ och $y = \frac{r}{s}$, kan vi sätta $\bar{X} = ps$, $\bar{Y} = rq$ och $\bar{Z} = qs$. Då uppfyller $(\bar{X}, \bar{Y}, \bar{Z})$ ekvationen $aX+bY-cZ = 0$. Genom att multiplicera med -1 ifall $Z < 0$ och genom att dela med $\text{sgd}(\bar{X}, \bar{Y}, \bar{Z})$ får vi en primitiv heltalslösning (X, Y, Z) och vi ser att vi får tillbaka (x, y) när vi beräknar $x = \frac{X}{Z}$ och $y = \frac{Y}{Z}$.

Detta visar att avbildningen mellan de rationella punkterna (x, y) och primitiva tripplar (X, Y, Z) med $Z > 0$ enligt påståendet är en bijektion.

Eftersom ett degenererat kägelsnitt består av antingen bara punkter eller en eller två linjer, kan vi på detta sätt beskriva alla rationella punkter på varje degenererat kägelsnitt.

Övning 7.7. (i) Planet $Z = 1$ skär käglan i cirkeln $X^2 + Y^2 = 1$ vilken är en ellips.

(ii) Välj $Z = Y + 1$. Vi finner att

$$X^2 + Y^2 = (Y + 1)^2$$

vilket är ekvivalent med att

$$X^2 = 2Y + 1.$$

Skärningen uppfyller därför ekvationen $Y = \frac{1}{2}X^2 - \frac{1}{2}$ vilken beskriver en parabel.

(iii) Till exempel $Y = 1$ ger skärningen $Z^2 - X^2 = 1$ som är en hyperbel i planet.

(iv) Till exempel $Y = 0$ ger ekvationen $Z^2 - X^2 = 0$, alltså de två linjerna $Z = X$ och $Z = -X$ i planet.

(v) Med till exempel $Z = Y$ får man linjen $X = 0$.

(vi) Välj $Z = 0$ och man får endast origo som skärningspunkt.

Övning 7.9. (i) Ekvationen blir $0 = aX^2 + bXZ + cZ^2 - YZ$.

(ii) Om vi sätter $X = 0$, $Y = 1$ och $Z = 0$, så är ekvationen uppfylld.

(iii) Villkoret att linjen passerar genom $(0 : 1 : 0)$ ger att $s = 0$, så att linjerna genom $(0 : 1 : 0)$ har formen $rX + tZ = 0$ där r och t är heltal.

(iv) Linjerna i planet får ekvationen $rx + t = 0$ vilket motsvarar de vertikala linjerna på formen $x = q$ där $q = -t/r$ är ett rationellt tal. De skär parabeln i punkterna $(x, y) = (q, f(q))$ för rationella tal q vilket är samma resultat som i övning 7.4.

Övning 8.1. Antag att (a, b, c, d) är en icketrivial heltalslösning sådan att summan $a^2 + b^2$ är minimal. Då måste $a^2 + b^2 > 0$ gälla. Enligt tabellen

$a \pmod 7$	$a^2 \pmod 7$
0	0
1	1
2	4
3	2
4	2
5	4
6	1

måste a^2 och b^2 vara $0, 1, 2$ eller 4. Men nu konstruerar vi följande additionstabell i $\mathbb{Z}_7$:

	0	1	2	4
0	0	1	2	4
1		2	3	5
2			4	6
4				1

(Varför behöver vi bara konstruera halva tabellen?) Men vi vet att 7 delar $a^2 + b^2$, så summan av a^2 och b^2 måste bli 0 modulo 7, och vi ser att det enda sättet som a^2 och b^2 kan summera till noll är om $a^2 \equiv b^2 \equiv 0 \pmod 7$. Eftersom 7 är ett primtal är detta i sin tur endast möjligt om $a \equiv b \equiv 0 \pmod 7$. Vi har alltså att $a = 7m$ och $b = 7n$ för heltal m och n. Vi sätter in i ekvationen och får att

$$c^2 + d^2 = 7\left(m^2 + n^2\right)$$

Alltså är (c, d, m, n) en heltalslösning. Men $c^2 + d^2 = \frac{1}{7}\left(a^2 + b^2\right) < a^2 + b^2$ vilket motsäger vårt antagande att $a^2 + b^2$ var minimalt.

Övning 8.3. Antag att (x, y, z) är en heltalslösning där $|x|$ är minimalt. Eftersom p delar $py^3 + p^2z^3$, så måste då p dela x^3 och därmed även x eftersom p är ett primtal. Vi kan skriva $x = pv$ där v är ett heltal. Vi sätter in detta i ekvationen och får att

$$p^3v^3 + py^3 + p^2z^3 = 0$$

Division med p ger att $y^3 + pz^3 + p^2v^3 = 0$. Vi upprepar samma argument och får att $y = pv$ och sedan $z = pw$ där v och w är heltal samt att (u, v, w) upppfyller

$$u^3 + pv^3 + p^2w^3 = 0$$

Detta är en motsats till vårt antagande att (x, y, z) var en heltalslösning med $|x|$ minimalt eftersom $|v| = |x|/p < |x|$.

Övning 8.5. Antag att Fermats sista sats gäller för $n = 4$ och udda primtal.

Antag desssutom att det finns en heltalslösning till ekvationen $x^n + y^n = z^n$ där $n \geqslant 3$. Vi hävdar att n har en delare d som är antingen 4 eller ett udda primtal. Ty antingen är talet n delbart med ett udda primtal d, eller så är n en tvåpotens. Men en tvåpotens större än 3 är delbar med 4. Vi kan skriva $n = dm$ där m är ett heltal. Det följer att

$$(x^m)^d + (y^m)^d = (z^m)^d$$

och därmed är (x^m, y^m, z^m) en icketrivial heltalslösning till ekvationen $x^d + y^d = z^d$ som enligt vårt antagande inte har några icketriviala heltalslösningar.

Förslag till vidare läsning

De tre böckerna nedan är alla klassiska läroböcker i grundläggande talteori. Silvermans bok är lite vänligare än de efterföljande, som är skrivna i en väldigt kompakt stil som kräver långsam läsning.

[1] Joseph H. Silverman: *A friendly introduction to number theory.* 3:e upplagan. Prentice Hall, 2005.

[2] Godfrey H. Hardy, Edward M. Wright: *An introduction to the theory of numbers.* 6:e upplagan. Oxford University Press, 2008.

[3] Ivan Niven, Herbert S. Zuckerman, Hugh L. Montgomery: *An introduction to the theory of numbers.* 5:e upplagan. Wiley, 1991.

Nästa lilla häfte handlar specifikt om diofantiska ekvationer, och hinner behandla förvånande mycket stoff på en liten mängd text.

[4] Alexander O. Gelfond: *The solution of equations in integers.* Översatt från den ryska originalutgåvan. P. Noordhoff, Ltd., 1960.

Nästa bok är avsevärt svårare än de andra böckerna i denna referenslista och kräver att läsaren har läst en grundkurs i abstrakt och linjär algebra, eller är beredd att göra en viss bredvidläsning.

[5] Kenneth Ireland, Michael Rosen: *A classical introduction to modern number theory.* 2:a upplagan. Graduate Texts in Mathematics, 84. Springer-Verlag, 1990.

Följande bok är en klassisk introduktion till projektiv geometri. Boken handlar enbart om projektiv geometri över de reella talen och i två dimension (precis som i detta kompendium), vilket gör den ovanligt lättläst och satserna kan illustreras med bilder.

[6] Harold S.M. Coxeter: *The real projective plane.* 3:e upplagan. Springer-Verlag, 1993.

Böckerna ovan, och många andra böcker, finns att låna på Matematikbiblioteket, Lindstedtsvägen 25 (bottenvåningen). Biblioteket är öppet för alla.