HÖGSKOLAN
I GÄVLE

AKADEMIN FÖR TEKNIK OCH MILJÖ
Avdelningen för elektronik, matematik och naturvetenskap

Entydig faktorisering och Fermats stora sats i fallet n = 3

Leroy Kermanshahani

2018

Examensarbete, Grundnivå (kandidatexamen), 15 hp
Matematik

Handledare: Rolf Källström

Examinator: Johan Björklund

Förord

Detta examensarbete i matematik på kandidatnivå har genomförts under handledning av Rolf Källström, professor vid Högskolan i Gävle. Jag har under skrivandets gång fått kommentarer och svar på frågor som uppkommit, respons som varit till värdefull vägledning i författandet av texten, och jag vill härmed rikta ett stort tack till Rolf Källström för detta.

Sammanfattning

I denna text presenteras, via begreppet entydig faktorisering, Aritmetikens fundamentalsats. Först originalversionen för de ”vanliga” heltalen och senare en anpassad version för så kallade imaginära kvadratiska talringar, där elementen utgör en mer generell form av heltal. Däremellan redogörs för viktiga begrepp som behövs vid studiet av kvadratiska talringar, egenskaper för några olika typer av ringar, speciellt euklidiska, samt sambanden dem emellan. Ett huvudresultat är att entydig faktorisering, vilken gäller för alla ”vanliga” heltal större än 1, snarare är undantag än regel för imaginära kvadratiska talringar. Avslutningsvis ges en tillämpning av Aritmetikens fundamentalsats i form av ett bevis för Fermats stora sats i fallet $n = 3$.

Innehåll

1 Introduktion

Vi börjar med att ge ett antal grundläggande definitioner och satser som är bland de mest centrala i talteorin och som är bra att känna till för att kunna ta till sig texten, men även för att relatera tal och begrepp som senare tas upp i nya sammanhang till de ”vanliga” heltalen som vi betraktar här. Det mesta av detta borde dock redan vara välbekant för läsaren. Därefter formuleras tre lemman [1] som möjliggör formuleringen av den viktiga satsen i kapitel 2 med tillhörande bevis.

Definition 1.1. Ett *primtal* p är ett naturligt tal större än 1 vars enda delare är 1 och p självt. Ett naturligt tal större än 1 som inte är ett primtal är ett *sammansatt tal*.

Definitionen innebär att varje sammansatt tal n, till skillnad från ett primtal, kan skrivas som en produkt av *två* naturliga tal skilda från 1 och n, det vill säga

$$n = n_1 n_2 \tag{1.1}$$

där $1 < n_1 < n$ och $1 < n_2 < n$. Vi säger att (1.1) är en *heltalsfaktorisering* av n, där n_1 och n_2 är *heltalsfaktorer* till n. Detta behöver dock inte vara de enda möjliga heltalsfaktorerna. Exempelvis har vi för $n = 12$ att $12 = 2 \cdot 6 = 3 \cdot 4$, vilket alltså är två *olika* par av heltalsfaktorer.

Man kan också dela upp talet n i ett (ändligt) antal *primtal*. Vi får då i vårt exempel att $12 = 2 \cdot 2 \cdot 3$. Detta är de *enda primtalsfaktorerna* och vi säger därför att *primtalsfaktoriseringen* av 12 är *entydig*. Det ska snart visa sig att primtalsfaktoriseringen för ett (vanligt) heltal större än 1 *alltid* är entydig, oavsett vilket sådant tal man väljer. Notera att man bortser från faktorernas ordningsföljd. I vårt exempel räknas alltså $2 \cdot 2 \cdot 3$ som samma primtalsfaktorisering som $2 \cdot 3 \cdot 2$ och $3 \cdot 2 \cdot 2$. Mer om primtalsfaktorisering tas upp i kapitel 2 och kapitel 3.

Innan vi ger några definitioner och satser om delbarhet för heltal formuleras först en viktig sats inom mängdläran som vi kommer ha användning av i kapitel 3. Beviset ges dock inte här då då satsen ligger utanför teorin för denna text.

Sats 1.2. *(Välordningsprincipen) Varje icke-tom mängd av positiva heltal innehåller ett minsta tal.*

Definition 1.3. Ett heltal $a \neq 0$ *delar* ett heltal b om det finns ett heltal c sådant att $b = ac$, vilket vi skriver som $a \mid b$ (”a *delar* b”). Om a *inte* delar b skriver vi i stället $a \nmid b$.

Definition 1.4. Heltalet a är en *gemensam delare* till heltalen b och c om $a \mid b$ och $a \mid c$. Om b och c inte båda är lika med 0, så finns ett största sådant heltal som kallas *största gemensamma delaren* (sgd) till b och c och som betecknas $\text{sgd}(b, c)$. På samma sätt betecknar vi största gemensamma delaren till heltalen $a_1, a_2, ..., a_n$, där inte alla är lika med 0, med $\text{sgd}(a_1, a_2, ..., a_n)$.

Ett annat sätt att uttrycka $\text{sgd}(a_1, a_2, ..., a_n)$ är som det minsta positiva heltalet d som kan skrivas på formen

$$d = m_1a_1 + m_2a_2 + ... + m_na_n, \tag{1.2}$$

där $m_1, m_2, ..., m_n$ är heltal. Vi utelämnar här beviset för detta, som återges i till exempel [1, s. 12].

Definition 1.5. Heltalen $a_1, a_2, ..., a_n$ sägs vara *relativt prima* om $\text{sgd}(a_1, a_2..., a_n) = 1$.

Nästa sats är en generalisering av *Bézouts identitet*, uppkallad efter den franske matematikern Étienne Bézout verksam på 1700-talet (den ursprungliga satsen gäller för *två* heltal a och b).

Sats 1.6. *Om $a_1, a_2, ..., a_n$ är heltal, så finns det heltal $m_1, m_2, ..., m_n$ sådana att*

$$\text{sgd}(a_1, a_2, ..., a_n) = m_1a_1 + m_2a_2 + ... + m_na_n.$$

Bevis. Satsen följer direkt av existensen av $d = \text{sgd}(a_1, a_2, ..., a_n)$ i (1.2). □

Korollarium 1.7. *Om heltalen $a_1, a_2, ..., a_n$ är relativt prima, så finns det heltal $m_1, m_2, ..., m_n$ sådana att*

$$m_1a_1 + m_2a_2 + ... + m_na_n = 1.$$

Bevis. Korollariet följer av Sats 1.6 med $d = 1$ då $a_1, a_2, ..., a_n$ är relativt prima. □

Lemma 1.8. *Antag att a, b och c är heltal och att $\text{sgd}(a, b) = 1$. Om $a \mid bc$ så gäller att $a \mid c$.*

Bevis. Att $\text{sgd}(a, b) = 1$ innebär enligt Korollarium 1.7 att det finns heltal m och n sådana att $ma + nb = 1$. Multiplikation med c ger $mac + nbc = c$. Eftersom $a \mid mac$, och enligt villkor $a \mid bc$, fås att $a \mid (mac + nbc)$, det vill säga $a \mid c$. □

Lemma 1.9. *(Euklides lemma) Om a och b är heltal, p är ett primtal och $p \mid ab$ så gäller att $p \mid a$ eller $p \mid b$.*

Bevis. Eftersom p är ett primtal så är de enda möjligheterna att $\text{sgd}(p, a) = p$ eller $\text{sgd}(p, a) = 1$. I första fallet gäller att $p \mid a$. I andra fallet gäller enligt Lemma 1.8 att $p \mid b$. □

Som en parentes kan redan nu nämnas att satsen omvänt säger att om $p \mid ab$ ger att $p \mid a$ eller $p \mid b$, för heltal a och b, så är p ett primtal. Detta har vi nytta av längre fram då vi generaliserar begreppet primtal till att gälla även andra system än de vanliga heltalen. Vi avslutar detta inledande kapitel med en generalisering av Euklides lemma.

Lemma 1.10. *Om primtalet p delar $a_1 a_2 \cdots a_n$ så delar p minst en av heltalsfaktorerna $a_1, a_2, ..., a_n$.*

Bevis. Om $p \mid a_1$ är vi klara. I annat fall gäller enligt Lemma 1.9 att $p \mid a_2 \cdots a_n$. Om $p \mid a_2$ är vi klara, i annat fall gäller att $p \mid a_3 \cdots a_n$. Fortsätter vi på detta sätt får vi att p delar minst en av heltalsfaktorerna $a_1, a_2, ..., a_n$. □

2 Om primtalsfaktorisering

Som vi snart kommer se har varje heltal $n > 1$ *primtalsfaktoriseringen*

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r}, \tag{2.1}$$

där baserna $p_1, p_2, ..., p_r$ är de olika primtalsfaktorerna och exponenterna $\alpha_1, \alpha_2, ..., \alpha_r$ är positiva heltal som anger antalet gånger respektive faktor förekommer i n [2]. Vi kallar i fortsättningen denna mycket grundläggande aritmetiska procedur helt enkelt för en *faktorisering* av n.

Exempel 2.1. Vi kan skriva talet 4950 som

$$4950 = 2 \cdot 3 \cdot 3 \cdot 5 \cdot 5 \cdot 11 = 2 \cdot 3^2 \cdot 5^2 \cdot 11$$

med primtalsfaktorerna 2, 3, 5 och 11, där 3 och 5 förekommer två gånger vardera samt 2 och 11 en gång vardera i faktoriseringen.

Vi kan också skriva (2.1) som en oändlig produkt, med hjälp av produktsymbolen, som

$$n = \prod_p p^{\alpha(p)} \tag{2.2}$$

där $\alpha(p)$ är icke-negativt och p löper över alla primtal [2]. Att $\alpha(p) = 0$ innebär då att primtalet p inte ingår i faktoriseringen av n. För kvadrattalet n^2 gäller att alla exponenter $\alpha(p)$ i (2.2) är jämna. Även omvändningen gäller, det vill säga om alla exponenter är jämna så är talet ett kvadrattal. Vidare har ett kvadratfritt heltal n, som inte är delbart med något kvadrattal större än 1, enbart exponenter $\alpha(p)$ som antar värdena 0 eller 1. Även här gäller omvändningen. Vi återkommer till begreppet kvadratfritt heltal senare i teorin om kvadratiska talkroppar och kvadratiska talringar.

Satsen som presenteras i nästa kapitel är så pass viktig att den fått namnet *Aritmetikens fundamentalsats*, och ligger till grund för en stor mängd andra matematiska satser och resultat. Den säger att för ett givet heltal $n > 1$ så finns det en och endast en kombination (uppsättning) av primtal vars produkt är lika med det givna talet, enligt (2.1). Intuitivt kan det tyckas att påståendet alltid borde gälla men att så är fallet är inte alls självklart och inte heller sant, till exempel då man studerar vissa så kallade kvadratiska talringar i stället för de ”vanliga” heltalen $\mathbf{Z} = \{0, \pm 1, \pm 2...\}$. I det följande kommer visas att satsen alltså inte gäller generellt för alla typer av heltal, och en speciell typ av sådana tal kommer att definieras och diskuteras med början i kapitel 5.

3 Aritmetikens fundamentalsats

Vi är nu redo att formulera och bevisa talteorins viktigaste sats. Notera att satsen inte bara gäller för sammansatta tal utan även för primtal. Vi betraktar nämligen ett primtal som en produkt bestående av en enda primtalsfaktor, nämligen primtalet självt. Observera också att med begreppet *faktorisering* i beviset avses primtalsfaktorisering, precis som nämndes i föregående kapitel.

Sats 3.1. *(Aritmetikens fundamentalsats) Varje heltal $n > 1$ har en entydig primtalsfaktorisering om man bortser från faktorernas ordningsföljd.*

Bevis. (se [1]) Vi delar upp beviset i två delar. Först visas att det finns (minst) en faktorisering och sedan att dessa faktoriseringar (om de antas vara flera) verkligen är identiska, sånär som på ordningen av faktorerna.

Existens av faktorisering: Antag att det existerar ett heltal $n > 1$ som inte har en faktorisering och låt H beteckna mängden av alla sådana heltal. Enligt välordningsprincipen (Sats 1.2) finns då i H ett minsta tal m. Detta tal är sammansatt eftersom det, om det vore ett primtal, i så fall skulle ha en faktorisering, nämligen talet självt. Sätt nu $m = ab$ där $1 < a < m$ och $1 < b < m$, vilket enligt definitionen av m ger att $a, b \notin H$. Enligt definitionen av talet m måste a och b ha en faktorisering, vilket ger att även $m = ab$ har en faktorisering. Detta innebär att m inte tillhör H, som därför är den tomma mängden. Antagandet är alltså falskt och vi har därmed visat faktoriseringens existens för varje heltal $n > 1$.

Faktoriseringens entydighet: Antag att heltalet $n > 1$ har två faktoriseringar

$$n = p_1 p_2 \cdots p_r = q_1 q_2 \cdots q_s \tag{3.1}$$

där $r \leq s$. Enligt Lemma 1.10 delar p_1 en av faktorerna $q_1, q_2, ..., q_s$, säg att $p_1 \,|\, q_1$ (efter en eventuell omnumrering av q-faktorerna). Men då båda dessa är primtal följer att $p_1 = q_1$, vilket gör att vi kan förkorta bort p_1 och q_1 i likheten (3.1). Vi får då att

$$p_2 p_3 \cdots p_r = q_2 q_3 \cdots q_s.$$

På samma sätt fås att $p_2 = q_2, p_3 = q_3, ..., p_r = q_r$ och vi kan förkorta bort faktorerna $p_2, p_3, ..., p_r$ samt $q_2, q_3, ..., q_r$ eftersom $r \leq s$ enligt villkor. Det återstår nu att visa att antalet p-faktorer är lika med antalet q-faktorer i (3.1). Om $r < s$ fås den uppenbart orimliga likheten $1 = q_{r+1} q_{r+2} \cdots q_s$. Alltså gäller att $r = s$ och vi har därmed visat faktoriseringens entydighet för varje heltal $n > 1$. □

4 Några grundläggande algebraiska strukturer

För att enklare kunna skapa sig en förståelse för begreppen som dyker upp i senare avsnitt följer här en kortfattad genomgång av teorin för tre så kallade *algebraiska strukturer*, här ibland kortare benämnda *strukturer*[3]. Förhoppningen är helt enkelt att läsaren ska komma in i ett slags "algebraiskt tänkande" som kan motivera användandet och underlätta tillgodogörandet av nämnda begrepp. Avsnittet kan samtidigt ses som en möjlighet till överblick över området. Först en enkel definition.

Definition 4.1. En *binär operator* eller en *kompositionsregel* $*$ på en icke-tom mängd M är en funktion från den kartesiska produkten $M \times M$, bestående av alla par av element i M, till M självt.

Varje par av element i M ger alltså ett element i M - den binära operatorn sägs därför vara *sluten*, det vill säga man hamnar aldrig "utanför mängden". En binär operator $*$ garanterar alltså *slutenhet* på mängden.

Exempel 4.2. Vanlig *addition, subtraktion* och *multiplikation* på heltalen $\mathbf{Z} = \{0, \pm 1, \pm 2, ...\}$, de rationella talen $\mathbf{Q} = \{\frac{a}{b} \,|\, a, b \in \mathbf{Z}, b \neq 0\}$, de reella talen $\mathbf{R}$ och de komplexa talen $\mathbf{C} = \{a + bi \,|\, a, b \in \mathbf{R}, i^2 = -1\}$ samt vanlig *division* på $\mathbf{Q}^*, \mathbf{R}^*$ och $\mathbf{C}^*$ (det vill säga på alla nollskilda element i respektive mängd) är exempel på binära operatorer - likaså matrisaddition och matrismultiplikation på mängden av alla reella kvadratiska matriser.

4.1 Grupper

Bland de mest grundläggande strukturerna återfinns grupper. Innan vi ger en formell definition ska nämnas att kännedom om dessa är en förutsättning för en förståelse av nästkommande strukturs definition.

Definition 4.3. En *grupp* $(G, *)$, eller betecknad enbart G, är en mängd G med en binär operator $*$ sådan att

1. (Associativitet) $*$ är *associativ*, det vill säga $a * (b * c) = (a * b) * c$ för alla $a, b, c \in G$,
2. (Existens av neutralt element) G innehåller ett *neutralt element* e, alltså ett element e sådant att $a * e = e * a = a$ för alla $a \in G$ samt
3. (Existens av inverser) varje element $a \in G$ har en *invers* $b \in G$, alltså ett element b sådant att $a * b = b * a = e$.

Om $*$ också är *kommutativ*, det vill säga om $a * b = b * a$ för alla $a, b \in G$, så säger vi att gruppen är *abelsk*. (Kommutativitet)

Exempel 4.4. Heltalen med (den binära operatorn) addition $(\mathbf{Z}, +)$ och de nollskilda rationella talen med (den binära operatorn) multiplikation $(\mathbf{Q}^*, \cdot)$ är två exempel på grupper. Anledningen till att heltalen med multiplikation, $(\mathbf{Z}, \cdot)$, däremot *inte* bildar en grupp är att det saknas (multiplikativa) inverser. För exempelvis $a = 4$ ges inversen av det tal b som är lösning till ekvationen $4 \cdot b = b \cdot 4 = 1$, nämligen $\frac{1}{4}$, som *inte* är ett heltal, det vill säga finns inte i mängden. Det tredje kravet i definitionen är alltså inte uppfyllt och $(\mathbf{Z}, \cdot)$ är därmed inte en grupp.

Exempel 4.5. För att visa att de rationella talen med multiplikation $(\mathbf{Q}, \cdot)$ inte utgör en grupp räcker det att konstatera att talet 0 inte har en (multiplikativ) invers. Ekvationen $0 \cdot b = b \cdot 0 = 1$ har helt enkelt ingen lösning för b. Genom att exkludera nollan får man dock en mängd där *alla* element har en invers, och $(\mathbf{Q}^*, \cdot)$ är därmed en grupp enligt ovan.

Vi nämner här kort ett viktigt begrepp inom gruppteorin.

Definition 4.6. Antag att $(G, *)$ är en grupp. En delmängd H till G kallas då en *delgrupp* till $(G, *)$ om även $(H, *)$ är en grupp. Vi skriver detta som $H \leq G$. Om det också gäller att $H \neq G$ säger vi att H är en *äkta delgrupp* till G, vilket betecknas $H < G$.

Om H är en *icke-tom* delmängd till G räcker det att H är sluten med avseende på $*$ och invertering $(a, b \in H \Rightarrow a * b, a^{-1} \in H)$ för att visa att H är en delgrupp till G. Beviset för detta kan hittas i till exempel [3, s. 77].

4.2 Ringar

I och med introduktionen av begreppet ring närmar vi oss det centrala i denna text. Det kan vara värdefullt att gå tillbaka hit från senare kapitel och avsnitt för att försäkra sig om att den presenterade teorin utgår från just detta. Även här gäller att kännedom om definitionen krävs för att tillgodogöra sig innehållet i nästkommande strukturs definition.

Definition 4.7. En *ring* $(R, +, \cdot)$, eller betecknad enbart R, är en mängd R med de binära operatorerna addition $(+)$ och multiplikation $(\cdot)$ sådan att

1. $(R, +)$ är en abelsk grupp med det neutrala elementet 0 och har en invers $b \in R$ till varje element $a \in R$,
2. $\cdot$ är associativ, det vill säga $(ab)c = a(bc)$ gäller för alla $a, b, c \in R$ samt
3. $\cdot$ är distributiv över $+$, det vill säga $a(b+c) = ab + ac$ och $(a+b)c = ac + bc$ gäller för alla $a, b, c \in R$.

Exempel 4.8. Det algebraiska begreppet ”ring” är ett mycket omfattande begrepp. Exempel på ringar är heltalsringen $\mathbf{Z}$, men även $\mathbf{Q}, \mathbf{R}$ och $\mathbf{C}$ (som också uppfyller definitionen för nästa struktur) samt så kallade polynomringar och matrisringar. Vi ska dock senare inrikta oss på en speciell typ av ringar med tal som element, och fokus kommer då att vara på dessa typer av ringar, *kvadratiska talringar*.

Motsvarande begrepp till delgrupper ovan, men för ringar, definieras här. Definitionen med tillhörande kriterier har vi nytta av i avsnitt 5.15.

Definition 4.9. Antag att $(R, +, \cdot)$ är en grupp. En delmängd S till R kallas då en *delring* till $(R, +, \cdot)$ om även $(S, +, \cdot)$ är en ring. Om det också gäller att $S \neq R$ säger vi att S är en *äkta delring* till R.

Man kan visa att en delmängd S till en ring $(R, +, \cdot)$ är en delring till R om och endast om S är sluten med avseende på $+$ och $\cdot$ samt innehåller det multiplikativt neutrala elementet 1. Beviset för detta tas dock inte med här.

4.3 Kroppar

Vi har nu kommit fram till den tredje och sista strukturen som presenteras här, som liksom ringar är nära besläktad med begrepp som introduceras senare i texten. Här följer definitionen.

Definition 4.10. En *kropp* $(K, +, \cdot)$, eller betecknad enbart K, är en kommutativ ring R med det multiplikativt neutrala elementet 1 och där alla element i R utom 0 har en multiplikativ invers (är *enheter*).

Exempel på kroppar är de rationella talen $\mathbf{Q}$, de reella talen $\mathbf{R}$ och de komplexa talen $\mathbf{C}$, var och en med de binära operatorerna addition och multiplikation. Ytterligare exempel ges i nästa kapitel, där vi kommer definiera en speciell typ av kroppar med tal som element, *kvadratiska talkroppar*.

5 Kvadratiska talringar

5.1 Några definitioner

Innan vi definierar de för denna text centrala begreppen *talkropp* och *talring* behöver vi förstå vad begreppet *kroppsutvidgning* innebär. Ett sätt att åskådliggöra vad som menas med detta begrepp är att jämföra det med ett annat mer konkret begrepp, som i själva verket är dess logiska "motsats".

Definition 5.1. En *delkropp* K till en kropp $(L, +, \cdot)$ är en icke-tom delmängd till L som är sluten med avseende på $+$ och $\cdot$ samt invertering. L är då (omvänt) en *kroppsutvidgning* av K, det vill säga en kropp som innehåller K.

Definition 5.2. En *talkropp* är en ändlig kroppsutvidgning av den rationella talkroppen $\mathbf{Q}$. Att kroppsutvidgningen är *ändlig* innebär att den har ändlig dimension betraktat som vektorrum över $\mathbf{Q}$.

Exempel 5.3. Ett exempel på en talkropp är mängden av rationella tal $\mathbf{Q} = \{\frac{a}{b} \mid a, b \in \mathbf{Z}, b \neq 0\}$, som är den minsta möjliga talkroppen. Mängden av de "vanliga" heltalen $\mathbf{Z} = \{0, \pm 1, \pm 2, ...\}$ utgör däremot *inte* en talkropp eftersom den inte är sluten med avseende på division. Till exempel gäller för heltalen 3 och 4, som ju båda tillhör $\mathbf{Z}$, att kvoten $\frac{3}{4}$ *inte* tillhör mängden då den inte är ett heltal. Jämför med definitionerna i kapitel 4.

Definition 5.4. Ett *algebraiskt heltal* $z \in \mathbf{C}$ är ett nollställe till ett polynom på formen

$$z^n + a_{n-1}z^{n-1} + ... + a_1 z + a_0,$$

där koefficienterna $a_i \in \mathbf{Z}$ för $i = 1, 2, ..., n$. Vi säger att z är *algebraiskt över* $\mathbf{Z}$.

Två enkla exempel på algebraiska heltal är $\sqrt{2} \in \mathbf{C}$ och $i = \sqrt{-1} \in \mathbf{C}$. I det första fallet har vi nämligen att $z = \sqrt{2}$ är ett nollställe till polynomet $z^2 - 2$ med koefficienten $a_0 = -2 \in \mathbf{Z}$, medan $z = i$ är ett nollställe till polynomet $z^2 + 1$ med koefficienten $a_0 = 1 \in \mathbf{Z}$. Båda är alltså algebraiska över $\mathbf{Z}$.

Utifrån ovanstående definierar vi nu följande viktiga begrepp.

Definition 5.5. En *talring* R tillhörande en talkropp K är ringen av algebraiska heltal i K, det vill säga mängden av alla element i K som är algebraiska över $\mathbf{Z}$.

Anmärkning 1: Talringar benämns ofta *heltalsringar* (på engelska *rings of integers*) i litteraturen, och det kommer i detta sammanhang att vara underförstått att vi med "talringar" menar just heltalsringar.

Anmärkning 2: Definitionen innebär att varje talring ingår i en motsvarande talkropp, så att talringen med andra ord definieras utifrån en given talkropp.

Exempel 5.6. Ett exempel på en talring är heltalen $\mathbf{Z} = \{0, \pm 1, \pm 2, ...\}$ i $\mathbf{Q}$, som är den enklast möjliga talringen. Ett annat exempel är de *gaussiska heltalen* som studeras i avsnitt 5.15.

5.2 Kvadratiska talkroppar och kvadratiska talringar

Inför följande definitioner konstateras att ett *heltal* $m \neq 1$ sägs vara *kvadratfritt* om det inte är delbart med kvadraten på något primtal [3]. Till exempel är -5, 6 och 14 kvadratfria heltal, medan heltalen -4, 8 och 18 (delbara med $4 = 2^2, 4$ respektive $9 = 3^2$) *inte* är kvadratfria.

Definition 5.7. En *kvadratisk talkropp* $\mathbf{Q}[\sqrt{m}]$ är en talkropp av grad 2, det vill säga en talkropp av dimension 2 betraktat som vektorrum över $\mathbf{Q}$. Dess element utgörs av mängden av alla komplexa nollställen till ett kvadratiskt polynom på formen $z^2 + a_1 z + a_0$, där $a_0, a_1 \in \mathbf{Q}$, det vill säga alla tal på formen $z = a_0 + a_1\sqrt{m}$ där $a_0, a_1 \in \mathbf{Q}$ och m är ett kvadratfritt heltal.

Definition 5.8. En *kvadratisk talring* $\mathbf{D}$ är ringen av algebraiska heltal $z \in \mathbf{Q}[\sqrt{m}]$, det vill säga mängden av alla $z \in \mathbf{Q}[\sqrt{m}]$ som är algebraiska över $\mathbf{Z}$. Vi säger att z är ett *algebraiskt heltal* (av grad 2) i $\mathbf{D}$.

Anmärkning 1: Beteckningen $\mathbf{D}$ kommer här från begreppet "(Integral) Domain" (se till exempel [2, s. 127]).

Anmärkning 2: Notera att varje kvadratisk talring $R = \mathbf{D}$ är definierad utifrån den motsvarande kvadratiska talkroppen $K = \mathbf{Q}[\sqrt{m}]$, med beteckningar som i Definition 5.5.

Vi kommer i fortsättningen benämna algebraiska heltal i $\mathbf{D}$ för bara *heltal*, så länge det inte finns risk för sammanblandning med de de "vanliga" heltalen $\mathbf{Z}$.

Ett element $z \in \mathbf{C}$ är alltså ett (algebraiskt) heltal i $\mathbf{D}$ om och endast om z är ett nollställe till ett kvadratiskt polynom (det vill säga ett polynom av grad 2) på formen $z^2 + a_1 z + a_0$, där $a_0, a_1 \in \mathbf{Z}$. Vilka dessa är framgår i avsnitt 5.4.

Vi uppehåller oss i fortsättningen, om inte annat framgår, vid kvadratiska tal*ringar* och konstaterar att den kvadratiska talringen $\mathbf{D}$ tillhörande $\mathbf{Q}[\sqrt{m}]$ kallas *imaginär* om $m < 0$, medan den kallas *reell* om $m > 0$.

5.3 Delare, enheter och associerade element

Definition 5.9. Ett nollskilt heltal $a \in \mathbf{D}$ är en *delare* till ett heltal $b \in \mathbf{D}$ om och endast om det finns ett heltal $c \in \mathbf{D}$ sådant att $b = ac$. Vi skriver i sådana fall $a \mid b$.

Begreppet *delare* definieras alltså formellt på samma sätt för heltal i $\mathbf{D}$ som för heltalen i $\mathbf{Z}$. Som vi ska se gäller denna "generalisering" - från talringen $\mathbf{Z}$ till kvadratiska talringar - även exempelvis följande centrala begrepp.

Definition 5.10. Varje delare till 1 i en kvadratisk talring $\mathbf{D}$ kallas en *enhet* i ringen. Nollskilda element a och b i denna ring kallas *associerade* om det gäller att $a = b\varepsilon$, där ε är en enhet i ringen.

Exempel 5.11. För heltalen i $\mathbf{Z}$ gäller exempelvis att associerade element till 20 är $1 \cdot 20 = 20$ och $(-1) \cdot 20 = -20$, eftersom det finns exakt två enheter i $\mathbf{Z}$, nämligen ± 1.

5.4 Vilka algebraiska heltal ingår i kvadratiska talringar?

Av följande sats framgår exakt vilka algebraiska heltal z som den kvadratiska talringen $\mathbf{D}$, definierad utifrån den kvadratiska talkroppen $\mathbf{Q}[\sqrt{m}]$, innehåller. Jämför med Definition 5.8.

Sats 5.12. *Antag att m är ett kvadratfritt heltal. Om $m \equiv 2, 3 \pmod 4$ kan den till $\mathbf{Q}[\sqrt{m}]$ motsvarande kvadratiska talringen skrivas på formen*

$$\mathbf{D} = \mathbf{Z}[\sqrt{m}] = \{a + b\sqrt{m} \mid a, b \in \mathbf{Z}\}.$$

Om istället $m \equiv 1 \pmod 4$ kan den till $\mathbf{Q}[\sqrt{m}]$ motsvarande kvadratiska talringen skrivas på formen

$$\mathbf{D} = \mathbf{Z}\left[\frac{1+\sqrt{m}}{2}\right] = \{a + b\sqrt{m} \mid a, b \in \mathbf{Z}\} \cup \left\{\frac{a+b\sqrt{m}}{2} \mid \textit{udda } a, b \in \mathbf{Z}\right\}.$$

Bevis. (se [2]) Vi ska identifiera algebraiska heltal som är på formen $z = \frac{a+b\sqrt{m}}{c}$, där $a, b, c \in \mathbf{Z}$ och $c > 0$. Vi antar, utan inskränkning, att $\text{sgd}(a, b, c) = 1$ så att z är uttryckt i sin enklaste form. Om $b = 0$ så gäller att $z \in \mathbf{Q}$ och att z är ett algebraiskt heltal om och endast om $z \in \mathbf{Z}$, det vill säga $c = 1$. Om i stället $b \neq 0$ så är $z \notin \mathbf{Z}$ och polynomekvationen är kvadratisk, det vill säga

$$\left(x - \frac{a+b\sqrt{m}}{c}\right)\left(x - \frac{a-b\sqrt{m}}{c}\right) = x^2 - \frac{2a}{c}x + \frac{a^2 - b^2 m}{c^2} = 0. \tag{5.1}$$

Enligt avsnitt 5.1 är z ett algebraiskt heltal i $\mathbf{D}$ om och endast om (5.1) har heltalskoefficienter och är *monisk*, det vill säga har 1 som högstagradskoefficient. Alltså är z ett heltal om och endast om

$$c \mid 2a, \quad c^2 \mid a^2 - b^2 m, \tag{5.2}$$

inklusive fallet $b = 0$ (eftersom $\text{sgd}(a, b, c) = 1$). Om $\text{sgd}(a, c) > 1$ och $c \mid 2a$ så har a, c en gemensam primtalsfaktor, säg p, och det gäller att $p \nmid b$ eftersom $sgd(a, b, c) = 1$. Vi får då att $p^2 \mid a^2, p^2 \mid c^2$ och om $c^2 \mid a^2 - b^2 m$ så skulle vi ha att $p^2 \mid b^2 m$ och alltså $p^2 \mid m$, vilket är omöjligt eftersom m är ett kvadratfritt heltal. Villkoren (5.2) gäller därför bara om $sgd(a, c) = 1$. Om $c \mid 2a$ och $c > 2$ så är $\text{sgd}(a, c) > 1$, så (5.2) kan gälla bara om $c = 1$, vilket är klart, eller $c = 2$.

Om $c = 2$ ger (5.2) att $a^2 \equiv b^2 m \pmod 4$ och a måste vara udda eftersom $\text{sgd}(a, c) = 1$. Villkoren (5.2) blir då $b^2 m \equiv a^2 \equiv 1 \pmod 4$, vilket ger att b är udda och därmed att $m \equiv b^2 m \equiv 1 \pmod 4$.

Sammanfattningsvis gäller alltså att (5.2) är uppfyllt om och endast om $c = 1$ eller $c = 2$, a, b är udda och $m \equiv 1 \pmod 4$. □

5.5 Normen

Eftersom heltalen i $\mathbf{D}$ till skillnad från de ”vanliga” heltalen $\mathbf{Z}$ inte direkt anger ”storleken” på talen, behöver vi ett slags motsvarande ”mått” på dessa som gör det möjligt att jämföra dem vad gäller olika egenskaper. Som vi ska se har detta en större praktisk betydelse än vad man vid första anblicken kan tro, speciellt när det kommer till faktorisering.

Definition 5.13. (se [3]) För varje $z \in \mathbf{D}$ kallas

$$N(z) = |z|^2 = z\overline{z}$$

för *normen* av z.

Anmärkning: För $z = a + b\sqrt{m}$ blir normen alltså

$$N(z) = (a + b\sqrt{m})(a - b\sqrt{m}) = a^2 - mb^2. \tag{5.3}$$

I fallet $m \equiv 1 \pmod 4$ ingår i $\mathbf{D}$, enligt Sats 5.12, algebraiska heltal på formen $z = \frac{a+b\sqrt{m}}{2}$. Normen blir alltså i detta fall

$$N(z) = \left(\frac{a + b\sqrt{m}}{2}\right)\left(\frac{a - b\sqrt{m}}{2}\right) = \frac{a^2 - mb^2}{4}. \tag{5.4}$$

Exempel 5.14. I den kvadratiska talringen $\mathbf{D} = \mathbf{Z}[\sqrt{2}]$ har vi att $N(6-2\sqrt{2}) = 6^2 - 2(-2)^2 = 28$ och $N(2 + 3\sqrt{2}) = 2^2 - 2 \cdot 3^2 = -14$. Normen av ett algebraiskt heltal kan alltså vara såväl positiv som negativ, vilket strax kommer framgå.

En del kan nu sägas om normen, och vi samlar ett antal av dessa egenskaper i en sats. Men först ett lemma.

Lemma 5.15. *För ett kvadratfritt heltal m gäller att $\sqrt{m} \notin \mathbf{Q}$.*

Bevis. (se [3]) Det inses lätt att satsen är sann för $m < 0$. Antag därför att $\sqrt{m} \in \mathbf{Q}$ för något $m > 0$, alltså att $\sqrt{m} = \frac{a}{b}$ för några $a, b \in \mathbf{Z}$. Detta ger att $a^2 = nb^2$. Aritmetikens fundamentalsats ger nu att varje primtalsfaktor i a^2 och b^2 finns med ett jämnt antal gånger i faktoriseringen av respektive tal. Men eftersom n är kvadratfritt finns det minst ett primtal som finns med ett udda antal gånger i faktoriseringen av nb^2. Men detta är en motsägelse, varför vårt antagande inte stämmer och det alltså gäller att $\sqrt{m} \notin \mathbf{Q}$. □

Sats 5.16.

1. $N(z) \in \mathbf{Z}$ *för alla* $z \in \mathbf{D}$
2. $N(z) = 0$ *för något* $z \in \mathbf{D} \Leftrightarrow z = 0$
3. $N(z) = N(\overline{z})$ *för alla* $z \in \mathbf{D}$
4. $N(wz) = N(w)N(z)$ *för alla* $w, z \in \mathbf{D}$
5. $N(z) = \pm 1$ *för något* $z \in \mathbf{D} \Leftrightarrow z$ *är en enhet i* $\mathbf{D}$.

Bevis. (se [2])

1. Enligt (5.3) och (5.4) ges normen av $a^2 - mb^2$ respektive $\frac{a^2-mb^2}{4}$ där a och b är heltal, vilket ger att även normen är ett heltal.
2. ($\Leftarrow$) Att $N(0) = 0$ inses enkelt genom insättning av $a = b = 0$ i (5.3) respektive (5.4).

($\Rightarrow$) Antag att $N(z) = 0$ där $z = a + b\sqrt{m}$ eller $z = \frac{a+b\sqrt{m}}{2}$. Då gäller $a^2 - mb^2 = 0$, det vill säga $mb^2 = a^2$. Om $b \neq 0$ gäller $m = \frac{a^2}{b^2}$, alltså är $\sqrt{m} = |\frac{a}{b}|$ ett rationellt tal. Enligt Lemma 5.15 är dock detta omöjligt så $b = 0$ måste gälla. Alltså är även $a = 0$ och därmed $z = 0$.

3. För $z = a + b\sqrt{m}$ gäller att $N(z) = a^2 - mb^2 = a^2 - m(-b)^2 = N(\overline{z})$.

4. Om $w = a + b\sqrt{m}$ och $z = c + d\sqrt{m}$ så gäller att

$$\begin{aligned} N(wz) &= N((a + b\sqrt{m})(c + d\sqrt{m})) \\ &= N(ac + ad\sqrt{m} + bc\sqrt{m} + bdm) \\ &= N((ac + bdm) + (ad + bc)\sqrt{m}) \\ &= (ac + bdm)^2 - m(ad + bc)^2 \\ &= a^2c^2 + 2abcdm + m^2b^2d^2 - m(a^2d^2 + 2abcd + b^2c^2) \\ &= a^2c^2 + m^2b^2d^2 - ma^2d^2 - mb^2c^2. \end{aligned}$$

   Men vi har också att

$$\begin{aligned} N(w)N(z) &= (a^2 - mb^2)(c^2 - md^2) \\ &= a^2c^2 - ma^2d^2 - mb^2c^2 + m^2b^2d^2. \end{aligned}$$

   Alltså är $N(wz) = N(w)N(z)$.

5. ($\Leftarrow$) Antag att $z \in \mathbf{D}$ är en enhet. Då finns ett heltal ε sådant att $z\varepsilon = 1$. Detta ger $N(z\varepsilon) = N(z)N(\varepsilon) = N(1) = 1$, vilket innebär att $N(z) = \pm 1$ eftersom $N(z)$ och $N(\varepsilon)$ är heltal.

   ($\Rightarrow$) Om $N(z) = \pm 1$ där z är ett heltal i $\mathbf{D}$ så gäller att $z\overline{z} = \pm 1$ och alltså att $z \,|\, 1$, vilket innebär att z är en enhet.

Därmed är satsen bevisad. □

Observera, som nämnts tidigare, att elementen i en godtycklig kvadratisk talring i fortsättningen oftast kommer att benämnas ”heltal”, trots att det inte längre är talringen $\mathbf{Z} = \{0, \pm 1, \pm 2...\}$ vi studerar, utan talringen $\mathbf{D} = \mathbf{Z}[\sqrt{m}]$ för ett kvadratfritt heltal m, enligt Sats 5.12. Vi kommer med andra ord att studera en annan typ av matematiskt ”system”, där den ordinära definitionen av ”heltal” inte längre gäller utan har generaliserats.

5.6 Vilka är enheterna i kvadratiska talringar?

Det kan ur bland annat faktoriseringssynpunkt vara intressant att känna till vilka enheter en given kvadratisk talring innehåller. Det visar sig att, för *imaginära* sådana, är enheterna maximalt sex till antalet. Vi går nu över till att bestämma samtliga enheter till varje kvadratisk talring, först de imaginära och därefter de reella. Av följande satser framgår exakt vilka dessa är.

Sats 5.17. *Antag att m är ett negativt kvadratfritt heltal. Den imaginära kvadratiska talringen $\mathbf{D} = \mathbf{Z}[\sqrt{m}]$ respektive $\mathbf{D} = \mathbf{Z}[\frac{1+\sqrt{m}}{2}]$ har då de två enheterna ± 1, förutom i fallen då $m = -1$ och $m = -3$: enheterna i $\mathbf{D} = \mathbf{Z}[\sqrt{-1}] = \mathbf{Z}[i]$ är ± 1 och $\pm i$, och enheterna i $\mathbf{D} = \mathbf{Z}[\frac{1+\sqrt{-3}}{2}]$ är ± 1 och $\frac{\pm 1 \pm \sqrt{-3}}{2}$.*

Bevis. (se [2]) Enligt Sats 5.16 ska vi söka de heltal $a \in \mathbf{D}$ för vilka gäller att $N(a) = \pm 1$. Enligt Sats 5.12 kan a skrivas som antingen $x + y\sqrt{m}$ eller $\frac{x+y\sqrt{m}}{2}$, där $x, y \in \mathbf{Z}$, och där x, y i det senare fallet är udda. Då gäller att $N(a) = x^2 - my^2$ respektive $N(a) = \frac{x^2-my^2}{4}$. Eftersom $m < 0$ så gäller att $N(a) \geq 0$, vilket innebär att det inte finns något a för vilket $N(a) = -1$ gäller. Den enda ekvation vi behöver lösa är alltså $N(a) = 1$. Vi delar nu upp beviset i fyra fall.

1. För $m < -1$ har vi att $1 = N(a) = x^2 - my^2 \geq -my^2 \geq 2y^2$ och de enda lösningarna är $y = 0, x = \pm 1$. Enheterna är alltså $a = \pm 1$.
2. För $m = -1$ får vi ekvationen $1 = N(a) = x^2 + y^2$ vilken har de fyra lösningarna $x = 0, y = \pm 1$ samt $x = \pm 1, y = 0$. Enheterna är alltså $a = \pm 1$ och $a = \pm i$.
3. För $m < -3$ och $m \equiv 1 (\mathrm{mod}\, 4)$ gäller att $x^2 - my^2 \geq 1 - m > 4$ för udda x och y. Det finns därför inga lösningar till $1 = N(a) = \frac{x^2-my^2}{4}$ i detta fall.
4. För $m = -3$ får vi ekvationen $1 = N(a) = \frac{x^2+3y^2}{4}$ med udda x och y, vilken har lösningarna $x = 1, y = \pm 1$ och $x = -1, y = \pm 1$. Dessa motsvarar enheterna $a = \frac{\pm 1 \pm \sqrt{-3}}{2}$.

Därmed är satsen bevisad. □

Vi övergår nu till vad som kan sägas om enheterna i en *reell* kvadratisk talring.

Sats 5.18. *Antag att m är ett positivt kvadratfritt heltal. Den reella kvadratiska talringen $\mathbf{D}$ har då oändligt många enheter. Dessa kan skrivas som $\pm u^n$ där $n \in \mathbf{Z}$ och där $u > 1$ är en enhet i $\mathbf{D}$. Vi kallar den minsta enheten $\varepsilon > 1$ för den fundamentala enheten i $\mathbf{D}$.*

Bevis. (se [2]) Enligt Sats 5.12 är heltal på formen $z = x + y\sqrt{m}$, där $x, y \in \mathbf{Z}$, också heltal i $\mathbf{D}$. För dessa gäller att $N(z) = x^2 - my^2$, det vill säga om $x^2 - my^2 = 1$ så är z en enhet. Då $m > 1$ har denna ekvation oändligt många lösningar, vilket ger att antalet enheter i $\mathbf{D}$ är oändligt[1]. För bevis av satsens andra del, se [4, s. 192]. □

5.7 Irreducibla element

Definition 5.19. Ett heltal $z \neq 0$, och ej en enhet, i den kvadratiska talringen $\mathbf{D}$ som inte kan faktoriseras som $z = ab$, där a och b ej är enheter, kallas ett *irreducibelt element*. I annat fall är z ett *reducibelt element*.

Det enda sättet att skriva det irreducibla elementet $z \in \mathbf{D}$ som en produkt av två heltal är alltså att låta en av faktorerna vara en enhet.

[1] Se även avsnitt 6.4.

I talringen $\mathbf{Z}$ är samtliga irreducibla element på formen $\pm p$, där p är ett primtal. Notera också att ett irreducibelt element p i talringen $\mathbf{Z}$ är delbart endast med enheterna ± 1 och dess associerade element, det vill säga $p \cdot 1 = p$ och $p(-1) = -p$.

Vi visar nu vad som kan sägas om heltal i en kvadratisk talring $\mathbf{D}$ vars norm är lika med ett (positivt eller negativt) primtal.

Sats 5.20. *Antag att $z \in \mathbf{D}$ och att p är ett primtal. Om $N(z) = \pm p$ så är z ett irreducibelt element i $\mathbf{D}$.*

Bevis. (se [3]) Antag motsatsen, det vill säga att $N(z) = \pm p$ för något primtal p och att z inte är irreducibelt, det vill säga reducibelt. Då finns $v, w \in \mathbf{D}$, nollskilda och ej enheter, sådana att $z = vw$. Detta ger att $\pm p = N(z) = N(vw) = N(v)N(w)$, det vill säga antingen $N(v)$ eller $N(w)$ är ± 1. Antingen v eller w är alltså en enhet, vilket motsäger antagandet. Därmed är z irreducibelt och satsen är bevisad. □

5.8 Primelement

Definition 5.21. Ett heltal $z \neq 0$, ej en enhet, i den kvadratiska talringen $\mathbf{D}$ sådant att om $z \mid ab$, där $a, b \in \mathbf{D}$, ger att $z \mid a$ eller $z \mid b$, kallas ett *primelement.*

Exempel 5.22. Talet 2 är ett primelement i $\mathbf{Z}$, men *inte* i $\mathbf{D} = \mathbf{Z}[\sqrt{-1}]$, eftersom vi kan skriva $2 = (1 + i)(1 - i)$ där 2 som synes delar produkten i högerledet men inte någon av de ingående faktorerna $1 \pm i$.

Notera att ett primelement z i talringen $\mathbf{Z}$, liksom varje irreducibelt element enligt ovan, är delbart endast med enheterna ± 1 och dess associerade element, det vill säga $z \cdot 1 = z$ och $z(-1) = -z$. Anledningen till att det förhåller sig så framgår av följande avsnitt.

5.9 Om sambandet mellan irreducibla element och primelement

De två senaste definitionerna tycks kanske beskriva ”vanliga” primtal, så som vi definierat och diskuterat dem i kapitel 1. Det gäller ju för dessa tal att de bara kan skrivas som produkten av 1 och sig självt, vilket ju liknar Definition 5.19. Å andra sidan är innebörden av Euklides lemma för primtal mycket likt Definition 5.21. Kan det då finnas någon koppling mellan dessa tre begrepp? Svaret är att det *finns* ett nära samband, nämligen att begreppet irreducibla element *i vissa fall* är ekvivalent med primelement, varav *ett* sådant fall är då vi håller oss till den ”vanliga” talringen $\mathbf{Z}$. Detta förklarar alltså varför de båda nya begreppen lätt kan relateras till varandra - de beskriver helt enkelt samma tal, benämnda primtal, i $\mathbf{Z}$. Vi ska dock enligt målet med detta kapitel inte stanna här utan kommer utvidga och titta närmare på dessa begrepp utifrån kvadratiska talringar $\mathbf{D}$, där de inte alls nödvändigtvis innebär precis samma tal. I avsnitt 5.11 ges ett exempel på detta. Först ett enkelt konstaterande, som speciellt gäller för kvadratiska talringar.

Sats 5.23. *Varje primelement i en talring är ett irreducibelt element i samma ring.*

5.10 Existens av faktorisering i irreducibla element

Vi visar nu att det alltid går att skriva ett heltal i en kvadratisk talring som en produkt av irreducibla element. Notera att denna egenskap motsvarar existensdelen i Aritmetikens fundamentalsats (Sats 3.1) i kapitel 3.

Sats 5.24. *Varje heltal $z \neq 0$, som ej är en enhet, i* $\mathbf{D}$ *kan faktoriseras i irreducibla element.*

Bevis. (se [2]) Om z inte är ett irreducibelt element så kan det faktoriseras som $z = ab$, där varken a eller b är enheter. Om a eller b inte är irreducibla fortsätter vi att faktorisera dessa. Detta går att göra endast ett ändligt antal gånger, eftersom vi annars skulle få att $z = a_1 a_2 \cdots a_n$ där n är godtyckligt stort och ingen faktor a_j är en enhet. Men detta skulle, för godtyckligt stort n, ge att

$$N(z) = \prod_{j=1}^{n} N(a_j)$$

och alltså att

$$|N(z)| = \prod_{j=1}^{n} |N(a_j)| \geq 2^n,$$

eftersom $|N(a_j)|$ är ett heltal > 1. Alltså går det att faktorisera z i (ett ändligt antal) irreducibla element. □

5.11 UFD och entydig faktorisering i irreducibla element

Målet för återstoden av detta kapitel är att generalisera Aritmetikens fundamentalsats till att gälla inte bara ringen av ”vanliga” heltal $\mathbf{Z}$, utan även *vissa* kvadratiska talringar $\mathbf{D}$. Som vi senare ska se är en sådan generalisering endast möjlig för imaginära sådana, det vill säga då $m < 0$. Vi vill alltså besvara frågan ”När gäller entydig faktorisering för imaginära kvadratiska talringar?”. Med detta mål i sikte börjar vi med att definiera hur egenskapen ”entydig primtalsfaktorisering” i den ursprungligt formulerade satsen ska omformuleras och tolkas i detta nya sammanhang, samt hur vi ska benämna dessa ringar, i det eller de fall egenskapen gäller.

Definition 5.25. En kvadratisk talring $\mathbf{D}$ sägs ha en *entydig faktorisering* om varje heltal $z \neq 0$, och ej en enhet, i $\mathbf{D}$ har en entydig faktorisering i irreducibla element om man bortser från faktorernas ordningsföljd och förekomster av associerade element. Vi säger då att $\mathbf{D}$ är en *ring med entydig faktorisering (UFD* efter engelskans *Unique Factorization Domain).*

Att man bortser från ”förekomster av associerade element” ska tolkas som att två faktoriseringar, som skiljer sig bara ifråga om irreducibla element som sinsemellan är associerade, räknas som *en* faktorisering. Detta innebär att de båda produkterna (med parvis associerade irreducibla element p_i och $\varepsilon_i p_i$)

$$z = p_1 p_2 ... p_n = (\varepsilon_1 p_1)(\varepsilon_2 p_2) \cdots (\varepsilon_n p_n), \tag{5.5}$$

där alla ε_i är enheter vars totala produkt är 1, ses som en och samma faktorisering av z. [2]

Definition 5.25 innebär alltså att ett godtyckligt element $z \neq 0$, som ej är en enhet, i en kvadratisk talring med entydig faktorisering kan skrivas som

$$z = \varepsilon p_1 p_2 ... p_n, \tag{5.6}$$

där p_i är irreducibla element i $\mathbf{D}$ och ε är en enhet. Mer generellt kan detta skrivas med produktsymbolen som

$$z = \varepsilon \prod_p p^{\alpha(p)}, \tag{5.7}$$

där ε är en enhet, p löper över alla *olika* irreducibla element i $\mathbf{Z}[\sqrt{m}]$ och (de icke-negativa) exponenterna $\alpha(p)$ anger antalet gånger respektive faktor förekommer i z. Med "olika" irreducibla element menas sådana som *inte* är associerade enligt (5.5). Lägg märke till att såväl enheten ε som exponenterna $\alpha(p)$ är entydigt bestämda av z eftersom vi har att göra med en UFD.

Definition 5.26. (se [3]) Den *största gemensamma delaren* (sgd) till heltalen $a_1, a_2, ..., a_n$ (minst ett skilt från 0) i en UFD är ett heltal d i UFD:n som är delare till samtliga a_i, sådant att om $d_1 \mid a_i$ för något d_1 i UFD:n så gäller att $d_1 \mid d$. Heltalet d är entydigt bestämt sånär som på multiplikation med en enhet. Heltalen a_i sägs vidare vara *relativt prima* om deras sgd är lika med en enhet.

Notera att det av definitionen följer att det i en UFD kan finnas fler än en största gemensamma delare till en uppsättning heltal och att dessa är sinsemellan associerade. Om exempelvis $1+\sqrt{-1}$ är en sgd i $\mathbf{D} = \mathbf{Z}[\sqrt{-1}]$ (som är en UFD) fås, enligt Sats 5.17, att även $1-\sqrt{-1}$, $-1+\sqrt{-1}$ och $-1-\sqrt{-1}$ är sgd:er. Definitionen säger också att det i en UFD *alltid* finns sgd:er till givna heltal.

I en UFD är i själva verket varje irreducibelt element ett primelement. Även omvändningen till Sats 5.23 gäller alltså i dessa fall, och vi har följande viktiga sats som vi ger utan bevis.

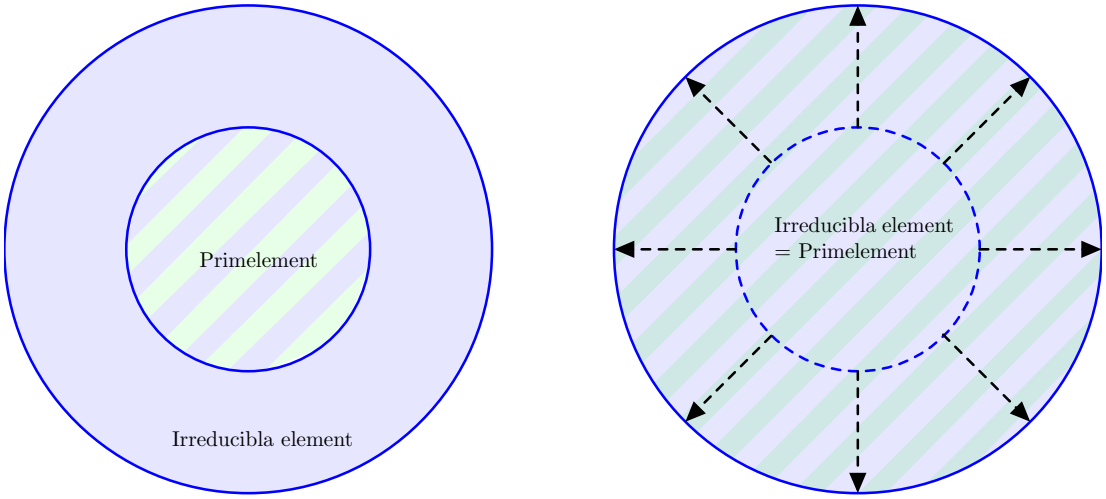
Sats 5.27. *I en UFD sammanfaller de irreducibla elementen med primelementen.*

I figur 1(a) på sidan 18 ses relationen mellan primelementen och de irreducibla elementen i en godtycklig talring enligt Sats 5.23. Sats 5.27 illustreras i figur 1(b).

Exempel 5.28. Ett enkelt exempel på en UFD är talringen $\mathbf{Z}$ där ju, som inledningsvis nämnts, irreducibla element och primelement är två olika namn på samma tal, nämligen primtalen. Vi ska nu undersöka andra typer av talringar med eventuellt samma egenskap som $\mathbf{Z}$, och ska som exempel titta närmare på den kvadratiska talringen $\mathbf{D} = \mathbf{Z}[\sqrt{-5}]$ med heltalen $z = a+b\sqrt{-5}$ där $a, b \in \mathbf{Z}$. Målet är alltså att avgöra huruvida den har en entydig faktorisering. Vi vet att det finns irreducibla element i denna talring och att varje element kan faktoriseras i sådana. Vidare känner vi också till begreppet norm, som i detta exempel är $N(z) = (a+b\sqrt{-5})(a-b\sqrt{-5}) = a^2+5b^2$. Att vi kan faktorisera z i icke-enheter innebär att vi, för $x_1, x_2, y_1, y_2 \in \mathbf{Z}$, kan skriva talet som

$$z = a + b\sqrt{-5} = (x_1 + y_1\sqrt{-5})(x_2 + y_2\sqrt{-5})$$

där normen av respektive faktor är större än 1. Här har vi uteslutit *triviala* faktoriseringar som $a + b\sqrt{-5} = (1)(a + b\sqrt{-5})$ och $a + b\sqrt{-5} = (-1)(-a - b\sqrt{-5})$, där faktorerna är någon av



(a) (b)

Figur 1: (a) I en talring är mängden av primelement en delmängd av mängden av irreducibla element. Varje primelement är alltså ett irreducibelt element. (b) I en UFD är mängden av irreducibla element lika med mängden av primelement. Ett element är alltså irreducibelt om och endast om det är ett primelement.

enheterna ± 1 enligt Sats 5.17 samt *associeringar* av z. Enligt Sats 5.16 är normen multiplikativ vilket innebär att

$$N(a+b\sqrt{-5}) = N((x_1+y_1\sqrt{-5})(x_2+y_2\sqrt{-5})) = N(x_1+y_1\sqrt{-5})N(x_2+y_2\sqrt{-5})$$

och alltså, eftersom respektive norm är större än 1,

$$1 < N(x_1+y_1\sqrt{-5}) < N(a+b\sqrt{-5}), 1 < N(x_2+y_2\sqrt{-5}) < N(a+b\sqrt{-5}).$$

Eftersom normen av $z = a+b\sqrt{-5}$ ges av $N(z) = a^2+5b^2$ ser vi att

$$N(z) \geq 5 \tag{5.8}$$

för icke-reella tal i $\mathbf{D} = \mathbf{Z}[\sqrt{-5}]$, det vill säga då $b \neq 0$.
Ett tal som inte kan faktoriseras enligt ovan kallas som nämnts ett *irreducibelt element* i $\mathbf{D}$. Talet 3 är till exempel ett irreducibelt element då talet inte kan faktoriseras enligt (5.28), eftersom en faktorisering i icke-enheter,

$$3 = (x_1+y_1\sqrt{-5})(x_2+y_2\sqrt{-5}), \tag{5.9}$$

efter normering skulle innebära att

$$9 = N(3) = N(x_1 + y_1\sqrt{-5})N(x_2 + y_2\sqrt{-5}),$$

som ger att de båda normerna i högerledet är lika med 3, vilket motsäger olikheten (5.8) som säger att faktorerna var och en ska vara minst lika med 5 för icke-reella tal (några reella lösningar finns inte eftersom 3 är ett primtal i $\mathbf{Z}$). Motsvarande argument ger att talet 2 också är ett irreducibelt element i denna kvadratiska talring.

För att visa att det finns heltal i $\mathbf{D} = \mathbf{Z}[\sqrt{-5}]$ för vilka det inte finns en entydig faktorisering i irreducibla element betraktar vi talet 6 som kan skrivas som

$$6 = 2 \cdot 3 = 1 + 5 = 1^2 + 1^2 \cdot 5 = N(1 + \sqrt{-5}) = (1 + \sqrt{-5})(1 - \sqrt{-5}), \tag{5.10}$$

alltså på *två olika sätt*, dels som $2 \cdot 3$ och dels som $(1+\sqrt{-5})(1-\sqrt{-5})$. Att 2 och 3 är irreducibla har vi redan visat. Det återstår alltså att visa att även $1 \pm \sqrt{-5}$ är irreducibla. Vi gör ansättningen

$$1 \pm \sqrt{-5} = (x_1 + y_1\sqrt{-5})(x_2 + y_2\sqrt{-5}), \tag{5.11}$$

vilket ger

$$6 = N(1 \pm \sqrt{-5}) = N(x_1 + y_1\sqrt{-5})N(x_2 + y_2\sqrt{-5}), \tag{5.12}$$

som inte har några lösningar i icke-enheter eftersom det inte finns heltal $x_1 + y_1\sqrt{-5}$ och $x_2 + y_2\sqrt{-5}$ vars norm (2 och 3) uppfyller (5.8).

Vi kan alltså konstatera att den kvadratiska talringen $\mathbf{D} = \mathbf{Z}[\sqrt{-5}]$ *inte* har en entydig faktorisering och är således *inte* en UFD. Vidare kan noteras att faktorerna 2 och 3, trots att de alltså är irreducibla, inte är primelement i $\mathbf{D}$, ty de delar varken $1 + \sqrt{-5}$ eller $1 - \sqrt{-5}$ trots att de båda enligt (5.10) delar produkten $(1 + \sqrt{-5})(1 - \sqrt{-5}) = 6$ (jämför med Definition 5.21).

5.12 Divisionsalgoritmen

Sats 5.29. *(Divisionsalgoritmen) Om $a, b \in \mathbf{Z}$ är positiva, så finns det entydigt bestämda $k, r \in \mathbf{Z}$ sådana att*

$$a = bk + r, \quad 0 \leq r < b. \tag{5.13}$$

Bevis. (se [4]) Vi visar först att (5.13) har åtminstone en lösning. Betrakta mängden H, säg, av heltal på formen $a - bu$ där u löper över alla heltal. Om vi väljer u sådant att

$$u = \begin{cases} -1, & a \geq 0 \\ a, & a < 0 \end{cases},$$

så blir talet $a - bu$ icke-negativt. Delmängden av H innehållande alla icke-negativa heltal har då enligt Sats 1.2 ett minsta element. Vi kallar detta tal för r, och tillhörande värde på u för k. Då gäller att

$$r = a - bk \geq 0, \quad r - b = a - bk - b = a - (k + 1)b < 0,$$

vilket ger att olikheterna i (5.13) är uppfyllda för detta val av k och r.

För att visa entydigheten, det vill säga att dessa heltal är de enda lösningarna, antar vi att (5.13) gäller även för andra val av heltal, säg k' och r', det vill säga

$$a = bk' + r', \quad 0 \leq r' < b.$$

Men eftersom det gäller, om $k' < k$, att

$$a - bk' = r' \geq a - (k-1)b = (a - bk) + b = r + b \geq b$$

medan vi, om $k' > k$, har att

$$a - bk' = r' \leq a - (k+1)b = (a - bk) - b = r - b < 0,$$

så följer att $k' = k$ måste gälla och och alltså även $r' = r$, vilket ger att valet av k och r är entydig. □

Notera att divisionsalgoritmen helt enkelt innebär en division av (*dividenden*) a med (*divisorn*) b som ger *kvoten* k och *resten* r, samt att $r = 0$ om och endast om $b \mid a$, det vill säga om divisionen "går jämnt upp".

Divisionsalgoritmen kan enkelt generaliseras till att gälla *alla* heltal $b \neq 0$, det vill säga även de icke-negativa. Detta görs genom att ändra olikheterna i satsen till $0 \leq r < |b|$. Satsen kan också generaliseras till att även gälla exempelvis vissa kvadratiska talringar (se avsnitt 5.13) och polynom. Ett exempel på en praktisk användning av divisionsalgoritmen i det förstnämnda fallet ges i avsnitt 5.15 då vi studerar den speciella kvadratiska talringen $\mathbf{Z}[i]$.

Det kan vara värt att poängtera att divisionsalgoritmen är en sats och, namnet till trots, inte en algoritm. Att den ändå benämns som en algoritm har att göra med att beviset ger en metod för att beräkna kvoten och resten utifrån givna värden på a och b. Denna beräkning, som praktiskt kan genomföras även på andra sätt, benämns *(Euklidisk) division.*

5.13 Euklides algoritm

Vi ska nu studera en algoritm som fått sitt namn efter den grekiske matematikern Euklides som levde på 300-200-talet f.Kr., mest känd för sin skrift Elementa bestående av tretton böcker där han sammanfattar den då kända geometrin i form av definitioner, satser och bevis, men där även denna algoritm finns beskriven. Men innan vi formulerar den centrala satsen i detta avsnitt ger vi följande lemma.

Lemma 5.30. *Antag att* $a, b \in \mathbf{Z}$ *och* $b > 0$. *Om* $a = bk + r$, *där* $0 \leq r < b$, *så är* $\text{sgd}(a, b) = \text{sgd}(b, r)$.

Bevis. (se [1]) Om heltalet d_1 delar a och b, så delar d_1 även b och r eftersom $r = a - bk$. Om heltalet d_2 omvänt delar b och r, så delar d_2 även a och b eftersom $a = bk + r$. De gemensamma delarna till a och b respektive b och r är alltså desamma, vilket speciellt ger att $\text{sgd}(a, b) = \text{sgd}(b, r)$. □

Sats 5.31. *(Euklides algoritm) Om vi för två givna positiva heltal $a, b \in \mathbf{Z}$ upprepar divisionsalgoritmen (Sats 5.29) så att vi får en följd av ekvationer på formen*

$$\begin{aligned}
a &= bk_1 + r_1, & 0 < r_1 < b,\\
b &= r_1k_2 + r_2, & 0 < r_2 < r_1,\\
r_1 &= r_2k_3 + r_3, & 0 < r_3 < r_2,\\
r_2 &= r_3k_4 + r_4, & 0 < r_4 < r_3,\\
&\vdots & \\
r_{j-2} &= r_{j-1}k_j + r_j, & 0 < r_j < r_{j-1},\\
r_{j-1} &= r_jk_{j-1} + r_{j+1}, & r_{j+1} = 0,
\end{aligned}$$

så är den sista nollskilda resten, r_j, lika med den största gemensamma delaren till a och b, det vill säga $\text{sgd}(a, b) = r_j$.

Bevis. (se [1]) Att talet r_j existerar, det vill säga att algoritmen ger en ändlig följd av ekvationer med nollskilda rester enligt satsen, inses genom att resterna bildar en avtagande följd av heltal $b > r_1 > r_2 > ... \geq 0$. Lemma 5.30 ger nu att

$$\begin{aligned}
\text{sgd}(a, b) = \text{sgd}(b, r_1) = \text{sgd}(r_1, r_2) = ... = \text{sgd}(r_{j-2}, r_{j-1}) = \text{sgd}(r_{j-1}, r_j)\\
= \text{sgd}(r_j, r_{j+1}) = \text{sgd}(r_j, 0) = r_j,
\end{aligned}$$

vilket visar att $\text{sgd}(a, b) = r_j$. □

Exempel 5.32. Vi ska bestämma $\text{sgd}(303, 36)$ med hjälp av Euklides algoritm. Vi sätter därför $a = 303$ och $b = 36$ och använder divisionsalgoritmen ett upprepat antal gånger vilket ger

$$\begin{aligned}
303 &= 36 \cdot 8 + 15, & 0 < 15 < 36,\\
36 &= 15 \cdot 2 + 6, & 0 < 6 < 15,\\
15 &= 6 \cdot 2 + 3, & 0 < 3 < 6,\\
6 &= 3 \cdot 2 + 0. &
\end{aligned}$$

Eftersom den sista nollskilda resten är 3 har vi enligt Sats 5.31 att $\text{sgd}(303, 36) = 3$.

I satsen förutsätts a och b vara positiva. I det fall a eller b i stället är negativt sätter vi $a := |a|$ respektive $b := |b|$ ty $\text{sgd}(a, b) = \text{sgd}(|a|, |b|)$ för alla $a, b \in \mathbf{Z}$. Lägg också märke till att Euklides algoritm avbryts så fort resten i divisionsalgoritmen för två heltal blir lika med 0, därav de strikta olikheterna i alla rader utom den sista.

Som synes är Euklides algoritm en enkel metod att bestämma den största gemensamma delaren till två heltal, och kan generaliseras till att gälla exempelvis även heltal i vissa kvadratiska talringar, så kallade *euklidiska ringar*. Ett nödvändigt (och tillräckligt) villkor för detta är att divisionsalgoritmen gäller, eftersom Euklides algoritm kräver att det i varje steg, utifrån två givna heltal, går att finna *två entydigt bestämda heltal* enligt just divisionsalgoritmen. Vi ska i nästa avsnitt titta närmare på denna speciella typ av kvadratiska talringar för vilka också gäller att heltalen, det vill säga elementen i talringen, har en entydig faktorisering.

5.14 Euklidiska ringar

Definition 5.33. (se [2]) En kvadratisk talring $\mathbf{D}$ sägs vara en *euklidisk ring* om heltalen i $\mathbf{D}$ uppfyller divisionsalgoritmen, alltså om det för $a, b \in \mathbf{D}$ där $b \neq 0$ gäller att det finns heltal $k, r \in \mathbf{D}$ sådana att $a = bk + r$ där $|N(r)| < |N(b)|$.

Här har alltså r och b i villkoret i (5.13) ersatts med $|N(r)|$ respektive $|N(b)|$. Vi säger att normen utgör den *euklidiska värderingen* på $\mathbf{D}$ och noterar detta som $f(z) = N(z)$. Notera att enligt Sats 5.31 och det efterföljande resonemanget är $\mathbf{D}$ en euklidisk ring om och endast om Euklides algoritm gäller i ringen.

Sats 5.34. *Varje euklidisk ring är en UFD.*

Bevis. (se [2]) Först visar vi att om α och β är två relativt prima heltal i $\mathbf{D}$, det vill säga om inga gemensamma faktorer förutom enheter finns, så finns det heltal λ_0 och μ_0 i $\mathbf{D}$ sådana att $\alpha\lambda_0 + \beta\mu_0 = 1$. Låt nu S beteckna mängden av heltal på formen $\alpha\lambda + \beta\mu$, där λ och μ antar alla heltal i $\mathbf{D}$. Normen av ett godtyckligt heltal i S är ett heltal, så vi väljer ett heltal $\varepsilon := \alpha\lambda_1 + \beta\mu_1$, för vilket gäller att beloppet av normen är mindre än beloppet av normen för varje annat heltal i S. Det gäller alltså att $|N(\varepsilon)| = \min |N(\alpha\lambda + \beta\mu)|$. Euklides algoritm ger nu att

$$\alpha = \varepsilon\gamma + \delta, \quad |N(\delta)| < |N(\varepsilon)|,$$

och alltså

$$\delta = \alpha - \varepsilon\gamma = \alpha - \gamma(\alpha\lambda_1 + \beta\mu_1) = \alpha(1 - \gamma\lambda_1) + \beta(-\gamma\mu_1).$$

vilket ger att $\delta \in S$. Detta kräver, enligt definitionen av ε, att $|N(\delta)| = 0$, och alltså att $\varepsilon = 0$ enligt Sats 5.16. Vi får nu att $\alpha = \varepsilon\gamma$, det vill säga $\varepsilon \,|\, \alpha$. På samma sätt visas att $\varepsilon \,|\, \beta$, vilket ger att ε är en enhet. Då är ε^{-1} också en enhet och vi får

$$1 = \varepsilon^{-1}\varepsilon = \varepsilon^{-1}(\alpha\lambda_1 + \beta\mu_1) = \alpha(\varepsilon^{-1}\lambda_1) + \beta(\varepsilon^{-1}\mu_1) = \alpha\lambda_0 + \beta\mu_0.$$

Vi ska nu visa att om p är ett irreducibelt element i $\mathbf{D}$ och $p \,|\, \alpha\beta$ så gäller att $p \,|\, \alpha$ eller $p \,|\, \beta$. Om vi antar att $p \nmid \alpha$ så har p och α inga gemensamma faktorer förutom enheter, och det finns därför, enligt Sats 1.6, heltal λ_0 och μ_0 för vilka gäller att $1 = p\lambda_0 + \alpha\mu_0$. Multiplikation med β ger nu att $\beta = p\beta\lambda_0 + \alpha\beta\mu_0$, och alltså att $p \,|\, \beta$ eftersom $p \,|\, \alpha\beta$ enligt förutsättning. Via induktion kan detta generaliseras till att konstatera att om $p \,|\, \alpha_1\alpha_2 \cdots \alpha_n$ så gäller att p delar minst en av produktens faktorer (jämför med Lemma 1.10). Den återstående delen av beviset är identiskt med beviset av Aritmetikens fundamentalsats (Sats 3.1). □

5.15 De gaussiska heltalen - ett exempel på en euklidisk ring

Vi introducerar nu en typ av heltal som fått sitt namn av den tyske matematikern och fysikern Carl Friedrich Gauss som var verksam på 1700- och 1800-talet. Han betraktas för övrigt som en av de främsta matematikerna någonsin och har blivit känd som ”matematikernas konung”.

Definition 5.35. (se [3]) Ett *gaussiskt heltal* är ett komplext tal på formen $z = a + bi$ där $a, b \in \mathbf{Z}$ och $i^2 = -1$. Mängden av alla gaussiska heltal, tillsammans med addition och multiplikation, bildar den kvadratiska talringen

$$\mathbf{Z}[i] = \{a + bi \,|\, a, b \in \mathbf{Z}\}. \tag{5.14}$$

Alltså är $\mathbf{D} = \mathbf{Z}[i] = \mathbf{Z}[\sqrt{-1}]$ den speciella kvadratiska talringen då $m = -1$ i Sats 5.12.

Att $\mathbf{Z}[i]$ är en ring inses genom att jämföra med Definition 4.7. Speciellt gäller att den är sluten med avseende på de binära operatorerna addition och multiplikation [4]. För $a + bi, c + di \in \mathbf{Z}[i]$ gäller nämligen att

$$(a + bi) + (c + di) = (a + c) + (b + d)i \in \mathbf{Z}[i],$$

respektive

$$(a + bi)(c + di) = ac + adi + bci + bdi^2 = (ac - bd) + (ad + bc)i \in \mathbf{Z}[i].$$

Vidare är de gaussiska heltalen som synes en delmängd av ringen av de komplexa talen, det vill säga $\mathbf{Z}[i] \subset \mathbf{C}$, och innehåller det multiplikativt neutrala elementet 1. Enligt avsnitt 4.2 bildar de gaussiska heltalen därmed en (äkta) delring av de komplexa talen.

I enlighet med Definition 5.13 är *normen* av ett gaussiskt heltal z lika med produkten av talet och dess konjugat, eller kvadraten på talets absolutbelopp, det vill säga

$$N(z) = |z|^2 = z\overline{z} = (a + bi)(a - bi) = a^2 + b^2. \tag{5.15}$$

Motsvarande egenskaper för normen som i det allmänna fallet gäller för de gaussiska heltalen, vilka vi formulerar i en sats.

Sats 5.36.
1. $N(z) \in \mathbf{Z}, N(z) \geq 0$ *för alla* $z \in \mathbf{Z}[i]$
2. $N(z) = 0$ *för något* $z \in \mathbf{Z}[i] \Leftrightarrow z = 0$
3. $N(z) = N(\overline{z})$ *för alla* $z \in \mathbf{Z}[i]$
4. $N(wz) = N(w)N(z)$ *för alla* $w, z \in \mathbf{Z}[i]$.
5. $N(z) = 1$ *för något* $z \in \mathbf{Z}[i] \Leftrightarrow z$ *är en enhet i* $\mathbf{Z}[i]$.

Bevis. (se [3])

1. Eftersom enligt (5.15) gäller att $N(z) = a^2 + b^2$, och kvadraten på ett heltal alltid är icke-negativt, följer att $N(z) \geq 0$ och är ett heltal för alla $z \in \mathbf{Z}[i]$.

2. ($\Leftarrow$) Insättning av $z = a + bi = 0$ i uttrycket för normen ger $N(z) = 0^2 + 0i = 0$.

   ($\Rightarrow$) Om $N(z) = 0$ så har vi att $a^2 + b^2 = 0$. De enda $a, b \in \mathbf{Z}$ som uppfyller ekvationen är $a = b = 0$, det vill säga $z = 0$.

3. För $z = a + bi$ gäller att $N(z) = a^2 + b^2 = a^2 + (-b)^2 = N(a - bi) = N(\overline{z})$.

4. Vi sätter $w = a + bi$ och $z = c + di$ vilket ger att $wz = (a + bi)(c + di)$ $= (ac - bd) + (ad + bc)i$. Då fås att

$$\begin{aligned} N(wz) &= N((ac - bd) + (ad + bc)i) \\ &= (ac - bd)^2 + (ad + bc)^2 \\ &= (ac)^2 - 2abcd + (bd)^2 + (ad)^2 + 2abcd + (bc)^2 \\ &= (ac)^2 + (ad)^2 + (bc)^2 + (bd)^2. \end{aligned}$$

   Å andra sidan gäller att

$$\begin{aligned} N(w)N(z) &= (a^2 + b^2)(c^2 + d^2) \\ &= a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2 \\ &= (ac)^2 + (ad)^2 + (bc)^2 + (bd)^2, \end{aligned}$$

   vilket innebär att $N(wz) = N(w)N(z)$ för alla $w, z \in \mathbf{Z}[i]$.

5. ($\Leftarrow$) Om $z = a + bi$ är en enhet så finns ett heltal ε sådant att $z\varepsilon = 1$. Detta ger $N(z\varepsilon) = N(z)N(\varepsilon) = N(1) = 1$, vilket innebär att $N(z) = 1$ eftersom $N(z)$ och $N(\varepsilon)$ enligt punkt 1 är heltal större än eller lika med 0.

   ($\Rightarrow$) Om $N(z) = 1$ där $z = a + bi$ så gäller enligt (5.15) att $z\overline{z} = 1$ och alltså att $z \mid 1$, vilket innebär att z är en enhet.

Därmed är satsen bevisad. □

Vi visar nu att de gaussiska heltalen, liksom de ”vanliga” heltalen $\mathbf{Z}$, bildar en euklidisk ring.

Sats 5.37. *Den kvadratiska talringen $\mathbf{Z}[i]$ är en euklidisk ring.*

Bevis. (se [4]) Sätt $\alpha = a + bi$ och $\beta = c + di$ och antag att $\beta \neq 0$. Antag vidare att $\frac{\alpha}{\beta} = r + si$, där $r, s \in \mathbf{Z}$. Vi väljer nu $m, n \in \mathbf{Z}$ sådana att $|r - m| \leq \frac{1}{2}$ och $|s - n| \leq \frac{1}{2}$. Om vi nu sätter $\delta = m + ni \in \mathbf{Z}[i]$ får vi $|\frac{\alpha}{\beta} - \delta| = (r - m)^2 + (s - n)^2 \leq \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$. Sätt $\rho = \alpha - \beta\delta \in \mathbf{Z}[i]$. Då gäller antingen att $\rho = 0$ eller att $|\rho| = \alpha - \beta\delta| = |\beta(\alpha/\beta - \delta)| = |\beta||\frac{\alpha}{\beta} - \delta| \leq \frac{1}{2}|\beta| < |\beta|$. Det följer nu att $\mathbf{Z}[i]$ är en euklidisk ring. □

Som vi sett i Sats 5.17 ges mängden av alla enheter i $\mathbf{Z}[i]$ av $\{\pm 1, \pm i\}$.

Definition 5.38. Det gaussiska heltalet z sägs vara ett *irreducibelt element* i $\mathbf{Z}[i]$ om det enda sättet att faktorisera z i gaussiska heltal är i enheter och z självt. I annat fall sägs z vara *reducibelt*.

Observera analogin mellan Definition 5.19 och Definition 5.38.

Exempel 5.39. Exempelvis är talet $1+i$ irreducibelt i $\mathbf{Z}[i]$. Antag nämligen att talet är reducibelt. Då skulle gälla att $1+i = wz$ för några tal $w, z \in \mathbf{Z}[i]$ som inte är enheter. Eftersom $N(1+i) = 2$ gäller då att $N(wz) = N(w)N(z) = 2$ vilket betyder att antingen $N(w) = 1$ eller $N(z) = 1$ då 2 är ett primtal. Vi har därmed en motsägelse och $1 + i$ är alltså irreducibelt.

Exempel 5.40. Eftersom $2 = (1+i)(1-i)$ och faktorerna inte är enheter så är 2 ett exempel på ett reducibelt gaussiskt heltal. Detta visar att ett tal som är ett primtal i $\mathbf{Z}$ mycket väl kan vara reducibelt i $\mathbf{Z}[i]$.

Sats 5.41. *Antag att $z \in \mathbf{Z}[i]$ och att $p \in \mathbf{Z}$ är ett primtal. Om $N(z) = p$ så är z ett irreducibelt element i $\mathbf{Z}[i]$.*

Bevis. (se [3]) Antag motsatsen, det vill säga att $N(z) = p$ för något primtal p och att z inte är irreducibelt. Då finns $v, w \in \mathbf{Z}[i]$, av vilka ingen är en enhet eller 0, så att $z = vw$. Detta ger att $p = N(z) = N(vw) = N(v)N(w)$, det vill säga $N(v) = 1$ eller $N(w) = 1$. Antingen v eller w är alltså en enhet vilket motsäger antagandet. Därmed är z irreducibelt och satsen är bevisad. □

Vi har till exempel att $N(3+2i) = (3+2i)(3-2i) = 3^2 + 2^2 = 13$ som är ett primtal. Alltså är $3+2i$ ett irreducibelt element i $\mathbf{Z}[i]$.

Eftersom de gaussiska heltalen är en UFD gäller enligt Sats 5.27 följande sats.

Sats 5.42. *Ett element $z \in \mathbf{Z}[i]$ är irreducibelt om och endast om det är ett primelement i $\mathbf{Z}[i]$. Vi säger då att z är ett gaussiskt primtal.*

Jämför med terminologin för heltalen i $\mathbf{Z}$ (som ju också är en UFD) där de irreducibla elementen, som sammanfaller med primelementen, som bekant benämns primtal.

Ovanstående är dock bara ett tillräckligt villkor för att z ska vara irreducibelt och därmed ett gaussiskt primtal. Det är alltså inget nödvändigt villkor vilket följande exempel visar. Vi har att $N(3) = 9$ som *inte* är ett primtal. Om vi antar att $3 = (a+bi)(c+di)$, där faktorerna ej är enheter, så gäller alltså att

$$9 = N(3) = N((a+bi)(c+di)) = N(a+bi)N(c+di) = (a^2+b^2)(c^2+d^2). \tag{5.16}$$

Faktorerna måste alltså båda vara lika med 3, men eftersom 3 inte kan skrivas som summan av två heltalskvadrater så är 3 irreducibelt och alltså ett gaussiskt primtal, trots att normen alltså inte är ett primtal.

Exemplet säger något om vilka heltal som är gaussiska primtal. Nedan ges de precisa villkoren för att ett heltal ska vara ett irreducibelt element och alltså ett gaussiskt primtal i $\mathbf{Z}[i]$.

Sats 5.43. *Ett gaussiskt heltal $a+bi$ är ett gaussiskt primtal om och endast om antingen*
1) $a = 0$ eller $b = 0$ och absolutbeloppet av det nollskilda talet är ett primtal på formen $4n+3$, där $n \geq 0$, eller
2) $a, b \neq 0$ och $N(a+bi) = a^2 + b^2 = p$, där p är ett primtal på formen $4n+1$, där $n \geq 0$.

Varje gaussiskt primtal z har åtta olika delare - enheterna ± 1 och $\pm i$ samt de fyra associeringarna av z, det vill säga $\pm z$ och $\pm iz$.

Exempel 5.44. Här följer ett exempel på hur division i $\mathbf{Z}[i]$ kan utföras med hjälp av den tidigare beskrivna divisionsalgoritmen (Sats 5.29).

Vi vill beräkna $\frac{a}{b}$ där $a = 8 + 9i$ och $b = 4 - i$. Vi får

$$\frac{a}{b} = \frac{8+9i}{4-i} = \frac{(8+9i)(4+i)}{(4-i)(4+i)} = \frac{23+44i}{17} = \frac{23}{17} + \frac{44}{17}i.$$

Heltalen närmast $\frac{23}{17}$ och $\frac{44}{17}$ är 1 respektive 3, varför vi sätter kvoten till $k = 1 + 3i$. Resten r blir då

$$r = a - bk = (8+9i) - (4-i)(1+3i) = 1 + 2i.$$

Alltså är $k = 1 + 3i$ och $r = 1 - 2i$ en kvot respektive rest vid divisionen. Här är $N(r) = 5$ och $N(b) = 17$, det vill säga $N(r) < N(b)$ enligt Definition 5.33.

Euklides algoritm är användbar inte bara för heltalen $\mathbf{Z}$ utan även för Euklidiska ringar såsom den kvadratiska talringen $\mathbf{Z}[i]$, alltså de gaussiska heltalen. Detta eftersom divisionsalgoritmen gäller i dessa ringar. Följande exempel får illustrera hur algoritmen kan användas i $\mathbf{Z}[i]$.

Exempel 5.45. Vi vill beräkna alla sgd:er till de gaussiska heltalen $a = 3 - 11i$ och $b = 7 - i$. Vi har att

$$\frac{a}{b} = \frac{3-11i}{7-i} = \frac{(3-11i)(7+i)}{(7-i)(7+i)} = \frac{16-37i}{25} = \frac{16}{25} - \frac{37}{25}i,$$

vilket, efter avrundning, ger kvoten $k_1 = 1 - i$ som i sin tur ger resten

$$r_1 = a - bk_1 = (3-11i) - (7-i)(1-i) = -3 - 3i.$$

I nästa steg har vi på motsvarande sätt att

$$\frac{b}{r_1} = \frac{7-i}{-3-3i} = \frac{(7-i)(-3+3i)}{(-3-3i)(-3+3i)} = \frac{-18+24i}{18} = -1 + \frac{4}{3}i,$$

vilket, efter avrundning, ger kvoten $k_2 = -1 + i$ och därmed resten

$$r_2 = b - r_1k_2 = (7-i) - (-3-3i)(-1+i) = 1 - i.$$

I tredje och sista steget fås

$$\frac{r_1}{r_2} = \frac{-3-3i}{1-i} = -3i,$$

vilket innebär att divisionen går jämnt upp, det vill säga $k_3 = -3i$ och $r_3 = 0$. Vi kan alltså konstatera att *en* största gemensamma delare till a och b är den sista nollskilda resten $r_2 = 1 - i$. Övriga sgd:er ges av de tre associerade heltalen $-r_2 = -1 + i, ir_2 = 1 + i$ och $-ir_2 = -1 - i$.

5.16 Principalidealringar

Definition 5.46. (se [3]) Antag att R är en ring. En delring I till R sägs vara ett *ideal* i R om och endast om $ar, ra \in I$ för alla $a \in I$ och $r \in R$. Om $I \neq R$ så sägs I vara ett *äkta ideal*.

Ett ideal I i ringen R är alltså, enkelt uttryckt, en ”ring i ringen” där man vid multiplikation av ett godtyckligt element i I med ett godtyckligt element i R ”stannar kvar” i idealet. För kommutativa ringar förenklas användandet av definitionen något, eftersom det för en given ring

R och en given delring I räcker att visa att *antingen* $ar \in I$ *eller* $ra \in I$ för att kunna säga att I verkligen är ett ideal i R.

Varje ring $R \neq \{0\}$ har minst två ideal - det så kallade *triviala idealet* $\{0\}$ och R självt.

Definition 5.47. (se [3]) Ett *principalideal* $\langle a \rangle$ i en ring är ett ideal som genereras av ett enda element i ringen, det vill säga kan skrivas på formen

$$\langle a \rangle = \{ar \,|\, r \in R\}$$

för något $a \in R$. Elementet a kallas för *generatorn* till $\langle a \rangle$.

Definition 5.48. En *principalidealring* (eller *PID* efter engelskans *Principal Ideal Domain*) är en ring där varje ideal är ett principalideal.

Exempel 5.49. Talringen $\mathbf{Z}$ är en principalidealring, eftersom varje ideal $n\mathbf{Z} = \{nz \,|\, z \in \mathbf{Z}\}$, där n är ett heltal, enligt Definition (5.47) är ett principalideal med generatorn n.

Vi ska nu se att ett annat typexempel på PID:er faktiskt är de i avsnitt 5.14 beskrivna euklidiska ringarna.

Sats 5.50. *Varje euklidisk ring är en PID.*

Bevis. (se [3]) Antag att R är en euklidisk ring, att f är en euklidisk värdering på R och att I är ett ideal i R. Vi ska nu visa att $I = \langle b \rangle$ för något $b \in R$. Det triviala idealet $I = \langle 0 \rangle$ är ett huvudideal. Vi antar därför att $I \neq \{0\}$ alltså att I innehåller minst ett nollskilt element. Av elementen i I väljer vi b så att $f(b)$ är minimalt.

Välj nu godtyckligt ett element $a \in I$. Om vi kan visa att $a = bk$ för något $k \in R$ så är vi klara. Eftersom R är en euklidisk ring så finns $k, r \in R$ sådana att

$$a = bk + r, \tag{5.17}$$

där antingen $r = 0$ eller $f(r) < f(b)$. Eftersom $a, b \in I$ och I är ett ideal så följer av (5.17) att

$$r = a - bk,$$

och alltså att $r \in I$. Olikheten $f(r) < f(b)$ kan då inte gälla, eftersom vi antog att $f(b)$ är minimalt. Vi har alltså en motsägelse och därmed gäller att $r = 0$, vilket ger $a = bk$ och beviset är klart. □

Observera att omvändningen till satsen *inte* gäller: en PID är inte nödvändigtvis en euklidisk ring. Principalidealringen $\mathbf{D} = \mathbf{Z}[\frac{1+\sqrt{-19}}{2}]$ kan i någon mening sägas vara det ”enklaste”exemplet på en sådan PID, övriga ges av de tre sistnämnda i Sats 5.53.

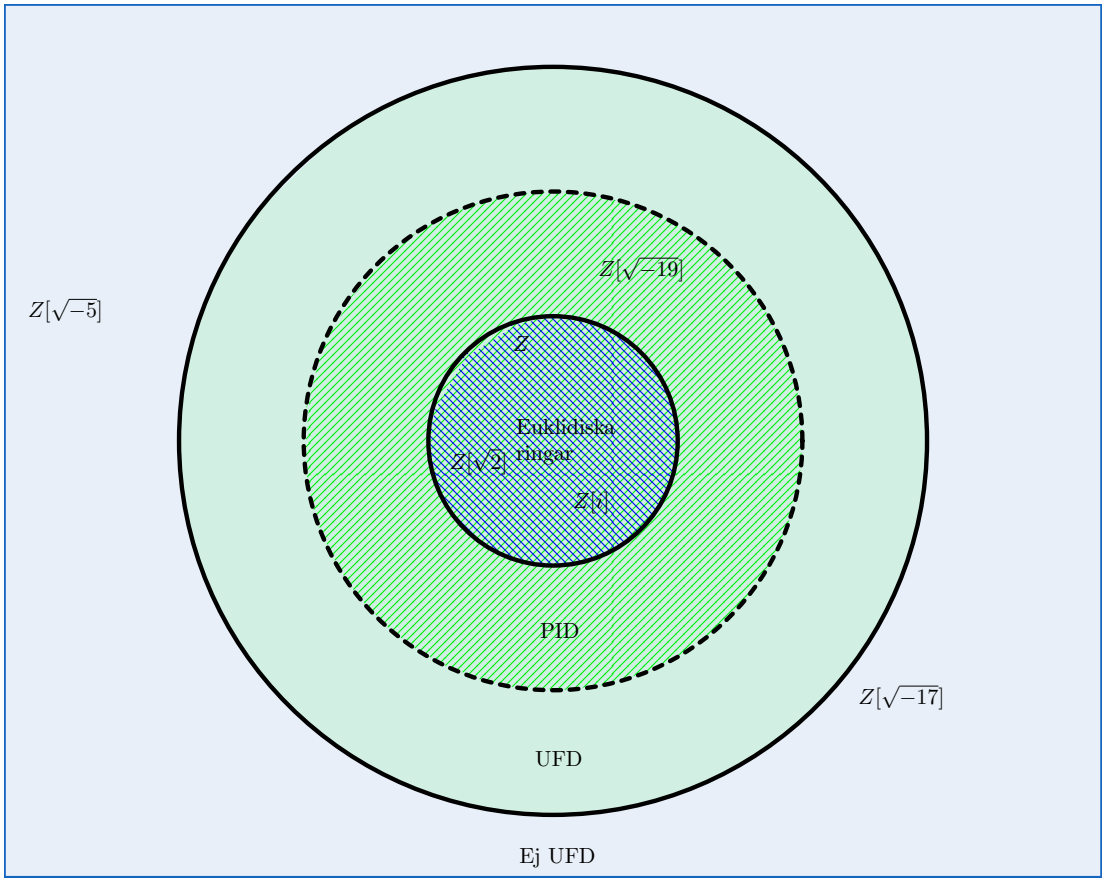
En PID har flera olika egenskaper som liknar de för heltalen $\mathbf{Z}$, till exempel att varje element i en PID entydigt kan faktoriseras i primelement samt att det alltid finns en största gemensam delare till två godtyckliga element i PID:en. Av detta kan man gissa sig till att det finns en nära koppling mellan PID:er och UFD:er, vilket också framgår av följande sats.

Sats 5.51. *Varje PID är en UFD.*

Beviset för Sats 5.51 är mer komplicerat än det för Sats 5.50 då det kräver mer ingående teori om ideal och tas därför inte upp här. För den intresserade kan dock hänvisas till [3, s. 410]. Vidare gäller speciellt för kvadratiska heltalsringar $\mathbf{D}$ även omvändningen till satsen, varav följer att $\mathbf{D}$ är en PID om och endast om den är en UFD.

Av Sats 5.50 och Sats 5.51 följer också att varje euklidisk ring är en UFD, helt i enlighet med Sats 5.34.

En sammanfattning av hur de olika typerna av ringar är relaterade till varandra ges i figur 2.



Figur 2: Relationen mellan ringar utan entydig faktorisering, med entydig faktorisering (UFD), principalidealringar (PID) och euklidiska ringar i form av ett Venndiagram samt exempel på talringar tillhörande dessa. Observera den streckade cirkeln, som anger att det för kvadratiska talringar råder ekvivalens mellan PID och UFD.

5.17 Något om klasstal

Vi presenterar nu kort ett begrepp som har en nära koppling till UFD:er och som används till exempel för att kunna visa vilka kvadratiska talringar som har entydig faktorisering.

Definition 5.52. *Klasstalet* för en talkropp K är ordningen av, det vill säga antalet element i, idealklassgruppen för K.

För en definition av begreppet *idealklassgrupp*, se till exempel [4, s. 177].

Värt att nämna är att klasstalet för K då K är en talring alltid är ändligt, vilket inom den algebraiska talteorin är ett viktigt faktum.

Enkelt uttryckt är klasstalet ett mått på hur mycket talringen tillhörande K avviker från att vara en PID, och därmed från att vara en UFD och alltså ha entydig faktorisering.

En talring har klasstalet 1 om och endast om den är en UFD eller, enligt Sats 5.51, en PID. Vidare är klasstalet för exempelvis $\mathbf{D} = \mathbf{Z}[\sqrt{-5}]$ lika med 2, eftersom idealklassgruppen är av ordning 2. Detta innebär alltså att denna kvadratiska talring *inte* är en UFD, vilket vi tidigare sett och också konstaterar i nästa avsnitt.

5.18 Vilka kvadratiska talringar har entydig faktorisering?

Som tidigare beskrivits är det bara *vissa* kvadratiska talringar som är UFD, för vilka det alltså gäller att varje heltal i ringen på ett entydigt sätt kan skrivas som en produkt av en enhet och primelement, om man bortser från faktorernas ordning och förekomster av associerade element. Eftersom teorin för detta faller utanför ramen för denna text rundar vi av detta kapitel med att nämna för vilka imaginära kvadratiska talringar denna egenskap gäller. Vi formulerar detta i en sats, vars fullständiga bevis är alltför komplicerat för att tas upp här.

Sats 5.53. *De enda imaginära kvadratiska talringar* $\mathbf{D} = \mathbf{Z}[\sqrt{m}]$ *respektive* $\mathbf{D} = \mathbf{Z}[\frac{1+\sqrt{m}}{2}]$ *som är principalidealdomäner, och alltså har entydig faktorisering, är de för vilka gäller att*

$$m = -1, -2, -3 - 7, -11, -19, -43, -67, -163.$$

Bevis. (se [2]) Vi visar här att $\mathbf{D} = \mathbf{Z}[\sqrt{-2}]$ har entydig faktorisering. För bevis av detta för några andra av m -värdena, se [2, s. 431].

Antag att $\alpha, \beta \in \mathbf{D}$ och att $\beta \neq 0$, vilket ger att $\frac{\alpha}{\beta} = u + v\sqrt{-2}$, där $u, v \in \mathbf{Z}$. Vi väljer nu heltal x och y närmast u och v, så att

$$0 \leq |u - x| \leq \frac{1}{2}, \quad 0 \leq |v - y| \leq \frac{1}{2}. \tag{5.18}$$

Sätt nu $\gamma := x + y\sqrt{-2}$ och $\delta := \alpha - \beta\gamma$ som då blir heltal i $\mathbf{D}$. Vi får nu att normen av δ är

$$\begin{aligned} N(\delta) = N(\alpha - \beta\gamma) = N(\beta)N\Big(\frac{\alpha}{\beta} - \gamma\Big) &= N(\beta)N((u - x) + (v - y)\sqrt{-2}) \\ &= N(\beta)((u - x)^2 + 2(v - y)^2), \end{aligned}$$

som efter normering av båda leden ger

$$|N(\delta)| = |N(\beta)||(u-x)^2 + 2(v-y)^2|. \tag{5.19}$$

Olikheterna i (5.18) ger nu att

$$0 \leq (u-x)^2 + 2(v-y)^2 \leq \frac{1}{4} + \frac{2}{4} = \frac{3}{4}.$$

Det sista beloppet i (5.19) är alltså mindre än 1, varav fås att $|N(\delta)| < |N(\beta)|$ och därmed att $\mathbf{D} = \mathbf{Z}[\sqrt{-2}]$ är en euklidisk ring och alltså enligt Sats 5.34 har entydig faktorisering. □

Notera att exakt de imaginära kvadratiska talringar som anges i sats 5.53, enligt diskussionen i avsnitt 5.16, har klasstal 1. Detta förmodades av den berömde tyske matematikern Carl Friedrich Gauss på 1800-talet och bevisades på 1960-talet av bland andra den amerikanske matematikern Harold Stark [4, s. 192].

Vi har nu nått vårt tidigare formulerade mål med detta kapitel och ger nu alltså en generaliserad version av Aritmetikens fundamentalsats som gäller för varje imaginär kvadratisk talring $\mathbf{D}$.

Korollarium 5.54. *(Generaliserad version av Aritmetikens fundamentalsats)*
Varje heltal $z \neq 0$, som ej är en enhet, i den imaginära kvadratiska talringen $\mathbf{D} = \mathbf{Z}[\sqrt{m}]$ respektive $\mathbf{D} = \mathbf{Z}[\frac{1+\sqrt{m}}{2}]$ har en entydig faktorisering i irreducibla element (då man bortser från faktorernas ordningsföljd och förekomster av associerade element) om och endast om $m \in \{-1, -2, -3, -7, -11, -19, -43, -67, -163\}$.

Bevis. Sats 5.53 tillsammans med Definition 5.25 ger korollariet. □

Slutsatsen blir alltså att det snarare är undantag än regel att en godtycklig imaginär kvadratisk talring har entydig faktorisering, i motsats till ringen av ”vanliga” heltal där faktoriseringen av ett heltal $n > 1$ *alltid* är entydig.

Vad gäller då för *reella* kvadratiska talringar i fråga om vilka som har entydig faktorisering? Till skillnad från det imaginära fallet känner man för närvarande till ett stort antal sådana talringar med klasstalet 1, och som alltså är UFD:er, men det är ännu inte känt om de är oändligt många. De tio första är $\mathbf{D} = \mathbf{Z}[\sqrt{m}]$ respektive $\mathbf{D} = \mathbf{Z}[\frac{1+\sqrt{m}}{2}]$ för $m = 2, 3, 5, 6, 7, 11, 13, 14, 17, 19$.

6 Den diofantiska ekvationen $x^3 + y^3 = z^3$

Vi inleder detta kapitel med ett begrepp som fått sitt namn efter den grekiske matematikern Diofantos verksam i Alexandria på 200-talet f.Kr. Han är känd för skriften Arithmetika bestående av tretton böcker där bland annat detta begrepp studeras.

Definition 6.1. En ekvation där endast heltalslösningar söks kallas en *diofantisk ekvation.*

Vilken ekvation som helst kan alltså sägas vara diofantisk, förutsatt att endast lösningar i form av heltal tillåts. Vi ska i detta kapitel studera en speciell diofantisk ekvation, med tre variabler och av grad 3. Men innan dess formulerar vi den mer generella, och mycket kända, diofantiska ekvationen av godtycklig grad. Men först en inledande historisk bakgrund.

6.1 Fermats stora sats

Den franske domaren Pierre de Fermat, som också verkade som matematiker, formulerade 1637 en sats som senare blev uppkallad efter honom - Fermats stora sats. Det var dock inte satsen i sig som i första hand blev förknippat med honom, utan det faktum att han påstod sig ha ett *bevis* för denna utan att, vad man hittills vet, ha publicerat det någonstans. I en anteckning skriver han nämligen, i anslutning till satsen, att han upptäckt "ett i sanning underbart bevis för detta påstående, men marginalen är alltför smal för att rymma detta"[2]. Denna formulering har alltsedan texten uppmärksammades inspirerat matematiker världen över att försöka komma fram till ett bevis för satsen.

Det skulle dröja ända till 1995, det vill säga mer än 350 år, innan det presenterades ett sådant bevis. Den som lyckades med detta var den engelske matematikern och professorn Andrew Wiles, vid denna tid verksam vid amerikanska Princeton University. Det slutgiltiga beviset, som är både väldigt avancerat och omfattande, hade föregåtts av många års intensivt arbete och resulterade för Wiles i ett stort antal matematiska och vetenskapliga priser och utmärkelser från olika delar av världen. Beviset hör till de mest kända i matematiken och det finns för den intresserade mycket att läsa om historien bakom lösandet av denna oerhört svårbevisade sats (se till exempel [6]). Huruvida Fermat verkligen hade funnit ett bevis för satsen redan på 1600-talet är det ingen som säkert vet, men det finns inget som tyder på detta då Wiles bevis innehåller mycket matematik som inte var utvecklad vid denna tid.

Sats 6.2. *(Fermats stora sats) Den diofantiska ekvationen*

$$x^n + y^n = z^n \tag{6.1}$$

har inga lösningar $x, y, z \neq 0$ för heltal $n > 2$.

6.2 Eulers bevis för Fermats stora sats i fallet $n = 3$

Även om det under flera sekler inte publicerades något bevis för Fermats stora sats så har arbetet på vägen dit kantats av ett stort antal matematiska upptäckter och till och med nya teoretiska forskningsfält inom matematiken, varav kan nämnas algebraisk talteori och ringteori [7]. Detta eftersom satsen inspirerat matematiker och andra till att studera och bevisa den för olika givna

[2] Översatt från https://en.wikiquote.org/wiki/Pierre_de_Fermat

exponenter n i (6.1), det vill säga för ett slags "delproblem" av Sats 6.2. Detta har ofta lyckats, och visat sig vara en betydligt enklare uppgift i jämförelse med att ta sig an den ursprungliga satsen. Det kan här nämnas att Fermats stora sats, bortsett från fallet $n = 4$, endast behöver visas i fallen då n är ett primtal, eftersom alla fall då exponenten antar värdet av någon multipel kn av n då också är bevisade. Detta följer av att

$$x^{kn} + y^{kn} = z^{kn} \iff (x^k)^n + (y^k)^n = (z^k)^n,$$

där den vänstra diofantiska ekvationen, med ett sammansatt tal som exponent, saknar icke-triviala lösningar givet att det visats att ekvationen saknar sådana för primtalet n. Detta inses genom att betrakta den högra ekvivalenta ekvationen.

Vi ska här gå igenom ett bevis för Fermats stora sats då $n = 3$, formulerad av den schweiziske matematikern Leonhard Euler på 1700-talet [8]. Satsen säger alltså i detta fall att den diofantiska ekvationen

$$x^3 + y^3 = z^3 \tag{6.2}$$

inte har några lösningar $x, y, z \neq 0$.

Beviset bygger på det faktum att om ekvationen har en lösning i form av positiva heltal x, y, z så finns det tre mindre positiva heltal som *också* är en lösning till ekvationen[3]. Då sådana lösningar inte kan genereras oändligt många gånger kan slutsatsen dras att det överhuvudtaget inte finns några lösningar till ekvationen. Denna form av bevis kallas på engelska *infinite descent*, infördes av Fermat och är nödvändigt för att kunna bevisa satsen med Eulers metod.

Vidare är Eulers bevis relativt långt och använder sig av ett flertal slutsatser som kan härledas till delbarhet och så kallad *paritet*, det vill säga egenskapen att ett heltal är antingen udda eller jämnt. Grundläggande kunskaper om dessa begrepp bör därför vara tillräckligt för att relativt väl hänga med i beviset, åtminstone till en början. För att förstå det fullt ut har man dock nytta av algebraisk talteori, och då speciellt Aritmetikens fundamentalsats och det faktum att den kvadratiska talringen $\mathbf{D} = \mathbf{Z}[\frac{1+\sqrt{-3}}{2}]$ har entydig faktorisering. Med detta sagt, låt oss börja från början.

Bevis. Antag motsatsen, det vill säga att den diofantiska ekvationen (6.2) har en lösning $x, y, z > 0$. Vi kan först konstatera att en eventuell gemensam faktor till två av variablerna x, y, z delar även den tredje. Detta tal kan därför förkortas bort och x, y, z är därmed *parvis relativt prima*, det vill säga $\text{sgd}(x, y) = \text{sgd}(x, z) = \text{sgd}(y, z) = 1$. Det inses nu lätt att det finns *högst ett* jämnt tal (två jämna tal ger att även det tredje är jämnt vilket innebär att de inte är relativt prima - en motsägelse), samtidigt som det finns minst ett jämnt tal (två udda tal ger att det tredje är jämnt). Alltså är *exakt ett* tal jämnt och de båda andra udda.

Antag först att det enda jämna talet är z. Då är $x + y$ och $x - y$ båda jämna (summan respektive differensen av två udda tal är jämn) och vi sätter $x + y = 2p$ och $x - y = 2q$ för heltal p, q, vilket ger $x = p + q$ och $y = p - q$. Vi bryter ut $x + y$ och får, efter insättning och förenkling,

$$\begin{aligned} z^3 &= x^3 + y^3 \\ &= (x + y)(x^2 - xy + y^2) \\ &= (p + q + p - q)((p + q)^2 - (p + q)(p - q) + (p - q)^2) \\ &= 2p(p^2 + 2pq + q^2 - p^2 + q^2 + p^2 - 2pq + q^2) \\ &= 2p(p^2 + 3q^2), \end{aligned}$$

[3] Om x, y eller z är *negativa* kan de tre motsvarande positiva lösningarna tas fram (se s. 33). Det räcker därför att studera heltal $x, y, z > 0$.

som alltså är en kub. En del kan nu sägas om p, q, nämligen att dessa

1. är av olika paritet, eftersom $x = p + q$ och $y = p - q$ är udda enligt antagandet,
2. är relativt prima, eftersom x och y är relativt prima, och
3. kan antas vara positiva. (Om $x < y$ byter vi plats på x, y så att $q > 0$. Fallet $x = y$ är omöjligt eftersom då följer att $q = 0$ vilket ger $x = y = p = 1$ enligt punkt 2, och därför $z^3 = 2$ vilken saknar lösning då z är ett heltal.)

Sammanfattningsvis ger antagandet alltså att (6.2) har en lösning med z jämnt och x, y udda, att det finns relativt prima positiva heltal p, q av olika paritet sådana att $2p(p^2 + 3q^2)$ är en kub [8, s. 40].

Antag nu i stället att det enda jämna talet är x eller y, säg x. Genom att flytta det udda talet y^3 till höger led i (6.2) fås

$$x^3 = z^3 - y^3 = (z - y)(z^2 + yz + y^2). \tag{6.3}$$

Sätt nu $z - y = 2p$ och $z + y = 2q$ vilket ger $z = q + p$ och $y = q - p$, som insatt i (6.3) ger

$$x^3 = 2p((q + p)^2 + (q + p)(q - p) + (q - p)^2) = 2p(p^2 + 3q^2). \tag{6.4}$$

Ekvation (6.4) ger nu alltså att talet $2p(p^2 + 3q^2)$ även i detta fall är en kub, med samma villkor på p och q som ovan [8, s. 41].

Vidare är ett villkor för att $2p(p^2 + 3q^2)$ är en kub, att produktens båda faktorer - om de är relativt prima - var och en är en kub. Detta villkor, som följer av Aritmetikens fundamentalsats (Sats 3.1), betecknar vi i fortsättningen med (*).

Att faktorerna är relativt prima är dock inte självklart. Att p och q har olika paritet ger att $p^2 + 3q^2$ är udda och varje gemensam delare till $2p$ och $p^2 + 3q^2$ är en gemensam delare till p och $p^2 + 3q^2$ och därmed också till p och $3q^2$. Den enda möjliga gemensamma delaren till dessa är 3 (eftersom p och q är relativt prima). Om 3 delar p inses lätt att 3 även delar $2p$ och $p^2 + 3q^2$ som alltså inte är relativt prima. Vi delar nu upp resonemanget i två fall, det första då 3 *inte* är en delare till p och det andra då 3 *är* en delare till p. För båda fallen ska visas att ”infinite descent” gäller.

Fall 1: Här gäller att $3 \nmid p$, vilket innebär att $2p$ och $p^2 + 3q^2$ är relativt prima och därför är båda kuber enligt (*). Vi vill nu hitta kuber på formen $p^2 + 3q^2$. Formeln

$$(a^2 + 3b^2)(c^2 + 3d^2) = (ac - 3bd)^2 + 3(ad + bc)^2$$

kan användas för detta ändamål [8, s. 41]. Vi får då

$$\begin{aligned}(a^2 + 3b^2)^3 &= (a^2 + 3b^2)((a^2 - 3b^2)^2 + 3(2ab)^2) \\ &= (a(a^2 - 3b^2) - 3b(2ab))^2 + 3(a(2ab) + b(a^2 - 3b^2))^2 \\ &= (a^3 - 9ab^2)^2 + 3(3a^2b - 3b^3)^2.\end{aligned}$$

Om vi nu för godtyckliga värden på $a, b \in \mathbf{Z}$ sätter

$$p = a^3 - 9ab^2, \quad q = 3a^2b - 3b^3, \tag{6.5}$$

så fås alltså att $(a^2 + 3b^2)^3$ är en kub på formen $p^2 + 3q^2$.

I själva verket kan *samtliga* kuber på denna form fås genom lämpliga värden på a och b i (6.5). Vi har alltså att $p^2 + 3q^2$ är en kub *om och endast om* (6.5) gäller för något a och b. Det är alltså fråga om en ekvivalens och inte enbart en implikation. Detta var något som Euler inte inkluderade i beviset, vilket vi återkommer till och studerar närmare i nästa avsnitt. Vi nöjer oss så länge med att bara konstatera att detta samband är sant för att kunna gå vidare i beviset.

Vi ser att p och q kan faktoriseras som

$$p = a(a^2 - 9b^2) = a(a - 3b)(a + 3b), \quad q = 3b(a^2 - b^2) = 3b(a - b)(a + b).$$

Någon gemensam delare till a och b kan inte finnas eftersom denna då även skulle vara en gemensam delare till såväl p som q, vilket strider mot antagandet att p och q är relativt prima. Alltså är a, b relativt prima. Enligt (*) har vi också att

$$2p = 2a(a - 3b)(a + 3b) \tag{6.6}$$

är en kub [8, s. 42].

Eftersom vi vet att p och q har olika paritet följer att också a och b har olika paritet (ty om a och b har samma paritet så skulle p och q båda vara jämna). Alltså är $a \pm 3b$ udda och varje gemensam delare till $2a$ och $a \pm 3b$ är en gemensam delare till a och $a \pm 3b$ och därmed till a och $\pm 3b$. Varje gemensam delare till $a + 3b$ och $a - 3b$ är på motsvarande sätt en gemensam delare till a och $3b$. Vi har alltså att

4. $c_1 \mid 2a$ och $c_1 \mid (a \pm 3b) \Longrightarrow c_1 \mid a$ och $c_1 \mid \pm 3b$, samt

5. $c_2 \mid (a + 3b)$ och $c_2 \mid (a - 3b) \Longrightarrow c_2 \mid a$ och $c_2 \mid 3b$,

för $c_1, c_2 \in \mathbf{Z}$ och a, b enligt ovan. Punkt 4 och punkt 5 ger nu att 3 är den enda möjliga gemensamma delaren till $2a$ och $a \pm 3b$. Men $3 \nmid a$ eftersom annars skulle gälla att $3 \mid p$ vilket strider mot antagandet. Vi har alltså att de tre faktorerna i (6.6) är relativt prima, och därför är samtliga kuber. Om vi nu sätter $2a = \alpha^3, a - 3b = \beta^3$ och $a + 3b = \gamma^3$ får vi att $\beta^3 + \gamma^3 = 2a = \alpha^3$ vilket ger en lösning α, β, γ till ekvationen $x^3 + y^3 = z^3$ där α, β, γ antar värden mindre än den antagna ursprungliga lösningen x, y, z.

Vi har vidare att $\alpha^3\beta^3\gamma^3 = 2a(a - 3b)(a + 3b) = 2p$ är positiv och en delare till z^3 alternativt x^3, om z respektive x är jämnt. Alltså är $\alpha^3\beta^3\gamma^3 < z^3$. Observera att inget hindrar α, β eller γ från att vara negativa kuber då man kan flytta över dessa till motsatt led i ekvationen och bilda positiva kuber, eftersom $(-\alpha)^3 = -\alpha^3$. Ekvationen blir då $X^3 + Y^3 = Z^3$ där X, Y, Z är positiva och $Z^3 < z^3$. Därmed är "infinite descent" bevisad för detta fall.

Fall 2: Här gäller att $3 \mid p$, vilket innebär att vi kan anta att $p = 3s$ där s är ett positivt heltal. Eftersom p och q är relativt prima gäller att $3 \nmid q$. Då fås enligt (6.4) att

$$2p(p^2 + 3q^2) = 2 \cdot 3s(9s^2 + 3q^2) = 3^2 \cdot 2s(3s^2 + q^2)$$

är en kub. Eftersom $3^2 \cdot 2s$ och $3s^2 + q^2$ är relativt prima så är båda kuber enligt (*). Om vi, i väntan på ett senare bevis av ett lemma och på samma sätt som i (6.5), antar att *samtliga* kuber på formen $3s^2 + q^2$ fås då

$$q = a^3 - 9ab^2, \quad s = 3b(a - b)(a + b),$$

för godtyckliga heltal a och b, så följer att $3^2 \cdot 2s = 3^3 \cdot 2b(a - b)(a + b)$ som alltså är en kub. Således är även $2b(a - b)(a + b)$ en kub. Eftersom faktorerna är relativt prima sätter vi nu $2b = \alpha^3, a - b = \beta^3, a + b = \gamma^3$, vilket ger att $\gamma^3 - \beta^3 = 2b = \alpha^3$. Vi får nu, efter eventuell

omflyttning av negativa kuber som i fall 1, en ekvation på formen $X^3 + Y^3 = Z^3$ där X, Y, Z är positiva och $Z^3 < z^3$. "Infinite descent" är därmed bevisad även för detta fall.

Vi har nu visat att "infinite descent" alltid gäller om det skulle finnas en lösning $x, y, z > 0$ till den diofantiska ekvationen (6.2), vilket vi från början antog. Alltså kan en sådan lösning inte finnas och Fermats stora sats i fallet $n = 3$ är därmed (nästan) bevisad enligt Eulers metod. Men som nämnts har vi hittills bara *antagit* det Euler inte inkluderade i beviset. Med andra ord är satsen först med denna komplettering fullständigt bevisad. □

6.3 Komplettering av Eulers bevis

Det som återstår av beviset för Fermats stora sats i fallet $n = 3$ är att visa följande[4]:

Om $a, b \in \mathbf{Z}$ är relativt prima och om $a^2 + 3b^2$ är en kub så gäller att

$$a + b\sqrt{-3} = (p + q\sqrt{-3})^3 \tag{6.7}$$

för några $p, q \in \mathbf{Z}$.

Det var detta samband som Euler inte tog med i sitt bevis, och som alltså på en väsentlig punkt gjorde beviset ofullständigt. Beviset för sambandet, i form av lemmat som avslutar detta avsnitt, fullbordar alltså Eulers bevis. För att teoretiskt nå dit faktoriserar vi stegvis $a + b\sqrt{-3}$ på fyra olika sätt. Värt att nämna är att (6.7), efter normering av båda leden, *medför* ekvationen

$$a^2 + 3b^2 = (p^2 + 3q^2)^3,$$

vilket ger det villkor vi känner igen sedan tidigare. Skillnaden är helt enkelt att vi nu betraktar (algebraiska) heltal i $\mathbf{D} = \mathbf{Z}[\frac{1+\sqrt{-3}}{2}]$ i stället för heltal i $\mathbf{Z}$ genom att använda oss av det faktum att $\mathbf{D}$ har entydig faktorisering.

1. Vi ska visa följande:

   (se [8, s. 52]) Om a och b är relativt prima och om $a^2 + 3b^2$ är jämnt, så fås att

   $$a + b\sqrt{-3} = (1 \pm \sqrt{-3})(u + v\sqrt{-3}), \tag{6.8}$$

   med lämpligt tecken och där $u, v \in \mathbf{Z}$.

   Eftersom $a^2 + 3b^2$ är jämnt så har a och b samma paritet och då de är relativt prima måste båda vara udda. De kan alltså båda skrivas på formen $4n \pm 1$, och antingen $a + b$ eller $a - b$ är delbart med 4.

   I det första fallet kan ekvationen

   $$4(a^2 + 3b^2) = (1^2 + 3 \cdot 1^2)(a^2 + 3b^2) = (a - 3b)^2 + 3(a + b)^2$$

   delas med $16 = 4^2$ i båda led varav fås

[4] Vi använder här beteckningarna i [8], där heltalen tidigare benämnda a och b ersatts med p respektive q, och vice versa.

$$\frac{a^2+3b^2}{4} = \frac{(a-3b)^2+3(a+b)^2}{4^2} = u^2+3v^2,$$

där $u := \frac{a-3b}{4}$ och $v := \frac{a+b}{4}$.

Vi kan nu uttrycka $a+b\sqrt{-3}$ i termer av u och v genom att utnyttja sambandet

$$u+v\sqrt{-3} = \frac{(a+b\sqrt{-3})(1+\sqrt{-3})}{4},$$

varefter vi, genom att förlänga med konjugatet, får att $a+b\sqrt{-3}$
$= (1-\sqrt{-3})(u+v\sqrt{-3})$, vilket skulle visas.

I det andra fallet kan på liknande sätt visas att $a+b\sqrt{-3} = (1+\sqrt{-3})(u+v\sqrt{-3})$ för lämpligt valda heltal u, v. Eftersom a och b är relativt prima gäller att även u och v är relativt prima. Observera också att $a^2+3b^2 = 4(u^2+3v^2)$.

2. Vi ska visa följande:

> (se [8, s. 52]) Om a och b är relativt prima och om a^2+3b^2 är delbart med det udda primtalet P, så kan P skrivas på formen $P = p^2+3q^2$ där q och p är positiva heltal, och $a+b\sqrt{-3}$ kan skrivas på formen
>
> $$a+b\sqrt{-3} = (p \pm q\sqrt{-3})(u+v\sqrt{-3}) \tag{6.9}$$
>
> med lämpligt tecken och där $u, v \in \mathbf{Z}$.

Beviset för det första påståendet återfinns i [8, s. 50]. Antingen $pb+aq$ eller $pb-aq$ är delbart med P. I det första fallet kan ekvationen $P(a^2+3b^2) = (p^2+3q^2)(a^2+3b^2) = (pa-3qb)^2+3(pb+aq)^2$ delas med P^2 i båda led varav fås

$$\frac{a^2+3b^2}{P} = \frac{(pa-3qb)^2+3(pb+aq)^2}{P^2} = u^2+3v^2,$$

där $u := \frac{pa-3qb}{P}$ och $v := \frac{pb+aq}{P}$.

Vi använder nu på motsvarande sätt som tidigare sambandet

$$u+v\sqrt{-3} = \frac{(p+q\sqrt{-3})(a+b\sqrt{-3}}{P},$$

vilket efter multiplikation med $p-q\sqrt{-3}$ ger

$$\begin{aligned}(p-q\sqrt{-3})(u+v\sqrt{-3}) &= \frac{(p-q\sqrt{-3})(p+q\sqrt{-3})(a+b\sqrt{-3}}{P} \\ &= \frac{(p^2+3q^2)(a+b\sqrt{-3})}{p^2+3q^2} \\ &= a+b\sqrt{-3},\end{aligned}$$

vilket skulle visas.

I det andra fallet kan på liknande sätt visas att $a + b\sqrt{-3} = (p + q\sqrt{-3})(u + v\sqrt{-3})$. Återigen, eftersom a och b är relativt prima gäller detta även för u och v. Observera också att $a^2 + 3b^2 = P(u^2 + 3v^2)$.

3. Vi ska visa följande:

> (se [8, s. 53]) Om a och b är relativt prima så fås att
>
> $$a + b\sqrt{-3} = \pm(p_1 \pm q_1\sqrt{-3})(p_2 \pm q_2\sqrt{-3}) \cdots (p_n \pm q_n\sqrt{-3}), \tag{6.10}$$
>
> där alla $p_i, q_i \in \mathbf{Z}^+$ och $p_i^2 + 3q_i^2$ är antingen lika med 4 eller ett udda primtal.

Om $a^2 + 3b^2$ är jämnt så är det delbart med 4, om det är skilt från 1 så finns en faktor P lika med 4 eller ett udda primtal. Antingen (6.8) eller (6.9) ger då $a + b\sqrt{-3} = (p \pm q\sqrt{-3})(u + v\sqrt{-3})$, där $P = p^2 + 3q^2$. Detta ger att u och v är relativt prima. Att stryka en faktor $p + q\sqrt{-3}$ från $u + v\sqrt{-3}$ är samma sak som att stryka en faktor från $a + b\sqrt{-3}$, dock är $u^2 + 3v^2 = \frac{a^2+3b^2}{P}$ mindre än $a^2 + 3b^2$. Genom att upprepa denna procedur får vi slutligen $a + b\sqrt{-3} = (p_1 \pm q_1\sqrt{-3}) \cdots (p_n \pm q_n\sqrt{-3})(u + v\sqrt{-3})$ där $u^2 + 3v^2 = 1$. Alltså är $u = \pm 1, v = 0$ vilket ger att $u + v\sqrt{-3} = \pm 1$. Faktoriseringen enligt (6.10) är därmed klar.

4. Vi ska visa följande:

> (se [8, s. 53]) Om a och b är relativt prima så är faktorerna i (6.10) entydigt bestämda, bortsett från valet av tecken, eftersom
>
> $$(p_1^2 + 3q_1^2)(p_2^2 + 3q_2^2) \cdots (p_n^2 + 3q_n^2) = a^2 + 3b^2 \tag{6.11}$$
>
> är en faktorisering av $a^2 + 3b^2$ i 4:or och udda primtal. Om dessutom faktorn $p + q\sqrt{-3}$ förekommer i faktoriseringen så förekommer inte faktorn $p - q\sqrt{-3}$ och vice versa.

Vi visar först att $P = p^2 + 3q^2$ ger entydigt bestämda p och q, bortsett från valet av tecken, om P är lika med 4 eller ett udda primtal. Detta är klart i fallet $P = 4$. Om P är ett udda primtal och om det kan skrivas som $P = a^2 + 3b^2$, så följer enligt punkt 2 att

$$a + b\sqrt{-3} = (p \pm q\sqrt{-3})(u + v\sqrt{-3})$$

och $P = P(u^2 + 3v^2)$ vilket ger att $u^2 + 3v^2 = 1$, det vill säga $u = \pm 1, v = 0$ och alltså $a + b\sqrt{-3} = \pm(p \pm q\sqrt{-3})$, vilket skulle visas.

Det andra påståendet kan visas genom att inse att $p + q\sqrt{-3}$ och $p - q\sqrt{-3}$ inte båda kan vara faktorer eftersom det då skulle finnas en faktor $p^2 + 3q^2$, vilket är omöjligt då a och b är relativt prima.

Vi är nu redo att formulera det lemma som behövs för att komplettera Eulers bevis för Fermats stora sats i fallet $n = 3$.

> **Lemma 6.3.** *Låt $a, b \in \mathbf{Z}$ vara relativt prima sådana att $a^2 + 3b^2$ är en kub. Det existerar då några $p, q \in \mathbf{Z}$ sådana att $a + b\sqrt{-3} = (p + q\sqrt{-3})^3$.*

Bevis. (se [8, s. 54]) Låt $a^2+3b^2 = P_1P_2\cdots P_n$ vara en faktorisering av 4:or och udda primtal enligt (6.11) ovan. Om det i denna faktorisering finns exakt k stycken 4:or så är 2^{2k} den största tvåpotensen som delar a^2+3b^2. Eftersom denna enligt villkor är en kub följer det att $2k$ och därmed också k är en multipel av 3. Dessutom måste varje udda primtal P i faktoriseringen finnas med i ett antal som är en multipel av 3. Detta sammantaget ger att n är delbart med 3. Vi ordnar nu faktorerna så att $P_{3k+1} = P_{3k+2} = P_{3k+3}$. De faktorer i (6.10) som då motsvarar varje gruppering av tre lika faktorer i P, är identiska eftersom det enda som skulle kunna skilja sig åt - tecknet i faktorn $p \pm q\sqrt{-3}$ - enligt punkt 4 inte kan vara olika samtidigt. Genom att ta en faktor från varje gruppering av tre lika faktorer i P och multiplicera ihop dem får vi talet $p+q\sqrt{-3}$ sådant att $a+b\sqrt{-3} = \pm(p+q\sqrt{-3})^3$. Vi har slutligen att

$$-(p+q\sqrt{-3})^3 = (-1)^3(p+q\sqrt{-3})^3 = (-p-q\sqrt{-3})^3$$

och lemmat är bevisat. □

I och med detta är alltså Fermats stora sats i fallet $n = 3$ fullständigt bevisad.

6.4 Något om Pells ekvation

Vi avslutar detta kapitel med att kort studera en speciell typ av ekvation. Denna har en nära koppling till teorin som studerats här, och det är just denna specifika aspekt av ekvationen som kommer lyftas fram. Att den tas upp som avslutning på kapitlet och texten beror på att den inte är nödvändig för att förstå genomgången teori, men också på att ekvationen är diofantisk och att antalet lösningar av den är av intresse, liksom för ekvation (6.1) i Fermats stora sats.

Definition 6.4. (se [2]) Den diofantiska ekvationen

$$x^2 - dy^2 = N$$

där d och N är givna heltal, kallas *Pells ekvation.*

Den engelske 1600-talsmatematikern John Pell, som ekvationen är uppkallad efter, är i själva verket mycket lite förknippad med analysen av denna. Att så ändå är fallet beror på ett misstag av Euler, som trodde att Pells diskussion av ett bevis för ekvationen innebar att Pell hade bevisat satsen, vilket alltså visade sig inte stämma.

Om $d < 0$ i Pells ekvation finns ett ändligt antal lösningar. Om d är ett kvadrattal, det vill säga $d = a^2$ för något heltal a, fås ekvationen $(x+ay)(x-ay) = N$, som också den har ett ändligt antal lösningar. I övriga fall är antalet lösningar oändligt och kan beräknas genom användning av exempelvis kedjebråk. Den som först bevisade att Pells ekvation i fallet $N = 1$ (och då $d > 0$ inte är ett kvadrattal) har oändligt många lösningar var den italienske matematikern och astronomen Joseph-Louis Lagrange på 1700-talet[7].

Det finns en nära koppling mellan Pells ekvation i fallet $N = 1$ och kvadratiska talringar, nämligen att ekvationens vänsterled är lika med normen för motsvarande algebraiska heltal. Det gäller alltså att

$$1 = x^2 - dy^2 = (x+y\sqrt{d})(x-y\sqrt{d}) = N(x+y\sqrt{d}),$$

vilket innebär att samtliga lösningar till ekvationen $x^2 - dy^2 = 1$ ges av de heltalspar (x, y) för vilka $x + y\sqrt{d}$ är en enhet, i motsvarande kvadratiska talring $\mathbf{D}$, med norm 1. För $d < 0$,

motsvarande de imaginära kvadratiska talringarna, finns därmed ett begränsat antal enheter medan det för de reella kvadratiska talringarna, då $d > 0$ inte är ett kvadrattal i Pells ekvation, finns ett oändligt antal enheter, i enlighet med de i avsnitt 5.6 nämnda resultaten.

Referenser

[1] Sjöberg, B. (1992). *Grundkurs i talteori*. Åbo: Sigma vid Åbo akademi.

[2] Niven, I., Zuckerman, H.S., Montgomery, H.L. (1991). *An introduction to the theory of numbers*. New York: Wiley.

[3] Svensson, P-A. (2001). *Abstrakt algebra*. Lund: Studentlitteratur.

[4] Ireland, K., Rosen, M. (1990). *A classical introduction to modern number theory*. Berlin: Springer.

[5] LeVeque, W.J. (1996). *Fundamentals of number theory*. New York: Dover.

[6] Singh, S. (2005). *Fermats gåta - så löstes världens svåraste matematiska problem*. Stockholm: Pan.

[7] Rosen, K.H. (2000). *Elementary number theory and its applications*. Reading, MA: Addison-Wesley.

[8] Edwards, H.M. (2000). *Fermat's last theorem - A genetic introduction to algebraic number theory*. New York: Springer.