Anders Thorup

Diofantiske ligninger

Algebra og talteori, 1999

Kvadratiske ligninger (KVADR)

Brøker (BRK)

Index

Kvadratiske ligninger

0. Den lineære ligning (Udkast).

(0.1) Setup. Hovedemnet i dette kapitel er den kvadratiske ligning,

$$ax^2 + bxy + cy^2 = k. \tag{0.1.1}$$

Her er a, b, c og k givne hele tal, og opgaven er at bestemme heltalsløsninger (x, y) til ligningen. Det antages sædvanligvis, at *diskriminanten* $D := b^2 - 4ac$ ikke er et kvadrat, idet opgaven ellers er trivial. Antag nemlig, at D er et kvadrat, $D = d^2$, hvor d er et helt tal. Da kan andengradspolynomiet $F(x, y)$ på venstresiden faktoriseres: For $a = 0$ har vi trivielt $F(x, y) = bxy + cy^2 = (bx + cy)y$, og for $a \neq 0$ får vi, efter multiplikation med a,

$$aF(x, y) = a(ax^2 + bxy + cy^2) = (ax + \tfrac{1}{2}(b - d)y)(ax + \tfrac{1}{2}(b + d)y),$$

hvor koefficienterne på højresiden er hele tal, idet b og d øjensynlig har samme paritet. I begge tilfælde er ligningen (0.1.1) altså ækvivalent med en ligning af formen,

$$(a_1x + b_1y)(a_2x + b_2y) = l.$$

Det er klart, hvordan en ligning af denne form løses: For hver faktorisering af l som produkt af to hele tal, $l = l_1l_2$, skal man bestemme heltalsløsningerne til det lineære ligningssystem,

$$\begin{aligned} a_1x + b_1y &= l_1 \\ a_2x + b_2y &= l_2 \end{aligned} \tag{0.1.2}$$

Ligningen (0.1.1) fører altså, når diskriminanten $D = b^2 - 4ac$ er et kvadrat, til et lineært ligningssystem af formen (0.1.2).

I dette indledende afsnit vil vi behandle den generelle lineære ligning: For en given $r \times s$-matrix A med hele koefficienter, altså $A \in \mathrm{Mat}_{r,s}(\mathbb{Z})$, og en given søjle b af r hele tal, altså $b \in \mathbb{Z}^r$, søges heltalsløsninger $x \in \mathbb{Z}^s$ til det lineære ligningssystem,

$$Ax = b. \tag{0.1.3}$$

Vi tænker på løsningen $x \in \mathbb{Z}^s$ som en søjle; af typografiske grunde vil ofte betragte den tilsvarende række x^{tr}.

Bemærk, at begrundelsen, via den kvadratiske ligning (0.1.1), for at behandle det helt generelle ligningssystem (0.1.3) er lidt tynd: Ligningssystemet (0.1.2) har jo en (kvadratisk) 2×2-koefficientmatrix, og det er nemt at bestemme determinanten Δ udtrykt ved a, b, c. Hvis $\Delta \neq 0$, har (0.1.2) præcis én rational løsning (x, y), og der er derfor heltalsløsninger netop når x og y er hele tal. Det er altså kun tilfældet $\Delta = 0$, der ikke er helt oplagt.

(0.2) Elementære operationer. Det generelle lineære ligningssystem kan løses ved at underkaste systemet elementære operationer. De *elementære rækkeoperationer* på en matrix er:

(1) gang den j'te række med -1, og ombyt den så med i'te række (for $j \neq i$),
(2) læg et heltalsmultiplum, q gange j'te række, til den i'te række (for $j \neq i$).

Elementære søjleoperationer defineres tilsvarende.

I visse forbindelser vil man som elementær rækkeoperation medtage, at gange en række med -1. Når denne operation medtages, kan operationen i (1) naturligvis erstattes med en ren ombytning af de to rækker. Vi vil *ikke* medtage multiplikation med -1 som en elementær operation, og holder altså fast ved vores version af (1). Vi vil omtale (1) som en rækkeombytning, idet det altså er underforstået, at en af de to rækker samtidig multipliceres med -1.

Det er i øvrigt klart, at operationen (1) kan udføres ved gentagne gange at udføre operationen (2). For to søjler (a, b) får vi, ved gentagne gange at anvende søjleoperationen (2) med $q = \pm 1$, at

$$(a, b) \mapsto (a, b - a) \mapsto (b, b - a) \mapsto (b, -a).$$

Ligningssystemet (0.1.3) beskrives ofte ved den udvidede koefficientmatrix (b, A), som er en $r \times (s + 1)$-matrix. Det er klart, at elementære rækkeoperationer på den udvidede matrix ikke ændrer løsningsmængden. Denne observation er som bekendt udgangspunktet for Gauss-elimination.

Søjleoperationer på matricen A modsvares af variabelskift: At ombytte i'te og og j'te søjle svarer til at „ombytte de variable" x_i og x_j. At lægge q gange j'te søjle til den i'te svare til at „indføre en ny variabel" $x_i := x_i + qx_j$. Det er – i princippet – klart, hvordan man ud fra løsninger til det transformerede ligningssystem bestemmer løsningerne til det oprindelige ligningssystem.

Rækkeoperationer på matricen A kan udføres ved at multiplicere A fra venstre med en *elementær $r \times r$-matrix*. Fx svarer ombytningen (1) til multiplikationen TA, hvor T har samme elementer som enhedsmatricen på nær $t_{ij} = -1$, $t_{ji} = 1$ og $t_{ii} = t_{jj} = 0$; bemærk, at den ekstra multiplikation af den ene række med -1 sikrer, at matricen har determinant $+1$. Tilsvarende svarer operationen (2) til multiplikationen TA, hvor matricen T har samme elementer som enhedsmatricen på nær $t_{ij} = q$; også denne matrix har determinant $+1$. De elementære matricer af de to typer er essentielt (for $i < j$) matricerne,

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 & q \\ 0 & 1 \end{pmatrix}, \tag{0.2.1}$$

hvor vi kun har angivet elementerne på pladserne ii, ij, ji, og jj. For $i > j$ skal de to angivne matricer transponeres. Den inverse til en elementær matrix er igen en elementær matrix.

Gentagne rækkeoperationer svarer til en multiplikation TA, hvor T er et produkt af elementære matricer. Et sådant produkt har determinant 1, og ligger altså specielt i gruppen $\mathrm{SL}_r(\mathbb{Z})$. Tilsvarende kan gentagne søjleoperationer udføres ved en multiplikation AS, hvor $s \times s$-matricen S er et produkt af elementære matricer.

(0.3) Elementardivisorsætningen. *Enhver matrix A i $\mathrm{Mat}_{r,s}(\mathbb{Z})$ kan ved elementære operationer på rækker og søjler omformes til en $r \times s$ matrix af normalformen,*

$$D = \begin{pmatrix} d_1 & & & & \\ & d_2 & & & \\ & & \ddots & & \\ & & & d_p & \end{pmatrix},$$

hvor $d_1, \ldots, d_p$ er hele tal forskellige fra 0 med $d_1 \mid d_2 \mid \cdots \mid d_p$, og hvor matricen har 0 på alle de øvrige pladser.

Bevis. Hvis $A = 0$, har A allerede den søgte normalform, med $p = 0$, så vi antager, at $A \neq 0$. Vi kan herefter, blandt de numeriske værdier af tallene forskellige fra 0 i A, betragte den mindste, betegnet $\nu(A)$. Normalformen fremkommer efter en række skridt. I hvert skridt sænkes ν-værdien, dvs A omformes til en matrix A' med $\nu(A') < \nu(A)$. Øjensynlig kan ν-værdien kun sænkes endelig mange gange. Når ν-værdien ikke længere kan sænkes, er vi godt på vej til at have omformet A til normalformen.

Ideen i det enkelte skridt er følgende: Antag, at der i en bestemt søjle, på pladserne i og j, står tallene a og d, hvor $|a| > |d| > 0$ og d ikke er divisor i a. Anvendes rækkeoperationen (2) med faktoren $-q$, erstattes elementet a af $a' := a - qd$, og her kan vi, ifølge Sætningen om division med rest, bestemme q således, at vi for resten a' har $0 < a' < |d|$. Hvis den fremkomne rest a' er mindre end $\nu(A)$, har vi altså sænket ν-værdien.

Skridt 1. Specielt kan vi altid vælge en søjle, hvori der findes et element d med $|d| = \nu(A)$. Det følger, at vi kan sænke ν-værdien, med mindre de øvrige elementer i denne søjle alle er multipla af d. Tilsvarende kan vi sænke ν-værdien, med mindre alle elementer i rækken, som indeholder d, er multipla af d. Efter endeligt mange skridt fremkommer altså en matrix A med den egenskab, at hvis d er et element i matricen med $|d| = \nu(A)$, så er hvert tal i rækken og i søjlen, der indeholder d, et multiplum af d.

Skridt 2. I den fremkomne matrix kan vi, efter en eventuel ombytning af rækker og af søjler, antage, at $d = a_{11}$. Elementet a_{i1}, for $i > 1$, er et multiplum af d, så ved at trække et passende multiplum af første række fra den i'te opnås en matrix med $a_{i1} = 0$. Tilsvarende kan vi opnå, at $a_{1j} = 0$ for $j > 1$. Antag nu, at den omformede matrix indeholder et element a_{ij}, som ikke er et multiplum af d. Læg første søjle til den j'te. Herved ændres a_{1j} til d, men a_{ij} ændres ikke, da $a_{i1} = 0$. Trækkes herefter q gange første række fra den i'te, ændres a_{ij} til $a_{ij} - qd$. Her kan q bestemmes således, at $0 < a_{ij} - qd < |d|$. Omformningen sænker derfor derfor ν-værdien.

Skridt 3. Det følger, at vi efter endelig mange gentagelser af skridt $1 + 2$ opnår en matrix A, hvor alle a_{ij} er multipla af $d_1 := a_{11}$, og hvor $a_{1j} = 0$ og $a_{i1} = 0$ for $i > 1$ og for $j > 1$. Hvis $r = 1$ eller $s = 1$ har matricen normalformen. Ellers har den formen,

$$\begin{pmatrix} d_1 & 0 \\ 0 & d_1 A' \end{pmatrix},$$

hvor A' er en $(r-1)\times(s-1)$-matrix. Herfra opnås normalformen induktivt, ved at fortsætte med matricen A'. ∎

(0.4) Løsningen. Når matricen A har normalformen D i Sætning (0.3), får det lineære ligningssystem formen,

$$d_1x_1 = b_1,\ \ldots,\ d_px_p = b_p,\ 0 = b_{p+1},\ \ldots,\ 0 = b_r.$$

Ligningen har altså løsninger, hvis og kun hvis b_i er et multiplum af d_i for $1 \leqslant i \leqslant p$ og $b_i = 0$ for $p < i \leqslant r$. Hvis betingelsen er opfyldt, bliver den fuldstændige løsning:

$$x^{\mathrm{tr}} = (b_1/d_1, \ldots, b_p/d_p, t_{p+1}, \ldots, t_r),$$

hvor $t_{p+1}, \ldots, t_r$ er hele tal. (Her er x^{tr} den transponerede af søjlen x.)

I det almindelige tilfælde findes, ifølge Elementardivisorsætningen, matricer S og T, produkter af elementære matricer, og en ligning,

$$TAS = D,$$

hvor D har normalformen. Øjensynlig gælder:

$$Ax = b \iff TASS^{-1}x = Tb \iff DS^{-1}x = Tb.$$

Løsningerne til $Ax = b$ bestemmes altså ved at løse ligningen $Dy = c$, hvor $c = Tb$, og så sætte $x = Sy$.

(0.5) Kogebogsopskrift. I praksis løses ligningen $Ax = b$ sådan: Anbring øverst i 4 kolonner $r\times r$-enhedsmatricen E_r, søjlen b, matricen A, og $s\times s$-enhedsmatricen E_s. Række- og søjleoperationerne, der fører fra A til $TAS = D$ udføres nu på matricerne i de 4 kolonner således: rækkeoperationer udføres i kolonnen svarende til A *og* i de to første kolonner, søjleoperationer udføres i kolonnen svarende til A *og* i den sidste kolonne. Indholdet, først og sidst, i de 4 kolonner er:

$$\begin{array}{c|c|c|c} E_r & b & A & E_s \\ \vdots & \vdots & \vdots & \vdots \\ T & Tb = c & TAS = D & S \end{array}$$

Specielt aflæses matricen D, søjlen c, og transformationsmatricen S. Herefter undersøges, om $Dy = c$ kan løses; i bekræftende fald fås løsningerne til $Ax = b$ som $x = Sy$. [Første kolonne kan naturligvis undværes!].

(0.6) Korollar. *Enhver matrix i gruppen* $\mathrm{SL}_r(\mathbb{Z})$ *er et produkt af elementære matricer. Gruppen* $\mathrm{SL}_2(\mathbb{Z})$ *er frembragt af de to matricer,*

$$S := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad T := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Bevis. Lad $A \in \mathrm{SL}_r(\mathbb{Z})$. Der findes da produkter S og T af elementære matricer således, at $D = TAS$ har formen i (0.3). Her har S, T, og A alle determinanten 1, så det $D = 1$.

Specielt følger det, at $p = r$, og at $d_1 \cdots d_r = 1$. Hvert d_i er altså ± 1. Dobbelt anvendelse af operationen (1) giver fortegnsskift på både i'te og j'te række. Vi kan derfor antage, at højst et af tallene d_i er lig med -1, og da produktet er lig med $+1$, må de alle være $+1$. Men så er D enhedsmatricen, og $A = T^{-1}S^{-1}$ er derfor et produkt af elementære matricer.

Når $r = 2$ er der to typer elementære matricer, nemlig matricerne beskrevet i (0.2.1) og deres transponerede. Vi får:

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = S, \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = S^3, \quad \begin{pmatrix} 1 & q \\ 0 & 1 \end{pmatrix} = T^q, \quad \begin{pmatrix} 1 & 0 \\ q & 1 \end{pmatrix} = ST^{-q}S^{-1}.$$

Heraf følger den sidste påstand. □

(0.7) Anvendelse på frie kommutative grupper. En kommutativ (additivt skrevet) gruppe M kaldes en *fri gruppe*, hvis M har en *basis*, dvs hvis der findes endelig mange elementer $u_1, \ldots, u_r$ i M således, at hvert element $m \in M$ har en fremstilling, som en linearkombination,

$$m = x_1u_1 + \cdots + x_ru_r, \tag{0.7.1}$$

med entydigt bestemte heltalskoefficienter $(x_1, \ldots, x_r)$. Sættet x af koefficienter, som søjle, er *koordinatsøjlen* for m mht basen $(u_1, \ldots, u_r)$. Ækvivalent betyder det, at M som gruppe er isomorf med gruppen $\mathbb{Z}^r$. Mere præcist, efter valget af basis er afbildningen,

$$(x_1, \ldots, x_r) \mapsto x_1u_1 + \cdots + x_ru_r\,,$$

en isomorfi $\mathbb{Z}^r \xrightarrow{\sim} M$.

Bemærk, at antallet af elementer i en basis er entydigt bestemt ved M. Er nemlig $(v_1, \ldots, v_s)$ endnu en basis, får vi, af eksistensen i (0.7.1), fremstillinger,

$$(v_1, \ldots, v_s) = (u_1, \ldots, u_r)S, \quad (u_1, \ldots, u_r) = (v_1, \ldots, v_s)T$$

[eller kort: $v = uS$, $u = vT$], med matricer $S \in \mathrm{Mat}_{r,s}(\mathbb{Z})$ og $T \in \mathrm{Mat}_{s,r}(\mathbb{Z})$. Heraf ses, at $v = vTS$ og $u = uST$, og nu får vi, af entydigheden i (0.7.1), at TS og ST er enhedsmatricer. Men heraf følger som bekendt, at T og S er kvadratiske matricer (og hinandens inverse).

Det følger altså specielt, at *rangen* af den fri gruppe M, defineret som $\mathrm{rk}\, M :=$ antal elementer i en basis $(u_1, \ldots, u_r)$, er veldefineret, og at de øvrige baser har formen $v = uS$ med en matrix $S \in \mathrm{GL}_r(\mathbb{Z})$.

For to frie grupper M og N med valgte baser u og v, altså $M = \mathbb{Z}^r$ og $N = \mathbb{Z}^s$, er det let at se, at homomorfierne $\varphi\colon N \to M$ er afbildningerne $x \mapsto Ax$, med en matrix $A \in \mathrm{Mat}_{r,s}(\mathbb{Z})$. Matricen A er bestemt ved ligningen,

$$(\varphi v_1, \ldots, \varphi v_s) = (u_1, \ldots, u_r)A, \quad \text{eller kort: } \varphi(v) = uA;$$

den j'te søjle i A er altså koordinatsøjlen for $\varphi(v_j)$. Sammenhængen mellem $n \in N$ og koordinatsøjlen x mht til basen v kan beskrives ved ligningen $n = vx$. Af ligningerne,

$$\varphi(n) = \varphi(vx) = \varphi(v)x = uAx,$$

hvor den midterste udnytter, at φ er en homomorfi, aflæses, at Ax er koordinatsøjlen for φ mht basen u for M.

Med ændrede baser, bestemt ved ligninger $u = u'T$ og $v = v'S$ (bemærk, at ligningerne udtrykker de „gamle“ basiselementer som linearkombinationer af de „nye“), ændres matricen A. Ændringen kan beskrives sådan: Af ligningerne $n = vx = v'Sx$ aflæses, at koordinatsøjlen x' for n mht basen v' er bestemt som $x' = Sx$. Af ligningerne,

$$\varphi(v') = \varphi(vS^{-1}) = \varphi(v)S^{-1} = uAS^{-1} = u'TAS^{-1},$$

aflæses, at matricen A' for φ mht til baserne u' og v' er bestemt som $A' = TAS^{-1}$.

Af Elementardivisorsætningen følger derfor: *Enhver homomorfi $\varphi\colon N \to M$, mellem kommutative frie grupper N og M af rang s og r, kan i passende valgte baser beskrives ved en matrix af normalformen i* (0.3).

Lemma. *Lad M være en kommutativ fri gruppe af rang r. Da er enhver undergruppe $N \subseteq M$ fri af rang højst lig med r.*

Bevis. Vi kan antage, at $M = \mathbb{Z}^r$. Påstanden vises ved induktion efter r. For $r = 0$ er den triviel. For $r = 1$ er $M = \mathbb{Z}$; altså er N en undergruppe af $\mathbb{Z}$, og dermed cyklisk, dvs af formen $\mathbb{Z}u$. Hvis $u \neq 0$, er u en basis for N, og N er derfor fri af rang 1. Hvis $u = 0$, så er $N = (0)$ fri af rang 0.

Nu bemærker vi, at hvis vi for en kommutativ gruppe N har en homomorfi $\varphi\colon N \to N'$, hvor billedet $\varphi(N)$ og kernen $\varphi^{-1}(0)$ er frie grupper af rang s og t, så er N fri af rang $s + t$. Vælges nemlig i N elementer $u_1, \ldots, u_s$, hvis billeder udgør en basis for $\varphi(N)$, og suppleres med en basis $v_1, \ldots, v_t$ for kernen $\varphi^{-1}(0)$, så fremkommer, som det let ses, en basis for N.

Denne bemærking udnyttes i induktionsskridtet: Vi kan antage, at $M = \mathbb{Z}^r$, og betragter projektionen $\mathbb{Z}^r \to \mathbb{Z}^{r-1}$ svarende til de første $r-1$ koordinater. Kernen består af r-sæt, som kun har noget forskelligt fra 0 på sidstepladsen. Specielt er kernen isomorf med $\mathbb{Z}$. Betragt videre restriktionen til undergruppen $N \subseteq \mathbb{Z}^r$,

$$\varphi\colon N \to \mathbb{Z}^{r-1}.$$

Billedet $\varphi(N)$ er en undergruppe i $\mathbb{Z}^{r-1}$, og dermed (induktivt) en fri gruppe af rang højst $r-1$. Kernen $\varphi^{-1}(0)$ er en undergruppe af $\mathbb{Z}$, og dermed (ifølge det viste) en fri gruppe af af rang højst 1. Af bemærkningen følger derfor, at N er en fri gruppe af rang højst lig med $(r-1)+1 = r$. ∎

Sætning. *Lad M være en kommutativ fri gruppe af rang r, og lad N være en undergruppe. Da findes en basis $u_1, \ldots, u_r$ for M, og tal $d_1 \mid d_2 \mid \cdots \mid d_s$ med $s \leqslant r$ således, at $d_1u_1, \ldots, d_su_s$ er en basis for N.*

Bevis. Ifølge lemmaet er N fri af rang $s \leqslant r$. Andvend resultatet om homomorfier på inklusionsafbildningen $N \to M$. Det følger, at baser kan vælges således, at inklusionen kan beskrives ved en matrix af normalformen. Men det betyder netop, at baserne har den angivne form. (Bemærk, at vi i denne situation nødvendigvis må have $p = s$.) ∎

Korollar. (Struktursætningen). *Enhver endeligt frembragt kommutativ gruppe er isomorf med en direkte sum af cykliske grupper:*

$$\mathbb{Z}/\mathbb{Z}d_1 \oplus \cdots \oplus \mathbb{Z}/\mathbb{Z}d_r.$$

Bevis. En kommutativ (additivt skrevet) gruppe P kaldes endeligt frembragt, hvis der findes endelig mange elementer $e_1, \ldots, e_r \in P$ således, at hvert element $p \in P$ har en fremstilling $p = x_1e_1 + \cdots + x_re_r$ med heltalskoefficienter x_i (ikke nødvendigvis entydigt bestemte). Ækvivalent betyder det, at der findes en surjektiv homomorfi:

$$\varphi\colon \mathbb{Z}^r \to P.$$

Anvend Sætningen med $M = \mathbb{Z}^r$ og $N := \varphi^{-1}(0)$. Via den nye basis for M har vi stadig $M = \mathbb{Z}^r$, og nu er N undergruppen,

$$N = \mathbb{Z}d_1 \oplus \cdots \oplus \mathbb{Z}d_s \oplus (0) \oplus \cdots \oplus (0),$$

Kvotienten M/N har derfor den angivne form (med $d_i = 0$ for $s < i \leqslant r$), og kvotienten er isomorf med P. ∎

Korollar. *Betragt en homomorfi $\varphi\colon N \to M$ mellem kommutative fri grupper af rang r, i givne baser for N og M af formen $x \mapsto Ax$ med en matrix $A \in \mathrm{Mat}_r(\mathbb{Z})$. Da er følgende betingelser ækvivalente:*

(i) *φ er injektiv.*
(ii) *Billedgruppen $\varphi(N)$ har endeligt index: $|M : \varphi(N)| < \infty$.*
(iii) $\det A \neq 0$.

Er betingelserne opfyldt, gælder ligheden $|M : \varphi(N)| = \det A$.

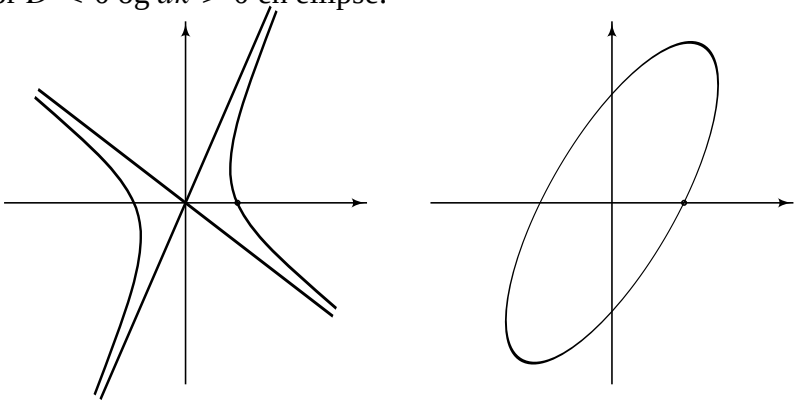
Bevis. Ændres baserne, ændres matricen A til en matrix af formen TAS^{-1}, hvor $S, T \in \mathrm{GL}_r(\mathbb{Z})$. Specielt har S og T determinant ± 1, så $|\det A|$ ændres ikke. Vi kan derfor antage, at A har normalformen i (0.3). Specielt er A så en diagonalmatrix. Lad $d_1, \ldots, d_r$ være diagonalelementerne (med notationen i (0.3) er det de første p elementer, der er forskellige fra 0). Når M identificeres med $\mathbb{Z}^r$, er billedgruppen undergruppen $d_1\mathbb{Z} \oplus \cdots \oplus d_r\mathbb{Z}$, og kvotienten er den direkte sum af de cykliske grupper $\mathbb{Z}/d_i\mathbb{Z}$. Det er herefter klart, at hver af betingelserne er ensbetydende med at alle tallene d_i er forskellige fra 0. Når betingelsen er opfyldt, er index, dvs ordenen af kvotientgruppen, lig med produktet af tallene $|d_i|$, og altså lig med $|\det A|$. ∎

1. Pell's ligning.

(1.1) Setup. Som nævnt betragtes i dette kapitel faste hele tal a, b, c og for et helt tal k den kvadratiske ligning,

$$ax^2 + bxy + cy^2 = k. \tag{1.1.1}$$

Det *antages* i det følgende, at *diskriminanten* $D := b^2 - 4ac$ ikke er et kvadrat. Specielt er altså a og c forskellige fra 0. De reelle løsninger til (1.1.1) udgør, for $D > 0$ og $k \neq 0$ en hyperbel, og for $D < 0$ og $ak > 0$ en ellipse:



Det er formålet, at løse ligningen (1.1.1) som diofantisk ligning, altså at undersøge heltalsløsninger (x, y). Ækvivalent svarer dette til at bestemme gitterpunkter, dvs punkter med heltalskoordinater, på det tilsvarende keglesnit.

Til tallene a, b, c knyttes polynomiet,

$$F(X, Y) = aX^2 + bXY + cY^2,$$

der er et homogent andengradspolynomium, også kaldet en *kvadratisk form*. Ligningen kan skrives $F(x, y) = k$. Videre sættes:

$$\eta := \frac{-b + \sqrt{D}}{2a} \quad \text{og} \quad \xi := a\eta = \frac{-b + \sqrt{D}}{2}. \tag{1.1.2}$$

Hvis $D > 0$ (det reelle tilfælde), fortolkes $\sqrt{D}$ som den positive kvadratrod, hvis $D < 0$ (det imaginære tilfælde), vælges kvadratroden i den øvre halvplan. Tallet η er øjensynlig rod i polynomiet $aX^2 + bX + c$. Den anden rod i dette polynomium betegnes η'. Tilsvarende er ξ rod i polynomiet $X^2 + bX + ac$; den anden rod i dette polynomium er $\xi' = a\eta'$. Tallene η' og ξ' fås af udtrykkene for η og ξ ved at skifte fortegn på $\sqrt{D}$. I det reelle tilfælde er $\xi > \xi'$, i det imaginære tilfælde ligger ξ i den øvre halvplan og ξ' er det komplekst konjugerede til ξ. Det tilsvarende gælder om beliggenheden af η og η', hvis $a > 0$.

Af $aX^2 + bX + c = a(X - \eta)(X - \eta')$ får vi faktoriseringen,

$$F(X, Y) = a(X - \eta Y)(X - \eta' Y). \tag{1.1.3}$$

Bemærk, at hyperblens asymptoter er linierne $x = y\eta'$ og $x = y\eta$. Øjensynlig er $\xi = \eta$, hvis $a = 1$.

(1.2) Pell's ligning. I denne paragraf betragter vi ligningen (1.1.1) for $a = 1$ og $k = \pm 1$, altså ligningen,

$$x^2 + bxy + cy^2 = \pm 1. \tag{1.2.1}$$

Som bekendt definerer tallet ξ en *kvadratisk talring* $\mathbb{Z}[\xi]$, bestående af alle tal af formen $x - y\xi$ for $x, y \in \mathbb{Z}$, og par (x, y) af hele tal svarer til tal $\alpha = x - y\xi$ i $\mathbb{Z}[\xi]$. For *normen*, $\mathrm{N}(\alpha)$, af $\alpha = x - y\xi$ gælder som bekendt ligningen,

$$\mathrm{N}(x - y\xi) = x^2 + bxy + cy^2. \tag{1.2.2}$$

Heltalsløsninger (x, y) til (1.2.1) svarer således bijektivt til tal $\alpha = x - y\xi$ i $\mathbb{Z}[\xi]$ således, at $\mathrm{N}(\alpha) = \pm 1$, altså til enhederne i ringen $\mathbb{Z}[\xi]$. I det imaginære tilfælde ($D < 0$) er det nemt (og velkendt!), hvordan løsningerne bestemmes: Der er kun endelig mange løsninger, og for $D < -4$ er der kun de trivielle løsninger $(x, y) = (\pm 1, 0)$, svarende til de trivielle enheder ± 1 i $\mathbb{Z}[\xi]$.

I det følgende *antager* vi, at diskriminanten $D = b^2 - 4c$ er positiv (og stadig: ikke et kvadrat); tallene i $\mathbb{Z}[\xi]$ er så relle tal. Et specialtilfælde er den klassiske Pell'ske ligning,

$$x^2 - dy^2 = 1, \tag{1.2.3}$$

hvor d er positiv og ikke et kvadrat, med diskriminanten $D = 4d$. Vi vil mere generelt kalde ligningen (1.2.1) for *Pell's ligning*. For $k = +1$, altså ligningen,

$$x^2 + bxy + cy^2 = 1. \tag{1.2.4}$$

taler vi om den *egentlige Pell'ske ligning*, for $k = -1$ siger vi, med et sprogligt misfoster, den *ikke-Pell'ske ligning*. Løsningerne til den egentlige Pell'ske ligning svarer bijektivt til de enheder ε i $\mathbb{Z}[\xi]$, for hvilke $\mathrm{N}(\varepsilon) = +1$.

(1.3) Lagrange's Sætning. *Den egentlige Pell'ske ligning* (1.2.4) *har altid uendelig mange løsninger.*

Bevis. Beviset herunder, der skyldes Dirichlet, er i en række skridt.

(1) For hvert $q \in \mathbb{N}$ findes hele tal (x, y) således, at

$$|x - y\xi| < 1/q \quad \text{og} \quad 1 \leqslant y \leqslant q. \tag{1.3.1}$$

Bestem nemlig for $i = 0, 1, \ldots, q$ det hele tal x_i, som opfylder, at $x_i - i\xi \in [0, 1[$. Der er $q + 1$ tal $x_i - i\xi$. Deles intervallet $[0, 1[$ i q lige store skuffer (intervaller) af længde $1/q$, følger det, at to af tallene $x_i - i\xi$ ligger i samme skuffe. Antag, at det indtræffer for i og for j, hvor $0 \leqslant i < j \leqslant q$. Da tallene ligger i samme skuffe, er deres afstand mindre end $1/q$, dvs,

$$|(x_j - x_i) - (j - i)\xi| < 1/q.$$

Altså er $x := x_j - x_i$ og $y := j - i$ brugbar i (1.3.1).

(2) Der findes uendelig mange par $(x, y) \in \mathbb{Z} \times \mathbb{N}$ således, at

$$|x - y\xi| < 1/y. \tag{1.3.2}$$

For givne endelig mange par (x_ν, y_ν) som opfylder (1.3.2) kan vi nemlig vælge $q \in \mathbb{N}$ således, at

$$1/q < \min |x_\nu - y_\nu \xi|.$$

Vælges dernæst (x, y), som opfylder (1.3.1), får vi

$$|x - y\xi| < 1/q \leqslant 1/y.$$

Altså vil parret (x, y) opfylde (1.3.2), og valget at q sikrer, at dette par er forskelligt fra de givne.

(3) Der findes et positivt tal K således, at for uendelig mange par $(x, y) \in \mathbb{Z} \times \mathbb{Z}$ er

$$|x^2 + bxy + cy^2| \leqslant K. \tag{1.3.3}$$

Hertil bemærkes først, at når y er positiv, er

$$|x - y\xi'| = |x - y\xi + y(\xi - \xi')| \leqslant |x - y\xi| + y\sqrt{D}.$$

Når (1.3.2) er opfyldt for (x, y), slutter vi derfor, at

$$|x^2 + bxy + cy^2| = |x - y\xi| \cdot |x - y\xi'| < \frac{1}{y}\Big(\frac{1}{y} + y\sqrt{D}\Big) = \frac{1}{y^2} + \sqrt{D} \leqslant 1 + \sqrt{D}.$$

Påstanden følger derfor af (2), med $K := 1 + \sqrt{D}$.

(4) Der findes et helt tal $k \neq 0$ således, at for uendelig mange par $(x, y) \in \mathbb{Z} \times \mathbb{Z}$ er

$$x^2 + bxy + cy^2 = k. \tag{1.3.4}$$

De uendelig mange par (x, y) fra (3) kan nemlig lægges i skuffer svarende til værdien $k := x^2 + bxy + cy^2$. Der er kun endelig mange skuffer, idet $|k| \leqslant K$. En af skufferne indeholder derfor uendelig mange par. Skuffen for $k = 0$ indeholder kun parret $(0, 0)$ (hvorfor?). Der må altså være en skuffe for $k \neq 0$, der indeholder uendelig mange par.

(5) Der findes et par (x_0, y_0) af hele tal således, at der for uendelig mange par $(x, y) \in \mathbb{Z} \times \mathbb{Z}$ gælder betingelserne:

$$x^2 + bxy + cy^2 = k, \quad x \equiv x_0, \quad y \equiv y_0 \pmod{k}. \tag{1.3.5}$$

Hertil lægges de uendelig mange par (x, y) fra (4) i skuffer svarende til restklasserne modulo k af (x, y). Der er k^2 par af restklasser. En af skufferne indeholder derfor uendelig mange par.

(6) Betragt for to par (x, y) og (x_1, y_1), som opfylder (1.3.5), de tilsvarende tal $\alpha := x - y\xi$ og $\alpha_1 := x_1 - y_1\xi$ i den kvadratiske talring $\mathbb{Z}[\xi]$. Da er $\alpha_1 = \varepsilon\alpha$, hvor $\varepsilon \in \mathbb{Z}[\xi]$ og $\mathrm{N}(\varepsilon) = 1$.

Af ligningen i (1.3.5) følger nemlig, at $\mathrm{N}(\alpha) = k$, og af kongruensen følger, at $x_1 \equiv x$ og $y_1 \equiv y$ modulo k. Altså er $\alpha_1 = \alpha + k\beta$ med et tal $\beta \in \mathbb{Z}[\xi]$. Herefter er

$$\frac{\alpha_1}{\alpha} = \frac{\alpha + k\beta}{\alpha} = 1 + k\frac{\beta}{\alpha} = 1 + k\frac{\beta\alpha'}{\mathrm{N}(\alpha)} = 1 + \beta\alpha'.$$

Det følger, at $\varepsilon := \alpha_1/\alpha$ tilhører $\mathbb{Z}[\xi]$. Af $\alpha_1 = \varepsilon\alpha$ og $\mathrm{N}(\alpha_1) = \mathrm{N}(\alpha)$ følger, at $\mathrm{N}(\varepsilon) = 1$.

(7) Nu bevises sætningen således: Der er en uendelig følge af forskellige par (x_i, y_i), der opfylder (1.3.5), og hertil svarer en følge af forskellige tal $\alpha, \alpha_1, \alpha_2, \ldots$ i $\mathbb{Z}[\xi]$. Ifølge (6) er $\alpha_i = \varepsilon_i\alpha$, hvor $\varepsilon_i \in \mathbb{Z}[\xi]$ og $\mathrm{N}(\varepsilon_i) = 1$. Skrives $\varepsilon_i = u_i - v_i\xi$ med hele tal u_i, v_i, er (u_i, v_i) altså en løsning til den egentlige Pell'ske ligning. Ligningen har derfor uendelig mange løsninger. ∎

(1.4) Enhederne. Vi antager stadig, at $a = 1$ og $D > 0$. Løsningerne til ligningen (1.2.1) svarer bijektivt til enhederne i ringen $\mathbb{Z}[\xi]$. Af Lagrange's Sætning følger specielt, at der er uendelig mange enheder. En enhed $\varepsilon = x - y\xi$ er øjensynlig ikke-triviel, dvs forskellig fra ± 1, netop når $y \neq 0$. I det følgende vil vi specielt betragte enheder ε i $\mathbb{Z}[\xi]$, som opfylder uligheden:

$$1 < \varepsilon'. \tag{1.4.1}$$

Øjensynlig gælder for hver ikke-trivel enhed ε, at også $-\varepsilon$ og $\pm\varepsilon^{-1}$ er enheder. Af de fire enheder $\pm\varepsilon$ og $\pm\varepsilon^{-1}$ vil netop én opfylde betingelsen (1.4.1). Specielt findes der, ifølge Lagrange's Sætning, enheder, som opfylder (1.4.1).

For en enhed $\varepsilon = x - y\xi$, som opfylder (1.4.1), gælder ulighederne,

$$y \geqslant 1, \quad x + \tfrac{b}{2}y > 0. \tag{1.4.2}$$

Antag nemlig, at $\varepsilon' > 1$. Da $|\varepsilon\varepsilon'| = 1$, følger det først, at $|\varepsilon| < 1$. Specielt er altså $\varepsilon < \varepsilon'$. Af $x - y\xi < x - y\xi'$ følger, da $\xi' < \xi$, at $y \geqslant 1$. Videre er $|\varepsilon| < |\varepsilon'|$, altså $|x - y\xi| < |x - y\xi'|$, og dermed $|x/y - \xi| < |x/y - \xi'|$. Brøken x/y ligger altså tættere ved ξ end ved ξ'. Altså er x/y større end $(\xi + \xi')/2 = -b/2$, hvoraf den anden ulighed i (1.4.2) fremgår. [Det er i øvrigt ikke svært at vise, for en enhed $\varepsilon = x - y\xi$, at den anden ulighed i (1.4.2) medfører uligheden (1.4.1).]

(1.5) Sætning. *Blandt enhederne ε i $\mathbb{Z}[\xi]$, som opfylder* (1.4.1)*, findes én, ε_1, for hvilken ε_1' er mindst. Med denne enhed er gruppen af alle enheder givet ved ligningen,*

$$\mathbb{Z}[\xi]^* = \{\pm\varepsilon_1^p \mid p \in \mathbb{Z}\}. \tag{1.5.1}$$

Bevis. Hyperblerne med ligningerne $x^2 + bxy + cy^2 = \pm 1$ har linierne $x = y\xi$ og $x = y\xi'$ som asymptoter. For et punkt (x, y) i planen er $x - y\xi'$ den „vandrette" afstand (regnet med fortegn) fra punktet (x, y) til linien med ligningen $x = y\xi'$. Uligheden $x - y\xi' > 1$

bestemmer en halvplan (punkterne til højre for linien $x - y\xi' = 1$). I denne halvplan ligger en „halv" gren af hver af hyperblerne, og de to halve grene har linien $x = y\xi$ som fælles asymtote. Heraf følger, at der blandt gitterpunkterne på de to halve hyperbelgrene findes et, (x_1, y_1), for hvilken den vandrette afstand til linien $x = y\xi'$ er numerisk mindst mulig.

Ækvivalent betyder det, at der blandt enhederne, som opfylder (1.4.1), findes en, ε_1, for hvilken ε_1' er mindst mulig. Øjensynlig er ε_1 entydigt bestemt. For at vise (1.5.1) skal det vises, at en vilkårlig enhed ε kan skrives på formen $\pm\varepsilon_1^p$ med $p \in \mathbb{Z}$. Erstattes eventuelt ε med $-\varepsilon$, kan vi antage, at $\varepsilon' > 0$. Vælg så $p \in \mathbb{Z}$ således, at $(\varepsilon_1')^p \leqslant \varepsilon' < (\varepsilon_1')^{p+1}$, altså

$$1 \leqslant \varepsilon'/(\varepsilon_1')^p < \varepsilon_1'.$$

Her sikrer minimaliteten af ε_1', at den første ulighed er en lighed. Altså er $\varepsilon' = (\varepsilon_1')^p$, og dermed er $\varepsilon = \varepsilon_1^p$. ∎

(1.6). Den reelle akse fraregnet 0 og ± 1 består af fire intervaller, og det er klart i hvilket af disse intervaller tallet $(\pm\varepsilon_1^p)' = \pm(\varepsilon_1')^p$ befinder sig. Specielt ses, at enhederne ε, der opfylder (1.4.1), netop er potenserne ε_1^n, for $n \geqslant 1$.

Enheden ε_1 kaldes *grundenheden* i den kvadratiske talring $\mathbb{Z}[\xi]$. Ud fra grundenheden ε_1 er samtlige enheder bestemt ved (1.5.1). For normen finder vi

$$\mathrm{N}(\pm\varepsilon_1^p) = \mathrm{N}(\varepsilon_1)^p.$$

Heraf fås følgende:

Resultat. *Hvis der for grundenheden ε_1 gælder $\mathrm{N}(\varepsilon_1) = -1$, så svarer løsningerne til den ikke-Pell'ske ligning til enhederne $\pm\varepsilon_1^p$, hvor p er ulige, og løsningerne til den egentlige Pell'ske ligning svarer til enhederne $\pm\varepsilon_+^p$, hvor $\varepsilon_+ := \varepsilon_1^2$.*

Hvis $\mathrm{N}(\varepsilon_1) = 1$, så har den ikke-Pell'ske ligning ingen løsninger; løsningerne til den egentlige Pell'ske ligning svarer til enhederne $\pm\varepsilon_1^p$.

For et punkt (x, y) i halvplanen bestemt ved $x - y\xi' > 1$ gælder øjensynlig, at den vandrette afstand til asymptoten, dvs $x - y\xi$, er positiv, hvis (x, y) ligger på hyperblen $x^2 + bxy + cy^2 = 1$, og negativ, hvis (x, y) ligger på hyperblen $x^2 + bxy + cy^2 = -1$. Det følger specielt, at grundenheden ε_1 er negativ, når $\mathrm{N}(\varepsilon_1) = -1$, og positiv, når $\mathrm{N}(\varepsilon_1) = 1$. Hvis $\mathrm{N}(\varepsilon_1) = 1$, sætter vi $\varepsilon_+ := \varepsilon_1$. Hermed er enheden ε_+ i alle tilfælde positiv, og den svarer til en grundløsning til den egentlige Pell'ske ligning. Enheden ε_+ kaldes også den *positive grundenhed*.

Svarende til grundenheden $\varepsilon_1 = x_1 - y_1\xi$ får vi *grundløsningen* (x_1, y_1) til den Pell'ske ligning. Fra potenserne $\varepsilon_1^n = x_n - y_n\xi$ får vi nye løsninger (x_n, y_n). Af $\varepsilon_1^{n+1} = \varepsilon_1\varepsilon_1^n$ fås en induktiv bestemmelse af (x_n, y_n). Vi har nemlig,

$$\varepsilon_1^{n+1} = (x_1 - y_1\xi)(x_n - y_n\xi) = (x_1x_n - cy_1y_n) - (x_1y_n + x_ny_1 + by_1y_n)\xi,$$

og altså,

$$x_{n+1} = x_1x_n - cy_1y_n, \qquad y_{n+1} = x_1y_n + x_ny_1 + by_1y_n. \tag{1.6.1}$$

(1.7) Lemma. *Idet* $\varepsilon_1^n = x_n - y_n\xi$, *for* $n \geqslant 1$, *gælder for* y_n *ulighederne,*

$$1 \leqslant y_1 \leqslant y_2 < y_3 < y_4 < \cdots , \tag{1.7.1}$$

og uligheden $y_1 \leqslant y_2$ *er skarp med mindre* $D = 5$. *For* $D = 5$ *er* $y_1 = y_2 = 1$ *og* $x_1/y_1 < \xi < x_2/y_2$; *videre er* $y_{n+1} = y_n + y_{n-1}$ *for* $n \geqslant 2$.

Bevis. Udtrykket for y_{n+1} i (1.6.1) kan skrives

$$y_{n+1} = (x_1 + \tfrac{b}{2}y_1)y_n + (x_n + \tfrac{b}{2}y_n)y_1.$$

Som bemærket i (1.4) vil potenserne ε_1^n for $n \geqslant 1$ opfylde (1.4.1). Ulighederne i (1.4.2) gælder derfor for (x_n, y_n) og (x_1, y_1). Hvis $x_1 + (b/2)y_1 \geqslant 1$ fremgår det af udtrykket, at $y_{n+1} > y_n$.

Da x_1, y_1, b er hele tal, og $x_1 + (b/2)y_1 > 0$, gælder altid $x_1 + (b/2)y_1 \geqslant 1/2$. Vi mangler således at betragte specialtilfældet, hvor $x_1 + (b/2)y_1 = 1/2$. Her er

$$\frac{1}{2} = x_1 + \frac{b}{2}y_1 = x_1 - y_1\xi + y_1\frac{\sqrt{D}}{2} = \varepsilon_1 + y_1\frac{\sqrt{D}}{2}\,. \tag{1.7.3}$$

Da $\varepsilon_1' > 1$, er $|\varepsilon_1| < 1$. Af (1.7.3) følger derfor, at $y_1\sqrt{D}/2 < 3/2$, altså at $y_1^2 D < 9$. Tallet b må være ulige, og følgelig er $D \equiv 1 \pmod 4$. Uligheden viser derfor, at $D = 5$ og $y_1 = 1$. Af (1.7.3) får vi

$$\varepsilon_1 = x_1 - \xi = \frac{1}{2} - \frac{\sqrt{5}}{2} = -\tau,$$

hvor $\tau := (-1 + \sqrt{5})/2$. Tallet τ er den største rod i $X^2 + X - 1$, og vi har $\tau = \xi - x_1$. Heraf ses, at $\mathbb{Z}[\xi] = \mathbb{Z}[\tau]$, og at koefficienterne y_n ikke ændres, hvis vi erstatter ξ med τ. Vi kan altså antage, at ξ er rod i $X^2 + X - 1$. Herefter er $\xi^2 = 1 - \xi$ og $\varepsilon_1 = -\xi$. Heraf fås $\varepsilon_1^2 = \xi^2 = 1 - \xi$. Altså er $(x_1, y_1) = (0, 1)$ og $(x_2, y_2) = (1, 1)$, og $x_1/y_1 < \xi < x_2/y_2$. Det er let at se, ved induktion efter n, at

$$\varepsilon_1^n = y_{n-1} - y_n\xi, \qquad \text{hvor } y_n = y_{n-1} + y_{n-2}.$$

Heraf følger specielt ulighederne (1.7.1) også for $D = 5$. ∎

(1.8) Sætning. *For en enhed* $\varepsilon = x - y\xi$ *med* $\varepsilon' > 1$ *gælder uligheden,*

$$\left|\frac{x}{y} - \xi\right| < \frac{1}{y^2}\,. \tag{1.8.1}$$

Følgen x_n/y_n *er netop delfølgen af Farey-følgen for tallet* ξ *bestående af de Farey-approximationer* x/y *for hvilke* $x^2 + bxy + cy^2 = \pm 1$. *Specielt gælder for grundenheden* $\varepsilon_1 = x_1 - y_1\xi$, *at* x_1/y_1 *er den første brøk* x/y *Farey-følgen, for hvilken* $x^2 + bxy + cy^2 = \pm 1$.

Bevis. Af (1.4.2) følger, at $y \geqslant 1$ og at $x/y > -b/2$. Af den sidste ulighed følger videre, at $x/y - \xi' > -b/2 - \xi' = \sqrt{D}/2$, og dermed er $|x/y - \xi'| > \sqrt{D}/2$. Videre er $|x/y - \xi||x/y - \xi'| = |\,\mathrm{N}(\varepsilon)|/y^2 = 1/y^2$. Heraf fås,

$$\left|\frac{x}{y} - \xi\right| = \frac{1}{y^2}\left|\frac{x}{y} - \xi'\right|^{-1} < \frac{1}{y^2}\frac{1}{\sqrt{D}/2}\,.$$

Vi har $D \geqslant 5$, og dermed $\sqrt{D}/2 > 1$. Altså gælder (1.8.1).

Af (1.8.1) følger som bekendt, at x/y er en Farey-approximation til ξ. Specielt er hver brøk x_n/y_n en Farey-approximation. Antag omvendt, at x/y er en Farey-approximation til ξ, og at $\varepsilon := x - y\xi$ er en enhed. Da er $y \geqslant 1$, og afstanden $|x/y - \xi|$ er specielt mindre end 1. Da $\xi - (-b/2) = \sqrt{D}/2$ er større end 1, følger det, at $x/y > -b/2$, og heraf følger videre, at $x/y - \xi' > \sqrt{D}/2 > 1$. Altså er $x - y\xi' > y \geqslant 1$, og dermed $\varepsilon' > 1$. Som nævnt i (1.4) følger heraf, at $\varepsilon = \varepsilon_1^n$ med $n \geqslant 1$, og dermed er $(x, y) = (x_n, y_n)$.

At brøkerne x_n/y_n kommer i samme rækkefølge i Farey-følgen, følger af Lemma (1.7). □

(1.9) Eksempel. For $x^2 - 13y^2$ er $D = 52$ og $\xi = \sqrt{13}$. Farey-approximationer beregnes ved skemaet:

p	3	4	7	11	18
q	1	1	2	3	5
$p^2 - 13q^2$	-4	3	-3	4	-1

Grundenheden er derfor $\varepsilon_1 = 18 - 5\sqrt{13}$, med norm -1. Altså har den ikke-Pell'ske ligning $x^2 - 13y^2 = -1$ grundløsningen $(18, 5)$. Den positive grundenhed er $\varepsilon_+ = \varepsilon_1^2 = 18^2 + 13 \cdot 5^2 - 2 \cdot 18 \cdot 5\sqrt{13} = 649 - 180\sqrt{13}$, og den egentlige Pell'ske ligning $x^2 - 13y^2 = 1$ har derfor grundløsningen $(649, 180)$.

For $x^2 - 7y^2$ er $D = 28$ og $\xi = \sqrt{7}$. Farey-approximationer beregnes ved skemaet:

p	2	3	5	8
q	1	1	2	3
$p^2 - 7q^2$	-3	2	-3	1

Grundenheden er derfor $\varepsilon_1 = 8 - 3\sqrt{7}$, med norm $+1$. Altså har den ikke-Pell'ske ligning $x^2 - 7y^2 = -1$ ingen løsninger. Vi har $\varepsilon_+ = 8 - 3\sqrt{7}$, og $(8, 3)$ er grundløsningen til den egentlige Pell'ske ligning $x^2 - 7y^2 = 1$.

(1.10) Opgave. Beviset for (1.7) er ikke helt fuldstændigt. Det blev vist, at hvis undtagelsestilfældet $x_1 + (b/2)y_1 = 1/2$ indtræffer, så er $D = 5$ og $\varepsilon_1 = -\tau$, hvor τ er den største rod i $X^2 + X - 1$. Er det omvendt klart, at hvis $D = 5$, så indtræffer undtagelsestilfældet?

2. Ækvivalens af løsninger.

(2.1) Setup. Vi bibeholder notation og forudsætningerne fra (1.1), og ser på den almindelige kvadratiske ligning,

$$ax^2 + bxy + cy^2 = k. \tag{2.1.1}$$

Venstresiden er formen $F(x, y)$ og tallene ξ og η defineret i (1.1). Tallet η er den ene rod i $aX^2 + bX + c$, og η' er den anden rod; specielt er $\eta + \eta' = -b/a$ og $\eta\eta' = c/a$. Vi kan antage, at $k \neq 0$, idet ligningen for $k = 0$ kun har løsningen $(0, 0)$.

Tallet $\xi = a\eta$ er øjensynlig rod i $X^2 + bX + ac$, så til ξ hører en kvadratisk talring $\mathbb{Z}[\xi]$. Til tallet η knyttes følgende mængde af komplekse tal:

$$\mathbb{Q}[\eta] := \{p + q\eta \mid p, q \in \mathbb{Q}\}.$$

Da tallet η er irrationalt, er fremstillingen $\lambda = p + q\eta$, af et tal $\lambda \in \mathbb{Q}[\eta]$, entydig. Ligningen,

$$\eta^2 = -\frac{c}{a} - \frac{b}{a}\eta, \tag{2.1.2}$$

viser, at η^2 ligger i $\mathbb{Q}[\eta]$, og herefter er det let at vise, at $\mathbb{Q}[\eta]$ er en delring af $\mathbb{C}$. *Konjugering*, dvs afbildningen $p + q\eta \mapsto p + q\eta'$, er en involutorisk ringautomorfi af $\mathbb{Q}[\eta]$. For $\lambda \in \mathbb{Q}[\eta]$ defineres *normen* af λ ved $\mathrm{N}(\lambda) := \lambda\lambda'$. Øjensynlig er normen en multiplikativ afbildning, og $\mathrm{N}(\lambda) \neq 0$, når $\lambda \neq 0$. Af (1.1.3) fremgår, at

$$a\,\mathrm{N}(p - q\eta) = a(p - q\eta)(p - q\eta') = ap^2 + bpq + cq^2 = F(p, q). \tag{2.1.3}$$

Det følger specielt, at $\mathrm{N}(\lambda)$ altid er et rationalt tal. Heraf følger videre, at delringen $\mathbb{Q}[\eta]$ faktisk er et dellegeme af $\mathbb{C}$: Hvis $\lambda \in \mathbb{Q}[\eta]$ er forskellig fra 0, så er det reciprokke,

$$\lambda^{-1} = \frac{\lambda'}{\lambda\lambda'} = \frac{\lambda'}{\mathrm{N}(\lambda)},$$

igen et tal i $\mathbb{Q}[\eta]$.

Erstattes η med ξ får vi det samme legeme: $\mathbb{Q}[\eta] = \mathbb{Q}[\xi]$, idet $\xi = a\eta$ og $\eta = \xi/a$. Legemet $\mathbb{Q}[\eta]$ omfatter altså den kvadratiske talring $\mathbb{Z}[\xi]$.

I dette kapitel betegner vi med I^* følgende delmængde af $\mathbb{Q}[\eta]$:

$$I^* := \{x - y\eta' \mid x, y \in \mathbb{Z}\}. \tag{2.1.4}$$

For tal λ i I^* er fremstillingen $\lambda = x - y\eta'$, med hele tal x, y, entydig. Par (x, y) af hele tal svarer herved bijektivt til tal $x - y\eta'$ i I^*. Af (2.1.3) følger:

Heltalsløsninger (x, y) *til ligningen* (2.1.1) *svarer bijektivt til de tal* $\lambda = x - y\eta'$ *i* I^*, *for hvilke*

$$\mathrm{N}(\lambda) = \frac{k}{a}. \tag{2.1.5}$$

Hvis diskriminanten D er positiv, er tallene i $\mathbb{Q}[\eta]$ reelle, og de kan sammenlignes mht størrelse. Det er værd at bemærke, og ikke svært at vise, at i dette tilfælde er følgende betingelser ækvivalente, med $k := F(x, y)$:

$$y > 0, \quad 2ax + by > 0, \tag{i}$$

$$|a(x - y\eta)| < a(x - y\eta'), \tag{ii}$$

$$y > 0, \quad |x - y\eta| < |x - y\eta'|, \tag{iii}$$

$$a(x - y\eta') > \sqrt{|ak|} \tag{iv}$$

Den anden ulighed i (iii) udsiger, at x/y ligger tættere ved η end ved η'.

(2.2) Ækvivalens af løsninger. Delmængden I^* beskrevet i (2.1.4) er øjensynlig en kommutativ gruppe. Den indeholder øjensynlig $\mathbb{Z}$, og den indeholder ξ, idet $\xi = -b-\xi' = -b-a\eta'$. Følgelig er $\mathbb{Z}[\xi] \subseteq I^*$. Yderligere er I^* stabil under multiplikation med elementer fra $\mathbb{Z}[\xi]$, dvs

$$\alpha \in \mathbb{Z}[\xi],\ \lambda \in I^* \implies \alpha\lambda \in I^*.$$

Det er øjensynlig nok at vise denne påstand for $\alpha = \xi$ og $\lambda = \eta'$, og her følger det af $\xi\eta' = a\eta\eta' 2 = a(c/a) = c$, jfr (2.1.2).

Denne overvejelse viser, at vi naturligt kan definere en ækvivalensrelation i mængden I^*: To tal λ og $\hat{\lambda}$ i I^* kaldes *ækvivalente*, hvis der findes en enhed ε i $\mathbb{Z}[\xi]$ med $\mathrm{N}(\varepsilon) = 1$ således, at $\hat{\lambda} = \varepsilon\lambda$, og to par af hele tal, (x, y) og $(\hat{x}, \hat{y})$, kaldes *ækvivalente*, hvis de tilsvarende tal $\lambda = x - y\eta'$ og $\hat{\lambda} = \hat{x} - \hat{y}\eta'$ i I^* er ækvivalente. Tallet $\lambda = 0$ i I^*, svarende til parret $(0, 0)$, udgør naturligvis én *triviel* ækvivalensklasse.

Ækvivalente tal i I^* har samme norm, idet $\mathrm{N}(\varepsilon\lambda) = N(\varepsilon)N(\lambda) = \mathrm{N}(\lambda)$. Af (2.1.3) følger derfor, at vi for ækvivalente talpar (x, y) og $(\hat{x}, \hat{y})$ har $F(x, y) = F(\hat{x}, \hat{y})$. Specielt ses, at hvis parret (x, y) er en løsning til ligningen (1.1.1), så vil ethvert ækvivalent par være løsning til ligningen. Løsningerne til ligningen falder altså naturligt i ækvivalensklasser.

(2.3) Bemærkning. I forbindelse med løsning af (2.1.1) kan det sædvanligvis antages, at formen $F(x, y)$ er en *primitiv form*, dvs at koefficienterne a, b, c er primiske. Lad nemlig d være den største fælles divisor for koefficienterne. Med en oplagt notation er så $F(x, y) = d\hat{F}(x, y)$, og $\hat{F}(x, y)$ er en primitiv form. Hvis d ikke går op i k, har ligningen $F(x, y) = k$ ingen heltalsløsninger. Hvis d går op i k, lad os sige $k = d\hat{k}$, er ligningen $F(x, y) = k$ ækvivalent med ligningen $\hat{F}(x, y) = \hat{k}$. Specielt har de to ligninger de samme heltalsløsninger. Bemærk, at overgangen fra $F(x, y)$ til $\hat{F}(x, y)$ ikke ændrer η, idet vi har $\hat{\eta} = \eta$, men den ændrer i almindelighed ξ, idet vi har $\xi = d\hat{\xi}$. For de tilhørende kvadratiske talringe gælder altså inklusionen,

$$\mathbb{Z}[\xi] \subseteq \mathbb{Z}[\hat{\xi}],$$

som er skarp, hvis $d > 1$. I almindelighed må vi derfor forvente, at der er flere enheder i $\mathbb{Z}[\hat{\xi}]$ end i $\mathbb{Z}[\xi]$, og dermed færre ækvivalensklasser af løsninger til $\hat{F}(x, y) = \hat{k}$.

(2.4) Det imaginære tilfælde. Antag, at diskriminanten $D = b^2 - 4ac$ er negativ. Da er tallene η og ξ ikke-reelle, normen er kvadratet på den numeriske værdi, og alle enheder i talringen $\mathbb{Z}[\xi]$ har normen $+1$. Som nævnt i Kapitel 1 er gruppen af enheder endelig, og lig med ± 1, når $D < -4$. Ækvivalens af løsninger er således ikke særlig interessant i det imaginære tilfælde. Ligningen (2.1.1) kan iøvrigt skrives $|x - y\eta|^2 = k/a$. Den har derfor ingen løsninger, hvis $k/a < 0$. Hvis $k/a > 0$, bestemmer ligningen en ellipse i (x, y)-planen, og ligningen er således kun opfyldt for en begrænset mængde af punkter. Specielt har ligningen kun endelig mange heltalsløsninger (x, y). Mere præcist kan vi om en løsning (x, y) sige følgende: Da imaginærdelen af η er $\sqrt{|D|}/(2|a|)$, er $|x - y\eta|^2 \geqslant y^2|D|/(2a)^2$, altså $k/a \geqslant y^2|D|/(4a^2)$. Forudsættes desuden, at $y \geqslant 0$, følger det, at

$$0 \leqslant y \leqslant 2\sqrt{\frac{|ak|}{|D|}}\,. \tag{2.4.1}$$

Hvis den diofantiske ligning har løsninger, ligger således i hver ækvivalensklasse en løsning (x, y), hvor y opfylder ulighederne (2.4.1).

(2.5) Det reelle tilfælde. Antag, at diskriminanten $D = b^2 - 4ac$ er positiv. Da er tallene η og ξ reelle, og I^* består af reelle tal. Enhederne ε i $\mathbb{Z}[\xi]$ med norm $\mathrm{N}(\varepsilon) = 1$ svarer til løsningerne til den egentlige Pell'ske ligning, $x^2 + bxy + acy^2 = 1$, behandlet i Kapitel 1. Der er uendelig mange sådanne enheder. Følgelig indeholder hver ækvivalensklasse uendelig mange løsninger.

Lad ε_+ være den positive grundenhed i $\mathbb{Z}[\xi]$, svarende til grundløsningen til den egentlige Pell'ske ligning. Da er $0 < \varepsilon_+ < 1$ og $1 < \varepsilon'_+ = \varepsilon_+^{-1}$. Enhederne ε i $\mathbb{Z}[\xi]$ med $\mathrm{N}(\varepsilon) = 1$ er netop tallene $\pm\varepsilon_+^p$. Ækvivalensklassen, der indeholder et givet tal $\lambda \neq 0$ i I^*, består altså af alle tal $\pm\varepsilon_+^p\lambda$, for $p \in \mathbb{Z}$. Følgelig gælder, for hvert givet positivt tal K, at der i hver ikke-triviel ækvivalensklasse af tal i I^* findes præcis ét tal λ således, at

$$K < a\lambda \leqslant K\varepsilon_+^{-1}\,. \tag{*}$$

Sætning. *Lad ε_+ være den positive grundenhed svarende til den egentlige Pell'ske ligning $x^2 + bxy + acy^2 = 1$. I hver ækvivalensklasse af heltalsløsninger (x, y) til ligningen* (2.1.1) *findes da præcis en løsning (x, y), som opfylder ulighederne,*

$$y > 0, \qquad 2ax + by > 0, \qquad 2ax + by + y\sqrt{D} \leqslant 2\sqrt{|ak|}\varepsilon_+^{-1}\,. \tag{2.5.1}$$

For en sådan løsning (x, y) gælder specielt, at

$$1 \leqslant y < 2\sqrt{\frac{|ak|}{D}} \cdot \varepsilon_+^{-1}\,. \tag{2.5.2}$$

Bevis. Ækvivalensklasser af heltalsløsninger (x, y) til (2.1.1) svarer til ækvivalensklasser af tal $\lambda = x - y\eta'$ i I^* med $\mathrm{N}(\lambda) = k/a$. Sæt $K := \sqrt{|ka|}$. I hver ækvivalensklasse, og specielt

i hver ækvivalensklasse af tal λ som opfylder $\mathrm{N}(\lambda) = k/a$, findes da præcis et tal λ, som opfylder ulighederne i (*). I hver ækvivalensklasse af løsninger (x, y) til $F(x, y) = k$ findes altså præcis en løsning (x, y) således, at ulighederne i (*) er opfyldt for $\lambda = x - y\eta'$.

Her er $K = \sqrt{|ak|}$ og $a\lambda = a(x - y\eta')$. Af ækvivalensen mellem betingelserne (i) og (iv) i (2.1) følger, at den første ulighed i (*) er opfyldt, hvis og kun hvis de to første uligheder i (2.5.1) er opfyldt. Videre er $a\lambda = ax - y\xi' = ax + y(b + \sqrt{D})/2$. Den anden ulighed i (*) er altså opfyldt, hvis og kun hvis

$$\frac{2ax + by + y\sqrt{D}}{2} \leqslant \sqrt{|ak|}\varepsilon_+^{-1},$$

altså hvis og kun hvis den tredie ulighed i (2.5.1) er opfyldt.

Øjensynlig er (2.5.2) en konsekvens af (2.5.1). □

Korollar. *Der er kun endelig mange* (*eventuelt ingen*) *ækvivalensklasser af løsninger til ligningen* (2.1.1)*, og de kan effektivt bestemmes.*

Bevis. For hvert af de endelig mange hele tal y, som opfylder (2.5.2) er der naturligvis højst to værdier af x for hvilke ligningen er opfyldt. Af Sætningen følger derfor, at der kun er endelig mange ækvivalensklasser. Mere præcist, for at bestemme én løsning i hver ækvivalensklasse, bestemmes de heltalsløsninger (x, y) til ligningen, hvor y opfylder (2.5.2) og herfra bortkastes de løsninger, som ikke opfylder de to sidste uligheder i (2.5.1). □

Eksempel. For polynomiet $x^2 - 13y^2$ er $D = 52$, $\eta = \xi = \sqrt{13}$, og den positive grundenhed er $\varepsilon_+ = 649 - 180\sqrt{13}$, jfr (1.9). I højresiden af (2.5.2) indgår faktoren $2/\sqrt{D} \cdot \varepsilon_+^{-1} = 1/\sqrt{13} \cdot (649 + 180\sqrt{13}) \approx 360$. For at indse, at den diofantiske ligning $x^2 - 13y^2 = 2$ ikke har løsninger, er det altså „nok“ at efterprøve med par (x, y), hvor $1 \leqslant y \leqslant 360\sqrt{2} \approx 509$.

Bemærkning. Det er ikke svært at vise følgende (for $D > 0$): Antag, at $|k| < \frac{1}{2}\sqrt{D}$. Betragt en løsning (x, y) til (2.1.1) i området, hvor $y > 0$ og $2ax + by > 0$. Da er x/y en Farey-approximation til η. Hvis $y \gg 0$, er x/y en konvergent mht til kædebrøksudviklingen af η.

(2.6) Opgave. Betragt en løsning (x, y) til (2.1.1), svarende til $\lambda = x - y\eta' \in I^*$, og en løsning (u, v) til den Pell'ske ligning $x^2 + bxy + acy^2 = 1$, svarende til enheden $\varepsilon = u - v\xi$ i $\mathbb{Z}[\xi]$. Hvilken løsning $(\hat{x}, \hat{y})$ til (2.1.1) svarer så til $\varepsilon\lambda$?

3. Klassetallet.

(3.1) Setup. I det følgende betragtes en kvadratisk talring $R := \mathbb{Z}[\xi]$, hvor ξ er rod i et normeret andengradspolynomium,

$$X^2 + bX + c,$$

med hele koefficienter således, at diskriminanten, $D = b^2 - 4c$, ikke er et kvadrat. Vi antager mere præcist, at roden er

$$\xi = \frac{-b + \sqrt{D}}{2},$$

valgt som den største af de to rødder i det reelle tilfælde og som roden i den øvre halvplan i det imaginære tilfælde. Som i det foregående kan par (x, y) af hele tal identificeres med tal $x - y\xi$ i R. Talringen R, som kommutativ gruppe, er altså et *gitter* med basis $(1, -\xi)$. Vi vil ofte regne i den ensorienterede basis $(\xi, 1)$:

$$R = (1, -\xi)\mathbb{Z} = (\xi, 1)\mathbb{Z}.$$

(3.2) Bemærkning. Det er værd at notere, at den kvadratiske talring $R = \mathbb{Z}[\xi]$ kun afhænger af diskriminanten D. Diskriminanten er nemlig kongruent med 0 eller 1 modulo 4. Hvis D er lige, har vi altså en ligning $D = -4c_0$, og hvis D er ulige har vi en ligning $D = 1 - 4c_0$. Vi kan altså entydigt skrive $D = b_0^2 - 4c_0$, hvor b_0 enten er 0 eller 1. Øjensynlig er $b \equiv b_0 \pmod 2$. Er $\xi_0 := (-b_0 + \sqrt{D})/2$, så er altså $\xi = \xi_0 + h$, hvor $h = (b_0 - b)/2$ er et helt tal. Det følger, at $\mathbb{Z}[\xi] = \mathbb{Z}[\xi_0]$; basen $(\xi_0, 1)$ er endda ensorienteret med basen $(\xi, 1)$.

(3.3) Lemma. *Ethvert ideal $\mathfrak{a} \neq (0)$ i $R = \mathbb{Z}[\xi]$ er et gitter i $\mathbb{C}$. Mere præcist gælder, at idealerne forskellige fra* (0) *netop er gitrene af formen,*

$$\mathfrak{a} = d(\xi - t, s)\mathbb{Z}, \tag{3.3.1}$$

hvor d, s, t er hele tal med $d \geqslant 1$, $s \neq 0$, og s er divisor i $\mathrm{N}(\xi - t)$, dvs

$$t^2 + bt + c \equiv 0 \pmod s. \tag{3.3.2}$$

Index af $\mathfrak{a}$ i R er bestemt ved ligningen,

$$|R : \mathfrak{a}| = d^2|s|. \tag{3.3.3}$$

Tallet $d \geqslant 1$ er entydigt bestemt, og s er entydigt bestemt på nær fortegn. I fremstillingen (3.3.1) *kan det antages, at $0 \leqslant t < |s|$; med denne indskrænkning er også t entydigt bestemt.*

Bevis. Som kommutativ gruppe er R fri af rang 2, med basis $(\xi, 1)$. Et ideal $\mathfrak{a} \neq (0)$ er specielt en undergruppe i R. Heraf følger som bekendt, at $\mathfrak{a}$ er en fri gruppe af rang højst 2;

mere præcist er det velkendt (og let at indse), at følgende fremgangsmåde fører til en basis for $\mathfrak{a}$: Betragt homomorfien $\mathfrak{a} \to \mathbb{Z}$ givet ved

$$u\xi + v \mapsto u. \tag{3.3.4}$$

Vi bemærker først, at der findes hele tal forskellige fra 0 i $\mathfrak{a}$. Da $\mathfrak{a} \neq (0)$, findes nemlig et tal $\alpha \neq 0$ i $\mathfrak{a}$, og så er $\mathrm{N}(\alpha) = \alpha'\alpha$ et helt tal forskelligt fra 0 i $\mathfrak{a}$. Heraf følger først, at homomorfien ikke er nul-afbildningen. Er nemlig $u \neq 0$ et helt tal i $\mathfrak{a}$, så ligger ξu i $\mathfrak{a}$, og $u\xi$ afbildes på u ved homomorfien.

Billedet ved homomorfien er således en undergruppe forskellig fra (0) i $\mathbb{Z}$, og dermed af formen $d\mathbb{Z}$, hvor $d \geqslant 1$. Tallet d er specielt billede af et tal $d\xi - t$ i $\mathfrak{a}$. Betragt kernen for homomorfien. Det er øjensynlig undergruppen $\mathfrak{a} \cap \mathbb{Z}$ af $\mathfrak{a}$. Ifølge bemærkningen er kernen altså ikke (0). Den er derfor af formen $s\mathbb{Z}$, hvor $s \neq 0$. Det følger nu, at $(d\xi - t, s)$ er en basis for $\mathfrak{a}$, altså at

$$\mathfrak{a} = (d\xi - t, s)\mathbb{Z}. \tag{3.3.5}$$

Da $\mathfrak{a}$ er et ideal, er ξs i $\mathfrak{a}$. Definitionen af d sikrer derfor, at $s \in d\mathbb{Z}$. Igen, da $\mathfrak{a}$ er et ideal, er $\xi'(d\xi - t) = cd + (-b - \xi)(-t) = t\xi + cd + bt$ i $\mathfrak{a}$. Igen sikrer definitionen af d, at $t \in d\mathbb{Z}$. Tallene s, t i (3.3.5) er altså delelige med d. Altså findes en fremstilling (3.3.1) af $\mathfrak{a}$.

Det påstås yderligere, at kongruensen i (3.3.2) er opfyldt. Da $\mathfrak{a}$ er et ideal, ligger produktet $(\xi' - t)d(\xi - t)$ i $\mathfrak{a}$. Produktet er $d\,\mathrm{N}(\xi - t) = d(t^2 + bt + c)$, og det ligger i $\mathfrak{a}$, hvis og kun hvis det er et multiplum af ds, altså hvis og kun hvis $t^2 + bt + c$ er et multiplum af s. Altså gælder kongruensen. Omvendt er det let at vise, at hvis kongruensen er opfyldt, så bestemmer højresiden af (3.3.1) et ideal.

Basen $d(\xi - t, s)$ for $\mathfrak{a}$ er bestemt ved basen $(\xi, 1)$ for R ved matrixligningen,

$$d\begin{pmatrix} \xi - t \\ s \end{pmatrix} = \begin{pmatrix} d & -dt \\ 0 & ds \end{pmatrix}\begin{pmatrix} \xi \\ 1 \end{pmatrix}. \tag{3.3.6}$$

Af Elementardivisorsætningen følger derfor, at index $|R : \mathfrak{a}|$ er den numeriske værdi af matricens determinant, og altså lig med $d^2|s|$.

I den fundne fremstilling (3.3.1) er $d \geqslant 1$. Fra den første basisvektor kan vi trække et multiplum af den anden. Herved kan vi opnå, at $0 \leqslant t < |s|$. Med denne indskrænkning følger den anførte entydighed af d, s og t let. ∎

(3.4) Definition. Talringen $R = \mathbb{Z}[\xi]$ er et gitter, og med valget af basen $(\xi, 1)$ er R endda et *orienteret* gitter. Som orienterede gitre har vi ligningerne,

$$R = (\xi, 1)\mathbb{Z} = (1, -\xi)\mathbb{Z} = (1, \xi')\mathbb{Z}.$$

Derimod bestemmer baserne $(1, \xi)$ og $(\xi', 1)$ den modsatte orientering. Af Lemma (3.3) følger specielt, at idealerne forskellige fra (0) ligeledes er gitre. Idealer forskellige fra (0) kan altså *orienteres*. For et orienteret ideal $\mathfrak{a}$ kan vi opnå ligningen,

$$\mathfrak{a} = d(\xi - t, s)\mathbb{Z}, \tag{3.4.1}$$

som en lighed mellem orienterede gitre ved blot at afpasse fortegnet på s. Herefter er $d \geqslant 1$, s, og t med $0 \leqslant t < |s|$ entydigt bestemt. Determinanten af matricen i beviset for Lemma (3.3) giver nu for det orienterede index,

$$(R : \mathfrak{a}) = d^2 s. \tag{3.4.2}$$

(3.5) Idealer af givet index. Det fremgår af (3.4), at der kun er endelig mange orienterede idealer af et givet index. Mere præcist, for at bestemme alle idealer af et givet index $k \neq 0$, betragtes de endelig mange faktoriseringer,

$$k = d^2 s, \tag{3.5.1}$$

hvor $d \geqslant 1$, og for hver værdi af s de endelig mange værdier t med $0 \leqslant t < |s|$ og

$$t^2 + bt + c \equiv 0 \pmod{s}. \tag{3.5.2}$$

De orienterede idealer af index k er så netop gitrene,

$$\mathfrak{a} = d(\xi - t, s)\mathbb{Z}. \tag{3.5.3}$$

Ofte er man mest interesseret i *primitive* idealer, dvs orienterede idealer $\mathfrak{a}$, hvor tallet d i fremstillingen (3.5.3) er lig med 1. Primitive idealer af index k svarer altså bijektivt til løsninger (modulo k) til kongruensen,

$$t^2 + bt + c \equiv 0 \pmod{k}. \tag{3.5.4}$$

For hver løsning t, er også $\tilde{t} := -b - t$ en løsning (den *konjugerede løsning*), og idealet $(\xi - \tilde{t}, k)$ har altså ligeledes index k. De to idealer $(\xi - t, k)$ og $(\xi - \tilde{t}, k)$ er i øvrigt ens, netop når $\tilde{t} \equiv t \pmod{k}$, dvs $2t + b \equiv 0 \pmod{k}$, og det medfører, at $D \equiv 0 \pmod{k}$. Det skal understreges, at hvis modulus k ikke (numerisk) er et primtal, kan kongruensen (3.5.4) have flere end to løsninger.

Øjensynlig er $4(t^2 + bt + ac) = (2t + b)^2 - D$. Hvis t er en løsning til (3.5.4), er altså $u := 2t + b$ en løsning til kongruensen,

$$u^2 \equiv D \pmod{4k}. \tag{3.5.5}$$

Omvendt, hvis u er en løsning til (3.5.5), så er specielt $u \equiv D \pmod{2}$, og dermed $u \equiv b \pmod{2}$. Altså er $u = 2t + b$ med $t \in \mathbb{Z}$, og t er en løsning til (3.5.4). Bemærk, at de to løsninger u og $u + 2k$ til (3.5.5) herved svarer til samme løsning til (3.5.4).

Hvis k er ulige, eller D er lige, eller, mere generelt, hvis kongruensen $2b_0 = b \pmod{k}$ har en løsning b_0, er kongruensen (3.5.4) øjensynlig ensbetydende med kongruensen,

$$(t + b_0)^2 \equiv b_0^2 - c \pmod{k}. \tag{3.5.6}$$

Specielt følger det, når k er ulige, at kongruensen (3.5.4) kan løses netop når D modulo k er et kvadrat, og at løsninger svarer til „kvadratrødder modulo k“ af diskriminanten D. Når modulus k numerisk er et primtal p, kan eksistensen af sådanne kvadratrødder afgøres ved hjælp af det såkaldte *reciprocitetssymbol* $\left(\frac{D}{p}\right)$, som vi behandler i et appendix.

(3.6) Eksempel. Lad os betragte de orienterede idealer af index -12 i $\mathbb{Z}[\sqrt{13}]$. Der er to faktoriseringer, $-12 = 1^2 \cdot (-12)$ og $-12 = 2^2 \cdot (-3)$. Til den første faktorisering betragtes kongruensen,

$$t^2 \equiv 13 \pmod{12}, \tag{3.6.1}$$

der er ækvivalent med kongruenserne,

$$t^2 \equiv 13 \pmod{4}, \quad t^2 \equiv 13 \pmod{3}.$$

Den sidste kongruens har løsningerne $t = 1, 2$, og kongruensen (3.6.1) har derfor løsningerne $t = 1, 5, 7, 11$. Der er altså 4 primitive idealer af index -12, med baser:

$$(\sqrt{13} - 1, -12), \quad (\sqrt{13} - 5, -12), \quad (\sqrt{13} - 7, -12), \quad (\sqrt{13} - 11, -12).$$

Desuden er der 2 ikke-primitive idealer af index -12, med baser:

$$2(\sqrt{13} - 1, -3), \quad 2(\sqrt{13} - 2, -3).$$

(3.7) Orienterede hovedidealer. Et ideal $\mathfrak{a}$ i R er som bekendt et *hovedideal*, hvis der findes et tal $\alpha \in R$ således, at

$$\mathfrak{a} = \alpha R. \tag{3.7.1}$$

Antag $\alpha \neq 0$. Da R er et orienteret gitter, er αR igen et orienteret gitter. Et orienteret ideal af formen αR kaldes et *orienteret hovedideal*. Bemærk sprogbrugen: et hovedideal $\mathfrak{a}$, som er orienteret, er ikke nødvendigvis et orienteret hovedideal; det er et krav, at der findes et tal $\alpha \in R$ således, at (3.7.1) er en lighed mellem orienterede gitre. Et sådant α kaldes også en *orienteret frembringer* for $\mathfrak{a}$.

Antag, at $\alpha = x - y\xi$ er forskellig fra 0. Det orienterede ideal αR har basis $(\alpha\xi, \alpha)$, og $\alpha\xi = (x - y\xi)\xi = (x + by)\xi + cy$. Vi har altså matrixligningen,

$$\begin{pmatrix} \alpha\xi \\ \alpha \end{pmatrix} = \begin{pmatrix} x + by & cy \\ -y & x \end{pmatrix} \begin{pmatrix} \xi \\ 1 \end{pmatrix}.$$

Determinanten af matricen er $(x + by)x + cy^2 = x^2 + bxy + cy^2$, altså netop normen af α. For det orienterede index får vi defor ligningen,

$$(R : \alpha R) = \mathrm{N}(\alpha). \tag{3.7.2}$$

For et orienteret ideal $\mathfrak{a}$ betegner vi med $\mathfrak{a}^{\mathrm{op}}$ idealet med den modsatte orientering. Naturligvis er $\mathfrak{a}$ og $\mathfrak{a}^{\mathrm{op}}$ det samme ideal, men for det orienterede index er $(R : \mathfrak{a}^{\mathrm{op}}) = -(R : \mathfrak{a})$. Antag, at $\mathfrak{a}$ er et orienteret hovedideal med orienteret frembringer α. De øvrige frembringere for idealet $\mathfrak{a}$ er da tallene $\varepsilon\alpha$, hvor ε er en enhed i R. Det følger, fx af (3.7.2), at $\mathfrak{a}^{\mathrm{op}} = \varepsilon\alpha R$, hvis og kun hvis $\mathrm{N}(\varepsilon) = -1$. Det modsat orienterede ideal $\mathfrak{a}^{\mathrm{op}}$ har altså en orienteret frembringer, hvis og kun hvis der findes en enhed $\varepsilon \in R$ med $\mathrm{N}(\varepsilon) = -1$, altså hvis og kun hvis den ikke-Pell'ske ligning har løsninger.

(3.8) Lemma. *To orienterede idealer* $\mathfrak{a}$ *og* $\mathfrak{b}$ *i* $R = \mathbb{Z}[\xi]$ *er similære, hvis og kun hvis der findes tal* α *og* β *forskellige fra* 0 *i* R *således, at*

$$\beta\mathfrak{a} = \alpha\mathfrak{b}. \tag{3.8.1}$$

Bevis. Vi bemærker først, at hvis $\mathfrak{a} \neq (0)$ er et ideal og λ er et komplekst tal således, at $\lambda\mathfrak{a} \subseteq R$, så ligger λ nødvendigvis i legemet $\mathbb{Q}[\xi]$. Idealet indeholder nemlig et tal $\alpha \neq 0$, og af den antagne inklusion følger specielt, at $\beta := \lambda\alpha$ ligger i R. Altså ligger $\lambda = \beta/\alpha$ i (brøk)legemet $\mathbb{Q}[\xi]$.

Antag nu først, at $\mathfrak{a}$ og $\mathfrak{b}$ er similære, altså at $\mathfrak{b} = \lambda\mathfrak{a}$ med et komplekst tal $\lambda \neq 0$. Ifølge bemærkningen kan vi skrive $\lambda = \beta/\alpha$, hvor α, β er tal forskellige fra 0 i R. Ligningen $\mathfrak{b} = \lambda\mathfrak{a}$ giver nu, ved multiplikation med α, ligningen (3.8.1). Omvendt er det klart, at (3.8.1) medfører, at $\mathfrak{a}$ og $\mathfrak{b}$ er similære. ∎

(3.9) Definition. Ækvivalensklasser af orienterede idealer i $R = \mathbb{Z}[\xi]$ modulo similaritet kaldes *orienterede idealklasser*. Antallet af orienterede idealklasser (vi viser om lidt, at det er endeligt) kaldes det *orienterede klassetal*, og det betegnes $h^+(D)$. Ækvivalensklasser af idealer modulo ikke-orienteret similaritet kaldes blot *idealklasser*. Antallet er *klassetallet*, betegnet $h(D)$.

I det imaginære tilfælde er orientering ikke interessant. Hvert ideal i R er nemlig et imaginært gitter, og det har derfor kanonisk en positiv og en negativ orientering. Da ξ er valgt i den øvre halvplan er basen $(\xi, 1)$ negativt orienteret; denne orientering vil for os være den *kanoniske orientering* af et imaginært gitter. Gitteret på højresiden af (3.4.1) er øjensynlig kanonisk orienteret, netop når $s > 0$. To imaginære gitre er som bekendt ikke-orienteret similære, hvis og kun hvis de med den kanoniske orientering er similære. Heraf ses, at i en given orienteret idealklasse er enten alle idealerne positivt orienterede eller alle idealerene er negativt orienterede. Hver idealklasse bestemmer således to orienterede idealklasser. I det imaginære tilfælde har vi altså trivielt ligningen,

$$h^+(D) = 2\,h(D).$$

En af de orienterede idealklasser består af de orienterede idealer, der er similære med R. Hvis $\lambda R \subseteq R$, så er specielt $\lambda 1 \in R$, altså $\lambda \in R$. Denne klasse består altså af de orienterede hovedidealer. Tilsvarende udgør hovedidealerne forskellige fra (0) én klasse. Ligningen $h(D) = 1$ udsiger altså, at alle idealer i R er hovedidealer, altså at R er et hovedidealområde.

For at bevise endeligheden af klassetallene betragtes delmængden $Q(D)$ af $\mathbb{Q}[\xi]$ bestående af alle tal af formen,

$$\frac{\xi - t}{s},$$

hvor t og $s \neq 0$ er hele tal, og s er divisor i normen af tælleren, dvs

$$t^2 + bt + c \equiv 0 \pmod{s}. \tag{3.9.1}$$

Tallene i $Q(D)$ vil vi kalde *rødder hørende til diskriminanten* D.

Det fremgår af Lemma (3.3), og (3.4), at der til hver rod $\tau = (\xi - t)/s$ hører et orienteret ideal i R, nemlig idealet $(\xi - t, s)\mathbb{Z}$; omvendt, hvis (τ, ω) er en basis som i (3.4.1) for et orienteret ideal $\mathfrak{a}$, så er $\tau/\omega = (\xi - t)/s$ en rod. Af entydighedsovervejelserne omkring d, s, t følger, at vi endda kan definere en „kanonisk rod" $\tau_\mathfrak{a}$ hørende til hvert orienteret ideal. Det følger af et velkendt (trivielt) resultat om gitre, at to orienterede idealer er similære, hvis og hvis de tilhørende rødder er similære. Tilsvarende er to idealer ikke-orienteret similære, hvis og kun hvis de tilsvarende rødder er svagt similære. Orienterede idealklasser svarer derfor bijektivt til ækvivalenklasser af rødder modulo similaritet, og idealklasser svarer bijektivt til klasser af rødder modulo svag similaritet.

Antag, at $\tau = (\xi - t)/s$. Det følger umiddelbart, at $\tau + x$, for et helt tal x, igen er en rod. Videre er

$$\frac{1}{\tau} = \frac{s}{\xi - t} = \frac{s(\xi' - t)}{\mathrm{N}(\xi - t)} = \frac{\xi + b + t}{-\mathrm{N}(\xi - t)/s}\,. \tag{3.9.2}$$

Da τ er en rod, er nævneren, $-\mathrm{N}(\xi - t)/s$, et helt tal, og tælleren $\xi + b + t = -(\xi' - t) = -(\xi - t)'$ har øjensynlig normen $N(\xi - t)$, som er delelig med nævneren $-N(\xi - t)/s$. Det reciprokke tal $1/\tau$ er altså igen en rod. Trivielt er også $-\tau$ en rod. Heraf følger specielt, at enhver kædebrøksrest τ_n af τ (både i det reelle og i det imaginære tilfælde) igen er en rod. Af Elementardivisorsætningen følger endda, at hvis S er en vilkårlig matrix i $\mathrm{GL}_2(\mathbb{Z})$, så er $S(\tau)$ en rod.

I det imaginære tilfælde kan vi nøjes med at betragte rødder τ i den øvre halvplan. Da er τ similær med enhver af sine rester τ_n (ved den irregulære kædebrøksudvikling), og blandt resterne findes én reduceret. I det reelle tilfælde er τ svagt similær med enhver af sine rester τ_n (ved den regulære kædebrøksudvikling), og similær med enhver af de „lige" rester τ_{2n}. Yderligere er rødderne τ naturligvis kvadratiske tal, og resten τ_n et reduceret kvadratiske tal, når $n \gg 0$. Endeligheden af klassetallene, både i det reelle og det imaginære tilfælde, er derfor en konsekvens af følgende:

Resultat. *Der er kun endelig mange reducerede rødder i* $Q(D)$.

I det følgende efterviser vi resultatet, og viser samtidig hvorledes klassetallene (for numerisk små værdier af D) kan beregnes.

(3.10) Det reelle tilfælde. I det reelle tilfælde kaldes et kvadratisk tal τ som bekendt reduceret, hvis $0 < \tau < 1$ og $\tau' < -1$. Antag, at tallet har formen $\tau = (\xi - t)/s$, hvor $s \neq 0$ og t er hele tal. Hvis τ er reduceret, så følger af $0 < \tau$ og $1 < -\tau'$, at $1 < \tau - \tau' = \sqrt{D}/s$, at nævneren s er positiv; herefter kan betingelserne på τ på multipliceres med s og omformes til følgende:

$$-b/2 < t < \xi, \quad \xi - t < s < t - \xi' \tag{3.10.1}$$

(Bemærk, at uligheden $-b/2 < t$ er en konsekvens af de øvrige: af $\xi - t < s < t - \xi'$ følger specielt, at $2t > \xi + \xi' = -b$). Omvendt, af (3.10.2) følger, at $0 < \xi - t < s$; altså er s positiv, og ulighederne i (3.10.2) kan omformes til betingelserne på τ. Tallet $\tau = (\xi - t)/s$ er altså reduceret, hvis og kun hvis (3.10.1) er opfyldt.

Bemærk, at ulighederne i (3.10.2), med grænserne udtrykkt ved $\sqrt{D}$, har formen,

$$-\frac{b}{2} < t < -\frac{b}{2} + \frac{\sqrt{D}}{2}, \qquad \frac{\sqrt{D}}{2} - \frac{b}{2} - t < s < \frac{\sqrt{D}}{2} + \frac{b}{2} + t\,. \tag{3.10.2}$$

Af betingelserne fremgår, at der kun er endelig mange reducerede tal af formen $\tau = (\xi - t)/s$: Den først betingelse i (3.10.1) giver endelig mange muligheder for t, og for hver af dem giver den anden betingelse endelig mange muligheder for s. For at et sådant bestemt tal $\tau = (\xi - s)/t$ skal være en rod, kræves også kongruensbetingelsen i (3.9.1), altså at s er en divisor i $t^2 + bt + c$.

I praksis bestemmes de reducerede rødder altså ved at søge de hele tal t, som opfylder den første betingelse i (3.10.1), og for enhvert sådant t at søge de positive divisorer s i $t^2 + bt + c$, som opfylder den anden betingelse.

Blandt de reducerede rødder er similaritet og svag similaritet simpelt: De reducerede rødder, som er similære med en given reduceret rod τ er netop de *lige* rester τ_{2n}. Delmængder af $Q(D)$ bestående af en reduceret rod og alle dens lige rester svarer altså bijektivt til orienterede idealklasser. Tilsvarende svarer idealklasser bijektivt til delmængder bestående af en reduceret rod og *alle* dens rester. Som bekendt er mængden af lige rester af en given reduceret rod τ er lig med mængden af alle rester af τ, hvis og kun hvis periodelængden for τ er ulige.

(3.11) Det imaginære tilfælde. I det imaginære tilfælde kan vi blot betragte rødder i den øvre halvplan. Et rod τ kaldes som bekendt reduceret, hvis den ligger i det velkendte fundamentalområde F, dvs hvis

$$-\tfrac{1}{2} \leqslant \Re e\,\tau < \tfrac{1}{2}, \qquad |\tau| \geqslant 1;$$

desuden kræves, hvis $|\tau| = 1$, at $\Re e\,\tau \leqslant 0$. For $\tau = (\xi - t)/s$ er $s \geqslant 1$, da τ ligger i den øvre halvplan. Videre er $|\xi - t|^2 = \mathrm{N}(\xi - t) = t^2 + bt + c$, så ulighederne er ensbetydende med følgende:

$$-\frac{s}{2} < t + \frac{b}{2} \leqslant \frac{s}{2}\,, \quad s^2 \leqslant t^2 + bt + c; \tag{3.11.1}$$

desuden kræves, hvis lighed gælder i den sidste ulighed, at $t + \frac{1}{2}b \geqslant 0$. Af omskrivningen $t^2 + bt + c = (t + \frac{1}{2}b)^2 - D/4$ følger, at $s^2 \leqslant s^2/4 - D/4$. Altså er

$$1 \leqslant s \leqslant \sqrt{\frac{|D|}{3}}\,, \tag{3.11.2}$$

og nu medfører de første uligheder i (3.11.1) at

$$\left|t + \frac{b}{2}\right| \leqslant \sqrt{\frac{|D|}{12}}\,. \tag{3.11.3}$$

Det er klart, at (3.11.2) og (3.11.3) kun er opfyldt for endelig mange heltalspar (s, t).

I praksis bestemmes de reducerede rødder ved at søge de tal t, som opfylder (3.11.3), og for hver af dem søge de positive divisorer s i t^2+b+c, som opfylder (3.11.1).

Blandt de reducerede rødder er similaritet blot lighed. Det orienterede klassetal $h^+(D)$ er lig med $2h(D)$, og $h(D)$ er antallet af reducerede rødder (i den øvre halvplan).

Bemærk, at betingelserne i (3.11.1) „næsten" er invariante under substitutionen $t \mapsto \tilde{t}$, hvor $\tilde{t} := -b-t$. Løsninger vil således ofte forekomme parvis: med (s,t) også $(s,\tilde{t})$. Hvis lighed gælder i en af de to „bløde" uligheder i (3.11.1), så er $(s,\tilde{t})$ dog ikke en løsning, og hvis $t+\frac{1}{2}b=0$ (her må D være lige), så er $\tilde{t}=t$.

(3.12) Eksempel. Polynomiet x^2-13 har diskriminant $D=52$, og $\xi=\sqrt{13}$. Betragt en reduceret rod $\tau=(\sqrt{13}-t)/s$. Ifølge den første ulighed i (3.10.1) er $0<t<\sqrt{13}$, altså $1\leqslant t\leqslant 3$, og for hver af disse tre værdier af t er

$$\sqrt{13}-t<s<\sqrt{13}+t.$$

Desuden er s divisor i t^2-13. For $t=1$ fås mulighederne $s=3,4$, som begge er divisorer i $1^2-13=-12$. For $t=2$ fås $s=2,3,4,5$; heraf er kun $s=3$ divisor i $2^2-13=-9$. For $t=3$ fås $s=1,\ldots,6$; heraf er kun $s=1,2,4$ divisorer i $3^2-13=-4$. Der er således 6 reducerede rødder, svarende til parrene $(3,1)$, $(4,1)$, $(3,2)$, $(1,3)$, $(2,3)$ og $(4,3)$. Det første par svarer til roden $\tau=(\sqrt{13}-1)/3$. Her finder vi,

$$\begin{aligned}
\tau=\tau_0&=\frac{\sqrt{13}-1}{3};\\
\frac{1}{\tau_0}&=\frac{3}{\sqrt{13}-1}=\frac{\sqrt{13}+1}{4}=1+\frac{\sqrt{13}-3}{4};\\
\frac{1}{\tau_1}&=\frac{4}{\sqrt{13}-3}=\frac{\sqrt{13}+3}{1}=6+\frac{\sqrt{13}-3}{1};\\
\frac{1}{\tau_2}&=\frac{1}{\sqrt{13}-3}=\frac{\sqrt{13}+3}{4}=1+\frac{\sqrt{13}-1}{4};\\
\frac{1}{\tau_3}&=\frac{4}{\sqrt{13}-1}=\frac{\sqrt{13}+1}{3}=1+\frac{\sqrt{13}-2}{3};\\
\frac{1}{\tau_4}&=\frac{3}{\sqrt{13}-2}=\frac{\sqrt{13}+2}{3}=1+\frac{\sqrt{13}-1}{3}.
\end{aligned}$$

Det ses, at $\tau_5=\tau_0$, så kædebrøksudviklingen er periodisk med periode 5. De fem rester $\tau=\tau_0,\tau_1,\tau_2,\tau_3,\tau_4$ er (nødvendigvis) fem af de seks reducerede rødder. Den sidste svarer til parret $(s,t)=(2,3)$ og er altså roden $\eta:=(\sqrt{13}-3)/2$. Her må vi nødvendigvis have $\eta=\eta_0=\eta_1=\cdots$, i overensstemmelse med udregningen,

$$\begin{aligned}
\eta=\eta_0&=\frac{\sqrt{13}-3}{2};\\
\frac{1}{\eta_0}&=\frac{2}{\sqrt{13}-3}=\frac{\sqrt{13}+3}{2}=3+\frac{\sqrt{13}-3}{2}.
\end{aligned}$$

Der er altså to klasser af svagt similære reducerede rødder. Da periodelængden i begge klasser er ulige, er de to klasser også de to klasser af similære reducerede rødder. Vi har altså $h^+(52) = h(52) = 2$. Det fremgår specielt, at $\mathbb{Z}[\sqrt{13}]$ ikke er et hovedidealområde.

(3.13) Eksempel. Polynomiet $x^2 + 6$ har diskriminant $D = -24$, og $\xi = \sqrt{-6}$. Betragt en reduceret rod $\tau = (\sqrt{-6} - t)/s$. Af uligheden (3.11.3) følger, at $|t| \leqslant \sqrt{2}$. Altså at $t = -1, 0, 1$. Videre skal s være divisor i $t^2 + 6$ og $s^2 \leqslant t^2 + 6$. For $t = 0$ fås mulighederne $s = 1, 2$, og for $t = \pm 1$ fås kun muligheden $s = 1$. Det sidste er udelukket ifølge den første ulighed i (3.11.1). De reducerede rødder er altså $\sqrt{-6}$ og $\sqrt{-6}/2$. Følgelig er $h(-24) = 2$.

(3.14) Eksempel. De lige negative diskriminanter $D = -4c$ hører til polynomierne $X^2 + c$ (hvor $b = 0$). Specielt ses, for $D = -4$ og $D = -8$, at uligheden (3.11.3) medfører $t = 0$; videre medfører den anden ulighed i (3.11.1), at $s^2 \leqslant c$ (og $c = 1, 2$). Altså er $s = 1$. For $D = -4$ og $D = -8$ er der er altså kun én reduceret rod. Følgelig er $h(-4) = 1$ og $h(-8) = 1$. Talringene $\mathbb{Z}[i]$ og $\mathbb{Z}[i\sqrt{2}]$ er altså hovedidealområder.

De ulige negative diskriminanter $D = 1 - 4c$ for $c = 1, 2, \ldots$ er diskriminanterne for polynomierne $X^2 + X + c$. Betragt specielt følgende diskriminanter,

$$-3, \quad -7, \quad -11, \quad -19, \quad -43, \quad -67, \quad -163,$$

svarende til

$$c = 1\,,\, 2\,,\, 3\,,\, 5\,,\, 11\,,\, 17\,,\, 41.$$

For hver af disse værdier af c er det let at efterprøve, at alle tallene $t^2 + t + c$, for et t der opfylder uligheden (3.11.3), er primtal. For eksempel er højresiden i (3.11.3) for $c = 41$ lig med $\sqrt{163/12} < 4$; uligheden er altså uligheden $|t + \frac{1}{2}| \leqslant \frac{7}{2}$, som giver mulighederne $t = -4, -3, -2, -1, 0, 1, 2, 3$. Værdien $t^2 + t + 41$ ændres ikke, når t erstattes med $-1 - t$. Det er altså nok at betragte $t = 0, 1, 2, 3$, hvor værdierne er $0 + 0 + 41 = 41$, $1 + 1 + 41 = 43$, $4 + 2 + 41 = 47$, og $9 + 3 + 41 = 53$, som alle er primtal. [Tilfældet $c = 1$ kræver en særbehandling!]

For hver af disse værdier af t søges en divisor s i $t^2 + t + c$, med $s^2 \leqslant t^2 + t + c$ ifølge (3.11.1). Da $t^2 + t + c$ er et primtal, må vi have $s = 1$. Herefter viser den første ulighed i (3.11.1), at $t = 0$. Der er således kun én reduceret rod, nemlig ξ. Heraf følger, for de angivne værdier af D, at $h(D) = 1$. De tilsvarende kvadratiske talringe er altså hovedidealområder. Man kan vise, at der ikke er andre imaginære kvadratiske talringe end de nævnte, som er hovedidealområder.

4. Kædebrøksmetoden.

(4.1) Setup. Vi betragter igen den diofantiske ligning,

$$F(x, y) = ax^2 + bxy + cy^2 = k, \tag{4.1.1}$$

og beskriver her en metode, som i et begrænset antal skridt leder til en bestemmelse af én løsning i hver ækvivalensklasse. I øvrigt bibeholder vi notationen fra (1.1). Tallet $\xi = a\eta$ er rod i polynomiet $X^2 + bX + ac$, og bestemmer altså en kvadratisk talring $R = \mathbb{Z}[\xi]$. Undergruppen I^* af $\mathbb{Q}[\eta]$ blev introduceret i (2.1); den består af tallene af formen,

$$\lambda = x - y\eta', \quad \text{med } x, y \in \mathbb{Z}.$$

Vi så, at løsninger (x, y) til den diofantiske ligning svarer bijektivt til tal $\lambda = x - y\eta'$ i I^*, som opfylder $\mathrm{N}(\lambda) = k/a$. Ækvivalenklasser af løsninger (x, y) svarer til ækvivalensklasser af tal λ i I^*.

Talringen R er et orienteret gitter,

$$R = (1, -\xi)\mathbb{Z} = (\xi, 1)\mathbb{Z}.$$

I det følgende betegner vi med I det orienterede gitter,

$$I := (a, -\xi)\mathbb{Z} = (\xi, a)\mathbb{Z}.$$

Øjensynlig er $I \subseteq R$. Normen af ξ er ac, som er delelig med a. Af (3.3) (eller direkte) følger, at I er et orienteret ideal i R.

(4.2) Lemma. *Lad λ være et komplekst tal forskelligt fra 0. Da gælder:*

$$\lambda \in I^* \iff \lambda I \subseteq R. \tag{4.2.1}$$

For et tal $\lambda = x - y\eta' \neq 0$ i I^* er λI et orienteret ideal i R af (orienteret) index,

$$(R : \lambda I) = F(x, y) = aN(\lambda). \tag{4.2.2}$$

Hvis koefficienterne a, b og c er primiske, da gælder for to tal $\hat{\lambda} \neq 0$ og $\lambda \neq 0$ i I^*, at de orienterede idealer $\hat{\lambda} I$ og λI er ens, hvis og kun hvis tallene $\hat{\lambda}$ og λ er ækvivalente.

Bevis. Idealet I har basen (ξ, a). Vi har derfor $\lambda I \subseteq R$, hvis og kun hvis $\lambda\xi \in R$ og $\lambda a \in R$. Hvis $\lambda I \subseteq R$, så ligger $\lambda = (\lambda a)/a$ specielt i legemet $\mathbb{Q}[\eta]$; altså findes en fremstilling,

$$\lambda = x - y\eta', \quad \text{hvor } x, y \in \mathbb{Q}. \tag{4.2.3}$$

Tallet λ ligger i I^*, hvis og kun hvis der findes en fremstilling (4.2.3) med $x, y \in \mathbb{Z}$. I beviset for (4.2.1) kan vi derfor antage, at λ er et tal af formen (4.2.3).

Da $a\eta' = \xi'$ og $\xi\xi' = ac$, er $\lambda\xi = (x - y\eta')\xi = x\xi - yc$, altså

$$\lambda\xi = x\xi - cy. \tag{4.2.4}$$

Da $a\eta' = \xi' = -b - \xi$, er $\lambda a = (x - y\eta')a = xa + yb + y\xi$, altså

$$\lambda a = y\xi + ax + by. \tag{4.2.5}$$

Tallene x og y er koefficienterne til ξ på højresiderne. Det fremgår derfor umiddelbart, at $\lambda I \subseteq R$, hvis og kun hvis tallene x og y er hele, dvs hvis og kun hvis $\lambda \in I^*$.

Antag, at $\lambda \in I^*$. Da er $\lambda I \subseteq R$, og λI er et ideal, fordi I er et ideal. Yderligere er basen $(\lambda\xi, \lambda a)$ for λI, ifølge (4.2.4) og (4.2.5), givet ud fra basen $(\xi, 1)$ for R ved matrixligningen,

$$\lambda \begin{pmatrix} \xi \\ a \end{pmatrix} = \begin{pmatrix} x & -cy \\ y & ax + by \end{pmatrix} \begin{pmatrix} \xi \\ 1 \end{pmatrix}. \tag{4.2.6}$$

Matricen har determinant $ax^2 + bxy + cy^2 = F(x, y) = aN(\lambda)$. Altså gælder (4.2.2).

Betragt nu to tal λ og $\hat{\lambda}$ forskellige fra 0 i I^*. Hvis tallene er ækvivalente, altså $\hat{\lambda} = \varepsilon\lambda$, hvor ε er en enhed med $\mathrm{N}(\varepsilon) = 1$, så er øjensynlig $\hat{\lambda}I = \lambda I$ som orienterede gitre. Antag omvendt, at $\hat{\lambda}I = \lambda I$, og at koefficienterne a, b, c er primiske. Kvotienten $\varepsilon := \hat{\lambda}/\lambda$ ligger i legemet $\mathbb{Q}[\eta] = \mathbb{Q}[\xi]$. Det er nok at vise, at $\varepsilon \in \mathbb{Z}[\xi]$, thi det følger af (4.2.2), at $\mathrm{N}(\lambda) = \mathrm{N}(\varepsilon\lambda)$, og heraf fås $\mathrm{N}(\lambda) = \mathrm{N}(\varepsilon)\,\mathrm{N}(\lambda)$, og altså $\mathrm{N}(\varepsilon) = 1$.

Da $\varepsilon \in \mathbb{Q}[\xi]$, findes en fremstilling,

$$\varepsilon = u - v\xi,$$

med rationale tal $u, v \in \mathbb{Q}$. Det skal vises, at $u, v \in \mathbb{Z}$. Ifølge antagelsen er $\varepsilon\lambda I = \lambda I$, og heraf følger $\varepsilon I = I$. Tallene $\varepsilon\xi$ og εa ligger altså i I. Vi har

$$\varepsilon\xi = u\xi - v\xi^2 = (u + bv)\xi + cva, \qquad \varepsilon a = -av\xi + ua.$$

Da $\varepsilon\xi$ og εa ligger i I, viser ligningerne, at koefficienterne, $u + bv$, cv, $-av$ og u, er hele tal. Heraf ses først, at u er et helt tal, og dernæst, at tallene av, bv og cv er hele. Da a, b, c er primiske, følger det, at v er et helt tal.

Hermed er Lemmaet bevist. ∎

(4.3) Strategi. Lemma (4.2) er basis for følgende løsningsstrategi: Vi kan antage, at $k \neq 0$, idet ligningen for $k = 0$ kun har løsningen $(x, y) = (0, 0)$. Yderligere kan det antages, at koefficienterne a, b, c er primiske, jfr (2.3). Af lemmaet følger, at der er en bijektiv forbindelse mellem løsninger til ligningen og komplekse tal λ for hvilke λI er et ideal i R, af orienteret index lig med k. Yderligere gælder, at λ og $\hat{\lambda}$ svarer til ækvivalente løsninger, hvis og kun hvis $\lambda I = \hat{\lambda} I$. Et orienteret ideal $\mathfrak{a}$ i R har formen λI, hvis og kun hvis $\mathfrak{a}$ er similært med I. Herefter er strategien følgende:

(1) *Bestem samtlige orienterede idealer* $\mathfrak{a}$ *i* R *af orienteret index* k.

(2) *Afgør for hvert af disse orienterede idealer* $\mathfrak{a}$, *om* $\mathfrak{a}$ *er similært med* I.

(3) *Bestem, når* $\mathfrak{a}$ *er similært med* I, *et tal* λ *således, at* $\mathfrak{a} = \lambda I$.

Det følger af Lemmaet, at hvert λ bestemt i (3) har formen $\lambda = x - y\eta'$, hvor (x, y) er en heltalsløsning til (4.1.1), og at der herved bestemmes en løsning i hver ækvivalensklasse. Specielt, hvis der ikke findes orienterede idealer af index k i R, eller hvis ingen orienterede idealer af index k er similære med I, har ligningen ingen løsninger.

(4.4) Første skridt. Det følger af Lemma (3.3), at for et givet $k \neq 0$ er der kun endelig mange orienterede idealer af index k. Mere præcist, for at bestemme dem betragtes de endelig mange faktoriseringer,

$$k = d^2 s, \tag{4.4.1}$$

hvor $d \geqslant 1$, og for hver værdi af s de endelig mange værdier af t med $0 \leqslant t < |s|$ og

$$t^2 + bt + ac \equiv 0 \pmod{s}. \tag{4.4.2}$$

De orienterede idealer af index k er så netop gitrene,

$$\mathfrak{a} = d(\xi - t, s)\mathbb{Z}. \tag{4.4.3}$$

Lemma. *Antag for $\lambda = x - y\eta'$ i I^*, at $\lambda I = \mathfrak{a} = d(\xi - s, t)\mathbb{Z}$. Da er d den største fælles divisor for x og y.*

Bevis. Det fremgår af (beviset for) Lemma (3.2), at tallet d, for et givet ideal $\mathfrak{a}$, er frembringeren for billedgruppen for homomorfien $\mathfrak{a} \to \mathbb{Z}$ bestemt ved $v\xi + u \mapsto v$. Af (4.2.4) og (4.2.5) følger, at for $\mathfrak{a} = \lambda I$ er billedgruppen frembragt af x og y. Billedgruppen er altså frembragt af den største fælles divisor for x og y. Hermed er Lemmaet bevist. ∎

Af lemmaet følger, at hvis idealet (4.4.3) er similært med I, så vil den tilhørende ækvivalensklasse bestå af løsninger (x, y), hvor d er den største fælles divisor for x, y. Specielt ses, at idealer af formen (4.4.3) med $d = 1$, dvs primitive idealer, svarer til eventuelle *primitive løsninger*, dvs løsninger hvor x og y er primiske. Eventuelle primitive løsninger bestemmes altså ved at løse kongruensen (4.4.2) for $s = k$,

$$t^2 + bt + ac \equiv 0 \pmod{k}. \tag{4.4.4}$$

Antag, at a er divisor i b, altså at tallet $\eta = -b/a - \eta'$ ligger i I^*. Det følger så, at for hvert tal $\lambda = x - y\eta'$ i I^* ligger også $\lambda' = x - y\eta = (x + yb/a) + y\eta'$ i I^*. Øjensynlig er $\mathrm{N}(\lambda) = \mathrm{N}(\lambda')$. For hver løsning (x, y) til ligningen er altså også $(x + yb/a, -y)$ en løsning, kaldet den *konjugerede løsning*.

Øjensynlig er $-\xi' = \xi + b = \xi + (b/a)a$. Det følger, at $I = (-\xi', a)\mathbb{Z}$. Altså er

$$\lambda' I = (-\lambda'\xi', \lambda' a)\mathbb{Z}.$$

En orienteret basis for $\lambda' I$ fås altså fra en orienteret basis for λI ved først at konjugere og dernæst skifte orientering. Hvis idealet $d(\xi - t, s)\mathbb{Z}$ svarer til en løsning (x, y), altså $d(\xi - t, s)\mathbb{Z} = \lambda I$ med $\lambda = x - y\eta'$, så fremkommer den konjugerede løsning altså fra idealet $\lambda' I = d(t - \xi', s)\mathbb{Z} = d(\xi - \tilde{t}, s)\mathbb{Z}$, hvor $\tilde{t} := -b - t$. Øjensynlig er $\tilde{t}$ den til t „konjugerede“ løsning til kongruensen (4.4.2).

(4.5) Andet skridt. Similaritet af orienterede idealer kan afgøres ud fra rødderne. Til idealet I hører roden

$$\tau_I = \xi/a = \eta,$$

og til idealet $\mathfrak{a}$ i (4.4.3) hører roden $\tau = \tau_{\mathfrak{a}} = (\xi - t)/s$. Af (3.9) fremgår, at $\mathfrak{a}$ og I er similære, hvis og kun hvis rødderne τ og η er similære. Det sidste kan afgøres ud fra kædebrøksresterne: I det reelle tilfælde er τ og η similære, hvis og kun hvis der blandt resterne findes en ligning $\tau_h = \eta_m$, hvor h og m har samme paritet. I det imaginære tilfælde er betingelsen, at τ og η har den samme reducerede rest, eller ækvivalent, at der findes en ligning $\tau_h = \eta_m$.

(4.6) Tredie skridt. Antag, at der er bestemt en matrix $U \in \mathrm{SL}_2(\mathbb{Z})$ således, at $\eta = U(\tau)$. Da er $\mathfrak{a}$ similær med I, og den søgte proportionalitetsfaktor λ bestemmes af ligningen,

$$\lambda \begin{pmatrix} \xi \\ a \end{pmatrix} = dU \begin{pmatrix} \xi - t \\ s \end{pmatrix}. \tag{4.6.1}$$

Bemærk, at λ er „overbestemt“ ved ligningen, idet λ fx kan bestemmes ved lighed mellem første- eller anden-koordinaterne. Alternativt, med $\lambda = x - y\eta'$, kan vi direkte bestemme x, y ved at sammenligne koefficienterne til ξ. På venstresiden får vi, jfr (4.2.6), søjlen $\binom{x}{y}$, og på højresiden får vi øjensynlig første søjle i matricen dU. Ligningen (4.6.1) bestemmer altså løsningen,

$$\begin{pmatrix} x \\ y \end{pmatrix} = dU \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{4.6.2}$$

Kædebrøksudviklingerne (den regulære i det reelle tilfælde og den irregulære i det imaginære tilfælde) giver også en sådan matrix U. Der findes nemlig relationer,

$$\begin{pmatrix} \tau \\ 1 \end{pmatrix} \sim T_h \begin{pmatrix} \tau_h \\ 1 \end{pmatrix}, \quad \begin{pmatrix} \eta \\ 1 \end{pmatrix} \sim S_m \begin{pmatrix} \eta_m \\ 1 \end{pmatrix}, \tag{4.6.3}$$

hvor T_h og S_m er kædebrøksmatricerne hørende til τ og η. Ligningen $\tau_h = \eta_m$ (hvor h og m har samme paritet i det reelle tilfælde) giver derfor relationen (4.6.1) med $U := S_m T_h^{-1}$, og altså formlen,

$$\begin{pmatrix} x \\ y \end{pmatrix} = dS_m T_h^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{4.6.4}$$

Matricen T_h har determinant ± 1. Det er derfor umiddelbart at bestemme den første søjle i T_h^{-1} udfra den anden række i T_h.

(4.7) Bemærkning. Det skal understreges, at bestemmelsen af løsningen (x, y) i Tredie Skridt, (4.6.2), alene bygger på en ligning $\eta = U(\tau)$, hvor $U \in \mathrm{SL}_2(\mathbb{Z})$. I (4.6.4) er en sådan matrix $U = S_m T_h^{-1}$ bestemt ved kædebrøksmatricerne for η og τ, men der kan naturligvis være andre muligheder.

I det reelle tilfælde kan man af og til skaffe sige en matrix med numerisk mindre koefficienter end dem der indgår i $S_m T_h^{-1}$. Metoden er følgende: Betragt et index m, med

$n_0 \leqslant m < n_1$, hvor $n_0 + 1$ er periodestarten og $n_1 - n_0$ er periodelængden for η. Vi har altså $\eta_{n_0} = \eta_{n_1}$. For kædebrøksmatricerne S_n for η har vi for $n \leqslant m$, at $S_n = S_m S_{m,n}$, hvor

$$S_{m,n} = \begin{pmatrix} 0 & 1 \\ 1 & x_{m+1} \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & x_n \end{pmatrix}.$$

Øjensynlig er $\begin{pmatrix} \eta_m \\ 1 \end{pmatrix} \sim S_{m,n} \begin{pmatrix} \eta_n \\ 1 \end{pmatrix}$. Da $\eta_{n_0} = \eta_{n_1}$, får vi, for $n := n_1$,

$$\begin{pmatrix} \eta \\ 1 \end{pmatrix} \sim S_{n_0} \begin{pmatrix} \eta_{n_0} \\ 1 \end{pmatrix} = S_{n_0} \begin{pmatrix} \eta_{n_1} \\ 1 \end{pmatrix} \sim S_{n_0} S_{m,n_1}^{-1} \begin{pmatrix} \eta_m \\ 1 \end{pmatrix} = \tilde{S}_m \begin{pmatrix} \eta \\ 1 \end{pmatrix}, \tag{4.7.1}$$

hvor $\tilde{S}_m := S_{n_0} S_{m,n_1}^{-1}$, altså

$$\tilde{S}_m := S_{n_0} S_{m,n_1}^{-1} = S_{n_0} \begin{pmatrix} -x_{n_1} & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} -x_{m+1} & 1 \\ 1 & 0 \end{pmatrix}.$$

Øjensynlig er det $\tilde{S}_m = (-1)^{n_0+n_1-m} = (-1)^l$ det S_m, hvor l er periodelængden af η.

Antag, at der er fundet en rest τ_h med $\eta_m = \tau_h$. Af (4.7.1) får vi, ganske som i (4.6), ligningen $\eta = U(\tau)$ med $U = \tilde{S}_m T_h^{-1}$. Heraf ses: Hvis l er lige, og h og m har samme paritet, så har matricen U determinant $+1$; i formlen (4.6.4) kan vi altså erstatte S_m med $\tilde{S}_m$. Hvis l er ulige, og h og m er af modsat paritet, så har matricen U ligeledes determinent $+1$, så også i dette tilfælde kan vi bruge (4.6.4) med $\tilde{S}_m$ for S_m. (Hvis l er lige, og h og m er af modsat paritet, kan vi hverken bruge $S_m T_h^{-1}$ eller $\tilde{S}_m T_h^{-1}$: begge matricerne har determinanten -1.)

For matricerne $\tilde{S}_m$ har vi

$$\tilde{S}_m = \begin{pmatrix} \tilde{p}_{m-1} & \tilde{p}_m \\ \tilde{q}_{m-1} & \tilde{q}_m \end{pmatrix} = \tilde{S}_{m+1} \begin{pmatrix} -x_{m+1} & 1 \\ 1 & 0 \end{pmatrix}.$$

Heraf får vi en „baglæns“ rekursionsformel for koefficienterne,

$$\tilde{p}_{i-1} = -x_{i+1}\tilde{p}_i + \tilde{p}_{i+1}, \quad \tilde{q}_{i-1} = -x_{i+1}\tilde{q}_i + \tilde{q}_{i+1}; \tag{4.7.2}$$

udgangsværdierne er $\tilde{S}_{n_1} = S_{n_0}$, dvs $(\tilde{p}_{n_1-1}, \tilde{p}_{n_1}) := (p_{n_0-1}, p_{n_0})$ og $(\tilde{q}_{n_1-1}, \tilde{q}_{n_1}) := (q_{n_0-1}, q_{n_0})$.

(4.8) Bemærkning. Som bekendt siges ringen R at være et *hovedidealområde*, hvis ethvert (uorienteret) ideal $\mathfrak{a}$ i R er af formen $\mathfrak{a} = \lambda R$ med et tal λ i R. Betingelsen er øjensynlig ækvivalent med følgende: Ethvert ideal $\mathfrak{a} \neq (0)$ er med en af sine to orienteringer similært med R. Antag, at ringen $R = \mathbb{Z}[\xi]$ er et hovedidealområde. Lad $d(\xi - t, s)$ være et ideal af index k, hvor altså $k = d^2 s$, og t tilfredstiller kongruensen,

$$t^2 + bt + ac \equiv 0 \pmod{s}.$$

Mindst et af de to orienterede idealer $\mathfrak{a} = d(\xi - t, s)\mathbb{Z}$ og $\mathfrak{a}^{\mathrm{op}} = d(\xi - t, -s)\mathbb{Z}$ vil så være similært med I. Enten vil altså det første ideal svare til en løsning til $F(x, y) = k$ eller det andet ideal (som har index $-k$) vil svare til en løsning til $F(x, y) = -k$. I det imaginære tilfælde er det klart, hvilken af de to muligheder, der foreligger. I det reelle tilfælde foreligger de begge, hvis den ikke-Pell'ske ligning $x^2 + bxy + acy^2 = -1$ har løsninger.

(4.9) Eksempel. Løs ligningen $4x^2 + 2xy - 3y^2 = 23$. Her er $D = 52$, og

$$\eta = \frac{\sqrt{13}-1}{4}, \quad \xi = \sqrt{13} - 1.$$

Der er kun én faktorisering: $23 = 1^2 \cdot 23$. Altså er $s = k = 23$. Kongruensen $t^2 + 2t - 12 = (t+1)^2 - 13 \equiv 0 \pmod{23}$ har de to løsninger $t + 1 = \pm 6$, altså $t = 5$ og $t = -7 \equiv 16$. Der er derfor 2 orienterede idealer af index 23 i $\mathbb{Z}[\xi]$, med baserne $(\xi - 5, 23)$ og $(\xi - 16, 23)$.

Andet skridt: Tallet η har kædebrøksudviklingen bestemt ved

$$\begin{aligned}
\eta &= \frac{\sqrt{13}-1}{4} && = 0 + \frac{\sqrt{13}-1}{4};\\
1/\eta_0 &= \frac{4}{\sqrt{13}-1} = \frac{\sqrt{13}+1}{3} && = 1 + \frac{\sqrt{13}-2}{3};\\
1/\eta_1 &= \frac{3}{\sqrt{13}-2} = \frac{\sqrt{13}+2}{3} && = 1 + \frac{\sqrt{13}-1}{3};\\
1/\eta_2 &= \frac{3}{\sqrt{13}-1} = \frac{\sqrt{13}+1}{4} && = 1 + \frac{\sqrt{13}-3}{4}.\\
1/\eta_3 &= \frac{4}{\sqrt{13}-3} = \frac{\sqrt{13}+3}{1} && = 6 + \frac{\sqrt{13}-3}{1};\\
1/\eta_4 &= \frac{1}{\sqrt{13}-3} = \frac{\sqrt{13}+3}{4} && = 1 + \frac{\sqrt{13}-1}{4}.
\end{aligned}$$

Det fremgår, at $\eta_5 = \eta = \eta_0$, så periodelængden er 5. For η får vi tabellen,

η :

n	-1	0	1	2	3	4	5	6
x_n		0	1	1	1	6	1	1
p_n	1	0	1	1	2	13	15	28
q_n	0	1	1	2	3	20	23	43

Det første ideal har roden $\tau = (\xi - 5)/23 = (\sqrt{13} - 6)/23$. Her fås:

$$\begin{aligned}
\tau &= \frac{\sqrt{13}-6}{23} && = -1 + \frac{\sqrt{13}+17}{23};\\
1/\tau_0 &= \frac{23}{\sqrt{13}+17} = \frac{\sqrt{13}-17}{-12} && = 1 + \frac{\sqrt{13}-5}{-12};\\
1/\tau_1 &= \frac{-12}{\sqrt{13}-5} = \frac{\sqrt{13}+5}{1} && = 8 + \frac{\sqrt{13}-3}{1}.
\end{aligned}$$

Det ses, at $\tau_2 = \eta_4$. Det første ideal er derfor similært med I.

Tredie skridt: Til bestemmelse af matricen T_2 hørende til τ har vi tabellen,

τ :

n	-1	0	1	2
x_n		-1	1	8
q_n	0	1	1	9

Nederste række i T_2 er altså (1 9), og matricen S_4 fremgår af tabellen for η. Formel (4.6.4) giver altså løsningen,

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 13 \\ 3 & 20 \end{pmatrix} \begin{pmatrix} 9 \\ -1 \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \end{pmatrix}.$$

Det andet ideal har roden $\tau = (\xi - 16)/23 = (\sqrt{13} - 17)/23$. Her fås:

$$\begin{aligned} \tau &= \frac{\sqrt{13} - 17}{23} && = -1 + \frac{\sqrt{13} + 6}{23}\,; \\ 1/\tau_0 &= \frac{23}{\sqrt{13} + 6} = \frac{\sqrt{13} - 6}{-1} && = 2 + \frac{\sqrt{13} - 4}{-1}\,; \\ 1/\tau_1 &= \frac{-1}{\sqrt{13} - 4} = \frac{\sqrt{13} + 4}{3} && = 2 + \frac{\sqrt{13} - 2}{3}\,. \end{aligned}$$

Det ses, at $\tau_2 = \eta_1 = \eta_6$. Også det andet ideal er derfor similært med I. For dette τ får vi tabellen,

τ :

n	-1	0	1	2
x_n		-1	2	2
q_n	0	1	2	5

Nederste række i T_2 er altså (2 5), og matricen S_6 fremgår af tabellen for η. Vi får løsningen,

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 15 & 28 \\ 23 & 43 \end{pmatrix} \begin{pmatrix} 5 \\ -2 \end{pmatrix} = \begin{pmatrix} 19 \\ 29 \end{pmatrix}$$

Ligningen har altså to ækvivalensklasser af løsninger, repræsenteret af $(5, 7)$ og $(19, 29)$.

(4.10) Eksempel. Løs ligningen $4x^2 + 2xy - 3y^2 = 17$. Formen er som i det foregående eksempel, hvor resterne η_n af η er bestemt. Tabellen i (4.9) kan suppleres med følgende tabel til bestemmelse af matricerne $\tilde{S}_m$, jfr (4.7):

η :

n	-1	0	1	2	3	4	5
x_n		0	1	1	1	6	1
$\tilde{p}_n$	-23	15	-8	7	-1	1	0
$\tilde{q}_n$	20	-13	7	-6	1	0	1

Nederste række er udfyldt baglæns via rekursionsformlen (4.7.2). Udgangsværdierne er matricen S_0 i søjlerne 4 og 5.

Der er kun én faktorisering: $17 = 1^2 \cdot 17$. Altså er $s = k = 17$. Kongruensen $t^2+2t-12 = (t+1)^2 - 13 \equiv 0 \pmod{17}$ har de to løsninger $t = 7$ og $t = 8$. Der er derfor 2 orienterede idealer af index 17 i $\mathbb{Z}[\xi]$, med baserne $(\xi - 7, 17)$ og $(\xi - 8, 17)$.

Det første ideal giver roden $\tau = (\xi - 7)/17 = (\sqrt{13} - 8)/17$. Her fås:

$$\begin{aligned}
\tau &= \frac{\sqrt{13}-8}{17} && = -1 + \frac{\sqrt{13}+9}{17}\,; \\
1/\tau_0 &= \frac{17}{\sqrt{13}+9} = \frac{\sqrt{13}-9}{-4} && = 1 + \frac{\sqrt{13}-5}{-4}\,; \\
1/\tau_1 &= \frac{-4}{\sqrt{13}-5} = \frac{\sqrt{13}+5}{3} && = 2 + \frac{\sqrt{13}-1}{3}\,.
\end{aligned}$$

Det ses, at $\tau_2 = \eta_2$. Det første ideal er derfor similært med I.

Matricen S_2 fremgår af tabellen for η i (4.9). For at bestemme nederste række i T_2 udfyldes tabellen svarende til roden τ,

τ :

n	-1	0	1	2
x_n		-1	1	2
q_n	0	1	1	3

Løsningsformlen (4.6.4) giver herefter, at

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Det andet ideal har roden $\tau = (\xi - 8)/17 = (\sqrt{13} - 9)/17$. Her fås:

$$\begin{aligned}
\tau &= \frac{\sqrt{13}-9}{17} && = -1 + \frac{\sqrt{13}+8}{17}\,; \\
1/\tau_0 &= \frac{17}{\sqrt{13}+8} = \frac{\sqrt{13}-8}{-3} && = 1 + \frac{\sqrt{13}-5}{-3}\,; \\
1/\tau_1 &= \frac{-3}{\sqrt{13}-5} = \frac{\sqrt{13}+5}{4} && = 2 + \frac{\sqrt{13}-3}{4}\,.
\end{aligned}$$

Det ses, at $\tau_2 = \eta_3$. Da periodelængden er ulige, er også det andet ideal altså similært med I. Matricen T_2 er den samme som for den foregående rod. I løsningsformlen anvender vi $\tilde{S}_3$, der fremgår af tabellen ovenfor for η. Vi får:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 7 & -1 \\ -6 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 22 \\ -19 \end{pmatrix}.$$

Der er altså to ækvivalensklasser af løsninger, repræsenteret af $(2, 1)$ og $(22, -19)$.

(4.11) Eksempel. Løs ligningen $4x^2 + 2xy + 3y^2 = 23$. Her er $D = -44$, og

$$\eta = \frac{\sqrt{-11} - 1}{4}, \quad \xi = \sqrt{-11} - 1.$$

Der er kun én faktorisering: $23 = 1^2 \cdot 23$. Altså er $s = k = 23$. Kongruensen $t^2 + 2t + 12 = (t+1)^2 + 11 \equiv 0 \pmod{23}$ har de to løsninger $t + 1 = \pm 9$, altså $t = 8$ og $t = 13$. Der er derfor 2 orienterede idealer af index 23, med baserne $(\xi - 8, 23)$ og $(\xi - 13, 23)$.

Tallet η ligger øjensynlig i strimmelen $-\frac{1}{2} \leqslant \Re e\, \eta < \frac{1}{2}$. Altså er $\eta = \eta_0$. Videre ligger

$$\frac{-1}{\eta} = \frac{-4}{\sqrt{-11} - 1} = \frac{\sqrt{-11} + 1}{3}$$

i fundamentalområdet F, så η_1 er den reducerede rod.

Det første ideal har roden $\tau = (\xi - 8)/23 = (\sqrt{-11} - 9)/23$. Vi finder:

$$\begin{aligned}
\tau &= \frac{\sqrt{-11} - 9}{23} && = 0 + \frac{\sqrt{-11} - 9}{23}; \\
\frac{-1}{\tau_0} &= \frac{-23}{\sqrt{-11} - 9} = \frac{\sqrt{-11} + 9}{4} && = 2 + \frac{\sqrt{-11} + 1}{4}; \\
\frac{-1}{\tau_1} &= \frac{-4}{\sqrt{-11} + 1} = \frac{\sqrt{-11} - 1}{3} && = 0 + \frac{\sqrt{-11} - 1}{3}.
\end{aligned}$$

Den reducerede rest er altså τ_2, som er forskellig fra η_1. Det første ideal er derfor ikke similært med I, og det svarer ikke til en løsning.

For det andet ideal er $\tau = (\xi - 13)/23 = (\sqrt{-11} - 14)/23$. Øjensynlig er $\tau = -1 + \tau_0$,

$$\frac{-1}{\tau_0} = \frac{-23}{\sqrt{-11} + 9} = \frac{\sqrt{-11} - 9}{4} = -2 + \frac{\sqrt{-11} - 1}{4} = -2 + \eta.$$

Heraf fremgår, at τ er similær med η. Regningerne viser, at

$$\eta = 2 + \frac{-1}{\tau_0} = 2 + \frac{-1}{\tau + 1} = \frac{2\tau + 1}{\tau + 1}.$$

Af Formel (4.6.2) får vi den tilhørende løsning,

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Ligningen har altså én ækvivalensklasse af løsninger, repræsenteret af $(2, 1)$. Samtlige løsninger er altså $(2, 1)$ og $(-2, -1)$.

(4.12) Opgave. Det fremgår af overvejelserne i (4.4), at hvis $F(x, y) = k$ har en primitiv løsning (x, y), så findes der en løsning t til kongruensen,

$$t^2 + bt + ac \equiv 0 \pmod{k}.$$

Vis, at hvis $ux + vy = 1$, så er $t := ucy - v(ax + by)$ en sådan løsning, idet vi eksplicit har $t^2 + bt + ac = F(v, -u)k$. [Vink: multiplicer (4.2.6) med $(u\ v)$, og tag normen.]

5. Similaritet af kvadratiske former.

(5.1) Setup. Et homogent andengradspolynomium,

$$F(x, y) = ax^2 + bxy + cy^2, \tag{5.1.1}$$

kaldes som bekendt også en kvadratisk form. Vi har

$$F(x, y) = (\, x \quad y \,) \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

og $F(x, y)$ siges at være den *kvadratiske form* hørende til den symmetriske matrix ovenfor. Denne matrix betegnes i det følgende ligeledes med F. Bemærk, at selvom formen F har hele koefficienter, vil matricen F kun have hele koefficienter, hvis tallet b er lige. Determinanten af matricen F er øjensynlig $ac - b^2/4$, altså lig med $-D/4$, hvor D er diskriminanten af formen F. Hvis $D > 0$, er determinanten negativ, og formen er indefinit. Hvis $D < 0$, er formen definit; den er positiv definit, hvis a er positiv, og negativ definit, hvis a er negativ.

I det følgende betragtes en fast (hel) værdi D af diskriminanten. Det forudsættes, at D er kongruent med 0 eller 1 modulo 4, og at D ikke er et kvadrat. For denne værdi af D betragtes heltalsformer F med diskriminant D. Tallene a, b, c i (5.1.1) er altså hele tal, men de er ikke faste, og specielt antager vi ikke, at $a > 0$. Svarende til D vælges en gang for alle en form $F_0 = x^2 + b_0xy + c_0y^2$ med diskriminant D, og vi sætter

$$\xi_0 = \frac{-b_0 + \sqrt{D}}{2} \; ;$$

som sædvanlig vælges kvadratroden positiv, hvis $D > 0$, og i den øvre halvplan ellers. (For ikke at forveksle ξ_0 med en kædebrøksrest, kunne vi sørge for, at ξ_0 *er* en kædebrøksrest. I det reelle tilfælde kræves hertil, at $0 < \xi_0 < 1$. Dette kan opnås med $b_0 = 2[\sqrt{D}/2]$ når D er lige, og $b_0 = 2[(\sqrt{D} - 1)/2] + 1$ når D er ulige. I det imaginære tilfælde kræves, at $-\frac{1}{2} \leqslant \Re e\, \xi_0 < \frac{1}{2}$, og dette kan opnås med $b_0 = 0$ når D er lige, og $b_0 = 1$ når D er ulige.)

Tallet ξ_0 definerer den kvadratiske talring $\mathbb{Z}[\xi_0]$ hørende til diskriminanten D, jfr (3.2). Den tilhørende mængde af rødder $Q(D)$, jfr (3.9), består af alle tal af formen,

$$\frac{\xi_0 - t}{s},$$

hvor t og $s \neq 0$ er hele tal, og s er divisor i normen af tælleren, dvs $t^2 + b_0t + c_0 \equiv 0 \pmod{s}$. Det er klart, at mængden $Q(D)$ kun afhænger af D; specielt afhænger den ikke af valget af formen F_0.

Til formen F i (5.1.1) knyttes tallet η bestemt ved

$$\eta = \frac{-b + \sqrt{D}}{2a}; \tag{5.1.2}$$

tallet η er den ene rod i polynomiet $aX^2 + bX + c$. Tallet $\xi := a\eta$ er rod i polynomiet $X^2 + bX + ac$; specielt er $\mathrm{N}(\xi) = ac$. Bemærk, at $\xi - \xi_0 = (b_0 - b)/2$ er et helt tal, så $\mathbb{Z}[\xi] = \mathbb{Z}[\xi_0]$. I brøken $\eta = \xi/a$ er nævneren altså divisor i normen af tælleren. Tallet η tilhører derfor mængden $Q(D)$ af rødder. Tallet η vil blive kaldt *roden* hørende til formen F.

(5.2) Lemma. *Ved forskriften $F \mapsto \eta$ defineres en bijektiv forbindelse mellem former F med diskriminant D og tal η i mængden $Q(D)$ af rødder hørende til diskriminanten D.*

Bevis. Ud fra en rod $\tau = (\xi_0 - t)/s$ bestemmes formen F således: Lad $\xi = \xi_0 - t$ være tælleren, og sæt

$$a := s, \quad b := -(\xi + \xi') = 2t + b_0, \quad c := \frac{N(\xi)}{s}.$$

Øjensynlig er a og b hele tal, og tallet c er helt, da τ er en rod. Videre er

$$b^2 - 4ac = (\xi + \xi')^2 - 4\xi\xi' = (\xi - \xi')^2 = (\xi_0 - \xi_0')^2 = D.$$

Heraf følger, at forskriften, der til τ knytter formen med koefficienterne a, b og c defineret ovenfor, er en afbildning „den modsatte vej". De to afbildninger er øjensynlig „hinandens inverse". Hermed er lemmaet bevist. ∎

(5.3) Definition. Lad T være en 2×2-matrix. Til hver symmetrisk 2×2-matrix F hører da en *transformeret* matrix,

$$F^T := T^{\mathrm{tr}} F T,$$

som igen er symmetrisk, med $\det F^T = (\det T)^2 \det F$. Specielt ændres determinanten ikke, hvis T har determinant ± 1. Det fremgår umiddelbart af definitionen, at hvis U er endnu en matrix, så er

$$F^{TU} = (F^T)^U. \tag{5.3.1}$$

Bemærk, at enhedsmatricen, 1, og dens modsatte, -1, begge transformerer en symmetrisk matrix F over i sig selv.

Formen hørende til den transformerede matrix er bestemt ved

$$F^T(x, y) = (x \;\; y) T^{\mathrm{tr}} F T \begin{pmatrix} x \\ y \end{pmatrix}.$$

Det ses, at den *transformerede form* $F^T(x, y)$ fremgår af formen $F(x, y)$ ved at substituere $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \mapsto T\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$. Af sammenhængen mellem determinant og diskriminant fremgår, at den transformerede form F^T har samme diskriminant som F, hvis transformationsmatricen T har determinant ± 1.

Hvis $F = \left(\begin{smallmatrix} a & b/2 \\ b/2 & c \end{smallmatrix}\right)$ og $T = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ (her er D ikke den betragtede diskriminant), har vi

$$F^T = \begin{pmatrix} A & C \\ B & D \end{pmatrix} \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

Diagonalelementet på plads $(1, 1)$ bliver

$$(\, A \quad C \,) \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} A \\ C \end{pmatrix} = F(A, C).$$

Tilsvarende er $F(B, D)$ det andet diagonalelement. Heraf følger, at hvis $F(x, y)$ er en heltalsform (hvilket var standardantagelsen fra (5.1)) og T har hele koefficienter, så har den transformerede form $F^T(x, y)$ igen hele koefficienter.

To former $\hat{F}$ og F kaldes *similære*, hvis der findes en ligning,

$$\hat{F} = F^T, \tag{5.3.2}$$

med en matrix $T \in \mathrm{SL}_2(\mathbb{Z})$, og de kaldes GL_2*-ækvivalente*, hvis der findes en ligning (5.3.2) med $T \in \mathrm{GL}_2(\mathbb{Z})$. De to former vil blive kaldt *svagt similære* eller blot *ækvivalente*, hvis

$$\hat{F} = (\det T)F^T, \quad \text{med } T \in \mathrm{GL}_2(\mathbb{Z}).$$

(5.4) Udregning. Transformation af former F (med given diskriminant) med matricer i $\mathrm{GL}_2(\mathbb{Z})$ modsvares af operationer med de tilhørende rødder η:

(1) Matricen $T = \left(\begin{smallmatrix} 1 & B \\ 0 & 1 \end{smallmatrix}\right)$ transformerer F til formen $\hat{F} := F^T$ bestemt ved

$$\hat{F}(x, y) = a(x + By)^2 + b(x + By)y + cy^2 = ax^2 + (b + 2Ba)xy + (\dots)y^2.$$

Den tilhørende rod $\hat{\eta}$ er derfor

$$\hat{\eta} = \frac{-b - 2Ba + \sqrt{D}}{2a} = \eta - B.$$

(2) Matricen $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ transformerer F til formen $\hat{F} := F^S$ bestemt ved

$$\hat{F}(x, y) = F(-y, x) = cx^2 - bxy + ay^2.$$

Den tilhørende rod er derfor

$$\hat{\eta} = \frac{b + \sqrt{D}}{2c} = \frac{b^2 - D}{2c(b - \sqrt{D})} = \frac{2a}{b - \sqrt{D}} = \frac{-1}{\eta}.$$

(3) Matricen $U = \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ transformerer F til formen $\hat{F} := F^U$ bestemt ved

$$\hat{F}(x, y) = F(-x, y) = ax^2 - bxy + cy^2.$$

Den tilhørende rod er derfor

$$\hat{\eta} = \frac{b + \sqrt{D}}{2a} = -\eta'.$$

(4) Lad os endelig bemærke, at fortegnsskift,

$$\hat{F} := -F = -ax^2 - bxy - cy^2,$$

ikke ændrer diskriminanten. For den tilhørende rod får vi

$$\hat{\eta} = \frac{b + \sqrt{D}}{-2a} = \eta'.$$

(5.5) Sætning. *Lad T være en matrix i $\mathrm{GL}_2(\mathbb{Z})$. Til transformationen $F \mapsto F^{T^{-1}}$ blandt former F, hvor der transformeres med den inverse matrix T^{-1}, svarer blandt rødderne η transformationen,*

$$\eta \mapsto \begin{cases} T(\eta), & \textit{hvis } \det T = 1, \\ T(\eta'), & \textit{hvis } \det T = -1. \end{cases}$$

Bevis. Udregningerne i (5.4) viser, at påstanden gælder, når T er en af matricerne,

$$T_B := \begin{pmatrix} 1 & B \\ 0 & 1 \end{pmatrix}, \quad S_0 := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad U_0 := \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{5.5.1}$$

Videre bemærker vi, at hvis påstanden gælder for matricer S, T, så gælder den også for produktet ST. Antag fx, at $S \in \mathrm{SL}_2(\mathbb{Z})$. Ifølge (5.3.1) er

$$F^{(ST)^{-1}} = F^{T^{-1}S^{-1}} = (F^{T^{-1}})^{S^{-1}},$$

og hertil svarer så, når determinanten $\det T$ er henholdsvis $+1$ og -1, roden,

$$S(T(\eta)) = ST(\eta) \quad \text{og} \quad S(T(\eta')) = ST(\eta').$$

Det følger af Elementardivisorsætningen, at enhver matrix S i $\mathrm{SL}_2(\mathbb{Z})$ kan skrives som et endeligt produkt af matricer af formen T_B eller S_0. Altså gælder påstanden for matricer S i $\mathrm{SL}_2(\mathbb{Z})$. Videre kan enhver matrix T i $\mathrm{GL}_2(\mathbb{Z})$ med determinant -1 skrives $T = SU_0$ med en matrix $S \in \mathrm{SL}_2(\mathbb{Z})$. Altså gælder påstanden for T. Hermed er lemmaet bevist. ∎

(5.6) Korollar. *Lad F og $\hat{F}$ være to former med diskriminant D, og lad η og $\hat{\eta}$ være de tilsvarende rødder. Da gælder: $\hat{F}$ er similær med F, hvis og kun hvis $\hat{\eta}$ er similær med η; $\hat{F}$ er GL_2-ækvivalent med F, hvis og kun hvis $\hat{\eta}$ er similær med η eller med $-\eta'$; endelig er $\hat{F}$ svagt similær med F, hvis og kun er $\hat{\eta}$ er svagt similær med η. Specielt fremhæves: $\hat{F}$ er GL_2-ækvivalent med F, hvis og kun hvis $\hat{F}$ er similær med F eller med formen $F^- = ax^2 - bxy + cy^2$; $\hat{F}$ er svagt similær med F, hvis og kun hvis $\hat{F}$ er similær med F eller med $-F^- = -ax^2 + bxy - cy^2$.*

Bevis. Den første påstand følger umiddelbart af Sætning (5.5) for matricer i $\mathrm{SL}_2(\mathbb{Z})$. Matricerne i $\mathrm{GL}_2(\mathbb{Z})$ med determinant -1 er netop matricerne $T = ST_0$, hvor T_0 er matricen i (5.5.1) og $S \in \mathrm{SL}_2(\mathbb{Z})$, og $F^{T^{-1}} = (F^{T_0^{-1}})^{S^{-1}} = (F^-)^{S^{-1}}$. Formen $\hat{F}$ er altså GL_2-ækvivalent med F, hvis og kun hvis den er similær med F eller med F^-, altså hvis og kun hvis $\hat{\eta}$ er similær med η eller med $-\eta'$.

For at bevise den tredie påstand bemærkes, at hvis $T \in \mathrm{GL}_2(\mathbb{Z})$ har determinant -1, så har formen $(\det T)F^T = -F^T$, ifølge Sætning (5.5) og Udregning (5.4)(4), roden $T(\eta')' = T(\eta)$. Heraf ses, at $\hat{F}$ er svagt similær med F, hvis og kun hvis $\hat{\eta}$ er svagt similær med η. Igen, ved at bruge fremstillingen $T = ST_0$, ses, at dette indtræffer hvis og kun hvis $\hat{F}$ er similær med F eller med $-F^-$.

Hermed er Korollaret bevist. ∎

(5.7) Definition. En form F siges at være *reduceret*, hvis den tilhørende rod η er et reduceret kvadratisk tal. I det reelle tilfælde betyder dette, at $0 < \eta < 1$ og $\eta' < -1$. I det imaginære tilfælde betyder det, at $-\frac{1}{2} \leqslant \Re e\, \eta < \frac{1}{2}$ og $|\eta| \geqslant 1$, og at $\Re e\, \eta \leqslant 0$ hvis $|\eta| = 1$.

(5.8) Korollar. *Med given diskriminant D er der kun endelig mange reducerede former (og de kan effektivt bestemmes). Enhver form er similær med en reduceret form. Specielt gælder: Antallet af ækvivalensklasser af similære former er endeligt, og lig med det orienterede klassetal $h^+(D)$. Antallet af ækvivalensklasser af svagt similære former er endeligt og lig med klassetallet $h(D)$. Antallet $h'(D)$ af ækvivalensklasser af* GL_2*-ækvivalente former er endeligt.*

Bevis. Påstandene følger af resultatet i (3.9). ∎

(5.9) Eksempel. Det er let at se, jfr (3.10), at for $D = 8$ er der kun én reduceret rod, nemlig $\tau = (\sqrt{2} - 1)/1$. Den tilhørende reducerede form er $x^2 + 2xy - y^2$. Enhver form med diskriminant 8, fx $x^2 - 2y^2$ er altså similær med $x^2 + 2xy - y^2$. Specielt er $h(8) = h^+(8) = h'(8) = 1$.

(5.10) Eksempel. For $D = 52$ har vi bestemt de 6 reducerede rødder i Eksempel (3.11). Der er 5 reducerede rødder $\tau_0, \ldots, \tau_4$ hørende til resterne af $\tau = \tau_0 = (\sqrt{13} - 1)/3$. De tilhørende former er følgende:

$$3x^2 + 2xy - 4y^2, \quad 4x^2 + 6xy - y^2, \quad x^2 + 6xy - 4y^2,$$
$$4x^2 + 2xy - 3y^2, \quad 3x^2 + 4xy - 3y^2.$$

Den sidste reducerede rod er $\eta = (\sqrt{13} - 3)/2$, svarende til formen,

$$2x^2 + 6xy - 2y^2.$$

Det følger, at $h^+(52) = 2$. Former, der er similære med den sidste form må have lige koefficienter. Formen $x^2 - 13y^2$ må derfor være similær med de fem første. På nær similaritet er der altså to former af diskriminant 52, nemlig

$$x^2 - 13y^2 \quad \text{og} \quad 2x^2 + 6xy - 2y^2.$$

Det er klart, at de to former ikke er GL_2-ækvivalente og ikke svagt similære. Vi har altså $h'(52) = h(52) = 2$.

(5.11) Eksempel. For $D = -24$ har vi bestemt de reducerede rødder i den øvre halvplan, nemlig de to rødder $\sqrt{-6}$ og $\sqrt{-6}/2$, svarende til de to former $x^2 + 6y^2$ og $2x^2 + 3y^2$. Der er tilsvarende to reducerede rødder i den nedre halvplan, og følgelig er $h^+(-24) = 4$. På nær similaritet er der altså to positive former med diskriminant -24, nemlig

$$x^2 + 6y^2 \quad \text{og} \quad 2x^2 + 3y^2.$$

Øjensynlig er $h(-24) = 2$, og $h'(-24) = 4$.

(5.12) Definition. I det følgende siges den kvadratiske form F at *fremstille* tallet $k \neq 0$, hvis den diofantiske ligning $F(x, y) = k$ har løsninger, hvor x og y er primiske hele tal. Hvis k kan fremstilles af formen F, kan k naturligvis fremstilles af enhver form, der er GL_2-ækvivalent med F.

Antag, at tallet t er en løsning til kongruensen,

$$t^2 + b_0 t + c_0 \equiv 0 \pmod{k}. \tag{5.12.1}$$

Venstresiden i kongruensen er normen $\mathrm{N}(\xi_0 - t)$. Tallet $\tau_t = (\xi_0 - t)/k$ er altså en rod, og hertil svarer en form $G = G_t$. Af (beviset for) (5.2) fremgår, at

$$G_t = kx^2 + \cdots .$$

Vi har altså $G_t(1, 0) = k$, så k fremstilles af formen G_t. Med t er også ethvert tal af formen $u = t + Bk$ en løsning til kongruensen. Øjensynlig er $\tau_u = \tau_t - B$, så formen G_u er similær med G_t. Desuden er tallet $\tilde{t} := -b - t$ en løsning (den konjugerede løsning), og $\tau_{\tilde{t}} = -\tau_t'$. Formen $G_{\tilde{t}}$ er derfor GL_2-ækvivalent med G_t.

At kongruensen (5.12.1) har løsninger, er altså en tilstrækkelig betingelse for at k kan fremstilles af en form med diskriminant D. At den også er nødvendig fremgår af det følgende resultat.

(5.13) Sætning. *En form F med diskriminant D fremstiller $k \neq 0$, hvis og kun hvis F er similær med en af formerne G_t, hvor t er en løsning til kongruensen* (5.12.1)*. Hvis $k = \pm p$, hvor p er et primtal, kan k højst fremstilles af en klasse af* GL_2*-ækvivalente former.*

Bevis. Det følger af overvejelserne i (4.3), at ligningen $F(x, y) = k$ har en løsning med primiske hele tal x, y, hvis og kun hvis det orienterede ideal $I_F = (\xi, a)\mathbb{Z}$ er similært med et primitivt orienteret ideal $\mathfrak{a}$ af (orienteret) index k i $\mathbb{Z}[\xi]$. Vi har $\mathbb{Z}[\xi] = \mathbb{Z}[\xi_0]$, og de primitive orienterede idealer af index k i $\mathbb{Z}[\xi_0]$ er netop idealerne $\mathfrak{a}_t = (\xi_0 - t, k)\mathbb{Z}$, hvor t opfylder kongruensen (5.12.1). Formen F fremstiller altså k, hvis og kun gitret $(\xi, a)\mathbb{Z}$ er similært med et af gitrene $(\xi_0 - t, k)\mathbb{Z}$, altså hvis og kun hvis kvotienten $\eta = \xi/a$ er similær med en kvotient $\tau_t = (\xi_0 - t)/k$. Det sidste indtræffer, hvis og kun hvis formen F er similær med en af formerne G_t.

Antag, at $k = \pm p$, hvor p er et primtal, og at k fremstilles af en form F med diskriminant D. Kongruensen (5.12.1) har da en løsning t, og da p er et primtal, har kongruensen modulo k højst to løsninger, nemlig med t også den konjugerede løsning $\tilde{t} = -b - t$. Altså er F similær med G_t eller med $G_{\tilde{t}}$. Da G_t og $G_{\tilde{t}}$ er GL_2-ækvivalente, følger entydigheden. □

(5.14) Bemærkning. Som nævnt i (3.5) er kongruensen (5.12.1) essentielt ækvivalent med følgende:

$$v^2 \equiv D \pmod{4k}. \tag{5.14.1}$$

Hvis k er ulige, er den sidste kongruens løsbar, hvis og kun hvis følgende kongruens er løsbar:

$$v^2 \equiv D \pmod{k}. \tag{5.14.2}$$

(Dette følger fx af Den kinesiske Restklassesætning, idet diskriminanten D modulo 4 er et kvadrat.)

Hvis $k = \pm p$, hvor p er et ulige primtal, kan løsbarhed af kongruensen (5.14.2) afgøres ved hjælp af Reciprocitetssætningen. Hvis p går op i D har kongruensen (5.14.2) altid én løsning, $v = 0$. Hvis p ikke går op i D, har kongruensen løsninger, hvis og kun hvis $\left(\frac{D}{p}\right) = 1$, altså hvis og kun hvis $\left(\frac{p}{D}\right) = 1$. Specielt bemærkes, at den sidste betingelse kun afhænger af p's restklasse modulo den givne diskriminant D.

Hvis $k = \pm 2$, har kongruensen (5.14.1) formen $v^2 \equiv D \pmod 8$. Kongruensen har altså løsninger, hvis og kun hvis D modulo 8 er kongruent med 1, 0 eller 4.

(5.15) Bemærkning. Lad $F_1, \ldots, F_h$ være repræsentanter for de endelig mange GL_2-ækvivalensklasser af former med diskriminant D. Antallet h er altså tallet $h = h'(D)$. Af Sætning (5.13) følger, at k kan fremstilles af en af formerne F_i, hvis og kun hvis kongruensen (5.14.1) har løsninger, og at formen F_i er entydigt bestemt, hvis $\pm k$ er et primtal. For numerisk små værdier af diskriminanten kan det ofte afgøres hvilken af formerne F_i, der fremstiller k.

(5.16) Sætning. *Lad p være et primtal. Da kan p fremstilles på formen,*

$$p = x^2 + 6y^2, \tag{5.16.1}$$

hvis og kun hvis p modulo 24 er kongruent med 1 eller 7, og p kan fremstilles på formen,

$$p = 2x^2 + 3y^2, \tag{5.16.2}$$

hvis og kun hvis $p = 2$ eller $p = 3$ eller p modulo 24 er kongruent med 5 eller 11. Primtal udover de nævnte kan ikke fremstilles af en form med diskriminant -24.

Bevis. De to anførte former har diskriminant -24, og $h'(-24) = 4$ ifølge Eksempel (5.11). De to former er netop repræsentanterne for de 2 klasser af positive former. Af overvejelserne i (5.14) følger derfor, at p kan fremstilles ved en af de to former, hvis og kun hvis enten p er divisor i -24, dvs $p = 2$ eller $p = 3$, eller hvis $\left(\frac{-24}{p}\right) = 1$. Da $-24 = 8 \cdot (-3)$ gælder ifølge Reciprocitetssætningen (og definitionen af Kronecker-symbolet), at

$$\left(\frac{-24}{p}\right) = \left(\frac{p}{-24}\right) = \left(\frac{p}{8}\right)\left(\frac{p}{3}\right).$$

Værdien af produktet er 1, når begge faktorer er 1, dvs når $p \equiv \pm 1 \pmod 8$ og $p \equiv 1 \pmod 3$, og når begge faktorer er -1, dvs når $p \equiv \pm 3 \pmod 8$ og $p \equiv 2 \pmod 3$.

Det er klart, at $p = 2$ og $p = 3$ fremstilles af (5.16.2) og ikke af (5.16.1). Når $p > 3$, vil p's restklasse modulo 3 afgøre, hvilken af formerne (5.16.1) og (5.16.2), der eventuelt kan fremstille p. Hvis nemlig en af formerne fremstiller p, så kan x ikke være delelig med 3. Modulo 3 følger derfor af (5.16.1), at $p \equiv x^2 \equiv 1$, og af (5.16.2), at $p \equiv 2x^2 \equiv 2$. Kombineret med de ovenfor fundne kongruensbetingelser følger påstanden i Sætningen let. ∎

(5.17) Sætning. *Lad p være et primtal. Da kan p fremstilles på formen,*

$$p = x^2 - 2y^2, \tag{5.17.1}$$

hvis og kun hvis $p = 2$ eller p modulo 8 er kongruent med 1 eller 7.

Bevis. Ifølge Eksempel (5.13) er formen i (5.17.1), på nær similaritet, den eneste form af diskriminant $D = 8$. Påstanden vises nu analogt med beviset for Sætning (5.16). De angivne ulige primtal er netop de primtal p, for hvilke $\left(\frac{p}{8}\right) = 1$. □

6. Appendix: Reciprocitetssætningen.

(6.1) Definition. Lad p være et primtal. Et helt tal a kaldes en *kvadratisk rest* modulo p, hvis a er primisk med p og kongruensen $x^2 \equiv a \pmod{p}$ har en løsning. Ofte kaldes a en *kvadratisk ikke-rest*, hvis a er primisk med p og kongruensen ikke har løsninger. Det er klart, at spørgsmålet om hvorvidt a er en kvadratisk rest modulo p kun afhænger af a's restklasse modulo p: Tallet a er kvadratisk rest, netop når a's restklasse $\overline{a}$ i $\mathbb{Z}/p$ tilhører delmængden af kvadrater på de primiske restklasser. Tilfældet $p = 2$ er uinteressant. For et ulige primtal p defineres *Legendre-symbolet*,

$$\left(\frac{a}{p}\right) := \begin{cases} 1, & \text{hvis } a \text{ er kvadratisk rest modulo } p, \\ -1, & \text{hvis } a \text{ er kvadratisk ikke-rest modulo } p, \\ 0, & \text{hvis } p \text{ går op i } a. \end{cases}$$

(6.2) Den generelle Reciprocitetssætning. *Legendre-symbolet har en udvidelse til et symbol $\left(\frac{a}{b}\right)$, defineret når nævneren b er enten en diskriminant eller ulige og positiv, med følgende egenskaber:* (1) *Værdien $\left(\frac{a}{b}\right)$ afhænger kun af restklassen af a modulo b. Værdien er* 0, *hvis a ikke er primisk med b. Som funktion af tal a, der er primiske med b, er symbolet en homomorfi $(\mathbb{Z}/b)^* \to \{\pm 1\}$.*

(2) *For en diskriminant D og et ulige positivt tal u gælder reciprocitetsformlen,*

$$\left(\frac{u}{D}\right) = \left(\frac{D}{u}\right). \tag{6.2.1}$$

(6.3) Definition. For et helt tal $b \neq 0$ kaldes en funktion $\chi : \mathbb{Z} \to \mathbb{C}$ en (restklasse-)*karakter modulo b*, hvis (1) værdien $\chi(a)$ kun afhænger af a's restklasse modulo b, (2) værdien er 0 netop hvis a ikke er primisk med b, og (3) χ er multiplikativ: $\chi(a_1 a_2) = \chi(a_1)\chi(a_2)$. Den sidste betingelse er ensbetydende med at χ, som funktion af de primiske restklasser, er en homomorfi $(\mathbb{Z}/b)^* \to \mathbb{C}^*$. Hvis værdierne kun er 0, 1, og -1, dvs hvis $\chi(a)^2 = 1$ når a er primisk med b, kaldes χ en *kvadratisk karakter*. Egenskaben (6.2)(1) udtrykker altså, at symbolet $\left(\frac{a}{b}\right)$ som funktion af a er en kvadratisk karakter modulo b.

De tilladte nævnere b i (6.2) er enten positive og ulige, eller diskriminanter, dvs tal forskellige fra 0 (eventuelt negative), som modulo 4 er kongruente med 0 eller 1. Symbolet kaldes *Jacobi-symbolet*, når nævneren b er ulige og positiv, og *Kronecker-symbolet*, når nævneren er en diskriminant.

(6.4) Bemærkning. Inden vi beviser Reciprocitetssætningen, vil vi udlede en række konsekvenser, og vi vil vise, at et symbol med egenskaberne (1) og (2) i sætningen er entydigt fastlagt ved værdien $\left(\frac{3}{-4}\right)$. Med den sidste værdi lig med $+1$ bestemmes det trivielle symbol, med værdien -1 bestemmes altså udvidelsen af Legendre symbolet.

For en ulige diskriminant D er

$$\left(\frac{a}{D}\right) = \left(\frac{a}{|D|}\right). \tag{6.4.1}$$

Ligningen er naturligvis triviel, hvis D er positiv, så vi antager $D < 0$. Tallet D er kongruent med 1 modulo 4, så ved at trække et passende multiplum af D fra a kan vi antage, at a er positiv og $a \equiv 1 \pmod 4$. Herefter er $\left(\frac{-1}{a}\right) = \left(\frac{-4}{a}\right) = \left(\frac{a}{-4}\right) = \left(\frac{1}{-4}\right) = 1$, og vi får den søgte lighed,

$$\left(\frac{a}{D}\right) = \left(\frac{D}{a}\right) = \left(\frac{-1}{a}\right)\left(\frac{-D}{a}\right) = \left(\frac{-D}{a}\right) = \left(\frac{a}{-D}\right).$$

For at vise entydigheden er det nok betragte værdierne $\left(\frac{a}{b}\right)$, når a er primisk med b. For $b = -4$ er symbolet $\left(\frac{a}{-4}\right)$, for ulige a, en homomorfi $(\mathbb{Z}/4)^* \to \pm 1$. Specielt er $\left(\frac{1}{-4}\right) = 1$, og symbolet $\left(\frac{a}{-4}\right)$ er derfor fastlagt ved værdien $\left(\frac{3}{-4}\right)$ på den anden primiske restklasse.

For $b = 8$ har vi

$$\left(\frac{3}{8}\right) = \left(\frac{8}{3}\right) = \left(\frac{-1}{3}\right) = \left(\frac{-4}{3}\right) = \left(\frac{3}{-4}\right),$$
$$\left(\frac{5}{8}\right) = \left(\frac{8}{5}\right) = \left(\frac{3}{5}\right) = \left(\frac{5}{3}\right) = \left(\frac{-1}{3}\right) = \left(\frac{3}{-4}\right).$$

Da vi videre har $\left(\frac{7}{8}\right) = \left(\frac{8}{7}\right) = \left(\frac{1}{7}\right) = 1$ og $\left(\frac{1}{8}\right) = 1$, er symbolet $\left(\frac{a}{8}\right)$ bestemt.

Betragt dernæst værdierne $\left(\frac{2}{b}\right)$. Hvis b er lige, er værdien 0. Hvis b er ulige og positiv, har vi $\left(\frac{2}{b}\right) = \left(\frac{8}{b}\right) = \left(\frac{b}{8}\right)$ som blev bestemt ovenfor. Hvis b er ulige og negativ, er b nødvendigvis en diskriminant, og derfor er $\left(\frac{2}{b}\right) = \left(\frac{2}{|b|}\right)$ ifølge (6.4.1). Symbolet $\left(\frac{2}{b}\right)$ er således bestemt i alle tilfælde.

Nu er det nemt at se, at følgende algoritme fører til bestemmelse af $\left(\frac{a}{b}\right)$. Algoritmen udnytter kun de allerede fastlagte værdier af $\left(\frac{2}{b}\right)$. Algoritmen initialiseres med $\mathbf{a} := a$ og $\mathbf{b} := b$, hvor b enten er en diskriminant, eller ulige og positiv; registret $\mathbf{s}$ indeholder, når algoritmen stopper, værdien af $\left(\frac{a}{b}\right)$.

(0) Hvis $\mathbf{a}$ og $\mathbf{b}$ begge er lige, så sæt $\mathbf{s} := 0$ og STOP; ellers sættes $\mathbf{s} := 1$.
(1) Hvis $\mathbf{b} = 1$, så STOP.
(2) Bestem den principale rest r af $\mathbf{a}$ ved division med $\mathbf{b}$, altså $\mathbf{a} = q\mathbf{b} + r$ med $0 \leqslant r < |\mathbf{b}|$. Hvis $r = 0$, så sæt $\mathbf{s} := 0$ og STOP. Ellers sættes $\mathbf{a} := r$.
(3) Faktoriser den største potens af 2: skriv $\mathbf{a} = 2^\nu u$, hvor u er ulige (og positiv). Sæt $\mathbf{a} := u$. Hvis ν er ulige, så sæt $\mathbf{s} := \mathbf{s} * \left(\frac{2}{\mathbf{b}}\right)$.
(4) Hvis $\mathbf{b} \equiv 3 \pmod 4$, så sæt $\mathbf{b} := -\mathbf{b}$.
(5) Ombyt og gentag: Sæt $\mathbf{a} := \mathbf{b}$, $\mathbf{b} := \mathbf{a}$, og GOTO (1).

Bemærk, at algoritmen, bortset fra særbehandlingen af primtallet 2, essentielt er Euklid's algoritme til bestemmelse af den største fælles divisor for a og b.

(6.5) Bemærkning. Symbolet i Reciprocitetssætningen er ikke-trivielt, idet vi for eksempel for Legendre-symbolet har $\left(\frac{2}{3}\right) = -1$. Det generelle symbol svarer altså til at værdien $\left(\frac{3}{-4}\right)$ er -1. Herefter er $\left(\frac{a}{-4}\right)$ den ikke-trivielle homomorfi $(\mathbb{Z}/4)^* \to \{\pm 1\}$. For ulige, positive tal u har vi $\left(\frac{-1}{u}\right) = \left(\frac{-4}{u}\right) = \left(\frac{u}{-4}\right)$, altså

$$\left(\frac{-1}{u}\right) = \left(\frac{u}{-4}\right) = \begin{cases} 1 & \text{hvis } u \equiv 1 \pmod 4, \\ -1 & \text{hvis } u \equiv 3 \pmod 4. \end{cases} \tag{6.5.1}$$

Videre er $\left(\frac{2}{u}\right) = \left(\frac{8}{u}\right) = \left(\frac{u}{8}\right)$, og af udregningerne i (6.4) følger, at

$$\left(\frac{2}{u}\right) = \left(\frac{u}{8}\right) = \begin{cases} 1, & \text{hvis } u \equiv \pm 1 \pmod 8, \\ -1, & \text{hvis } u \equiv \pm 3 \pmod 8. \end{cases} \tag{6.5.2}$$

Endelig fremhæver vi, at for primiske, positive, ulige tal u, v er

$$\left(\frac{v}{u}\right) = \begin{cases} \left(\frac{u}{v}\right), & \text{når } u \text{ eller } v \text{ er } \equiv 1 \pmod 4, \\ -\left(\frac{u}{v}\right), & \text{når } u \text{ og } v \text{ er } \equiv 3 \pmod 4. \end{cases} \tag{6.5.3}$$

I det første tilfælde kan vi nemlig antage, at $u \equiv 1 \pmod 4$, og så følger resultatet direkte af (6.2.1). I det andet tilfælde er $u \equiv 3 \pmod 4$. Følgelig er $-u \equiv 1 \pmod 4$, så $-u$ er en ulige diskriminant. Af (6.4.1) ses, at $\left(\frac{v}{u}\right) = \left(\frac{v}{-u}\right)$, og så er

$$\left(\frac{v}{u}\right) = \left(\frac{v}{-u}\right) = \left(\frac{-u}{v}\right) = \left(\frac{-1}{v}\right)\left(\frac{u}{v}\right) = -\left(\frac{u}{v}\right),$$

idet det sidste lighedstegn følger af (6.5.1), da $v \equiv 3 \pmod 4$.

(6.6) Eksempel. Af (6.5.2) fås $\left(\frac{2}{15}\right) = 1$, $\left(\frac{2}{7}\right) = 1$, og $\left(\frac{2}{3}\right) = -1$; algoritmen giver altså

$$\begin{aligned} \left(\frac{15}{89}\right) = \left(\frac{89}{15}\right) = \left(\frac{14}{15}\right) = \left(\frac{2}{15}\right)\left(\frac{7}{15}\right) = 1 \cdot \left(\frac{-15}{7}\right) \\ = \left(\frac{6}{7}\right) = \left(\frac{2}{7}\right)\left(\frac{3}{7}\right) = 1 \cdot \left(\frac{-7}{3}\right) = \left(\frac{2}{3}\right) = -1. \end{aligned}$$

Inddrages (6.5.1) fås mere direkte: $\left(\frac{15}{89}\right) = \left(\frac{89}{15}\right) = \left(\frac{-1}{15}\right) = -1$.

(6.7) Bemærkning. Det skal understreges, at de tre formler i (6.5) er udledt som konsekvenser af egenskaberne ved det generelle symbol $\left(\frac{a}{b}\right)$. De tre formler, for ulige primtal $u = p$ og $v = q$, udgør *Gauss's Reciprocitetsformler*. De vedrører alene Legendre-symbolet. Som vi skal se, er de essentielle i beviset for Den generelle Reciprocitetssætning. I det følgende ser vi nærmere på Legendre-symbolet, vi beviser Gauss's Reciprocitetsformler, og vi viser hvorledes det generelle symbol $\left(\frac{a}{b}\right)$ kan defineres, så at den generelle Reciprocitetssætning er opfyldt.

Vi bemærker først, for et ulige primtal p, at de kvadratiske restklasser udgør en undergruppe af index 2 i gruppen $(\mathbb{Z}/p)^*$ af primiske restklasser; specielt er det netop halvdelen af de primiske restklasser, der er kvadratiske. De kvadratiske restklasser udgør nemlig billedmængden Q ved afbildningen $(\mathbb{Z}/p)^* \to (\mathbb{Z}/p)^*$ bestemt ved $x \mapsto x^2$. Denne afbildning er øjensynlig en homomorfi, og dens kerne består af de restklasser x modulo p, som opfylder $x^2 = 1$. Da p er et ulige primtal, er denne ligning opfyldt for præcis to restklasser, nemlig 1 og -1. Kernen er derfor en undergruppe af orden 2. Det følger, at billedet Q er en undergruppe, hvis orden er halvdelen af ordenen af $(\mathbb{Z}/p)^*$. Men det betyder netop, at Q har index 2.

(6.8) Lemma. *For et ulige primtal p er Legendre-symbolet $\left(\frac{a}{p}\right)$ en ikke-triviel kvadratisk karakter modulo p. Yderligere gælder Euler's Kriterium:*

$$a^{\frac{p-1}{2}} \equiv \left(\frac{a}{p}\right) \pmod{p}. \tag{6.8.1}$$

Bevis. Som nævnt i (6.7) udgør de kvadratiske rester en undergruppe Q af index 2 i gruppen $(\mathbb{Z}/p)^*$. Kvotientgruppen af $(\mathbb{Z}/p)^*$ modulo Q har altså orden 2, og kan derfor identificeres med gruppen $\{\pm 1\}$. Under denne identifikation er Legendre-symbolet, som funktion på $(\mathbb{Z}/p)^*$, den kanoniske homomorfi på kvotienten. Altså er Legendre-symbolet en homomorfi, og den er surjektiv, og altså ikke triviel.

For at eftervise Euler's Kriterium noterer vi følgende ligning i $\mathbb{F}_p[X]$:

$$X^{p-1} - 1 = \left(X^{\frac{p-1}{2}} - 1\right)\left(X^{\frac{p-1}{2}} + 1\right).$$

Ifølge Fermat's lille sætning gælder for hvert $x \neq 0$ i $\mathbb{F}_p$, at $x^{p-1} = 1$. Hvert $x \neq 0$ er altså rod i $X^{p-1} - 1$, og dermed også rod i et af de to polynomier på højresiden. For $a \in Q$ er $a = x^2$, og altså $a^{(p-1)/2} = x^{p-1} = 1$. Hvert af de $(p-1)/2$ elementer $a \in Q$ er derfor rod i den første faktor. Da graden er $(p-1)/2$, kan den første faktor ikke have yderligere rødder. De resterende elementer i $(\mathbb{Z}/p)^*$, dvs de kvadratiske ikke-rester, må derfor være rødder i den anden faktor. Heraf følger (6.8.1). ∎

(6.9) Bemærkning. For et ulige primtal p er Legendre-symbolet $\left(\frac{a}{p}\right)$ altså en ikke-triviel karakter modulo p. Den betegnes også χ_p. Det er i øvrigt den eneste ikke-trivielle kvadratiske karakter modulo p. For en kvadratisk karakter $\chi : (\mathbb{Z}/p)^* \to \{\pm 1\}$ er jo $\chi(x^2) = \chi(x)^2 = 1$. Kernen for χ vil derfor indeholde alle kvadrater. Da kvadraterne udgør en undergruppe af index 2, vil kernen for χ altså enten bestå af kvadraterne (og så er $\chi = \chi_p$) eller den vil være hele $(\mathbb{Z}/p)^*$ (og så er $\chi = \chi_1$ den trivielle karakter modulo p).

For $n = 2$ er der kun én kvadratisk karakter modulo n, idet der kun er én primisk restklasse modulo 2. For $n = 4$ har vi to primiske restklasser, nemlig 1 og -1, så der er én ikke-triviel karakter. Det er øjensynlig karakteren defineret ved højresiden af (6.5.1); vi betegner den χ_{-4}. For $n = 8$ er der fire primiske restklasser, ± 1 og ± 3. Gruppen $(\mathbb{Z}/8)^*$ er Klein's Vier-gruppe, idet alle (ulige) kvadrater modulo 8 er kongruente med 1. Udover den trivielle karakter χ_1 er der er altså 3 ikke-trivielle karakterer modulo 8. Den ene er øjensynlig χ_{-4}. En anden er karakteren defineret ved højresiden af (6.5.2); den betegner vi χ_8. Den tredie er herefter produktet $\chi_{-4}\chi_8$, som vi betegner χ_{-8}. De fire karakterer er bestemt ved tabellen,

a	1	3	5	7
χ_1	1	1	1	1
χ_{-4}	1	-1	1	-1
χ_8	1	-1	-1	1
χ_{-8}	1	1	-1	-1

(6.10) Gauss's Lemma. *Lad p være et ulige primtal, og antag, at p ikke går op i a. Da er*

$$\left(\frac{a}{p}\right) = (-1)^n, \tag{6.10.1}$$

hvor n er antallet af negative blandt de numerisk mindste rester modulo p af tallene xa for $1 \leqslant x \leqslant (p-1)/2$.

Bevis. Tallene xa for $1 \leqslant x \leqslant (p-1)/2$ er ikke delelige med p, så deres numerisk mindste rester er tal r med $1 \leqslant |r| \leqslant (p-1)/2$. Betragt to tal x_1 og x_2 med $1 \leqslant x_1, x_2 \leqslant (p-1)/2$, og lad r_1 og r_2 være de numerisk mindste rester af x_1a og x_2a. Antag, at $|r_1| = |r_2|$. Modulo p er så $0 = r_1 \pm r_2 \equiv (x_1 \pm x_2)a$; da $|x_1 \pm x_2| \leqslant p-1$, følger det først, at $x_1 \pm x_2 = 0$, og dernæst, at $x_1 = x_2$.

De numeriske værdier af de numerisk mindste rester af tallene xa for $1 \leqslant x \leqslant (p-1)/2$ er altså forskellige. Der er $(p-1)/2$ tal og $(p-1)/2$ muligheder for de numeriske værdier. De numeriske værdier må derfor være tallene $1, 2, \ldots, (p-1)/2$. Produktet af de numerisk mindste rester er derfor $1 \cdot 2 \cdots (p-1)/2$ multipliceret med $(-1)^n$, hvor n er antallet af negative faktorer. Modulo p har vi derfor kongruensen,

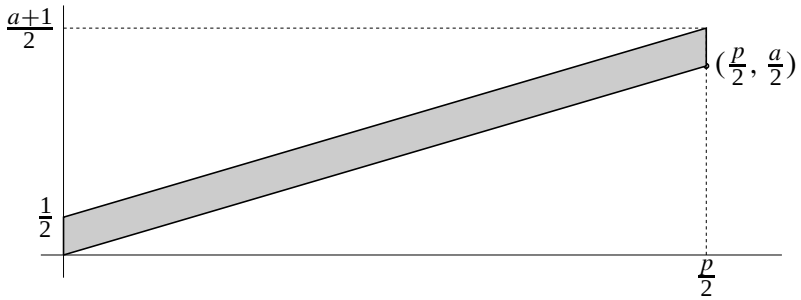
$$1a \cdot 2a \cdots \tfrac{p-1}{2}a \equiv (-1)^n 1 \cdot 2 \cdots \tfrac{p-1}{2},$$

og heraf følger $a^{(p-1)/2} \equiv (-1)^n$. Ligning (6.10.1) følger nu af Euler's Kriterium (6.8.1). ∎

(6.11) Bevis for Gauss's Reciprocitetsformler. En geometrisk fortolkning af tallet n fås på følgende måde: Øjensynlig er n antallet af tal x, med $1 \leqslant x \leqslant (p-1)/2$, for hvilke der findes et tal y med $-(p-1)/2 \leqslant xa - yp \leqslant -1$. Et sådant y er entydigt bestemt. Da p er ulige, er ulighederne for y ensbetydende med at $-p/2 < xa - yp < 0$. Tallet n er altså antallet af heltalspar (x, y) (gitterpunkter), som opfylder ulighederne,

$$0 < x < \tfrac{p}{2}, \qquad \tfrac{a}{p}x < y < \tfrac{a}{p}x + \tfrac{1}{2}.$$

Ulighederne bestemmer et parallellogram i planen:

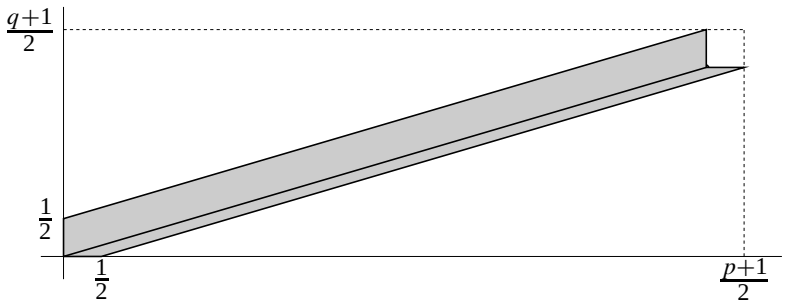


og tallet n er altså antallet af gitterpunkter i det indre af parallellogrammet.

Reciprocitetsformlen (6.5.3), for ulige primtal $p \neq q$, er ækvivalent med ligningen,

$$\left(\frac{q}{p}\right)\left(\frac{p}{q}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}. \tag{6.11.1}$$

Ifølge Gauss's Lemma er venstresiden $(-1)^{n+m}$, hvor n er antallet af gitterpunkter i parallellogrammet ovenfor, med $a := q$, og m er antallet af gitterpunkter i et tilsvarende parallellogram. Spejles dette sidste parallelogram i linien $x = y$ ses, at $n+m$ er antallet af gitterpunkter i det indre af den markerede figur herunder (da p og q er primiske, er der ingen gitterpunkter på linien fra $(0,0)$ til $(p/2, q/2)$).

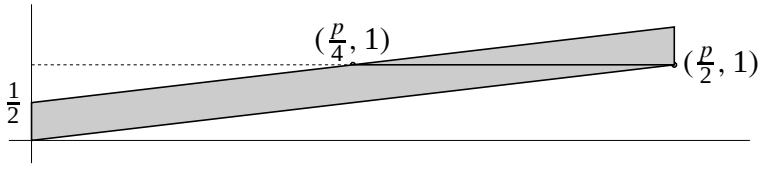


Det (åbne) rektangel består af den markerede figur, to trekanter, og et kvadrat med sidelængde $\frac{1}{2}$. I kvadratet findes ingen gitterpunkter. De to trekanter er kongruente, og indeholder derfor samme antal gitterpunkter. Modulo 2 er antallet, $n+m$, af gitterpunkter i den markerede figur altså lig med antallet af gitterpunkter i det åbne rektangel, dvs lig med $\frac{p-1}{2}\,\frac{q-1}{2}$. Heraf følger øjensynlig Formel (6.11.1).

Reciprocitetsformlen (6.5.2), for et ulige primtal p, er ligningen,

$$\left(\frac{2}{p}\right) = \chi_8(p), \tag{6.11.2}$$

hvor χ_8 er karakteren defineret ved højresiden af (6.5.2). Ifølge Gauss's Lemma er $\left(\frac{2}{p}\right) = (-1)^n$, hvor n er antallet af gitterpunkter i det indre af parallellogrammet (med $a := 2$):



I parallelogrammet er der øjensynlig kun gitterpunkter på linien, hvor $y = 1$, og antallet er $n = [\frac{p}{2}] - [\frac{p}{4}]$. Øjensynlig er

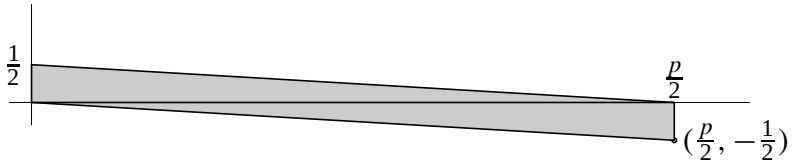
$$\left[\frac{p}{2}\right] - \left[\frac{p}{4}\right] = \begin{cases} 4h - 2h = 2h, & \text{hvis } p = 8h+1, \\ (4h-1) - (2h-1) = 2h, & \text{hvis } p = 8h-1, \\ (4h+1) - 2h = 2h+1, & \text{hvis } p = 8h+3, \\ (4h-2) - (2h-1) = 2h-1, & \text{hvis } p = 8h-3. \end{cases}$$

Heraf ses, at $(-1)^n = \chi_8(p)$, hvormed (6.11.2) er bevist.

Betragt endelig formel (6.5.1),

$$\left(\frac{-1}{p}\right) = \chi_{-4}(p). \tag{6.11.3}$$

Da p er ulige, er $\chi_{-4}(p) = (-1)^{(p-1)/2}$. Formlen følger derfor umiddelbart af Euler's Kriterium (6.8.1). Dette kriterium indgik også i beviset for Gauss's Lemma. Lad os alligevel bemærke, at Gauss's Lemma medfører (6.11.3). Vi har $\left(\frac{-1}{p}\right) = (-1)^n$, hvor n antallet af gitterpunkter i det indre af parallellogrammet (med $a := -1$):



Her er der kun gitterpunkter på linien hvor $y = 0$, og antallet er $n = [\frac{p}{2}] = (p-1)/2$.

Hermed er Gauss's reciprocitetsformler bevist. ∎

(6.12) Definition. Legendre-symbolet udvides nu til det generelle symbol $\left(\frac{a}{b}\right)$ nævnt i Sætning (6.2) på følgende måde: Et tal D kaldes en *primdiskriminant*, hvis enten $D = p$ er et primtal kongruent med 1 modulo 4, eller $D = -p$, hvor p er et primtal kongruent med 3 modulo 4, eller D er et af tallene $-4, 8, -8$. For et ulige tal u sætter vi $u^* = (-1)^{(u-1)/2}u$. De ulige primdiskriminanter er altså tallene af formen p^*, hvor p er et ulige primtal, og de lige primdiskriminanter er tallene 2^*, hvor 2^* (helt upræcist) betegner et af tallene $-4, 8, -8$.

For en primdiskriminant p^* defineres *Kronecker-symbolet* ved ligningerne,

$$\left(\frac{a}{p^*}\right) := \chi_p(a), \quad \left(\frac{a}{-4}\right) := \chi_{-4}(a), \quad \left(\frac{a}{8}\right) := \chi_8(a), \quad \left(\frac{a}{-8}\right) := \chi_{-8}(a),$$

hvor p er et ulige primtal i den første ligning. Enhver diskriminant D kan entydigt faktoriseres:

$$D = (\text{kvadrat}) \cdot p_1^* \cdots p_t^*, \tag{6.12.1}$$

hvor faktorerne p_i^* er forskellige primdiskriminanter og højst én er lige. Kronecker-symbolet $\left(\frac{a}{D}\right)$, hvor D er en diskriminant, defineres herefter som 0, hvis a ikke er primisk med D, og ellers som produktet af symbolerne $\left(\frac{a}{p_i^*}\right)$. *Jacobi-symbolet* $\left(\frac{a}{u}\right)$, hvor u er positiv og ulige, defineres tilsvarende: værdien er 0, hvis a ikke er primisk med u, og ellers som produktet af symbolerne $\left(\frac{a}{q_j}\right)$ for en „primopløsning",

$$u = (\text{kvadrat}) \cdot q_1 \cdots q_r. \tag{6.12.2}$$

Det er klart, at de to definitioner af $\left(\frac{a}{b}\right)$ stemmer overens, når b både er en diskriminant og positiv og ulige.

(6.13) Bevis for Den generelle Reciprocitetssætning. Det skal vises, at det udvidede symbol har egenskaberne (1) og (2) i (6.2). Egenskaben (1) er triviel, idet $\left(\frac{a}{b}\right)$, ud fra primopløsningen af b, er defineret som et produkt af karakterer. Betragt Reciprocitetsformlen (6.2.1). Skriv D som produkt på formen i (6.12.1), og skriv u som produkt af formen i (6.12.2). Begge sider af formlen er 0, hvis D og u ikke er primiske, så vi kan antage, at D og u er primiske. Under brug af de multiplikative egenskaber ses, at det er nok at vise formlen når $u = q$ er et ulige primtal og $D = p^*$ er en primdiskriminant. Det skal altså vises, når q ikke går op i p^*, at

$$\left(\frac{q}{p^*}\right) = \left(\frac{p^*}{q}\right).$$

Denne ligning følger let af Gauss's Reciprocitetsformler. ∎

(6.14) Tilføjelse. *Jacobi-symbolet $\left(\frac{a}{u}\right)$, for ulige positive u, er også multiplikativt i u, og der gælder formlerne:*

$$\left(\frac{-1}{u}\right) = \chi_{-4}(u) = \left(-1\right)^{(u-1)/2},$$
$$\left(\frac{2}{u}\right) = \chi_{8}(u) = \left(-1\right)^{(u^2-1)/8},$$
$$\left(\frac{u_1}{u_2}\right) = \pm\left(\frac{u_2}{u_1}\right),$$

hvor fortegnet i den sidste formel er -1, hvis $u_1 \equiv u_2 \equiv 3 \pmod 4$, og ellers $+1$.

Kronecker-symbolet $\left(\frac{a}{D}\right)$, for diskriminanter D, er også multiplikativt i D, og der gælder formlerne:

$$\left(\frac{-1}{D}\right) = \begin{cases} 1 & \text{når } D > 0, \\ -1 & \text{når } D < 0. \end{cases} \tag{6.14.1}$$

$$\left(\frac{2}{D}\right) = \chi_8(D) = \left(-1\right)^{(D^2-1)/8}, \quad \text{når } D \text{ er ulige}, \tag{6.14.2}$$

$$\left(\frac{D_1}{D_2}\right) = \pm\left(\frac{D_2}{D_1}\right), \tag{6.14.3}$$

hvor fortegnet i den sidste formel er -1, hvis D_1 og D_2 begge er negative, og ellers $+1$.

Bevis. Det følger umiddelbart af definitionen, at Jacobi-symbolet er multiplikativt i u, og formlerne for Jacobi-symbolet blev vist i (6.5).

For at vise, at Kronecker-symbolet er multiplikativt,

$$\left(\frac{a}{D_1 D_2}\right) = \left(\frac{a}{D_1}\right)\left(\frac{a}{D_2}\right), \tag{6.14.4}$$

bemærkes, at opløsningen (6.12.1) for $D_1 D_2$ fås udfra de tilsvarende opløsninger af D_1 og D_2. Det skal vises, at hver primdiskriminant p^*, som forekommer i D_1 og/eller D_2, bidrager

med samme faktor på begge sider af (6.14.4). Det er trivielt for en ulige primdiskriminant. For en lige primdiskriminant reduceres til tilfældet, hvor D_1 og D_2 er lige og forskellige primdiskriminanter. Mulighederen for D_1D_2 er så essentielt følgende:

$$(-4)\cdot 8 = (2^2)\cdot(-8), \qquad (-4)\cdot(-8) = (2^2)\cdot 8, \qquad 8\cdot -8 = 4^2\cdot(-4);$$

den påståede ligning (6.14.4) reduceres til definitionen: $\chi_{-8} = \chi_{-4}\chi_8$.

Betragt ligning (6.14.1). Begge sider er multiplikative i D, så det er nok at vise ligningen, når D er en primdiskriminant. Når $D = p \equiv 1 \pmod 4$, er begge sider 1. Når $D = -p \equiv 3 \pmod 4$, er begge sider lig med -1. Endelig, for en lige primdiskriminant følger påstanden af at $\chi_8(-1) = 1$ og $\chi_{-4}(-1) = \chi_{-8}(-1) = -1$.

I (6.14.2) er D en ulige diskriminant. Under brug af (6.4.1) får vi,

$$\Big(\frac{2}{D}\Big) = \Big(\frac{8}{D}\Big) = \Big(\frac{8}{|D|}\Big) = \Big(\frac{|D|}{8}\Big) = \chi_8(|D|) = \chi_8(D);$$

i den sidste ligning er det brugt, at $\chi_8(a) = \chi_8(-a)$ for alle a.

Endelig, i ligning (6.14.3) er begge sider 0, hvis D_1 og D_2 ikke er primiske. Antag altså, at D_1 og D_2 er primiske. Specielt er så et af tallene D_1 og D_2 ulige. Af symmetrigrunde kan vi antage, at D_2 er ulige. Under brug af (6.4.1) får vi,

$$\Big(\frac{D_1}{D_2}\Big) = \Big(\frac{D_1}{|D_2|}\Big) = \Big(\frac{|D_2|}{D_1}\Big).$$

Hvis D_2 er positiv, er dette den søgte formel. Hvis $D_2 < 0$, er højresiden lig med $\big(\frac{-D_2}{D_1}\big) = \big(\frac{-1}{D_1}\big)\big(\frac{D_2}{D_1}\big)$, og nu følger den søgte formel af (6.14.1). □

(6.15) Opgave. Vis, at følgende algoritme bestemer symbolet $\mathbf{s} = \big(\frac{a}{b}\big)$ uden at faktorisere potenser af 2. Initialiser med $\mathbf{a} := a$, $\mathbf{b} := b$.
(0) Sæt $\mathbf{s} := 1$. Hvis $\mathbf{b} \equiv 3 \pmod 4$, så sæt $\mathbf{b} := -\mathbf{b}$.
(1) Hvis $\mathbf{b} = 1$, så STOP.
(2) Bestem den principale rest r af $\mathbf{a}$ ved division med $\mathbf{b}$, altså $\mathbf{a} = q\mathbf{b} + r$ med $0 \leqslant r < |\mathbf{b}|$. Hvis $r = 0$, så sæt $\mathbf{s} := 0$ og STOP. Ellers sættes $\mathbf{a} := r$.
(3) Hvis $\mathbf{a} \equiv 3 \pmod 4$, så sæt $\mathbf{a} := -\mathbf{a}$. Hvis $\mathbf{a} \equiv 2 \pmod 4$: Hvis $\mathbf{b} > 0$, så sæt $\mathbf{a} := \mathbf{a} - \mathbf{b}$ og hvis $\mathbf{b} < 0$, så sæt $\mathbf{s} := -\mathbf{s}$ og $\mathbf{a} := -\mathbf{a} - \mathbf{b}$.
(4) Ombyt og gentag: Sæt $\mathbf{a} := \mathbf{b}$, $\mathbf{b} := \mathbf{a}$, og GOTO (1).

(6.16) Definition. I det følgende giver vi et alternativt bevis for Gauss's Reciprocitetsformler. Lad χ være en kvadratisk karakter modulo n. Lad der videre være givet et legeme L, og i L^* et element ζ, hvis orden netop er n. Under disse forudsætninger defineres den tilhørende *Gauss-sum* $G := G_{\chi,\zeta}$ som summen,

$$G := \sum_a \chi(a)\zeta^a,$$

hvor index a – her og i det følgende – gennemløber et repræsentantsystem for restklasserne primiske med n. Gauss-summen G er naturligvis element i det givne legeme L.

(6.17) Sætning. *For kvadratet på Gauss-summen $G_{\chi,\zeta}$ gælder ligningen,*

$$G_{\chi,\zeta}^2 = A_\chi,$$

hvor

$$A_\chi := \chi(-1)\sum_b \chi(b)W(b), \quad \textit{og} \quad W(b) := \sum_a \zeta^{a(b-1)};$$

begge summer er over de primiske restklasser modulo n. Leddene $W(b)$ tilhører primlegemet i L, og derfor kan de, og altså også summen A_χ, opfattes som hele tal bestemt modulo karakteristikken af L. Tallet $W(b)$ er bestemt ved ligningen,

$$W(b) = \frac{\varphi(n)}{\varphi(d)}\,\mu(d)$$

hvor $d = n/(b-1,n)$ er ordenen af ζ^{b-1} (og $\mu(d)$ er Möbius' funktion).

Bevis. Da funktionen χ er multiplikativ, følger det, at

$$G^2 = \sum_{a,b} \chi(ab)\zeta^{a+b},$$

hvor de to summationsindices a og b gennemløber de primiske restklasser modulo n. Erstattes i dobbeltsummen summationsindices a, b med $-a, ab$ ses, at

$$G^2 = \sum_{a,b} \chi(-a^2b)\zeta^{a(b-1)} = \chi(-1)\sum_b \chi(b)\sum_a \zeta^{a(b-1)} = \chi(-1)\sum_b \chi(b)W(b),$$

hvor $W(b) := \sum_a \zeta^{a(b-1)}$. Hermed er formlen for G^2 bevist.

Lad nu b være fast og primisk med n. Lad d, som i sætningen, betegne ordenen af ζ^{b-1}. Når a gennemløber restklasserne primiske med n, vil ζ^a gennemløbe de $\varphi(n)$ elementer af orden n i L^*, og $\zeta^{a(b-1)}$ vil gennemløbe de $\varphi(d)$ elementer ξ af orden d, idet hvert sådant ξ rammes $\varphi(n)/\varphi(d)$ gange. Følgelig er

$$W(b) = \frac{\varphi(n)}{\varphi(d)} \times \Big(\sum_\xi \xi\Big),$$

hvor ξ gennemløber de $\varphi(d)$ elementer af orden d i L^*. Det er klart, at disse $\varphi(d)$ elementer ξ netop er rødderne i L for cirkeldelingspolynomiet Φ_d; summen $\sum \xi$ er altså lig med -1 gange koefficienten til næsthøjestegradsleddet i Φ_d, og dermed, som det let ses, lig med $\mu(d)$. Heraf følger den i sætningen angivne formel for $W(b)$. ∎

(6.18) Udregning. Lad $n = q$ være et ulige primtal. I (6.17) antages ζ altså at have orden q, så ζ^{b-1} har også orden q med mindre $b - 1 = 0$. I undtagelsestilfældet $b = 1$ har $\zeta^0 = 1$ orden 1, og vi finder $W(1) = \varphi(q) = q - 1$. Cirkeldelingspolynomiet Φ_q er $\Phi_q = (X^q - 1)/(X - 1)$, så koefficienten til næsthøjestegradsleddet er 1. For $b \neq 1$ er altså $W(b) = -1$. Derfor er

$$A_\chi = \chi(-1)\Big(\chi(1)(q-1) - \sum_{b\neq 1}\chi(b)\Big) = \chi(-1)\Big(\chi(1)q - \sum_{b}\chi(b)\Big).$$

Her er naturligvis $\chi(1) = 1$. Hvis χ er triviel, dvs er konstant 1 på alle primiske restklasser, så er den sidste sum lig med antallet af primiske restklasser, altså lig med $q - 1$, og følgelig er $A = 1$. Antag, at χ er ikke-triviel, dvs også antager værdien -1. Da antages værdierne 1 og -1 lige mange gange på de primiske restklasser, og summen $\sum_b \chi(b)$ er derfor lig med 0. Altså gælder:

Hvis χ er en ikke-triviel karakter modulo et ulige primtal q, så er

$$A_\chi = \chi(-1)q. \tag{6.18.1}$$

Som nævnt i (6.9) findes der modulo q præcis to kvadratiske karakterer, nemlig den trivielle karakter og karakteren bestemt ved Legendre symbolet, $\chi_q(a) := \left(\frac{a}{q}\right)$. Karakteren χ_q er altså den eneste, som opfylder forudsætningerne for (6.18.1).

(6.19) Udregning. Antag, at $n = 4$ og at ζ har orden 4. De primiske restklasser b er 1 og $3 \equiv -1$. Øjensynlig er $W(1) = \zeta^0 + \zeta^0 = 2$ og $W(3) = \zeta^2 + \zeta^2 = -2$. For tallet A_χ har vi derfor

$$A_\chi = \chi(-1)\Big(\chi(1)2 - \chi(3)2\Big) = 2\chi(-1)\Big(1 - \chi(-1)\Big).$$

Hvis χ er triviel, får vi $A_\chi = 0$. Og videre:

Hvis χ er en kvadratisk karakter modulo 4, med $\chi(-1) = -1$, så er

$$A_\chi = 4\chi(-1) = -4. \tag{6.19.1}$$

Som nævnt i (6.9) er χ_{-4} den eneste karakter, som opfylder forudsætningen for (6.19.1).

(6.20) Udregning. Antag, at $n = 8$ og at ζ har orden 8. De primiske restklasser b er da 1, 3, 5 og 7, og de tilsvarende værdier af d, dvs ordenerne af ζ^0, ζ^2, ζ^4 og ζ^6, er henholdsvis 1, 4, 2 og 4. Cirkeldelingspolynomierne Φ_d for $d = 1, 2, 4$ er øjensynlig $\Phi_1 = X - 1$, $\Phi_2 = X + 1$, og $\Phi_4 = X^2 + 1$, så koefficienterne til næsthøjestegradsleddet er henholdsvis -1, 1 og 0. Følgelig bidrager kun $b = 1$ og $b = 5$ til formlen for A_χ. Det ses, at

$$A_\chi = \chi(-1)\Big(\chi(1)4 - \chi(5)4\Big) = 4\chi(-1)\Big(1 - \chi(5)\Big).$$

Tallet A_χ afhænger således kun af værdierne $\chi(-1)$ og $\chi(5)$. Hvis $\chi(5) = 1$, så er $A_\chi = 0$. Og:

Hvis χ er en kvadratisk karakter modulo 8, med $\chi(5) = -1$, så er

$$A_\chi = 8\chi(-1). \tag{6.20.1}$$

Karaktererne modulo 8 blev bestemt i (6.9). Det er netop karaktererne χ_8 og χ_{-8}, som opfylder forudsætningen for (6.20.1).

(6.21) Note. For en given kvadratisk karakter χ modulo n kan vi naturligvis som L vælge legemet $\mathbb{C}$ af komplekse tal og som ζ en primitiv n'te enhedsrod. (Som vi skal se i det følgende, er andre valg af L dog mere interessante for vores anvendelse af Sætning (6.17).) Antag, at $L = \mathbb{C}$. Gauss-summen $G = G_\zeta$ er da et komplekst tal. Dets kvadrat $A = G^2$ er ifølge sætningen et helt tal, uafhængigt af den valgte enhedsrod ζ. Hvis $A \geqslant 0$, er altså $G_\zeta = \pm\sqrt{A}$ et reelt tal og hvis $A < 0$ er $G_\zeta = \pm i\sqrt{|A|}$ rent imaginært. Fortegnet afhænger af valget af enhedsrod ζ. Vælg specielt enhedsroden $\zeta_n := \exp(2\pi i/n)$.

For $n = 4$ er $\zeta_4 = i$. Gauss-summen G svarende til karakteren χ_{-4} er så

$$G = i - i^3 = 2i,$$

i overensstemmelse med at $G^2 = -4$ ifølge (6.19.1).

For $n = 8$ er $\zeta_8 = (1+i)/\sqrt{2}$. For karakteren χ_8 finder vi for den tilhørende Gauss-sum,

$$G = \zeta_8 - \zeta_8^3 - \zeta_8^5 + \zeta_8^7 = 2\sqrt{2},$$

og for χ_{-8},

$$G = \zeta_8 + \zeta_8^3 - \zeta_8^5 - \zeta_8^7 = i2\sqrt{2},$$

begge resultater i overensstemmelse med udregningen af G^2 i (6.20.1).

Antag endelig, at $n = q$ er et ulige primtal og at $\chi(a) = \left(\frac{a}{q}\right)$ er Legendre symbolet. Af (6.18.1) følger, at $G^2 = q$ hvis q er kongruent med 1 modulo 4, og at $G^2 = -q$, hvis q er kongruent med 3. Man kan vise, for $\zeta = \zeta_q$, at der faktisk gælder ligningen,

$$G = \begin{cases} \sqrt{q} & \text{når } q \equiv 1 \pmod 4, \\ i\sqrt{q} & \text{når } q \equiv 3 \pmod 4, \end{cases}$$

men det er et dybtliggende resultat.

(6.22) Lemma. *Lad χ være en kvadratisk karakter modulo n, og lad p være et ulige primtal. Antag, at n og tallet $A := A_\chi$, defineret i Sætning* (6.17)*, er primiske med p. Da er*

$$\left(\frac{A}{p}\right) = \chi(p).$$

Bevis. Det er velkendt, at når n er primisk med p findes et legeme (endda et endeligt legeme), som har karakteristik p og indeholder et element ζ af orden n. Vi kan derfor i L betragte Gauss-summen $G = G_{\chi,\zeta}$. Af Sætning (6.17) følger, at $G^2 = A$. Heraf fås, at

$$G^p = (G^2)^{(p-1)/2}G = A^{(p-1)/2}G = \left(\frac{A}{p}\right)G,$$

hvor det sidste lighedstegn følger af Euler's Kriterium (6.8.1). På den anden side gælder, da L har karakteristik p, at afbildningen $L \to L$, bestemt ved $x \mapsto x^p$, bevarer addition og

multiplikation. Da G er en heltalslinearkombination af potenser ζ^a med koefficienter ± 1 får vi, at

$$G^p = \sum_a \chi(a)\zeta^{ap} = \chi(p)\sum_a \chi(ap)\zeta^{ap} = \chi(p)\sum_a \chi(a)\zeta^a = \chi(p)G.$$

Sammenligning af de to udtryk for G^p viser, at i legemet L gælder ligningen,

$$\Big(\frac{A}{p}\Big)G = \chi(p)G.$$

Ifølge forudsætningen er $G^2 = A$ forskellig fra 0 i L, og følgelig er $G \neq 0$. Ved division med G ses derfor, at i L er $\big(\frac{A}{p}\big) = \chi(p)$. Følgelig gælder den påståede ligning modulo karakteristikken p. Da ligningens to sider begge er ± 1, følger det, at ligningen gælder. ∎

(6.23) Bevis for Gauss's Reciprocitetsformler. Formlerne er de tre formler i (6.11). De følger af Lemmaet. Betragt nemlig først for $n = 4$ karakteren $\chi = \chi_{-4}$, og det tilhørende tal $A = A_\chi$. Ifølge Udregning (6.19) er $A = -4$. Specielt er A primisk med p. Af (6.22) fås derfor, at $\big(\frac{-4}{p}\big) = \chi_{-4}(p)$. Da Legendre-symbolet er multiplikativt, får vi $\big(\frac{-1}{p}\big) = \big(\frac{-4}{p}\big) = \chi_{-4}(p)$, som ønsket. Det skal dog straks understreges, at vi allerede har set, at dette resultat følger af Euler's Kriterium (6.8.1), og at dette kriterium indgik i beviset for Lemma (6.22).

Betragt dernæst for $n = 8$ karakteren $\chi = \chi_8$ og det tilhørende tal $A = A_\chi$. Ifølge Udregning (6.20) er $A = 8$, og altså $\big(\frac{8}{p}\big) = \chi_8(p)$. Da Legendre-symbolet er multiplikativt, får vi, som ønsket,

$$\Big(\frac{2}{p}\Big) = \Big(\frac{8}{p}\Big) = \chi_8(p).$$

Betragt endelig, for et ulige primtal $q \neq p$, karakteren $\chi(a) = \chi_q(a) = \big(\frac{a}{q}\big)$, og det tilhørende tal $A = A_\chi$. Ifølge Udregning (6.18) er $A = \chi(-1)q = (-1)^{(q-1)/2}q$, idet vi allerede har vist, at $\chi(-1) = \big(\frac{-1}{q}\big) = (-1)^{(q-1)/2}$. Idet $q \neq p$, er $A \not\equiv 0 \pmod p$. Vi får

$$\Big(\frac{(-1)^{\frac{q-1}{2}}q}{p}\Big) = \Big(\frac{A}{p}\Big) = \chi_q(p) = \Big(\frac{p}{q}\Big).$$

Da $\big(\frac{-1}{p}\big) = (-1)^{(p-1)/2}$, følger den ønskede ligning (6.11.1). ∎

(6.24) Bemærkning. Som en anvendelse af reciprocitetssætningen viser vi følgende om Mersenne-tallene $M_q = 2^q - 1$.

Sætning. *Lad q være et ulige primtal. Da er $2q+1$ divisor i M_q, hvis og kun hvis $2q+1$ er et primtal og $q \equiv 3 \pmod 4$.*

Bevis. Sæt $p := 2q + 1$. Da er p ulige, $p > 3$, og $p - 1 = 2q$.

„kun hvis": Antag, at $p \mid M_q$, altså $p \mid 2^q - 1$. Idet vi regner modulo p, er altså $2^q \equiv 1$. Følgelig er $(-2)^q \equiv -1$, og dermed er $(-2)^{2q} \equiv 1$. Den sidste kongruens viser, at modulo

p har -2 en orden, som er divisor i $2q$, og den første viser, at orden ikke kan være q. Da q er et primtal, og vi trivielt har $(-2)^2 \not\equiv 1$, følger det, at restklassen af -2 i gruppen $(\mathbb{Z}/p)^*$ har orden $2q$. Denne gruppe indeholder altså (mindst) $2q = p - 1$ elementer. Restklassen af 0 er derfor den eneste restklasse i $\mathbb{Z}/p$, som ikke er invertibel. Altså må p være et primtal. Yderligere følger det af kongruensen $(-2)^q \equiv -1$, at -2 ikke kan være et kvadrat modulo p. Værdien af Legendre symbolet $\left(\frac{-2}{p}\right)$ er altså -1. Modulo 8 er p derfor kongruent med 5 eller 7. Da $p = 2q + 1$, med q ulige, følger det, at $q \equiv 3 \pmod 4$.

„hvis“: Antag, at p er et primtal og at $q \equiv 3 \pmod 4$. Da er $p \equiv 7 \pmod 8$. Følgelig er 2 et kvadrat modulo p, altså $2 \equiv x^2 \pmod p$. Heraf ses, at $2^q = x^{2q} \equiv 1 \pmod p$. Følgelig går p op i $2^q - 1$. ∎

Af sætningen følger, at Mersenne-tallene $M_{11}, M_{23}, M_{83}, \ldots$ er delelige med, henholdsvis, 23, 47, 167, ... Det første sammensatte Mersenne-tal, som ikke står i denne liste er i øvrigt M_{29}.

Farey-brøker og kædebrøker

1. Farey-brøker.

(1.1) Definition. Lad N være et naturligt tal. Ved *Farey-brøkerne* af *orden* N forstås de rationale tal af formen a/s, hvor a og s er hele tal og $1 \leqslant s \leqslant N$. Mængden af Farey-brøker af orden N betegnes $\mathcal{F}_N$. Når vi taler om en brøk a/s i $\mathcal{F}_N$ vil det næsten altid være underforstået, at brøken er udforkortelig, altså at a og s er primiske, og at $s \geqslant 1$. Brøkerne i $\mathcal{F}_N$ kan naturligt ordnes i en følge: Hver Farey-brøk af orden N har en *forgænger* og en *efterfølger*, idet afstanden mellem to forskellige brøker a/s og b/t i $\mathcal{F}_N$ er

$$\left|\frac{a}{s} - \frac{b}{t}\right| = \frac{|ta - sb|}{st} \geqslant \frac{1}{st} \geqslant \frac{1}{N^2}.$$

Det er klart, at Farey-brøkerne af orden N kan identificeres med de linier gennem $(0, 0)$ i planen, som går gennem et *gitterpunkt* (x, y) (dvs et punkt med heltalskoordinater) i strimmelen bestemt ved $1 \leqslant y \leqslant N$. Bemærk, at linien gennem punktet (a, s) i strimmelen skærer linien $y = 1$ i punktet $(a/s, 1)$.

(1.2) Eksempel. For $N = 7$ fås $\mathcal{F}_7 =$

$$\left\{ \ldots, \frac{0}{1}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{2}{7}, \frac{1}{3}, \frac{2}{5}, \frac{3}{7}, \frac{1}{2}, \frac{4}{7}, \frac{3}{5}, \frac{2}{3}, \frac{5}{7}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{6}{7}, \frac{1}{1}, \ldots \right\}$$

(1.3) Lemma. *Lad a/s være en brøk i $\mathcal{F}_N$. Efterfølgeren b/t i $\mathcal{F}_N$ kan da bestemmes således: Der findes par (x, y) af hele tal, som tilfredsstiller betingelserne*

$$(1) \quad sx - ay = 1 \qquad og \qquad (2) \quad 1 \leqslant y \leqslant N,$$

og blandt disse par er det, som har y størst mulig, netop parret (b, t).

Bevis. Brøken a/s er antaget uforkortelig, så a og s er primiske. Den diofantiske ligning (1) har altså løsninger. Øjensynlig er (x, y) en løsning, hvis og kun hvis $(x + a, y + s)$ er en løsning. Ligningen har derfor en entydigt bestemt løsning (x, y) således at y ligger i et givet interval af længde s. Specielt, da $s \leqslant N$, findes der løsninger, der opfylder (2). Blandt disse løsninger (x, y) vælger vi den, som har y størst mulig.

Brøken x/y ligger i $\mathcal{F}_N$, og $a/s < x/y$. Antag, indirekte, at x/y ikke er efterfølgeren b/t. Da er

$$\frac{a}{s} < \frac{b}{t} < \frac{x}{y}.$$

Afstanden mellem første og anden brøk er mindst $1/(st)$ og afstanden mellem de to yderste brøker er $1/(sy)$ ifølge (1). Altså er $1/(sy) > 1/(st)$, og dermed er $t > y$. Følgelig er $1 \leqslant t - y < N$, og specielt er $(b-x)/(t-y)$ en Farey-brøk i $\mathcal{F}_N$. Det påstås, at

$$\frac{a}{s} < \frac{b-x}{t-y} < \frac{b}{t}. \tag{*}$$

Den sidste ulighed følger nemlig af uligheden $b/t < x/y$, og den første er, ifølge (1), ækvivalent med følgende:

$$1 < sb - at. \tag{**}$$

Højresiden er positiv, da $a/s < b/t$, og højresiden er altså mindst 1. Da $N \geqslant t > y$, sikrer maksimaliteten af y, at højresiden ikke kan være 1. Altså gælder (**), og dermed (*). Øjensynlig er (*) i modstrid med, at b/t var efterfølgeren til a/s. ∎

(1.4) Sætning. (1) *Hvis to brøker $a/s < b/t$ er på hinanden følgende i $\mathcal{F}_N$, så er*

$$sb - at = 1 \quad (\textit{og specielt er } \frac{b}{s} - \frac{a}{t} = \frac{1}{st}).$$

(2) *Hvis de tre brøker $a/s < c/u < b/t$ er på hinanden følgende i $\mathcal{F}_N$, så er*

$$\frac{c}{u} = \frac{a+b}{s+t}.$$

(3) *Lad to brøker $a/s < b/t$ være på hinanden følgende i $\mathcal{F}_N$. Da er $s+t > N$. Videre gælder, at brøkerne er på hinanden følgende i $\mathcal{F}_M$, når $N \leqslant M < s+t$, og blandt Farey-brøkerne i $\mathcal{F}_{s+t}$ er de tre brøker,*

$$\frac{a}{s}, \quad \frac{a+b}{s+t}, \quad \frac{b}{t},$$

på hinanden følgende.

Bevis. (1) følger af beskrivelsen i det foregående Lemma. (2) følger af (1), thi af $sc - ua = 1$ og $ub - tc = 1$ fås ved subtraktion, at $u(a+b) = c(s+t)$. I (3) er uligheden $s+t > N$ en konsekvens af, at brøken $(a+b)/(s+t)$ altid ligger mellem brøkerne a/s og b/t. Den sidste påstand i (3) følger af opskriften i Lemma (1.3) til bestemmelse af efterfølgeren til a/s i $\mathcal{F}_M$: For $M = N$ fører opskriften til $(x, y) = (b, t)$. Den næste mulighed for et par (x, y) med $y > t$ er $(x, y) := (a+b, s+t)$, og denne mulighed indtræffer første gang når $M = s+t$. Efterfølgeren til a/s i $\mathcal{F}_M$ er altså b/t, når $N \leqslant M < s+t$, og $(a+b)/(s+t)$ når $M = s+t$. Tilsvarende er $(a+b)/(s+t)$ forgængeren til b/t i $\mathcal{F}_{s+t}$. ∎

(1.5) Definition. Lad $\tau \in \mathbb{R}$ være et irrationalt tal. En brøk a/s (uforkortelig, med $s \geqslant 1$), vil her blive kaldt en *Farey-approximation* til τ, hvis der blandt alle brøker, hvis nævner højst er s, ikke findes nogen, der ligger mellem τ og a/s.

Med nævner 1 er der er øjensynlig to Farey-approximationer til τ, nemlig $a_1/1 < \tau$, hvor a_1 er den hele del af τ, og $\tau < (a_1+1)/1$. Med nævner $s > 1$ kan der højst være én Farey

approximation til τ. Antag nemlig, at $a/s < b/s$ er Farey-approximationer. Da må a/s og b/s specielt være naboer i $\mathcal{F}_s$, og afstanden mellem dem er derfor $1/s^2$. På den anden side er $a/s < (a+1)/s \leqslant b/s$, og det følger, at $b = a+1$; afstanden er derfor $1/s$. Følgelig er $s = 1$. Heraf ses, at Farey-approximationerne til τ naturligt kan opskrives i en følge, idet vi sætter $a_1/s_1 = a_1/1$, hvor a_1 er den hele del af τ, dernæst $a_2/s_2 := (a_1+1)/1$, og herefter ordner approximationerne efter voksende nævnere. Følgen af brøker a_n/s_n kaldes *Farey-følgen* bestemt ved τ.

(1.6) Lemma. *Hvis*

$$\left|\frac{x}{y} - \tau\right| < \frac{1}{y^2},$$

så er x/y en Farey-approximation til τ.

Bevis. Det kan antages, at $y \geqslant 1$, og at brøken x/y er uforkortelig. Hvis brøken a/s ligger mellem τ og x/y, så er

$$\frac{1}{y^2} > \left|\frac{x}{y} - \tau\right| \geqslant \left|\frac{x}{y} - \frac{a}{s}\right| = \frac{|sx - ya|}{sy} \geqslant \frac{1}{sy},$$

og følgelig er $s > y$. I $\mathcal{F}_y$ ligger der altså ingen brøker mellem x/y og τ. ∎

(1.7) Sætning. *Følgen a_n/s_n af Farey-approximationer til τ kan bestemmes induktivt således: Antag, at $n \geqslant 3$, og betragt brøkerne a_i/s_i for $i = 1, \ldots, n-1$. Lad a'/s' være den største, som er mindre end τ, og lad a''/s'' være den mindste, som er større end τ. Da er $a_n = a' + a''$ og $s_n = s' + s''$, og specielt er*

$$\frac{a_n}{s_n} = \frac{a' + a"}{s' + s"}.$$

Bevis. Vi lader b_n/t_n være følgen af brøker defineret ved den induktive forskrift, med begyndelsesbetingelserne $b_1/t_1 = a_1/1$ og $b_2/t_2 = (a_1+1)/1$, hvor a_1 er den hele del af τ. Det skal vises, at brøken b_n/t_n er uforkortelig og at $b_n/t_n = a_n/s_n$. Påstanden er triviel for $n = 1$ og $n = 2$. For at vise den for $n > 2$ kan vi antage, at den er vist for alle $i < n$. Sæt $N := s_{n-1}$. Da nævnerne s_i er voksende, ligger alle brøkerne a_i/s_i for $i = 1, \ldots, n-1$ i $\mathcal{F}_N$. Lad $a'/s' < a''/s''$ være de to på hinanden følgende brøker i $\mathcal{F}_N$ således, at τ ligger mellem a'/s' og a''/s''. Begge brøkerne må være Farey-approximationer, med nævner højst $N = s_{n-1}$, og dermed af formen a_i/s_i for $i < n$; en af de to brøker må naturligvis være a_{n-1}/s_{n-1}. Blandt brøkerne a_i/s_i med $i < n$ er a'/s' derfor den største mindre end τ og a''/s'' er den mindste større end τ. Induktivt følger det derfor, at brøkerne a'/s' og a''/s'' netop er de to brøker, der indgår i den induktive definition af b_n/t_n. Altså er $b_n/t_n = (a'+a'')/(s'+s'')$. Brøken ligger mellem a'/s' og a''/s''. Den må være uforkortelig, thi hvis den kunne forkortes, ville den få en nævner, som højst var den største af s' og s'', og altså højst N, i modstrid med at a'/s' og a''/s'' er på hinanden følgende i $\mathcal{F}_N$.

Vi mangler at vise, at $b_n/t_n = a_n/s_n$. Brøken a_n/s_n er en Farey-approximation med nævner større end N. Den ligger derfor mellem a'/s' og a''/s''. Sæt $s := s'+s'' = t_n$. Ifølge

Sætning (1.4) er $s > N = s_{n-1}$ og $a'/s' < b_n/s < a''/s''$ er på hinanden følgende i brøker i $\mathcal{F}_s$. Specielt er også b_n/s en Farey-approximation, og da $s > s_{n-1}$, er $s \geqslant s_n$. Følgelig ligger brøken a_n/s_n i $\mathcal{F}_s$, og nu følger det, at den må være lig med b_n/s.

Hermed er Sætningen bevist. □

(1.8) Eksempel. Betragt en kvadratisk form,

$$F(x, y) = ax^2 + bxy + cy^2,$$

hvor a, b, c er hele tal, og hvor diskriminanten $D = b^2 - 4ac$ er positiv og ikke et kvadrattal. Lad η være roden,

$$\eta = \frac{-b + \sqrt{D}}{2a}.$$

i andengradspolynomiet $aX^2 + bX + c$, og lad η' være den anden rod. Da er $aX^2 + bX + c = a(X - \eta)(X - \eta')$, så vi får faktoriseringen,

$$F(x, y) = a(x - \eta y)(x - \eta' y). \tag{1.8.1}$$

Idet vi antager, at $a > 0$, er $\eta > \eta'$, og $\eta - \eta' = \sqrt{D}/a$. Antag, at $\sqrt{D} > a$. Da er den hele del af η større end η'. Altså gælder for enhver Farey-approximation x/y til η, at $x/y > \eta'$, altså at $x > \eta' y$. I (1.8.1) er faktoren $x - \eta' y$ altså positiv, så det er fortegnet for $F(x, y)$ der afgør, om x/y er større eller mindre end η.

Herefter kan Farey-følgen for η bestemmes i et skema, som vi blot forklarer med et eksempel: Til $\eta = \sqrt{13}$ svarer $F(x, y) = x^2 - 13y^2$, og skemaet

p	3	4	7	11	18	29
q	1	1	2	3	5	8
$p^2 - 13q^2$	-4	3	-3	4	-1	-3

I første søjle er $q_1 = 1$ indsat, og p_1 er bestemt som det største hele tal for hvilket $F(p_1, 1)$ er negativ; p_1 er med andre ord den hele del af η. I næste søjle er $q_2 = 1$ og $p_2 = p_1 + 1$ indsat, og $F(p_2, q_2)$ er beregnet. For at bestemme (p_n, q_n) betragtes fortegnet for $F(p_{n-1}, q_{n-1})$; derefter vælges det største $i < n - 1$ for hvilket $F(p_i, q_i)$ har modsat fortegn. Herefter sættes $(p_n, q_n) := (p_i + p_{n-1}, q_i + q_{n-1})$. Endelig beregnes værdien $F(p_n, q_n)$.

Ved den sidste beregning kan det i øvrigt være nyttigt at observere, at

$$F(p' + p'', q' + q'') = F(p', q') + F(p'', q'') + 2ap'p'' + b(p'q'' + p''q') + 2cq'q''.$$

I forbindelse med den diofantiske ligning,

$$F(x, y) = k. \tag{1.8.2}$$

betragtes ofte området bestemt ved ulighederne:

$$y > 0, \qquad 2ax + by > 0. \tag{1}$$

Vi antager ikke her, at $a > 0$. Valget af η, med det positive fortegn på $\sqrt{D}$, sikrer, at $a\eta' < a\eta$.

Det er værd at bemærke, at for et par (x, y) af hele (eller blot reelle) tal, som opfylder (1.8.2), gælder, at (x, y) ligger i området, hvis og kun hvis en af følgende ækvivalente betingelser er opfyldt

$$|a(x - y\eta)| < a(x - y\eta'), \tag{2}$$

$$y > 0 \text{ og } |x - y\eta| < |x - y\eta'|, \tag{3}$$

$$a(x - y\eta') > \sqrt{|ak|}. \tag{4}$$

(Betingelsen (3) udsiger, at brøken x/y ligger tættere på η end på η'.)

Uligheden (2), af formen $|\alpha| < \alpha'$, er ensbetydende med de to uligheder $\alpha < \alpha'$ og $-\alpha < \alpha'$. Den første udsiger, at $-ya\eta < -ya\eta'$, og den gælder derfor, hvis og kun hvis $y > 0$. Den anden udsiger, at $2ax - ya(\eta + \eta') > 0$, dvs $2ax + by > 0$. Altså er (1) og (2) ækvivalente. Specielt følger det af (2), at $y > 0$, og det er derfor klart, at (2) medfører (3). Omvendt, hvis (3) gælder, så ligger brøken x/y ligger tættere ved η end ved η'. Ved multiplikation med a følger, at brøken ax/y ligger tættere ved $a\eta$ end ved $a\eta'$. Her er $a\eta > a\eta'$, så specielt følger det, at $ax/y > a\eta'$, altså at $a(x - y\eta')$ er positiv.

Endelig fremgår det af (1.8.1), at produktet af tallene $a(x - y\eta)$ og $a(x - y\eta')$ er lig med ka, og altså numerisk lig med kvadratet på $\sqrt{|ak|}$. Den anden faktor er derfor større end $\sqrt{|ak|}$, hvis og kun hvis den er større end den numeriske værdi af den første faktor. Altså er (4) og (2) ækvivalente.

2. Kædebrøker.

(2.1) Definition. Lad $n \geqslant 0$ være et helt tal, og lad der være givet to følger: $x_0, x_1, \ldots, x_n$ og $t_1, \ldots, t_n$. Ved den tilhørende *kædebrøk* af *længde* n forstås for $n = 0$ blot x_0, og for $n = 1, 2$ og 3 udtrykkene,

$$x_0 + \frac{t_1}{x_1} \quad , \quad x_0 + \frac{t_1}{x_1 + \dfrac{t_2}{x_2}} \quad , \quad x_0 + \frac{t_1}{x_1 + \dfrac{t_2}{x_2 + \dfrac{t_3}{x_3}}} \, ,$$

og, for n i almindelighed, udtrykket,

$$x_0 + \cfrac{t_1}{x_1 + \cfrac{t_2}{x_2 + \ddots + \cfrac{t_n}{x_n}}} \quad . \tag{2.1.1}$$

Brøkens *tællere* er følgen t_i, dens *nævnere* er følgen x_i (for $i \geqslant 1$). Under ét kaldes de to følger for kædebrøkens *koefficienter*. Den *ægte kædebrøk* af længde n fås i tilfældet, hvor $x_0 = 0$ (og $n \geqslant 1$). Den har formen

$$\cfrac{t_1}{x_1 + \cfrac{t_2}{x_2 + \ddots + \cfrac{t_n}{x_n}}} \, . \tag{2.1.2}$$

Nævneren i den ægte kædebrøk (2.1.2) er en kædebrøk af længde $n - 1$.

(2.2) Udregning. Ved at forlænge og addere kan man øjensynlig omforme kædebrøken (2.1.1) til en brøk p_n/q_n, hvor p_n og q_n er polynomier i x_i'erne og t_i'erne. For at bestemme p_n/q_n betragtes først den tilsvarende omformning af den ægte kædebrøk (2.1.2) til en brøk p'_n/q'_n. Nævneren i (2.1.2) er summen af x_1 og en ægte kædebrøk af orden $n-1$. Hvis p''/q'' betegner omformningen af denne sidste kædebrøk, har vi altså

$$\frac{p'_n}{q'_n} = \frac{t_1}{x_1 + p''/q''} = \frac{t_1 q''}{x_1 q'' + p''} \, ;$$

mere præcist har vi matrixligningen,

$$\begin{pmatrix} p'_n \\ q'_n \end{pmatrix} = \begin{pmatrix} 0 & t_1 \\ 1 & x_1 \end{pmatrix} \begin{pmatrix} p'' \\ q'' \end{pmatrix} .$$

som udtrykker tæller og nævner i den ægte kædebrøk af orden n ved tæller og nævner i den ægte kædebrøk af orden $n-1$ svarende til følgerne $x_2, \ldots, x_n$ og $t_2, \ldots, t_n$. Ved gentagen anvendelse af ligningen får vi,

$$\begin{aligned}\begin{pmatrix} p_n' \\ q_n' \end{pmatrix} &= \begin{pmatrix} 0 & t_1 \\ 1 & x_1 \end{pmatrix}\begin{pmatrix} 0 & t_2 \\ 1 & x_2 \end{pmatrix}\cdots\begin{pmatrix} 0 & t_{n-1} \\ 1 & x_{n-1} \end{pmatrix}\begin{pmatrix} t_n \\ x_n \end{pmatrix} \\ &= \begin{pmatrix} 0 & t_1 \\ 1 & x_1 \end{pmatrix}\begin{pmatrix} 0 & t_2 \\ 1 & x_2 \end{pmatrix}\cdots\begin{pmatrix} 0 & t_{n-1} \\ 1 & x_{n-1} \end{pmatrix}\begin{pmatrix} 0 & t_n \\ 1 & x_n \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix}. \qquad (2.2.1)\end{aligned}$$

Den reciprokke af den almindelige kædebrøk (2.1.1) er en ægte kædebrøk af orden $n+1$. Hvis denne sidste omformes til $\tilde{p}/\tilde{q}$, har vi altså $p_n/q_n = \tilde{q}/\tilde{p}$, eller

$$\begin{pmatrix} p_n \\ q_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \tilde{p} \\ \tilde{q} \end{pmatrix}.$$

Anvendes udtrykket (2.2.1) til bestemmelse af $\tilde{p}, \tilde{q}$, får vi den endelige formel:

$$\begin{pmatrix} p_n \\ q_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & x_0 \end{pmatrix}\begin{pmatrix} 0 & t_1 \\ 1 & x_1 \end{pmatrix}\begin{pmatrix} 0 & t_2 \\ 1 & x_2 \end{pmatrix}\cdots\begin{pmatrix} 0 & t_n \\ 1 & x_n \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix}. \qquad (2.2.2)$$

(2.3) Definition. Formlen ovenfor er udgangspunktet, når kædebrøken (2.1.1) skal tillægges en værdi. Ud fra de givne følger $x_0, x_1, \ldots, x_n$ og $t_1, \ldots, t_n$ defineres p_n og q_n ved formlen. Det er sædvane at tillægge kædebrøken en værdi, når blot $(p_n, q_n) \neq (0,0)$. Hvis $q_n \neq 0$ defineres værdien som p_n/q_n, og hvis $q_n = 0$ (og altså $p_n \neq 0$) defineres værdien som ∞.

I det følgende betegnes med S_n produktet af de kvadratiske matricer på højresiden af (2.2.2),

$$S_n := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & x_0 \end{pmatrix}\begin{pmatrix} 0 & t_1 \\ 1 & x_1 \end{pmatrix}\begin{pmatrix} 0 & t_2 \\ 1 & x_2 \end{pmatrix}\cdots\begin{pmatrix} 0 & t_n \\ 1 & x_n \end{pmatrix}. \qquad (2.3.1)$$

Ifølge (2.2.2) er søjlen $\left(\begin{smallmatrix} p_n \\ q_n \end{smallmatrix}\right)$ lig med den anden søjle i matricen S_n.

Bemærk, at definitionen af S_n også har mening for $n = -1$ og $n = 0$: Vi har

$$S_{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad S_0 = \begin{pmatrix} 1 & x_0 \\ 0 & 1 \end{pmatrix}.$$

I de følgende formler er det bekvemt at definere $(p_{-2}, q_{-2}) := (0,1)$, $(p_{-1}, q_{-1}) := (1,0)$ og $t_0 := 1$.

(2.4) Sætning. *Med betegnelserne i* (2.3) *gælder for* $n \geqslant 0$, *at*

$$S_n = \begin{pmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{pmatrix}, \qquad \det S_n = p_{n-1}q_n - p_n q_{n-1} = (-1)^n t_1 \cdots t_n, \qquad (2.4.1)$$

$$\begin{pmatrix} p_n \\ q_n \end{pmatrix} = \begin{pmatrix} p_{n-2} & p_{n-1} \\ q_{n-2} & q_{n-1} \end{pmatrix}\begin{pmatrix} t_n \\ x_n \end{pmatrix}, \quad \begin{pmatrix} p_{-2} & p_{-1} \\ q_{-2} & q_{-1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \qquad (2.4.2)$$

Bevis. Antag, at $n \geqslant 0$. Ifølge Ligning (2.2.3) er $\left(\begin{smallmatrix} p_n \\ q_n \end{smallmatrix}\right)$ den anden søjle i matricen S_n. Af definitionen i (2.3) følger videre, at

$$S_n = S_{n-1} \begin{pmatrix} 0 & t_n \\ 1 & x_n \end{pmatrix}. \tag{2.4.3}$$

(Bemærk, at det er definitionen $t_0 := 1$, der sikrer, at denne ligning også gælder for $n = 0$.) Den første søjle i S_n er derfor søjlen

$$S_n \begin{pmatrix} 1 \\ 0 \end{pmatrix} = S_{n-1} \begin{pmatrix} 0 & t_n \\ 1 & x_n \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = S_{n-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} p_{n-1} \\ q_{n-1} \end{pmatrix};$$

her bruges, for $n = 0$, den særlige definition af (p_{-1}, q_{-1}). Hermed er den første ligning i (2.4.1) bevist. Desuden gælder denne ligning også for $n = -1$, ifølge den særlige definition af (p_{-2}, q_{-2}). Determinanten af S_n kan nu bestemmes, dels ved det fundne udtryk, dels ved definitionen af S_n som produkt af $n + 2$ matricer.

Den anden ligning i (2.4.2) er blot definitionen af S_{-1}; den første følger af (2.4.3) ved at sammenligne anden søjle i de to matricer.

Hermed er formlerne bevist. ∎

(2.5) Definition. Til uendelige følger $x_0, x_1, x_2, \ldots$ og $t_1, t_2, \ldots$ hører en *uendelig kædebrøk*

$$x_0 + \cfrac{t_1}{x_1 + \cfrac{t_2}{x_2 + \cfrac{t_3}{x_3 + \ddots}}}. \tag{2.5.1}$$

I analogi med hvad der kendes fx fra uendelige rækker opfattes den uendelige kædebrøk (2.5.1) som en følge, nemlig følgen p_0/q_0, p_1/q_1, $p_2/q_2, \ldots$, der også kaldes følgen af *konvergenter*. (Af og til omtales også parrene (p_n, q_n) som *konvergenter*.) Såfremt følgen p_n/q_n er konvergent, siges den uendelige kædebrøk at være *konvergent*, og følgens grænseværdi kaldes da brøkens (*grænse-*)*værdi*. Ofte bruges (2.5.1) også som betegnelse for denne værdi.

(2.6) Konvention. Kædebrøker spiller en rolle i matematikken udover talteorien. Af særlig interesse er tilfældet, hvor koefficienterne, dvs x_i'erne og t_i'erne er funktioner, fx analytiske funktioner. Vi vil udelukkende betragte kædebrøker, hvor koefficienterne er tal, og her vil vi endda kun betragte kædebrøker, hvor alle tællerne t_i er lig med 1; de kaldes også *regulære kædebrøker*. Det er altså kædebrøker (endelige eller uendelige) af formen,

$$x_0 + \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cfrac{1}{x_3 + \ddots}}}. \tag{2.6.1}$$

Bemærk, at definitionen i (2.3.1) her giver

$$S_n := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & x_0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & x_1 \end{pmatrix}\cdots\begin{pmatrix} 0 & 1 \\ 1 & x_n \end{pmatrix}, \quad S_{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

og at formlerne i (2.4) har formen,

$$S_n = \begin{pmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{pmatrix}, \qquad \det S_n = p_{n-1}q_n - p_n q_{n-1} = (-1)^n, \tag{2.6.2}$$

$$\begin{pmatrix} p_n \\ q_n \end{pmatrix} = \begin{pmatrix} p_{n-2} & p_{n-1} \\ q_{n-2} & q_{n-1} \end{pmatrix}\begin{pmatrix} 1 \\ x_n \end{pmatrix}, \quad \begin{pmatrix} p_{-2} & p_{-1} \\ q_{-2} & q_{-1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{2.6.3}$$

I resten af dette kapitel betragtes kun kædebrøker, hvor tallene x_i er positive for $i \geqslant 1$; hvis koefficienterne x_i er hele, antages altså, at $x_0 \in \mathbb{Z}$ og $x_i \in \mathbb{N}$ for $i \geqslant 1$. Matrixligningerne i (2.6.3) er rekursionsformler til bestemmelse af (p_n, q_n). Af $q_n = q_{n-2} + x_n q_{n-1}$, med begyndelsesbetingelserne $q_{-1} = 0$ og $q_0 = 1$, følger, idet $x_i > 0$ for alle $i \geqslant 1$, at $q_i > 0$ for alle $i \geqslant 0$. Konvergenterne p_i/q_i, for $i \geqslant 0$, er altså reelle tal (og ikke ∞).

(2.7) Sætning. *Antag, at kædebrøken* (2.6.1) *har hele koefficienter, altså at* $x_0 \in \mathbb{Z}$ *og* $x_i \in \mathbb{N}$ *for* $i \geqslant 1$. *Da gælder: For hvert* i *er tallene* p_i, q_i *hele, primiske tal, og*

$$1 = q_0 \leqslant q_1 < q_2 < \cdots .$$

Hvis kædebrøken er endelig, af længde k, *så er dens værdi,* τ, *et rationalt tal, lig med den sidste konvergent* p_k/q_k. *Hvis kædebrøken er uendelig, så er den konvergent, og dens grænseværdi,* τ, *er et irrationalt tal. For konvergenterne* (*bortset fra den sidste i det endelige tilfælde*) *gælder uligheder:*

$$p_0/q_0 < p_2/q_2 < \cdots < \tau < \cdots < p_3/q_3 < p_1/q_1\,,$$
$$1 > |p_0 - q_0\tau| > |p_1 - q_1\tau| > \cdots .$$

Bevis. Af rekursionsformlen (2.6.3) fremgår, at p_n og q_n er hele tal, og af den sidste ligning i (2.6.2) følger umiddelbart, at de er primiske. Af rekursionsformlen,

$$q_n = x_n q_{n-1} + q_{n-2},$$

hvor startværdierne er $q_{-1} = 0$ og $q_0 = 1$, følger umiddelbart, at $1 = q_0 \leqslant q_1 < q_2 < \cdots$. Specielt er tallene q_n positive for $i \geqslant 0$. Hver konvergent p_n/q_n er altså et rationalt tal (dvs ikke ∞).

For alle $n \geqslant 1$ følger det af (2.6.2), at

$$\frac{p_{n-1}}{q_{n-1}} - \frac{p_n}{q_n} = \frac{(-1)^n}{q_{n-1}q_n}\,. \tag{2.7.1}$$

Af rekursionsformlen fås $\binom{p_n}{q_n} = S_{n-1}\binom{1}{x_n}$. Da første søjle i matricen S_{n-1} er søjlen $\binom{p_{n-2}}{q_{n-2}}$, gælder matrixligningen,

$$\begin{pmatrix} p_{n-2} & p_n \\ q_{n-2} & q_n \end{pmatrix} = S_{n-1}\begin{pmatrix} 1 & 1 \\ 0 & x_n \end{pmatrix}.$$

Lighed mellem determinanterne medfører, for $n \geqslant 2$, at

$$\frac{p_{n-2}}{q_{n-2}} - \frac{p_n}{q_n} = \frac{(-1)^{n-1}x_n}{q_{n-2}q_n}. \tag{2.7.2}$$

Af ligning (2.7.2) følger straks, at de „lige konvergenter" $p_0/q_0, p_2/q_2, \ldots$ udgør en strengt voksende følge, og at de „ulige konvergenter" $p_1/q_1, p_3/q_3, \ldots$ udgør en strengt aftagende følge. Af (2.7.1) ses, at enhver ulige konvergent er større end den efterfølgende lige konvergent. Ulighederne vedrørende de numeriske værdier $|p_n - q_n\tau|$ vises i den efterfølgende opskrift.

Antag nu, at kædebrøken er uendelig, således at vi har en uendelig følge af konvergenter p_n/q_n. Følgen q_n går mod uendelig. Det kan derfor sluttes af (2.7.1), at de to følger, af lige og ulige konvergenter, har en fælles grænseværdi. Endvidere sluttes, at grænseværdien τ ligger mellem de to brøker på venstresiden af (2.7.1). Heraf følger videre, at der gælder uligheden,

$$\left|\frac{p_n}{q_n} - \tau\right| < \frac{1}{q_n^2}.$$

Af denne ulighed, for alle n, følger let, at τ må være irrational. ∎

(2.8) Resterne. Under forudsætningen i (2.7) defineres den n'te *rest*, τ_n, af kædebrøken som (grænse-)værdien af den ægte kædebrøk med koefficienter $x_{n+1}, x_{n+2}, \ldots$. Hvis kædebrøken er af endelig længde k, er resten kun defineret, når $n < k$. I dette tilfælde forudsætter vi yderligere, at $x_k > 1$, og vi sætter $\tau_k := 0$. (Tilfældet $k = 0$, hvor kædebrøken blot er et helt tal x_0, kræver en helt triviel særbehandling.) Vi har

$$\tau = x_0 + \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cfrac{1}{\ddots + \cfrac{1}{x_n + \tau_n}}}}, \qquad \tau_n = \cfrac{1}{x_{n+1} + \cfrac{1}{x_{n+2} + \cfrac{1}{x_{n+3} + \ddots}}}.$$

idet den anden ligning blot er definitionen af resten, og den første følger af velkendte regler for regning med grænseværdier. Af ulighederne $p_0/q_0 < \tau < p_1/q_1$ for den almindelige kædebrøk følger, for resten τ_n, at $0 < \tau_n < 1/x_{n+1}$. Specielt er $0 < \tau_n < 1$. Hvis brøken har endelig længde k, har vi $\tau_{k-1} = 1/x_k < 1$ (idet vi har antaget $x_k > 1$), og $\tau_k = 0$. Vi har altså altid

$$0 \leqslant \tau_n < 1,$$

og uligheden er skarp, med mindre kædebrøken har endelig længde k og $n = k$.

Betragt den første (endelige) kædebrøk af længde n. Dens første konvergenter er øjensynlig (p_i, q_i) for $i = 0, \ldots, n-1$. Lad (π_n, λ_n) være den n'te konvergent. Kvotienten π_n/λ_n er da kædebrøkens værdi, altså lig med τ. Af (2.6.3) fremgår, at

$$\begin{pmatrix} \pi_n \\ \lambda_n \end{pmatrix} = S_{n-1} \begin{pmatrix} 1 \\ x_n + \tau_n \end{pmatrix}.$$

Da $\begin{pmatrix} 1 \\ x_n+\tau_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & x_n \end{pmatrix} \begin{pmatrix} \tau_n \\ 1 \end{pmatrix}$, er højresiden lig med $S_n \begin{pmatrix} \tau_n \\ 1 \end{pmatrix}$, og $\begin{pmatrix} \pi_n \\ \lambda_n \end{pmatrix} = \lambda_n \begin{pmatrix} \tau \\ 1 \end{pmatrix}$, da $\tau = \pi_n/\lambda_n$. Altså gælder ligningen,

$$\lambda_n \begin{pmatrix} \tau \\ 1 \end{pmatrix} = S_n \begin{pmatrix} \tau_n \\ 1 \end{pmatrix}. \tag{2.8.1}$$

Det er nok at huske, at ligningens to søjler er proportionale: $\begin{pmatrix} \tau \\ 1 \end{pmatrix} \sim S_n \begin{pmatrix} \tau_n \\ 1 \end{pmatrix}$, idet proportionalitetsfaktoren λ_n bestemmes ved lighed mellem andenkoordinaterne:

$$\lambda_n = q_{n-1}\tau_n + q_n. \tag{2.8.2}$$

Af den sidste ligning i (2.6.2) følger, at $p_{n-1}\lambda_n - \pi_n q_{n-1} = (-1)^n$. Igen, da $\tau = \pi_n/\lambda_n$, følger det, at

$$\lambda_n(p_{n-1} - q_{n-1}\tau) = (-1)^n. \tag{2.8.3}$$

For $n = 0$ er $\lambda_0 = q_0 = 1$, og i det endelige tilfælde, for $n = k$, er $\lambda_k = q_k$. For alle andre værdier af n er $q_n < q_n + \tau_n q_{n-1} < q_n + q_{n-1} \leqslant x_{n+1}q_n + q_{n-1} = q_{n+1}$. Specielt er

$$1 = q_0 = \lambda_0 \leqslant q_1 < \lambda_1 < q_2 < \lambda_2 < \cdots. \tag{2.8.4}$$

Af (2.8.3) følger, at

$$|p_n - q_n\tau| = 1/\lambda_{n+1}. \tag{2.8.4}$$

De sidste uligheder i Sætning (2.7) følger derfor af (2.8.4). Da $\lambda_{n+1} > q_{n+1} = x_{n+1}q_n + q_{n-1}$, får vi vurderingerne,

$$|p_n - q_n\tau| < \frac{1}{q_{n+1}} < \frac{1}{x_{n+1}q_n}. \tag{2.8.5}$$

(2.9) Opskrift. Lad τ være et reelt tal. Definer rekursivt en følge $\tau_0, \tau_1, \ldots$ af tal i intervallet $[0, 1[$ og en følge $x_0, x_1, x_2, \ldots$ af hele tal, hvor $x_i > 0$ for $i \geqslant 1$, således: Skriv

$$\tau = x_0 + \tau_0, \qquad \text{hvor } x_0 \in \mathbb{Z} \text{ og } 0 \leqslant \tau_0 < 1.$$

Stop, hvis $\tau_0 = 0$. I modsat fald er $0 < \tau_0 < 1$, og derfor kan man skrive

$$\frac{1}{\tau_0} = x_1 + \tau_1, \qquad \text{hvor } x_1 \in \mathbb{N} \text{ og } 0 \leqslant \tau_1 < 1.$$

Stop, hvis $\tau_1 = 0$. Induktivt, hvis $\tau_{n-1} \neq 0$, skrives

$$\frac{1}{\tau_{n-1}} = x_n + \tau_n, \qquad \text{hvor } x_n \in \mathbb{N} \text{ og } 0 \leqslant \tau_n < 1.$$

Processen stopper i det k'te skridt, hvis $\tau_k = 0$. Er dette tilfældet bliver de to følger $\tau_0, \tau_1, \ldots$ og $x_0, x_1, \ldots$ endelige; yderligere er $x_k > 1$, idet $x_k = 1/\tau_{k-1}$, og $\tau_{k-1} < 1$.

Processen definerer en kædebrøk af formen i (2.7), og af konstruktionen fremgår, for hvert n, at τ er værdien af den endelige kædebrøk angivet i (2.8). Hvis processen stopper, i det k'te skridt med $\tau_k = 0$, er det klart at kædebrøken er endelig og lig med τ; i dette tilfælde må τ øjensynlig være et rationalt tal. Hvis processen ikke stopper, bliver kædebrøken uendelig. Af udregningerne i (2.8) følger let, at kædebrøkens grænseværdi netop er τ. Af Sætning (2.7) følger, at τ i dette tilfælde må være irrational.

Heraf ses: *Processen stopper, hvis og kun hvis tallet τ er rationalt.*

Det er klart, at tallene τ_n defineret i opskriften netop er resterne af kædebrøken. Tallet x_n kaldes den n'te *kædebrøkskoefficient* og tallet τ_n kaldes den n'te *rest* af tallet τ. Kædebrøkens konvergenter p_n/q_n kaldes også *konvergenter* for τ.

(2.10) Eksempel. Anvendt på $\tau = \sqrt{13}$ giver opskriften:

$$\begin{aligned}
\tau = \sqrt{13} = \qquad\qquad &= 3 + \sqrt{13} - 3\,, \\
1/\tau_0 = \frac{1}{\sqrt{13}-3} = \frac{\sqrt{13}+3}{4} &= 1 + \frac{\sqrt{13}-1}{4}\,, \\
1/\tau_1 = \frac{4}{\sqrt{13}-1} = \frac{\sqrt{13}+1}{3} &= 1 + \frac{\sqrt{13}-2}{3}\,, \\
1/\tau_2 = \frac{3}{\sqrt{13}-2} = \frac{\sqrt{13}+2}{3} &= 1 + \frac{\sqrt{13}-1}{3}\,, \\
1/\tau_3 = \frac{3}{\sqrt{13}-1} = \frac{\sqrt{13}+1}{4} &= 1 + \frac{\sqrt{13}-3}{4}\,, \\
1/\tau_4 = \frac{4}{\sqrt{13}-3} = \frac{\sqrt{13}+3}{1} &= 6 + \sqrt{13} - 3\,.
\end{aligned}$$

Højresiden i linie i, for $i = 0, 1, \ldots$, er $x_i + \tau_i$. Det fremgår specielt, at $\tau_5 = \tau_0$. Altså er $\tau_6 = \tau_1$, $\tau_7 = \tau_2$, osv. Følgen τ_0, τ_1, τ_2 er altså periodisk med periode 5. Heraf ses, at også følgen $x_1, x_2, \ldots$ er periodisk med periode 5. Vi får således tabellen:

n	-2	-1	0	1	2	3	4	5	6	7	8	9	10
x_n			3	1	1	1	1	6	1	1	1	1	6
p_n	0	1	3	4	7	11	18	119	137	256	393	649	
q_n	1	0	1	1	2	3	5	33	38	71	109	180	

Rækken for x_n fremgår af udregningerne. Rækkerne for konvergenterne p_n, q_n er fremkommet ved at indsætte de faste værdier, for $n = -2$ og $n = -1$, og så udfylde resten af pladserne ved rekursionsformlen (2.6.3).

(2.11) Lemma. *Lad T_n være en matrix af formen i* (2.6)*, for et $n \geqslant 0$,*

$$T_n = \begin{pmatrix} 1 & y_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & y_1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & y_2 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & y_n \end{pmatrix}, \tag{2.11.1}$$

hvor $y_0 \in \mathbb{Z}$ og $y_i \in \mathbb{N}$ for $i > 0$. Lad τ være et reelt tal, og antag, at der findes en relation,

$$\begin{pmatrix} \tau \\ 1 \end{pmatrix} \sim T_n \begin{pmatrix} \epsilon \\ 1 \end{pmatrix}, \tag{2.11.2}$$

hvor $0 < \epsilon < 1$ (eller eventuelt $\epsilon = 0$ og enten $n = 0$ eller $n \geqslant 1$ og $y_n \geqslant 2$). Da er, med notationen i Opskrift (2.9)*, $y_i = x_i$ for $i = 0, \ldots, n$ og $\epsilon = \tau_n$.*

Bevis. Påstanden vises ved induktion efter n. For $n = 0$ følger det af relationen, at $\tau = y_0 + \epsilon$. Da $0 \leqslant \epsilon < 1$, er $y_0 = x_0$ og $\epsilon = \tau_0$. Antag, at $n \geqslant 1$ og at påstanden gælder for $n - 1$. Relationen kan skrives,

$$\begin{pmatrix} \tau \\ 1 \end{pmatrix} \sim T_{n-1} \begin{pmatrix} 1 \\ y_n + \epsilon \end{pmatrix} \sim T_{n-1} \begin{pmatrix} \epsilon' \\ 1 \end{pmatrix},$$

hvor $\epsilon' := 1/(y_n + \epsilon)$. Øjensynlig er $0 < \epsilon' < 1$. Induktivt er altså $y_i = x_i$ for $i < n$, og $\epsilon' = \tau_{n-1}$. Ligningen $1/\tau_{n-1} = y_n + \epsilon$ viser nu, at $y_n = x_n$ og at $\epsilon = \tau_n$. ∎

(2.12) Sætning. *Hvert irrationalt tal er grænseværdi for netop én uendelig kædebrøk med hele koefficienter. Hvert rationalt tal er værdi af netop to endelige kædebrøker med hele koefficienter, nemlig en af længde k af formen*

$$x_0 + \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cfrac{1}{\ddots + \cfrac{1}{x_k}}}}, \tag{2.12.1}$$

hvor $x_k \geqslant 2$, og den omformning (2.12.1)*, af længde $k + 1$, der fås ved at erstatte x_k med*

$$(x_k - 1) + \frac{1}{1}.$$

Bevis. Lad τ være det givne tal. Eksistensen er vist i Opskrift (2.9). Bemærk, at hvis τ er rational, således at processen stopper, fx i det k'te skridt, så må x_k være større end 1, idet $x_k = 1/\tau_{k-1}$ og $\tau_{k-1} < 1$ (tilfældet $k = 0$, som svarer til at τ er et helt tal, kræver en særovervejelse, som overlades til læseren).

Antag nu, at τ er værdien af en kædebrøk med hele koefficienter $y_0, y_1, y_2, \ldots$. Hvis kædebrøken har endelig længde k, kan det antages at $x_k \geqslant 2$, idet man ellers anvender omformningen nævnt i Sætningen. Det skal vises, at tallene y_i netop er tallene x_i bestemt i Opskrift (2.9). Denne påstand følger af Lemma (2.11), idet relationen (2.11.2), hvor ϵ lig med den n'te rest af kædebrøken, fremgår af (2.8.1). ∎

(2.13) Lemma. *Lad τ være et irrationalt tal. Enhver konvergent p_n/q_n for τ er da en Farey-approximation til τ. Mere præcist gælder, at konvergenterne netop er de brøker p/q i Farey-følgen for hvilke p/q og den efterfølgende brøk i følgen ligger på hver sin side af τ.*

Bevis. Den første konvergent er, med den valgte nummerering, brøken $p_0/q_0 = x_0/1$, hvor x_0 er det største hele tal, der er mindre end τ. Det er også den første brøk i Farey-følgen for τ. Den næste Farey-approximation er $(x_0+1)/1$, der ligger på den anden side af τ. Konvergenten p_0/q_0 er altså en Farey-approximation med den egenskab, at den efterfølgende approximation ligger på den anden side af τ. Vi viser, ved induktion efter n, at konvergenten p_n/q_n er en Farey-approximation med denne egenskab, og at de foregående Farey-approximationer med denne egenskab netop er konvergenterne p_i/q_i for $i \leqslant n$. I beviset for, at påstanden gælder for $n+1$ (med $n \geqslant 0$), antager vi, at den gælder for n og for $n-1$; det er ikke svært at modificere det efterfølgende således, at argumentet også virker for $n = 0$, med $p_{-1} = 1$ og $q_{-1} = 0$. Vi vil yderligere antage, at n er lige, idet argumentationen forløber tilsvarende, hvis n er ulige.

Af overvejelserne i (2.7) følger, at τ ligger mellem p_{n-1}/q_{n-1} og p_n/q_n. Idet vi har antaget, at n er lige, gælder altså ulighederne,

$$\frac{p_n}{q_n} < \tau < \frac{p_{n-1}}{q_{n-1}}.$$

Af induktionsantagelsen følger, at brøkerne p_n/q_n og p_{n-1}/q_{n-1} ligger i Farey-følgen, at efterfølgeren til p_{n-1}/q_{n-1} ligger til venstre for τ, og at p_n/q_n herefter er den første brøk i følgen, for hvilken efterfølgeren ligger til højre for τ. Blandt de foregående Farey-approximationer findes altså ingen mellem p_n/q_n og p_{n-1}/q_{n-1}. Heraf følger, jfr (1.7), at den næste Farey-approximation er følgende brøk for $x = 1$:

$$\frac{xp_n + p_{n-1}}{xq_n + q_{n-1}}\,. \tag{2.13.1}$$

Da den til p_n/q_n efterfølgende approximation er større end τ, er brøken (2.13.1) for $x = 1$ større end τ. Den følgende approximation er derfor brøken (2.13.1) for $x = 2$. Er også denne approximation større end τ bliver den næste approximation brøken (2.13.1) for $x = 3$. Det ses, at den første af de følgende Farey-approximationer, der har den anførte egenskab, er brøken (2.13.1), hvor x er størst mulig så at brøken er større end τ. Bemærk, at brøken (2.13.1) for $x \to \infty$ konvergerer mod p_n/q_n, der er mindre end τ; der findes altså en størst mulig værdi af x. Denne værdi er altså det største naturlige tal x, så at brøken (2.13.1) er større end τ. Da brøkens nævner er positiv, er den større end τ, hvis og kun hvis $xp_n + p_{n-1} > \tau(xq_n + q_{n-1})$, dvs hvis og kun hvis

$$x(q_n\tau - p_n) < -q_{n-1}\tau + p_{n-1}. \tag{2.13.2}$$

Da $p_n/q_n < \tau$, er $q_n\tau - p_n > 0$. Uligheden (2.13.2) er derfor ensbetydende med følgende:

$$x < \frac{-q_{n-1}\tau + p_{n-1}}{q_n\tau - p_n}\,, \tag{2.13.3}$$

og x er således det største hele tal, der opfylder denne ulighed.

På den anden side bestemmes konvergenterne ud fra resterne τ_n. Af (2.9.2) følger, at søjlerne $\begin{pmatrix}\tau\\1\end{pmatrix}$ og og $S_n\begin{pmatrix}\tau_n\\1\end{pmatrix}$ er proportionale. Da n er lige, er det $S_n = 1$. Heraf fås relationen,

$$\begin{pmatrix}\tau_n\\1\end{pmatrix} \sim S_n^{-1}\begin{pmatrix}\tau\\1\end{pmatrix} = \begin{pmatrix}q_n & -p_n\\ -q_{n-1} & p_{n-1}\end{pmatrix}\begin{pmatrix}\tau\\1\end{pmatrix}.$$

Heraf fremgår, at højresiden i (2.13.3) er lig med $1/\tau_n$. Med andre ord (jfr. Opskrift (2.9)), tallet x er koefficienten x_{n+1} i kædebrøken for τ. Med denne værdi af x følger det af rekursionsformlen (2.6.3), at brøken (2.13.1) netop er konvergenten p_{n+1}/q_{n+1}.

Hermed er induktionsbeviset fuldført. ∎

(2.14) Bemærkning. Af beviset for Lemma (2.13) fremgår, at koefficienterne $x_0, x_1, x_2, \ldots$ i kædebrøken for τ svarer til følgende opførsel af Farey-følgen for τ: Den første brøk i følgen er $x_0/1$, som er mindre end τ. Herefter kommer x_1 brøker, som er større end τ, efterfulgt af x_2 brøker, som er mindre end τ, efterfulgt af x_3 brøker, som er større end τ, osv.

(2.15) Sætning. *Lad τ være et irrationalt tal, og lad a/s være en brøk (uforkortelig, med $s \geqslant 1$). Lad a'/s' være den af de to naboer til a/s blandt Farey-brøkerne i $\mathcal{F}_s$, der ligger til samme side for a/s som τ. Da er a/s en konvergent for τ, hvis og kun hvis følgende ulighed er opfyldt:*

$$\left|\frac{a}{s} - \tau\right| < \frac{1}{s(s+s')}, \tag{2.15.1}$$

eller $s = 1$ og a er den hele del af τ.

Bevis. Konvergenten p_0/q_0 er $p_0/1$, hvor p_0 er den hele del af τ. Vi kan antage, at a/s ikke er denne konvergent.

Afstanden mellem naboerne a/s og a'/s' er $1/ss'$, der øjensynlig er større end højresiden af (2.15.1). Hvis (2.15.1) gælder, så følger det derfor af valget af a/s, at τ ligger mellem a/s og a'/s'. Hvis a/s er en konvergent, så er a/s en Farey-approximation ifølge det foregående lemma, og igen sikrer valget af a'/s', at τ ligger mellem a/s og a'/s'. I resten af beviset for ækvivalensen af de to betingelser kan det derfor forudsættes, at τ ligger mellem a/s og a'/s'. Specielt er a/s og a'/s' da Farey-approximationer til τ, og den Farey-approximation, der følger efter a/s, er brøken $(a+a')/(s+s')$. Af antagelsen følger, at i Farey-følgen kommer brøken a'/s' før brøken a/s.

Højresiden af uligheden (2.15.1) er netop afstanden mellem approximationen a/s og den følgende approximation $(a+a')/(s+s')$. Uligheden er derfor opfyldt, netop hvis τ ligger mellem disse to approximationer. Ækvivalensen følger derfor af karakteriseringen i Lemma (2.13). ∎

(2.16) Korollar. *Hvis følgende ulighed er opfyldt:*

$$\left|\frac{a}{s} - \tau\right| < \frac{1}{2s^2}, \tag{2.16.1}$$

så er brøken a/s en konvergent for τ.

Bevis. Med betegnelserne i sætningen er nemlig $s' \leqslant s$. Uligheden (2.16.1) medfører derfor uligheden (2.15.1). ∎

(2.17). Med notationen fra Eksempel (1.8) betragtes den kvadratiske form $F(x, y)$ og tallet $\eta = (-b + \sqrt{D})/(2a)$ (vi forudsætter ikke, at $a > 0$, og altså ikke at $\eta' < \eta$). Betragt videre den diofantiske ligning,

$$F(x, y) = ax^2 + bxy + cy^2 = k,$$

og hertil heltalsløsninger (x, y), som opfylder ulighederne i (2.9)(1):

$$y > 0, \quad 2ax + by > 0,$$

eller, ækvivalent, at brøken x/y ligger tættere ved η end ved η'.

Af faktoriseringen $F(x, y) = a(x - y\eta)(x - y\eta')$ fås ligningen,

$$\left|\frac{x}{y} - \eta\right| \cdot \left|\frac{x}{y} - \eta'\right| = |k/a| \cdot \frac{1}{y^2},$$

og altså, med $h := |x/y - \eta'|$,

$$\left|\frac{x}{y} - \eta\right| = \frac{|k/a|}{h} \cdot \frac{1}{y^2}. \tag{2.17.1}$$

Af antagelserne følger specielt, at afstanden $|x/y - \eta'|$ er større end halvdelen af afstanden $|\eta - \eta'| = \sqrt{D}/|a|$, altså at $h \geqslant \frac{1}{2}\sqrt{D}/|a|$. . Faktoren $|k/a|/h$ i (2.17.1) er altså mindre end $2|k|/\sqrt{D}$. Af (1.6) og (2.16) følger derfor:

Hvis $|k| < \frac{1}{2}\sqrt{D}$, så er x/y en Farey-approximation til η, og hvis $|k| < \frac{1}{4}\sqrt{D}$, så er x/y en konvergent for η.

Når y er stor, kan vi få en bedre vurdering af $h = |x/y - \eta'|$: Antag først, at brøken x/y ligger mellem η' og η. Da er $\sqrt{D}/|a| = |\eta - \eta'| = |x/y - \eta| + |x/y - \eta'|$, og altså

$$\sqrt{D}/|a| = |k/a|/h \cdot 1/y^2 + h.$$

Løsning af denne ligning (af grad 2 i h og med $h > \sqrt{D}/(2|a|)$) giver

$$h = \sqrt{D}/(2|a|) + \sqrt{D/(4a^2) - |k/a|/y^2}.$$

Antages i stedet, at brøken x/y ikke ligger mellem η' og η, gælder trivielt, at $h \geqslant \sqrt{D}/|a|$. Denne antagelse er øjensynlig ækvivalent med uligheden $a(x/y - \eta) > 0$. Heraf ses:

Sætning. *Antag, at* $|k| < \frac{1}{2}\sqrt{D}$, *og lad* (x, y) *være en løsning i området. Hvis*

$$\sqrt{D} + \sqrt{D - 4|ak|/y^2} > 4|k|, \tag{2.17.2}$$

eller $a(x/y - \eta) > 0$, *så er* $|x/y - \eta| < 1/(2y^2)$, *og følgelig er* x/y *en konvergent for* η. *Specielt er* x/y *en konvergent, hvis* $y \gg 0$.

Bevis. Hvis $a(x/y - \eta) > 0$, følger som nævnt, at $h > \sqrt{D}/|a|$, og dermed er $|k/a|/h < |k|/\sqrt{D} < \frac{1}{2}$. Hvis $a(x/y-\eta) < 0$, og (2.17.2) er opfyldt, så følger af udtrykket for h ovenfor, at $h|a| > 2|k|$. I begge tilfælde er altså $|k/a|/h < \frac{1}{2}$. Uligheden $|x/y - \eta| < 1/(2y^2)$ følger derfor af (2.17.1).

Da $|k| < \frac{1}{2}\sqrt{D}$, er det klart, at (2.17.2) er opfyldt, når $y \gg 0$. □

(2.18) Opgaver. (1) Bestem alle Farey-approximationer til tallet π med nævner højst 200. Er $3\frac{10}{71}$ en konvergent? Vis, at de tre første konvergenter, p_0/q_0, p_1/q_1, p_2/q_2 er 3/1, 22/7, 355/113.

(2) Vis påstanden sidst i beviset for (2.7): τ er irrational, når uligheden $|p_n/q_n - \tau| < 1/q_n^2$ er opfyldt for uendelig mange uforkortelige brøker p_n/q_n.

(3) Vis, at den ægte kædebrøk med $x_1 = x_2 = \cdots = 1$ fremstiller tallet $\tau = (\sqrt{5} - 1)/2$. [Vink: $\tau = \tau_0 = 1 + 1/\tau_0$.] Hvilket tal fremstilles, når $x_1 = x_2 = \cdots = n$?

(4) Lad τ være et irrationalt tal. Som nævnt i (2.16) vil uligheden $|a/s - \tau| < 1/(2s^2)$ medføre, at a/s er lig med en konvergent p_n/q_n. Vis omvendt, at uligheden altid gælder for mindst en af to på hinanden følgende konvergenter p_n/q_n og p_{n+1}/q_{n+1}. [Vink: udnyt, at τ ligger mellem p_n/q_n og p_{n+1}/q_{n+1}.] Vis, at uligheden gælder for konvergenten p_n/q_n, hvis $x_{n+1} \geqslant 2$.

3. Kædebrøker for kvadratiske tal.

(3.1) Notation. I det følgende betragtes et fast andengradspolynomium,

$$F(X,Y) = aX^2 + bXY + cY^2,$$

hvor koefficienterne a, b, c er hele tal, og hvor diskriminanten $D := b^2 - 4ac$ er positiv og ikke et kvadrat. Med η og ξ betegnes tallene,

$$\eta := \frac{-b+\sqrt{D}}{2a} \quad \text{og} \quad \xi := \frac{-b+\sqrt{D}}{2}.$$

Tallet η er den ene rod i polynomiet $aX^2 + bX + c$ (hvis $a > 0$, er η den største rod); den anden rod er det konjugerede tal η'. Tallet $\xi := a\eta$ er den største rod i det normerede polynomium $X^2 + bX + ac$. Tallene η og ξ er *kvadratiske* (irrationale) tal, dvs rødder i andengradspolynomier med hele koefficienter, og ξ er endda et *helt kvadratisk* tal, idet polynomiet med roden ξ er normeret. Øjensynlig er $F(x,y) = a(x-y\eta)(x-y\eta')$. Når x og y er rationale tal, vil de to sidste faktorer være konjugerede kvadratiske tal, og deres produkt vil derfor være normen af $x - y\eta$. For rationale (og specielt for hele) tal x og y gælder altså ligningen,

$$a\,\mathrm{N}(x-y\eta) = F(x,y). \tag{3.1.1}$$

I det følgende betragtes kædebrøksudviklingen af tallet η. Kædebrøkskoefficienterne x_n, konvergenterne p_n, q_n, matricerne S_n, osv refererer altså til tallet η, når intet andet er nævnt.

I undersøgelsen vil vi møde (kvadratiske) tal af formen,

$$\tau = \frac{\xi - t}{s}, \tag{3.1.2}$$

hvor t og $s \neq 0$ er hele tal, og hvor nævneren s er divisor i normen $N(\xi - t)$ af tælleren. Det er let at se, at mængden af sådanne tal kun afhænger af D, og vi betegner den $Q(D)$. Elementerne i $Q(D)$ vil vi kalde *rødderne* (hørende til den faste diskriminant D).

Vi bemærker først, at hvis τ er en rod i denne forstand og $x \in \mathbb{Z}$, så er også $\tau - x$ en rod, endda med samme nævner s. Vi har nemlig $\tau - x = (\xi - t - sx)/s$, og for normen af tælleren finder vi modulo s, at

$$N(\xi - t - sx) = (t+sx)^2 + b(t+sx) + ac \equiv t^2 + bt + ac = N(\xi - t) \equiv 0.$$

Videre bemærker vi, at hvis τ er en rod, så er også $1/\tau$ en rod, med fremstillingen

$$\frac{1}{\tau} = \frac{\xi - t_1}{s_1}, \quad \text{hvor } ss_1 = -N(\xi - t). \tag{3.1.3}$$

Da $\xi + \xi' = -b$, har vi nemlig

$$\frac{1}{\tau} = \frac{s}{\xi - t} = \frac{\xi' - t}{N(\xi - t)/s} = \frac{-(\xi' - t)}{-N(\xi - t)/s} = \frac{\xi + t + b}{-N(\xi - t)/s},$$

som er en fremstilling af den ønskede form, idet nævneren $s_1 := -N(\xi - t)/s$ er et helt tal og tælleren $-(\xi' - t)$ har normen $N(-(\xi' - t)) = N(\xi - t) = -ss_1$, som er et multiplum af nævneren s_1.

Endelig bemærker vi, at der kun er endelig mange rødder τ, som opfylder de følgende uligheder:

$$0 < \tau < 1, \quad \tau' < -1. \tag{3.1.4}$$

Antag nemlig, at ulighederne gælder for $\tau = (\xi - t)/s$. Af $0 < \tau$ og $1 < -\tau'$ følger, at $1 < \tau - \tau' = \sqrt{D}/s$. Altså er

$$1 \leqslant s < \sqrt{D}. \tag{3.1.5}$$

Specielt er s positiv, og så følger det af $0 < \tau < 1$, at $0 < \xi - t < s$, altså at

$$\xi - s < t < \xi. \tag{3.1.6}$$

Der er altså højst $\sqrt{D}$ muligheder for s, og for hver af dem højst $s < \sqrt{D}$ muligheder for t. Antallet af mulige par (s, t) er altså mindre end $\sqrt{D} \cdot \sqrt{D} = D$, og altså højst lig med $D - 1$. [Det er ikke så svært at finde en mindre overgrænse.]

(3.2) Euler's formel. *Resterne $\eta_0, \eta_1, \ldots$ af tallet η er kvadratiske tal af formen,*

$$\eta_n = \frac{\xi - e_n}{a_n}, \tag{3.2.1}$$

hvor e_n og $a_n \neq 0$ er hele tal og hvor nævneren a_n er divisor i normen $\mathrm{N}(\xi - e_n)$ af tælleren. For $n = 0$ er $a_0 = a$. For alle n gælder formlerne,

$$\mathrm{N}(\xi - e_n) = -a_n a_{n+1}, \tag{3.2.2}$$

$$\mathrm{N}(\eta_n) = -a_{n+1}/a_n, \qquad \mathrm{N}(q_n + q_{n-1}\eta_n) = (-1)^n a/a_n, \tag{3.2.3}$$

$$F(p_{n-1}, q_{n-1}) = \frac{a}{\mathrm{N}(q_n + q_{n-1}\eta_n)} = (-1)^n a_n. \tag{3.2.4}$$

Bevis. Den første påstand udtrykker, at resterne η_n er rødder. Denne påstand følger af overvejelserne i (3.1). Det er nemlig klart, at $\eta = \xi/a$ er en rod, idet nævneren a er divisor i normen, $\mathrm{N}(\xi) = ac$, af tælleren. Videre er $\eta_0 = \eta - x_0$ og for $n \geqslant 0$ er $\eta_{n+1} = 1/\eta_n - x_{n+1}$. Altså er alle resterne rødder. Desuden følger ligningen i (3.2.2) af (3.1.3).

Den første ligning i (3.2.3) følger umiddelbart af (3.2.2) og (3.2.1), idet normen er multiplikativ.

For at eftervise den anden ligning i (3.2.3) udnyttes formlerne i (2.8):

$$\lambda \begin{pmatrix} \eta \\ 1 \end{pmatrix} = S_n \begin{pmatrix} \eta_n \\ 1 \end{pmatrix}, \text{ hvor } \lambda = q_n + q_{n-1}\eta_n. \tag{3.2.5}$$

Af udtrykket $\eta_n = (\xi - e_n)/a_n = (a\eta - e_n)/a_n$ følger, at

$$\begin{pmatrix} \eta_n \\ 1 \end{pmatrix} = \begin{pmatrix} a/a_n & -e_n/a_n \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \eta \\ 1 \end{pmatrix},$$

og indsættelse i (3.2.5) giver derfor ligningen

$$\lambda \begin{pmatrix} \eta \\ 1 \end{pmatrix} = S_n \begin{pmatrix} a/a_n & -e_n/a_n \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \eta \\ 1 \end{pmatrix}.$$

Af denne ligning fremgår, at λ er egenværdi for produktet af de to matricer på højresiden. Følgelig er λ rod i det karakteristiske polynomium for produktmatricen. Dette polynomium er øjensynlig et normeret polynomium med rationale koefficienter, og den anden rod i dette polynomium må derfor være det til λ konjugerede tal λ'. Normen af λ, dvs $\lambda\lambda'$, er altså konstantleddet i det karakteristiske polynomium, og derfor lig med determinanten af produktmatricen. Da det $S_n = (-1)^n$, fås ligningen,

$$\mathrm{N}(\lambda) = (-1)^n a/a_n, \tag{3.2.6}$$

som kombineret med udtrykket for λ netop er den anden ligning i (3.2.3).

For at vise sætningens sidste formel udnyttes, at $\lambda(p_{n-1} - q_{n-1}\eta) = (-1)^n$, jfr (2.8.3). Anvendes normafbildningen herpå fås ligningen,

$$\mathrm{N}(p_{n-1} - q_{n-1}\eta) = 1/\mathrm{N}(\lambda). \tag{3.2.7}$$

Kombineres (3.1.1), (3.2.6) og (3.2.7), fås den søgte formel.

Hermed er sætningen bevist. ∎

(3.3) Sætning. *Lad τ være er irrationalt tal. Da er følgende betingelser ækvivalente:*

(i) *Tallet τ er et kvadratisk tal, dvs rod i et andengradspolynomium med hele koefficienter.*

(ii) *Følgen af kædebrøkskoefficienter $y_0, y_1, y_2, \ldots$ for τ er periodisk fra et vist trin, dvs der findes tal $n_0 \geqslant 0$ og $l > 0$ således, at $y_{i+l} = y_i$ for alle $i > n_0$.*

(iii) *Følgen af rester $\tau_0, \tau_1, \ldots$ indeholder en gentagelse, dvs der findes tal $n_0 \geqslant 0$ og $l > 0$ således, at $\tau_{n_0+l} = \tau_{n_0}$.*

Bevis. Resten τ_i er den ægte kædebrøk med koefficienter $y_{i+1}, y_{i+2}, \ldots$. Det er derfor klart, at (ii) og (iii) er ensbetydende.

Antag, at (iii) gælder. Det skal vises, at τ er et kvadratisk tal. For hvert n gælder en ligning af formen $\tau = (A\tau_n + B)/(C\tau_n + D)$, hvor A, B, C, D er koefficienterne i matricen S_n hørende til τ. Det er derfor nok at vise, at τ_n er et kvadratisk tal for en passende værdi af n. Vi kan således i resten af beviset erstatte τ med τ_{n_0}. Forudsætningen er da, at $\tau = \tau_l$. Med betegnelser hørende til kædebrøken for τ gælder ifølge (2.8.1) matrixligningen

$$\lambda \begin{pmatrix} \tau \\ 1 \end{pmatrix} = S_l \begin{pmatrix} \tau_l \\ 1 \end{pmatrix}, \text{ hvor } \lambda = q_{l-1}\tau_l + q_l.$$

Da $\tau = \tau_l$, viser matrixligningen, at λ er egenværdi for matricen S_l. Følgelig er λ rod i det karakteristiske polynomium for S_l. Dette polynomium er et normeret polynomium med hele koefficienter, så λ må være et (helt) kvadratisk tal. Af udtrykket ovenfor for λ følger nu, at også $\tau = \tau_l$ må være et kvadratisk tal. Altså gælder (i).

Antag omvendt, at (i) gælder. Det kan antages, at $\tau = \eta$ er det kvadratiske tal fra (3.1). Af formel (3.2.4) følger da, at

$$F(p_{n-1}, q_{n-1}) = \frac{a}{q_n + q_{n-1}\eta_n} \cdot \frac{1}{q_n + q_{n-1}\eta_n'} . \tag{3.3.1}$$

Her er venstresiden et helt tal forskelligt fra 0. Den første faktor på højresiden går mod 0 for $n \to \infty$, idet $q_n \to \infty$ og $\eta_n > 0$. Da ligningen gælder, slutter vi, at nævneren i den anden faktor må være numerisk mindre end 1, når n er tilstrækkelig stor. Da $0 < q_{n-1} < q_n$ for $n \geqslant 2$, følger det specielt, at der findes et index $r \geqslant 2$, så at der for alle $n \geqslant r$ gælder, at

$$\eta_n' < -1. \tag{3.3.2}$$

Mere præcist gælder uligheden (3.3.2), når den første faktor i (3.3.1) numerisk er mindre end 1 og $q_n > q_{n-1}$. Vælges specielt $r \geqslant 2$ så stor, at $q_r \geqslant |a|$, så gælder (3.3.2) altså, når $n \geqslant r$.

Antag nu, at $n \geqslant r$. Da er $\eta_n' < -1$, og vi har $0 < \eta_n < 1$. Med $\tau := \eta_n$ gælder altså ulighederne (3.1.4). Af overvejelserne i (3.1) følger derfor, at der kun er endelig mange muligheder for parret (a_n, e_n), endda højst $D - 1$ muligheder.

Blandt de D par (e_n, a_n) for $r \leqslant n \leqslant r + D - 1$ må der altså være to, der er ens. Altså findes n_0 og n_1 således, at $r \leqslant n_0 < n_1 < r + D$ og $(a_{n_0}, e_{n_0}) = (a_{n_1}, e_{n_1})$. Og så er $\eta_{n_0} = \eta_{n_1}$, og betingelsen (iii) er opfyldt.

Hermed er sætningen bevist. ∎

(3.4) Definition. Det kvadratiske tal η kaldes *reduceret*, hvis

$$0 < \eta < 1 \text{ og } \eta' < -1. \tag{3.4.1}$$

Resterne af et tal opfylder altid den første betingelse. I beviset for Sætningen, jfr (3.3.2), så vi, at resten η_n er reduceret, når $n \gg 0$. Mere præcist så vi, at η_n er reduceret, når $q_n \geqslant |a|$ med $n \geqslant 2$. Beviset for Sætningen byggede så videre på, at resterne η_n er af formen $(\xi - t)/s$, og at der kun er endelig mange reducerede tal af denne form. Mere præcist: antallet er højst $D - 1$.

Til det kvadratiske tal η findes ifølge sætningen et tal $n_0 \geqslant 0$ og et naturligt tal l, så at

$$\eta_{n_0} = \eta_{n_0+l}.$$

Vælg n_0 og l mindst mulige. Tallet l kaldes *periodelængden* for η. Følgen η_n for $n \geqslant n_0$ er øjensynlig periodisk med periode l. Derfor er også følgen a_n (og følgen e_n) for $n \geqslant n_0$ periodisk med periode l. Tallet η_n bestemmer koefficienten x_{n+1}, så følgen x_n for $n \geqslant n_0 + 1$ er også periodisk med periodelængden l. Tallet $n_0 + 1$, kaldet *periodestarten*, er altså det

trin, fra hvilket følgen $x_1, x_2, \ldots$ (med x_0 udeladt) er periodisk. Af beviset for sætning (3.3) følger, at periodelængden l er mindre end D.

For $i \geqslant n_0$ er $\eta_i = \eta_{i+l} = \cdots$, og η_n er reduceret for $n \gg 0$. Heraf følger, at η_i er reduceret for alle $i \geqslant n_0$. Det er en konsekvens af Galois's Sætning herunder, at tallet n_0 altid er det første index n for hvilket resten η_n er reduceret. Specielt ses, at hvis $q_n \geqslant |a|$ med $n \geqslant 2$, så er $n_0 \leqslant n$.

For $n \geqslant n_0$ er følgen a_n periodisk. Specielt antager følgen a_n kun endelig mange værdier. Af Euler's formel fremgår derfor, at følgen $F(p_n, q_n)$ kun antager endelig mange værdier. Specielt findes altså et helt tal k således, at $F(x, y) = k$ har uendelig mange løsninger (blandt konvergenterne (p_n, q_n)). Dette resultat er som bekendt et vigtigt skridt i Dirichlet's bevis for, at den egentlige Pell'ske ligning har uendelig mange løsninger. Som vi skal se herunder, giver kædebrøksudviklingen en mere direkte analyse af den Pell'ske ligning.

Bemærk i øvrigt, at uligheden (3.1.5), for $\tau := \eta_n$, medfører, at $1 \leqslant a_n < \sqrt{D}$ for $n \geqslant n_0$.

(3.5) Galois's sætning (1828)**, første del.** *Tallet η er reduceret, hvis og kun hvis der findes et naturligt tal l, så at $\eta = \eta_l$, altså hvis og kun hvis $x_0 = 0$ og følgen $x_1, x_2, \ldots$ er periodisk med perioden l.*

Bevis. „hvis": er trivielt, idet η_n er reduceret, når $n \gg 0$, jfr (3.4).

„kun hvis": Vi viser først, at hvis η er reduceret, så er alle rester η_n reducerede. Antag altså, at η er reduceret. Da er specielt $\eta = \eta_0$. Induktivt er det altså nok at vise, at η_1 er reduceret. Trivielt er $0 < \eta_1 < 1$. Den anden ulighed i (3.4.1) følger af vurderingen,

$$\eta_1' = -x_1 + \frac{1}{\eta_0'} < -x_1 \leqslant -1.$$

Dernæst bemærker vi, at hvis η er reduceret (og specielt lig med η_0), og $\eta_1 = \eta_{l+1}$, så er $\eta_0 = \eta_l$. Antag nemlig, at $\eta_1 = \eta_{l+1}$. Så følger det af $1/\eta_0 = x_1 + \eta_1$ og $1/\eta_l = x_{l+1} + \eta_{l+1}$, at $1/\eta_0 - 1/\eta_l = x_1 - x_{l+1}$. Efter konjugering fås ligningen,

$$\frac{1}{\eta_0'} - \frac{1}{\eta_l'} = x_1 - x_{l+1}.$$

Da η_0 og dermed også η_l (ifølge det allerede viste) er reducerede, ligger de to brøker på venstresiden i intervallet $]-1, 0[$. Differensen er derfor numerisk mindre end 1. Da højresiden er et helt tal, følger det, at differensen må være 0. Altså er $\eta_0 = \eta_l$.

Nu viser vi sætningen. Når $n \gg 0$, er $\eta_n = \eta_{n+l}$. Betragt det mindste n, for hvilket $\eta_n = \eta_{n+l}$. Hvis $n > 0$, følger det af bemærkningen, anvendt med $\eta := \eta_{n-1}$, at $\eta_{n-1} = \eta_{n-1+l}$, i modstrid med valget af n. Altså er $n = 0$, hvilket er påstanden i sætningen. ∎

(3.6) Korollar. *Hvis $1 \leqslant a < \frac{1}{2}\sqrt{D}$, så er resten η_0 reduceret, og følgerne $x_1, x_2, \ldots, e_0, e_1, e_2, \ldots$ og $a = a_0, a_1, a_2, \ldots$ er periodiske med periode l, og $a_n \geqslant 1$ for alle n. Hvis $a = 1$, gælder yderligere, at $a_i = 1$, hvis og kun hvis i er et multiplum af l.*

Bevis. Af $a \geqslant 1$ følger, at $\eta' < \eta$, og af $a < \frac{1}{2}\sqrt{D}$ følger, at $\eta - \eta' = \sqrt{D}/a > 2$. Altså er også $\eta_0 - \eta_0' > 2$. Da $\eta_0 < 1$, følger det, at $\eta_0' < \eta_0 - 2 < -1$. Altså er η_0 reduceret.

Påstanden om periodiciteten følger derfor af Galois's sætning. Da η_n er reduceret, følger det af (3.1.5) (som bemærket tidligere), at $a_n \geqslant 1$.

Antag, at $a = 1$ (og altså $\eta = \xi$). Da er forudsætningerne om a opfyldt, da $D = 5$ er den mindste diskriminant. Antag, at $a_i = 1$. Da er $\eta_0 = \xi - e_0$ og $\eta_i = \xi - e_i$. Da begge rester ligger i intervallet $]0, 1[$, er differensen $\eta_0 - \eta_i$ numerisk mindre end 1. På den anden side er differensen det hele tal $e_i - e_0$. Altså er $e_i = e_0$, og følgelig er $\eta_i = \eta_0$. Tallet i er derfor et multiplum af periodelængden for følgen η_n. □

(3.7) Den diofantiske ligning. Ofte betragtes med polynomiet $F(x, y)$ den diofantiske ligning $F(x, y) = k$. Højresiden k er altså et givet helt tal, og der søges heltalsløsninger til ligningen. Betragt først specialtilfældet, hvor $a = 1$. Med højresiden $k = \pm 1$ er det *Pell's ligning*,

$$x^2 + bxy + cy^2 = \pm 1; \tag{3.7.1}$$

med $k = +1$ er det den *egentlige Pell'ske ligning*, med $k = -1$ den *ikke-Pell'ske ligning*.

Løsninger (x, y) til Pell's ligning svarer til enheder $\varepsilon = x - y\xi$ i den kvadratiske talring $\mathbb{Z}[\xi]$. Med hver enhed ε er også $-\varepsilon$, ε^{-1} og $-\varepsilon^{-1}$ en enhed. Oftest er det derfor nok at betragte enheder $\varepsilon = x - y\xi$, som opfylder, at $1 < \varepsilon'$; som nævnt i (1.8) gælder denne betingelse på ε, hvis og kun hvis (x, y) tilhører området, hvor $y > 0$ og $2x + by > 0$.

Pell's ligning. *Løsningerne til Pell's ligning i det betragtede område er netop parrene af formen (p_{il-1}, q_{il-1}), hvor l er periodelængden for ξ. Mere præcist er $F(p_{il-1}, q_{il-1}) = (-1)^{li}$. For de tilhørende enheder $\varepsilon_i = p_{il-1} - q_{il-1}\xi$ gælder ulighederne,*

$$1 < \varepsilon_1' < \varepsilon_2' < \cdots .$$

(*Specielt er ε_1 den såkaldte grundenhed, og $\varepsilon_i = (\varepsilon_1)^i$*). *Den ikke-Pell'ske ligning har løsninger, hvis og kun hvis periodelængden l er ulige.*

Bevis. Lad (x, y) være en løsning i det betragtede område. Vi anvender overvejelserne i (2.17). Uligheden $|k| < \frac{1}{2}\sqrt{D}$ er opfyldt, da $k = \pm 1$ og $D = 5$ er den mindste positive diskriminant. Uligheden i (2.17.2) har formen,

$$\sqrt{D} + \sqrt{D - 4/y^2} > 4.$$

Det ses, at uligheden er opfyldt på nær når $D = 5$ og $y = 1$. På nær i undtagelsestilfældet følger det derfor, at x/y er en konvergent. I undtagelsestilfældet er $D = 5$ og $y = 1$. Da $y = 1$, er $x^2 + bx + c = \pm 1$, og $b^2 - 4c = 5$. Med $k = -1$ er altså $x = -b/2 \pm \sqrt{5-4}/2 = -b/2 \pm 1/2$; af de to løsninger ligger kun $(x_0, y_0) = (-b/2 + 1/2, 1)$ i området. Med $k = 1$ er $x = -b/2 \pm \sqrt{5+4}/2 = -b/2 \pm 3/2$; af de to løsninger ligger kun $(x_1, y_1) = (-b/2 + 3/2, 1)$ i området. Af de to brøker $x_0/y_0 = x_0$ og $x_1/y_1 = x_0 + 1$ er den første øjensynlig den hele del af $\xi = (-b + \sqrt{D})/2$, og derfor lig med konvergenten p_0/q_0. Den anden brøk er større end ξ, og den er derfor en konvergent ifølge sætningen i (2.17).

I alle tilfælde gælder altså, at x/y er en konvergent p_m/q_m for ξ. Da $F(x, y) = \pm 1$, må brøken x/y være uforkortelig, og da $y \geqslant 1$ følger det, at $(x, y) = (p_m, q_m)$. Værdierne

$F(p_m, q_m)$ er givet ved Euler's formel (3.2.4). Det ses, at værdien er ± 1, netop når $a_{m+1} = 1$. Af Korollar (3.6) følger, at dette indtræffer netop, når $m+1$ er et multiplum af l.

Hermed er de første påstande bevist.

De resterende påstande ses således: Konvergenten p_m/q_m er mindre en ξ, netop når m er lige. Heraf ses, at tallet ε_i har fortegnet $(-1)^{il}$. Da $\varepsilon_i \varepsilon_i' = (-1)^{li}$, følger det, at $\varepsilon_i' > 0$. Yderligere gælder ifølge (2.7), at $1 > |\varepsilon_1| > |\varepsilon_2| > \cdots$, og heraf følger, da $|\varepsilon_i \varepsilon_i'| = 1$, de påståede uligheder for ε_i'.

Enhederne ε med $\varepsilon' > 1$ er, ifølge det viste, netop enhederne ε_i. Heraf følger let, at ε_1 er grundenheden.

Den ikke-Pell'ske ligning har løsninger, hvis og kun hvis $(-1)^{il}$ for en passende værdi af i er lig med -1. Dette viser sætningens sidste påstand. □

Den almindelige diofantiske ligning,

$$ax^2 + bxy + cy^2 = k, \tag{3.7.2}$$

kan, for små værdier af k, analyseres direkte ved kædebrøksudviklingen for η:

Sætning. *Antag, at* $|k| < \frac{1}{2}\sqrt{D}$. *Da har ligningen* (3.7.2) *løsninger, hvis og kun hvis* k *har formen* $k = (-1)^n d^2 a_n$, *hvor* $d \geqslant 1$. *Ligningen har primitive løsninger, dvs løsninger med primiske tal* x, y, *hvis og kun hvis* k *forekommer i følgen* $(-1)^n a_n$.

Bevis. Hvis k har den angivne form, følger det af Euler's formel, at $(x, y) = (dp_{n-1}, dq_{n-1})$ er en løsning. Antag omvendt, at ligningen har løsninger (x, y). Under brug af, at den egentlige Pell'ske ligning har undelig mange løsninger, er det let at se, at (3.2.7) så har uendelig mange løsninger, endda uendelig mange løsninger i området, hvor $y > 0$ og $2ax + by > 0$. Der findes specielt løsninger i området, med $y \gg 0$. For en sådan løsning (x, y) gælder ifølge sætningen i (2.17), at x/y er lig med en konvergent p_m/q_m. Altså er $(x, y) = (dp_n, dq_n)$, hvor d er den største fælles divisor for x, y. Af Euler's formel følger nu, at $k = (-1)^{m+1} d^2 a_{m+1}$.

Hvis ligningen har primitive løsninger, følger det tilsvarende, at ligningen har primitive løsninger (x, y) i området, med $y \gg 0$. I dette tilfælde medfører ligningen $x/y = p_m/q_m$, at $(x, y) = (p_m, q_m)$, og dermed at $k = (-1)^{m+1} a_{m+1}$. □

Bemærk, at følgen a_n, for $n \geqslant n_0$, er periodisk med periode l. Følgen $(-1)^n a_n$, for $n \geqslant n_0$, har derfor periode $2l$, og den har periode l, hvis l er lige. I anvendelserne af sætningen er det derfor tilstrækkeligt at undersøge tallene $(-1)^i a_i$ for $i = 0, \ldots, n_0 + 2l - 1$ (og for lige l blot for $i = 0, \ldots, n_0 + l - 1$). Det fremgår i øvrigt af sætningen, at enhver værdi af $(-1)^i a_i$, hvor $|a_i| < \frac{1}{2}\sqrt{D}$, må have formen $(-1)^n d^2 a_n$ for $n \gg 0$.

(3.8) Galois's sætning, anden del. *Antag, at tallet* η *er reduceret, og lad* $0, x_1, x_2 \ldots$ *være følgen af kædebrøkskoefficienter, med periodelængden* l. *Da er tallet* $-1/\eta'$ *ligeledes reduceret, og den hertil hørende følge af kædebrøkskoefficienter er følgen*

$$0, x_l, x_{l-1}, \ldots, x_2, x_1, x_l, x_{l-1}, \ldots ,$$

ligeledes med med periodelængden l.

Bevis. I følge den første del, Sætning (3.5), er alle resterne η_n reducerede. For at bevise sætningen udnyttes relationen (2.8.1):

$$\begin{pmatrix}\eta\\1\end{pmatrix} \sim S_l \begin{pmatrix}\eta_l\\1\end{pmatrix} = S_l \begin{pmatrix}\eta\\1\end{pmatrix}. \tag{3.8.1}$$

Med $S := \left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right)$ gælder øjensynlig, at $\left(\begin{smallmatrix}-1/\eta\\ 1\end{smallmatrix}\right) \sim S\left(\begin{smallmatrix}\eta\\ 1\end{smallmatrix}\right)$. Af (3.8.1) fås derfor, at

$$\begin{pmatrix}-1/\eta\\1\end{pmatrix} \sim SS_l \begin{pmatrix}\eta\\1\end{pmatrix} \sim SS_lS^{-1} \begin{pmatrix}-1/\eta\\1\end{pmatrix}. \tag{3.8.2}$$

Da η er reduceret, er $x_0 = 0$. Matricen S_l er derfor et produkt af matricer af formen $\left(\begin{smallmatrix}0 & 1\\ 1 & x\end{smallmatrix}\right)$. For en sådan matrix finder man

$$S\begin{pmatrix}0 & 1\\1 & x\end{pmatrix}S^{-1} = \begin{pmatrix}x & -1\\-1 & 0\end{pmatrix} \sim \begin{pmatrix}-x & 1\\1 & 0\end{pmatrix} = \begin{pmatrix}0 & 1\\1 & x\end{pmatrix}^{-1}.$$

Ved gentagen anvendelse heraf ses, at relationen (3.8.2) kan skrives,

$$\begin{pmatrix}-1/\eta\\1\end{pmatrix} \sim \begin{pmatrix}0 & 1\\1 & x_1\end{pmatrix}^{-1} \cdots \begin{pmatrix}0 & 1\\1 & x_l\end{pmatrix}^{-1} \begin{pmatrix}-1/\eta\\1\end{pmatrix}.$$

Multipliceres med den inverse til produktmatricen, fås efter konjugering følgende relation,

$$\begin{pmatrix}0 & 1\\1 & x_l\end{pmatrix} \cdots \begin{pmatrix}0 & 1\\1 & x_1\end{pmatrix} \begin{pmatrix}-1/\eta'\\1\end{pmatrix} \sim \begin{pmatrix}-1/\eta'\\1\end{pmatrix}.$$

Da $\eta' < -1$, er $0 < -1/\eta' < 1$. Af Lemma (2.11) følger derfor, at $(-1/\eta')_l = -1/\eta'$, og at kædebrøkskoefficienterne for $-1/\eta'$ er følgen angivet i sætningen.

Hermed er sætningen bevist. ∎

(3.9) Sætning. *Antag, at* $1 \leqslant a < \frac{1}{2}\sqrt{D}$ *og at* a *er divisor i* b. *Da er følgerne* $x_1, x_2, \ldots$, $a = a_0, a_1, a_2, \ldots$ *og* $e_0, e_1, \ldots$ *periodiske med periode* l, *hvor* l *periodelængden for* η. *Yderligere gælder:*

(1) *For* $i = 1, \ldots, l-1$ *er* $x_{l-i} = x_i$. *Desuden er* $x_l = 2x_0 + (b/a)$.
(2) *For* $i = 0, \ldots, l-1$ *er* $\eta_{l-1-i} = -1/\eta_i'$.
(3) *For* $i = 0, \ldots, l-1$ *er* $\eta_{l-i-1} = (\xi - e_i)/a_{i+1}$, *dvs* $a_{l-1-i} = a_{i+1}$ *og* $e_{l-1-i} = e_i$.
(4) *For hvert* $i = 0, \ldots, l-1$ *gælder:* $a_i = a_{i+1}$, *hvis og kun hvis* $l = 2i+1$, *og* $e_i = e_{i+1}$, *hvis og kun hvis* $l = 2i+2$ *eller* $i = l-1$.

Hvis $a = 1$, *og altså* $\xi = \eta$, *så er begge antagelser opfyldt. I dette tilfælde gælder yderligere, at* $a_n = 1$, *hvis og kun hvis* n *er et multiplum af* l.

Bevis. Som nævnt i (3.6) er η_0 reduceret, så periodiciteten følger af første del af Galois's sætning. Ifølge anden del er $-1/\eta_0'$ ligeledes reduceret med kædebrøkskoefficienter

$x_l, \ldots, x_2, x_1, \ldots$, og perioden l. Specielt er den l'te rest af $-1/\eta_0'$ lig med $-1/\eta_0'$, så vi har relationen,

$$\begin{pmatrix} -1/\eta_0' \\ 1 \end{pmatrix} \sim \begin{pmatrix} 0 & 1 \\ 1 & x_l \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & x_1 \end{pmatrix} \begin{pmatrix} -1/\eta_0' \\ 1 \end{pmatrix}. \tag{3.9.1}$$

Sæt $B := -b/a - x_0$. Da er B et helt tal, og $\eta = -b/a - \eta' = B - \eta_0'$. Altså er

$$\begin{pmatrix} \eta \\ 1 \end{pmatrix} = \begin{pmatrix} B & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -\eta_0' \end{pmatrix} \sim \begin{pmatrix} B & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1/\eta_0' \\ 1 \end{pmatrix}.$$

Videre er

$$\begin{pmatrix} B & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & x_l \end{pmatrix} = \begin{pmatrix} 1 & B + x_l \\ 0 & 1 \end{pmatrix}.$$

Muliplikation af (3.9.1) med $\left(\begin{smallmatrix} B & 1 \\ 1 & 0 \end{smallmatrix}\right)$ giver derfor relationen,

$$\begin{pmatrix} \eta \\ 1 \end{pmatrix} \sim \begin{pmatrix} 1 & B + x_l \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & x_{l-1} \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & x_1 \end{pmatrix} \begin{pmatrix} -1/\eta_0' \\ 1 \end{pmatrix} \tag{3.9.2}$$

Ifølge antagelsen er B et helt tal. På højresiden er $0 < -1/\eta_0' < 1$. Af relationen (3.9.2) kan vi derfor aflæse koefficienterne $x_0, \ldots, x_{l-1}$ i kædebrøken for η, jfr (2.11). Vi får ligningerne $x_0 = B + x_l$, $x_1 = x_{l-1}, \ldots, x_{l-1} = x_1$. Da $B = -b/a + x_0$, er det netop ligningerne i (1).

Antag nu, at $0 \leqslant i < l$. For at eftervise (2) betragtes tallene η_{l-1-i} og $-1/\eta_i'$. Idet vi anvender Galois's sætning (anden del) på det andet tal, ser vi, at begge tal er reducerede med periodelængde l, og at de første l kædebrøkskoefficienter (den 0'te fraregnet) er de to følger:

$$x_{l-i}, x_{l-i+1}, \ldots x_{l-1}, x_l, x_1, , \ldots, x_{l-1-i}, \qquad x_i, x_{i-1}, \ldots, x_1, x_l, x_{l-1}, \ldots x_{i+1}.$$

Af ligningerne i (1) fremgår, at de to følger er identiske. Derfor er $\eta_{l-1-i} = -1/\eta_i'$.

Påstanden i (3) er en konsekvens af (2), idet vi får ligningerne,

$$\eta_{l-1-i} = \frac{-1}{\eta_i'} = \frac{-a_i}{\xi' - e_i} = \frac{-a_i(\xi - e_i)}{\mathrm{N}(\xi - e_i)} = \frac{\xi - e_i}{a_{i+1}},$$

hvor den sidste ligning gælder ifølge (3.2.2).

Betragt nu (4). Af ligning (3) fremgår, at $a_i = a_{i+1}$, hvis og kun hvis $\eta_{l-1-i} = \eta_i$. Da l er periodelængden, er η_j'erne for $j = 0, \ldots, l-1$ indbyrdes forskellige. Følgelig gælder den sidste ligning, hvis og kun hvis $l - 1 - i = i$, dvs hvis og kun hvis $l = 2i + 1$. Tilsvarende ses, at $e_i = e_{i+1}$, hvis og kun hvis $\eta_{l-1-i} = \eta_{i+1}$, og at det indtræffer, hvis og kun hvis $l = 2i + 2$ eller $i = l - 1$. Hermed er (4) bevist.

Sætningens sidste påstand fremgår af (3.6) ∎

(3.10) Note. Det følger af Sætning (3.5), for $1 \leqslant a < \frac{1}{2}\sqrt{D}$, at kædebrøksudviklingen af η bestemmes ud fra de første koefficienter $x_0, x_1, \ldots, x_l$. Hvis a desuden er divisor i b (bemærk, at denne betingelse er opfyldt, hvis $a = 1$ eller $b = 0$), så følger det af sætningen, at kendskab til den første „halvdel" af disse koefficienter er tilstrækkeligt. Under disse antagelser kan periodelængden l bestemmes udfra e_i'erne og a_i'erne. Ifølge sætningen findes nemlig enten et i, så at $a_i = a_{i+1}$, eller et i, så at $e_i = e_{i+1}$. Er i mindst mulig, og indtræffer det første, så er $l = 2i + 1$ (og periodelængden er altså ulige), indtræffer det andet, så er $l = 2i + 2$ (og periodelængden lige).

Hvis $b = 0$, har vi $F(x, y) = ax^2 - dy^2$, hvor $d := -c$, og vi kan antage, at a og dermed d er positive. Diskriminanten er $D = 4ad$, og vi har $a < \frac{1}{2}\sqrt{D}$, netop når $a < d$. I forbindelse med den diofantiske ligning $ax^2 - dy^2 = k$, kan vi skifte fortegn og ombytte x og y. Det er derfor sædvanligvis ingen indskrænkning at antage, at $a < d$.

(3.11) En anvendelse. *Antag, at $1 < a < \frac{1}{2}\sqrt{D}$, og at a er divisor i b. Da gælder: Af de tre ligninger $F(x, y) = 1$, $F(x, y) = -1$ og $x^2 + bxy + acy^2 = -1$ er der højst én, der har heltalsløsninger.*

Bevis. Den sidste ligning er den „ikke-Pell'ske". Som nævnt ovenfor har den løsninger, hvis og kun hvis perioden for ξ er ulige.

Af Sætning (3.6) fremgår, at følgen a_n for $n \geqslant 0$ er periodisk med periode l, og $a_n \geqslant 1$. Antag, at en af de to første ligninger har en løsning. Det følger da af resutatet i (3.7), at ± 1 forekommer i følgen $(-1)^n a_n$. Altså findes et $i < l$, så at $a_i = 1$. Da $a > 1$, er $i > 0$. For dette i er $\eta_i = \xi - e_i$. Ifølge (3.9)(3) er $\eta_{l-i} = \xi - e_{i-1}$. Specielt ses, at differensen mellem resterne, $\eta_{l-i} - \eta_i$, er et helt tal. Da resterne er reducerede, følger det, at de er ens, altså $\eta_i = \eta_{l-i}$. Da både i og $l - i$ er mindre end l, følger det, at $i = l - i$. Altså er $l = 2i$. Specielt ses, at l er lige, og at i var entydigt bestemt. Af udtrykket $\eta_i = \xi - e_i$ fremgår, at η_i og ξ har samme kædebrøksudvikling, bortset fra det første hele tal x_0 i udviklingen. Specielt har også ξ perioden l. Da l er lige, har den ikke-Pell'ske ligning altså ingen løsninger. Igen, da periodelængden l er lige, viser entydigheden af i, at højst en af de to første ligninger har løsninger.

Hermed er korollaret vist. ∎

(3.13) Eksempel. Polynomiet $F = x^2 - 13y^2$ har diskriminanten $D = 52$. Her er $\xi = \eta = \sqrt{13}$, og $b = 0$, så antagelserne i (3.9) er opfyldt. Man finder

$$\begin{aligned}
\eta = \sqrt{13} &= \frac{\sqrt{13} - 0}{1} &&= 3 + \frac{\sqrt{13} - 3}{1}\,,\\
1/\eta_0 &= \frac{1}{\sqrt{13} - 3} = \frac{\sqrt{13} + 3}{4} &&= 1 + \frac{\sqrt{13} - 1}{4}\,,\\
1/\eta_1 &= \frac{4}{\sqrt{13} - 1} = \frac{\sqrt{13} + 1}{3} &&= 1 + \frac{\sqrt{13} - 2}{3}\,,\\
1/\eta_2 &= \frac{3}{\sqrt{13} - 2} = \frac{\sqrt{13} + 2}{3} &&= \ldots\,.
\end{aligned}$$

Hermed er x_i, a_i, og e_i bestemt for $i = 0, 1, 2$. Vi behøver ikke fuldføre omskrivningen af $1/\eta_2$, idet det allerede fremgår, at $a_2 = a_3$. Periodelængden er derfor $l = 5$. Ved brug af (3.9) kan vi udfylde første og sidste række i skemaet herunder. De to mellemste rækker er udfyldt ved brug af rekursionsformlen for konvergenterne.

n	-2	-1	0	1	2	3	4	5	6	7	8	9	10
x_n			3	1	1	1	1	6	1	1	1	1	6
p_n	0	1	3	4	7	11	18	119	137	256	393	649	
q_n	1	0	1	1	2	3	5	33	38	71	109	180	
a_n			1	4	3	3	4	1	4	3	3	4	1

Af Euler's formel følger, at $p_{n-1}^2 - 13q_{n-1}^2 = (-1)^n a_n$. Specielt er $18^2 - 13 \cdot 5^2 = -1$ og $649^2 - 13 \cdot 180^2 = 1$. Af sætningen i (3.7) følger, at ligningen $x^2 - 13y^2 = \pm 2$ ikke har heltalsløsninger.

(3.14) Eksempel. Polynomiet $F = x^2 - 19y^2$ har diskriminant $D = 4 \cdot 19$. Igen er $a = 1$ og $b = 0$, så antagelserne i (3.9) er opfyldt. Man finder

$$\begin{aligned}
\eta = \sqrt{19} &= \frac{\sqrt{19} - 0}{1} &&= 4 + \frac{\sqrt{19} - 4}{1}, \\
1/\eta_0 &= \frac{1}{\sqrt{19} - 4} = \frac{\sqrt{19} + 4}{3} &&= 2 + \frac{\sqrt{19} - 2}{3}, \\
1/\eta_1 &= \frac{3}{\sqrt{19} - 2} = \frac{\sqrt{19} + 2}{5} &&= 1 + \frac{\sqrt{19} - 3}{5}, \\
1/\eta_2 &= \frac{5}{\sqrt{19} - 3} = \frac{\sqrt{19} + 3}{2} &&= 3 + \frac{\sqrt{19} - 3}{2}.
\end{aligned}$$

Heraf fremgår, at $e_2 = e_3$. Periodelængden er derfor $l = 6$. Vi kan nu udfylde første og sidste række i skemaet herunder. De to mellemste rækker er udfyldt ved brug af rekursionsformlen for konvergenterne.

n	0	1	2	3	4	5	6
x_n	4	2	1	3	1	2	8
p_n	4	9	13	48	61	170	
q_n	1	2	3	11	14	39	
a_n	1	3	5	2	5	3	1

Af Euler's formel følger, at $p_{n-1}^2 - 19q_{n-1}^2 = (-1)^n a_n$. Specielt er $170^2 - 19 \cdot 39^2 = 1$. Da periodelængden er lige, har den ikke-Pell'ske ligning $x^2 - 19y^2 = -1$ ingen løsninger. Betingelsen $|k| < \frac{1}{2}\sqrt{D}$ er opfyldt for $|k| \leqslant 4$. Af sætningen i (3.7) følger, at ingen af ligningerne $F(x, y) = k$, hvor $k = 2$, 3, eller -4, har heltalsløsninger, og at der med $k = 4$ ikke er primitive løsninger.

(3.15) Eksempel. Polynomiet $F = 3x^2 - 17y^2$ har diskriminanten $D = 2^2 \cdot 3 \cdot 17$. Her er $\eta = \sqrt{51}/3$ og $\xi = \sqrt{51}$. Antagelserne i (3.9) er opfyldt. Man finder

$$\begin{aligned}
\eta &= \frac{\sqrt{51}}{3} && = 2 + \frac{\sqrt{51}-6}{3}\,,\\
1/\eta_0 &= \frac{3}{\sqrt{51}-6} = \frac{\sqrt{51}+6}{5} && = 2 + \frac{\sqrt{51}-4}{5}\,,\\
1/\eta_1 &= \frac{5}{\sqrt{51}-4} = \frac{\sqrt{51}+4}{7} && = 1 + \frac{\sqrt{51}-3}{7}\,,\\
1/\eta_2 &= \frac{7}{\sqrt{51}-3} = \frac{\sqrt{51}+3}{6} && = 1 + \frac{\sqrt{51}-3}{6}\,.
\end{aligned}$$

Heraf fremgår, at $e_2 = e_3$. Periodelængden er derfor $l = 6$. Som ovenfor udfyldes skemaet:

n	0	1	2	3	4	5	6
x_n	2	2	1	1	1	2	4
p_n	2	5	7	12	19	50	
q_n	1	2	3	5	8	21	
a_n	3	5	7	6	7	5	3

Specielt ses, at $3 \cdot 50^2 - 17 \cdot 21^2 = 3$. Betingelsen $|k| < \frac{1}{2}\sqrt{D}$ er opfyldt for $|k| \leqslant 7$. Af sætningen i (3.7) følger, at ligningen $F(x, y) = k$, for $k = \pm 1$, ± 2, -3, ± 4, 5, 6, eller -7 ikke har heltalsløsninger.

(3.16) Eksempel. Ligningen $2x^2 - 3y^2 = -1$ har en løsning. Af (3.11) fremgår, at ingen af ligningerne $2x^2 - 3y^2 = 1$ og $x^2 - 6y^2 = -1$ har heltalsløsninger.

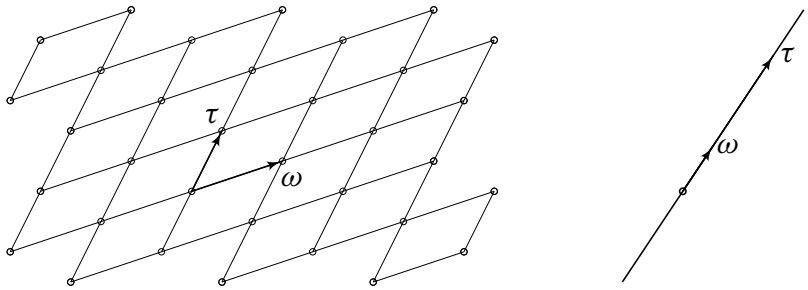
4. Similaritet af gitre.

(4.1) Definition. Ved et *gitter* forstås i det følgende en undergruppe $\Omega \subseteq \mathbb{C}$ af de komplekse tals additive gruppe således, at Ω er fri af rang 2. Som bekendt siges Ω at være fri af rang 2, hvis Ω har en *basis* (τ, ω) med 2 elementer, dvs hvis der findes to tal τ og ω i Ω således, at ethvert tal z i Ω entydigt er en heltalslinearkombination af τ og ω,

$$z = x\tau + y\omega, \qquad \text{med } x, y \in \mathbb{Z}. \tag{4.1.1}$$

Fremstillingen (4.1.1) er øjensynlig entydig, netop når ω og τ er forskellige fra 0 og kvotienten τ/ω ikke er et rationalt tal.

For en basis (τ, ω) for Ω er tallet τ/ω enten reelt (og irrationalt) eller imaginært. I det første tilfælde ligger alle tallene i Ω på en linie (nemlig på linien gennem 0 og ω), og vi taler om et *reelt gitter*; det er let at se, at tallene i Ω ligger overalt tæt på denne linie. I det andet tilfælde udspænder τ og ω den komplekse plan (som vektorrum over $\mathbb{R}$), og vi taler om et *imaginært gitter*; her kan tallene i Ω „ses“ som „gitterpunkter“ i planen. Bemærk modstriden med virkeligheden: de imaginære gitre kan „ses“, de reelle gittere kan ikke ses:



(4.2) Definition. Hvis (τ, ω) er en basis for et gitter Ω, så er de øvrige baser for Ω af formen,

$$\begin{pmatrix} \hat{\tau} \\ \hat{\omega} \end{pmatrix} = T \begin{pmatrix} \tau \\ \omega \end{pmatrix}, \qquad T = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{4.2.1}$$

hvor *skiftematricen* T er en matrix i $\mathrm{GL}_2(\mathbb{Z})$ (en heltalsmatrix med determinant ± 1). Det følger, at baserne for Ω falder i 2 klasser: to baser for Ω ligger i samme klasse, og siges at være *ensorienterede*, hvis skiftematricen har determinant $+1$; de ligger i hver sin klasse, og siges at være *modsat orienterede*, hvis matricen har determinant -1. [Bemærk, at baserne i (4.2.1) står som søjler; det er altså T^{tr}, der indgår ved koordinatskift.]

Gitteret Ω kaldes *orienteret*, hvis der er valgt en af de to klasser af baser. Ved en basis for et orienteret gitter Ω forstås altid en basis, som tilhører den udvalgte klasse.

Et givet gitter Ω kan orienteres ved at fremhæve en basis (τ, ω) for Ω; man skriver da også

$$\Omega = (\tau, \omega)\mathbb{Z}.$$

Lighedstegnet er her en lighed mellem orienterede gitre.

Bemærk, at baserne (τ, ω) og (ω, τ) er modsat orienterede. Som orienterede gitre er altså $(\tau, \omega)\mathbb{Z} \neq (\omega, \tau)\mathbb{Z}$. Derimod er $(\tau, \omega)\mathbb{Z} = (-\omega, \tau)\mathbb{Z}$.

(4.3) Definition. Betragt et imaginært gitter Ω. En basis (τ, ω) for Ω kaldes *positiv*, hvis tallet ω/τ har positiv imaginærdel (dvs ligger i den øvre halvplan), altså hvis omløbsretningen fra den første basisvektor, τ, til den anden, ω, er den sædvanlige positive omløbsretning.

Med den notation, vi har valgt, er det mest naturligt at se på kvotienten τ/ω. Hvis $(\hat{\tau}, \hat{\omega})$ er en anden basis for Ω, bestemt ved skiftematricen i (4.2.1), får man let ligningen,

$$\Im m(\hat{\tau}/\hat{\omega}) = \frac{\det T}{|C(\tau/\omega) + D|^2} \Im m(\tau/\omega). \tag{4.3.1}$$

Heraf fremgår, at determinanten af T er positiv, hvis og kun hvis imaginærdelene af $\hat{\tau}/\hat{\omega}$ og τ/ω har samme fortegn, dvs hvis og kun hvis de to baser enten begge er positive eller begge er negative. Baserne (τ, ω) for Ω, med τ/ω i den øvre halvplan, er altså ensorienterede. De bestemmer en orientering af Ω, som vi vil kalde den *kanoniske* orientering.

(4.4) Definition. Antag, at Ω og Ω_1 er gitre med $\Omega_1 \subseteq \Omega$. For givne baser (τ, ω) og (τ_1, ω_1) for Ω og Ω_1 kan kan tallene τ_1 og ω_1 udtrykkes som heltalslinearkombinationer af τ og ω. Der findes altså en relation,

$$\begin{pmatrix} \tau_1 \\ \omega_1 \end{pmatrix} = U \begin{pmatrix} \tau \\ \omega \end{pmatrix}, \tag{4.4.1}$$

hvor U er en heltalsmatrix. Af Elementardivisorsætningen følger, at der findes matricer $S, T \in \mathrm{SL}_2(\mathbb{Z})$ således, at SUT er en diagonalmatrix,

$$SUT = \begin{pmatrix} d & 0 \\ 0 & e \end{pmatrix}. \tag{4.4.2}$$

De to ligninger,

$$\begin{pmatrix} \hat{\tau} \\ \hat{\omega} \end{pmatrix} := T^{-1} \begin{pmatrix} \tau \\ \omega \end{pmatrix}, \qquad \begin{pmatrix} \hat{\tau}_1 \\ \hat{\omega}_1 \end{pmatrix} := S \begin{pmatrix} \tau_1 \\ \omega_1 \end{pmatrix}, \tag{4.4.3}$$

definerer nye baser for henholdvis Ω og Ω_1. Af de foregående ligninger følger, at

$$\begin{pmatrix} \hat{\tau}_1 \\ \hat{\omega}_1 \end{pmatrix} = SUT \begin{pmatrix} \hat{\tau} \\ \hat{\omega} \end{pmatrix} = \begin{pmatrix} d & 0 \\ 0 & e \end{pmatrix} \begin{pmatrix} \hat{\tau} \\ \hat{\omega} \end{pmatrix}.$$

Med hensyn til den nye basis $(\hat{\tau}, \hat{\omega})$ for Ω er den nye basis $(\hat{\tau}_1, \hat{\omega}_1)$ for Ω_1 altså bestemt ved

$$\hat{\tau}_1 = d\hat{\tau}, \qquad \hat{\omega}_1 = e\hat{\omega}.$$

Heraf fremgår først, da Ω_1 er et gitter, at d og e er forskellige fra 0. Videre følger det, at kvotientgruppen Ω/Ω_1 er isomorf med $\mathbb{Z}/|d|\mathbb{Z} \times \mathbb{Z}/|e|\mathbb{Z}$. Specielt har Ω_1 altså endeligt index i gruppen Ω, idet vi har $|\Omega : \Omega_1| = |de|$. Bemærk, at dette index er den numeriske værdi af determinanten af diagonalmatricen. Da matricerne S og T i (4.4.2) har determinant 1, har også U determinanten $|de|$. Vi har altså formlen,

$$|\Omega : \Omega_1| = |\det U|.$$

Antag nu, at de givne gitre Ω og Ω_1 er orienterede, og vælg baserne (τ, ω) og (τ_1, ω_1) i overensstemmelse hermed. Definer det *orienterede index* $(\Omega : \Omega_1)$ ved ligningen,

$$(\Omega : \Omega_1) := \det U. \tag{4.4.4}$$

Nye baser for de orienterede gitre Ω og Ω_1 fremgår af de gamle ved ligninger af formen (4.4.3), hvor matricerne S og T har determinant $+1$. Med hensyn til de nye baser ændres matricen U i (4.4.1) til SUT. Heraf følger, at højresiden i (4.4.4) er uafhængig af valg af basis for de orienterede gitre. Det orienterede index er altså veldefineret.

Det er klart, at ved orienteringsskift i et af gitrene Ω eller Ω_1 skifter det orienterede index fortegn. Det orienterede index $|\Omega : \Omega_1|$ er lig med 1, hvis $\Omega_1 = \Omega$ (som orienterede gitre), og lig med -1, hvis Ω_1 er gitteret Ω med den modsatte orientering.

(4.5) Definition. Lad λ være et komplekst tal forskelligt fra 0. Ved multiplikationen $z \mapsto \lambda z$ afbildes gitteret Ω øjensynlig på et gitter $\lambda\Omega$: hvis (τ, ω) er en basis for Ω, så er $(\lambda\tau, \lambda\omega)$ en basis for $\lambda\Omega$. Hvis $(\hat{\tau}, \hat{\omega})$ er en anden basis for Ω, bestemt ved skiftematricen T i (4.2.1), så er T også skiftematricen mellem $(\lambda\hat{\tau}, \lambda\hat{\omega})$ og $(\lambda\tau, \lambda\omega)$. Ensorienterede baser for Ω afbildes altså på ensorienterede baser for $\lambda\Omega$. Er Ω et orienteret gitter, så er billedet $\lambda\Omega$ derfor igen et orienteret gitter.

To gitre Ω og $\hat{\Omega}$ kaldes *similære*, hvis der findes et komplekst tal $\lambda \neq 0$ således, at

$$\hat{\Omega} = \lambda\Omega. \tag{4.5.1}$$

Er Ω og $\hat{\Omega}$ orienterede gitre, siges Ω og $\hat{\Omega}$ at være *similære*, hvis der findes et komplekst tal λ således, at ligningen (4.5.1) er opfyldt som en ligning melllem orienterede gitre. Når det skal understreges, at to gitre er similære som ikke-orienterede gitre, siger vi også, at de er *ikke-orienteret* similære.

Bemærk, at vi for et givet orienteret gitter $\Omega = (\tau, \omega)\mathbb{Z}$ kan vælge $\lambda := 1/\omega$: Gitteret $(\tau, \omega)\mathbb{Z}$ er altså similært med gitteret $(\tau/\omega, 1)\mathbb{Z}$. Specielt er ethvert orienteret gitter similært med et gitter af formen $(\tau, 1)\mathbb{Z}$.

Lad $\hat{\Omega}$ og Ω være orienterede gitre: $\Omega = (\tau, \omega)\mathbb{Z}$ og $\hat{\Omega} = (\hat{\tau}, \hat{\omega})\mathbb{Z}$. Da gælder ligningen (4.5.1), hvis og kun hvis $(\lambda\tau, \lambda\omega)$ er en basis for $\hat{\Omega}$, altså hvis og kun hvis,

$$\begin{pmatrix} \hat{\tau} \\ \hat{\omega} \end{pmatrix} = T \begin{pmatrix} \lambda\tau \\ \lambda\omega \end{pmatrix}, \quad \text{med } T = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}). \tag{4.5.2}$$

At ligningen (4.5.2) er opfyldt med et tal $\lambda \neq 0$ i $\mathbb{C}$ er ensbetydende med, at de to sider er *proportionale*, altså ensbetydende med relationen,

$$\begin{pmatrix} \hat{\tau} \\ \hat{\omega} \end{pmatrix} \sim T \begin{pmatrix} \tau \\ \omega \end{pmatrix}. \tag{4.5.3}$$

De to orienterede gitre, med givne baser, er altså similære, hvis og kun hvis der findes en relation (4.5.3) med en matrix $T \in \mathrm{SL}_2(\mathbb{Z})$. Tilsvarende er gitrene ikke-orienteret similære, hvis og kun hvis der findes en relation (4.5.3) med en matrix $T \in \mathrm{GL}_2(\mathbb{Z})$.

På de to sider af (4.5.3) er søjlernes anden-koordinat forskellig fra 0. Relationen er derfor ækvivalent med følgende lighed:

$$\hat{\tau}/\hat{\omega} = \frac{A\tau + B\omega}{C\tau + D\omega}.$$

(4.6) Definition. Lad τ være et irrationalt (komplekst) tal. For en matrix $T \in \mathrm{GL}_2(\mathbb{Z})$ skriver vi

$$T(\tau) := \frac{A\tau + B}{C\tau + D}, \quad T = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

Brøkens nævner er ikke 0, da τ ikke er et rationalt tal, så $T(\tau)$ er veldefineret. Ligningen $\hat{\tau} = T(\tau)$ kan alternativt udstrykkes ved relationen,

$$\begin{pmatrix} \hat{\tau} \\ 1 \end{pmatrix} \sim T \begin{pmatrix} \tau \\ 1 \end{pmatrix}.$$

Heraf følger let, at vi for to matricer $S, T \in \mathrm{GL}_2(\mathbb{Z})$ har ligningen $ST(\tau) = S(T(\tau))$.

To irrationale tal $\hat{\tau}$ og τ kaldes *svagt similære*, hvis der findes en matrix $T \in \mathrm{GL}_2(\mathbb{Z})$ således, at

$$\hat{\tau} = T(\tau). \tag{4.6.1}$$

De kaldes *similære*, hvis ligningen (4.6.1) er opfyldt med en matrix i $\mathrm{SL}_2(\mathbb{Z})$.

(4.7) Lemma. *Lad $\hat{\Omega} = (\hat{\tau}, \hat{\omega})\mathbb{Z}$ og $\Omega = (\tau, \omega)\mathbb{Z}$ være orienterede gitre. Da er $\hat{\Omega}$ og Ω similære, hvis og kun hvis tallene $\hat{\tau}/\hat{\omega}$ og τ/ω er similære. Gitrene er ikke-orienteret similære, hvis og kun hvis tallene er svagt similære.*

Bevis. Påstanden følger af definitionerne og udregningerne i (4.5). □

(4.8) Bemærkning. Spørgsmålet om similaritet af to gitre, med givne baser, „reduceres“ ifølge Lemmaet til spørgsmålet om similaritet af to givne irrationale tal $\hat{\tau}$ og τ: hvordan afgør man om $\hat{\tau}$ og τ er similære, og hvordan bestemmes, hvis de er similære, en matrix $T \in \mathrm{SL}_2(\mathbb{Z})$ således, at $\hat{\tau} = T(\tau)$? Som vi skal se, er dette ækvivalent med et problem om kædebrøker. I det reelle tilfælde kan problemet afgøres ved at kigge på kædebrøks-resterne af de to tal. Det imaginære tilfælde kan faktisk behandles tilsvarende, men vi vil her betragte en variant (defineret nedenfor) af den sædvanlige kædebrøksudvikling. I det reelle tilfælde er problemet endeligt, hvis tallene er kvadratiske tal, i det imaginære tilfælde er problemet altid endeligt.

Lad os iøvrigt bemærke, at i det imaginære tilfælde er pointen med at skelne mellem similaritet og svag similaritet uinteressant. To similære ikke-reelle tal ligger i samme halvplan (øvre eller nedre), og to svagt similære tal er similære, hvis og kun hvis de ligger i samme halvplan.

Lad os først betragte det reelle tilfælde:

(4.9) Lemma. *For en matrix $T \in \mathrm{GL}_2(\mathbb{Z})$ er følgende betingelser ækvivalente:*

(i) *T er en kædebrøksmatrix, dvs der findes $h \geqslant 0$ og hele tal $y_0, y_1, \ldots, y_h$, med $y_i \geqslant 1$ for $i \geqslant 1$, således, at T har formen,*

$$T = T_h = \begin{pmatrix} 1 & y_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & y_1 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & y_h \end{pmatrix}.$$

(ii) *T har en af følgende former,*

$$T = \begin{pmatrix} 1 & P \\ 0 & 1 \end{pmatrix}, \quad T = \begin{pmatrix} P-1 & P \\ 1 & 1 \end{pmatrix}, \quad T = \begin{pmatrix} P' & P \\ Q' & Q \end{pmatrix}, \quad \textit{hvor} \quad 0 < Q' < Q.$$

Bevis. Hvis T har formen i (i), så er, med den sædvanlige betegnelser for konvergenter,

$$T = T_h = \begin{pmatrix} p_{h-1} & p_h \\ q_{h-1} & q_h \end{pmatrix},$$

og $\det T_h = (-1)^h$. Videre er $0 = q_{-1} < 1 = q_0 \leqslant q_1 < q_2 < \ldots$. Hvis $q_{h-1} = 0$, så er $h = 0$ (og altså $q_h = 1$); da $\det T_0 = 1$, følger det, T er af den første form i (ii). Hvis $0 < q_{h-1} = q_h$, så er nødvendigvis $h = 1$ og $q_0 = q_1 = 1$; da $\det T_1 = -1$, følger det, at T er af den anden form i (ii). I alle andre tilfælde er T af den tredie form.

Antag omvendt, at (ii) er opfyldt, og lad (Q', Q) være den anden række i T. Da T har determinant ± 1, er Q' og Q primiske. Hvis $(Q', Q) = (0, 1)$, er $T = T_0$, med $y_0 := P$. Hvis $(Q', Q) = (1, 1)$, er $T = T_1$, med $y_0 := P - 1$ og $y_1 := 1$. Antag nu, at $0 < Q' < Q$, altså at T er af den tredie form i (ii). Da kan Q'/Q fremstilles ved to kædebrøker, en af lige længde og en af ulige længde, og da $0 < Q'/Q < 1$, er kædebrøkerne ægte. Vælg den fremstilling, hvis længde, h, opfylder, at $(-1)^h = \det T$. Idet $x_0 = 0, x_1, \ldots, x_h$ er koefficienterne i denne kædebrøk, og S_h er den tilhørende matrix, har vi

$$S_h := \begin{pmatrix} 0 & 1 \\ 1 & x_1 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & x_h \end{pmatrix}.$$

Den h'te konvergent er lig med Q'/Q. På den anden side er tæller og nævner i den h'te konvergent netop indholdet af anden søjle i matricen S_h. Da Q' og Q er primiske, og $Q > 0$, følger det, at anden søjle i S_h netop indeholder Q' og Q. Matricen S_h har altså formen,

$$S_h = \begin{pmatrix} p' & Q' \\ p & Q \end{pmatrix},$$

med hele tal p', p. Ifølge valget af h er $\det T = (-1)^h = \det S_h$. Altså er,

$$p'Q - pQ' = P'Q - PQ'.$$

Da Q og Q' er primiske, følger det af den sidste ligning, at

$$\begin{pmatrix} P' \\ P \end{pmatrix} = \begin{pmatrix} p' \\ p \end{pmatrix} + x \begin{pmatrix} Q' \\ Q \end{pmatrix},$$

hvor $x \in \mathbb{Z}$. Altså gælder matrixligningen,

$$T^{\mathrm{tr}} = \begin{pmatrix} P' & Q' \\ P & Q \end{pmatrix} = \begin{pmatrix} p' & Q' \\ p & Q \end{pmatrix} \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix} = S_h \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix},$$

eller på transponeret form,

$$T = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} S_h^{\mathrm{tr}}. \tag{4.9.1}$$

Faktorerne i S_h er symmetriske, men rækkefølgen ombyttes ved transponeringen. Af (4.9.1) følger derfor, at T er en kædebrøksmatrix, med $y_0 := x$, $y_1 := x_h$, $\ldots$ $y_h := x_1$.

Hermed er Lemmaet bevist. ∎

(4.10) Sætning. *Betragt to reelle irrationale tal $\hat{\tau}$ og τ og deres følger af rester. Da er τ og $\hat{\tau}$ svagt similære, hvis og kun hvis der findes indices m og h således, at $\tau_m = \hat{\tau}_h$, og de er similære, hvis og kun hvis indices m og h her kan vælges af samme paritet.*

Bevis. Antag først, at τ og $\hat{\tau}$ opfylder betingelsen, altså at $\tau_m = \hat{\tau}_h$. Med sædvanlig notation for kædebrøker gælder relationerne,

$$\begin{pmatrix} \tau \\ 1 \end{pmatrix} \sim S_m \begin{pmatrix} \tau_m \\ 1 \end{pmatrix}, \qquad \begin{pmatrix} \hat{\tau} \\ 1 \end{pmatrix} \sim \hat{S}_h \begin{pmatrix} \hat{\tau}_h \\ 1 \end{pmatrix}. \tag{4.10.1}$$

De to søjler, der står på højresiderne i relationerne i (4.10.1), er ifølge antagelsen den samme. Relationerne medfører derfor, at

$$\begin{pmatrix} \hat{\tau} \\ 1 \end{pmatrix} \sim \hat{S}_h S_m^{-1} \begin{pmatrix} \tau \\ 1 \end{pmatrix}. \tag{4.10.2}$$

Matricen $\hat{S}_h S_m^{-1}$ har determinant $(-1)^{h-m}$, så relationen (4.10.2) viser, at tallene $\hat{\tau}$ og τ er svagt similære. Hvis h og m har samme paritet, viser relationen endda, at $\hat{\tau}$ og τ er similære.

Antag omvendt, at $\hat{\tau}$ og τ er svagt similære. Da findes en relation, med $T \in \mathrm{GL}_2(\mathbb{Z})$,

$$\begin{pmatrix} \hat{\tau} \\ 1 \end{pmatrix} \sim T \begin{pmatrix} \tau \\ 1 \end{pmatrix}. \tag{4.10.3}$$

For hvert m gælder relationen $\left(\begin{smallmatrix} \tau \\ 1 \end{smallmatrix}\right) \sim S_m \left(\begin{smallmatrix} \tau_m \\ 1 \end{smallmatrix}\right)$, og indsættelse i (4.10.3) giver relationen,

$$\begin{pmatrix} \hat{\tau} \\ 1 \end{pmatrix} \sim T S_m \begin{pmatrix} \tau_m \\ 1 \end{pmatrix}. \tag{4.10.4}$$

Betragt nu koefficienterne i T og i TS_m for $m = 0, 1, \ldots,$

$$T = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \qquad TS_m = \begin{pmatrix} P'_m & P_m \\ Q'_m & Q_m \end{pmatrix}.$$

Det er nok at vise, at der findes et m således, at $Q_m \neq 0$ og $0 < Q'_m/Q_m < 1$. Antag nemlig, at et sådant m er fundet. Relationen i (4.10.4) ændres ikke, når matricen TS_m erstattes med $-TS_m$. Vi kan derfor antage, at $Q_m > 1$. Herefter følger det af Lemma (4.9), at TS_m er en kædebrøksmatrix. Relationen i (4.10.4) har altså formen,

$$\begin{pmatrix} \hat{\tau} \\ 1 \end{pmatrix} \sim \begin{pmatrix} 1 & x_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & x_1 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & x_h \end{pmatrix} \begin{pmatrix} \tau_m \\ 1 \end{pmatrix}. \tag{4.10.5}$$

Nu var $0 < \tau_m < 1$. Det følger derfor af Lemma (2.11), at $\hat{\tau}_h = \tau_m$. Desuden er $\hat{S}_h = \pm TS_m$, og specielt er $(-1)^h = \det(TS_m) = \det T(-1)^m$. Hvis matricen T i (4.10.3) har determinant 1, får m og h altså samme paritet.

For at bestemme et tal m med den ønskede egenskab, bemærkes først, at tallet Q_m er bestemt ved $Q_m = Cp_m + Dq_m = q_m(Cp_m/q_m + D)$ (og $Q'_m = Q_{m-1}$). Tallet $Cp_m/q_m + D$ konvergerer mod $C\tau + D$, som er forskellig fra 0, da τ er irrational. Specielt er altså $Q_m \neq 0$, når m er tilstrækkelig stor. For forholdet Q'_m/Q_m finder man

$$\frac{Q'_m}{Q_m} = \frac{Cp_{m-1} + Dq_{m-1}}{Cp_m + Dq_m} = \frac{q_{m-1}}{q_m} \cdot \frac{Cp_{m-1}/q_{m-1} + D}{Cp_m/q_m + D}.$$

Den anden faktor på højresiden konvergerer mod 1 for $m \to \infty$, idet både tæller og nævner konvergerer mod $C\tau + D$. Den første faktor q_{m-1}/q_m er altid mindre end 1. For at indse, at produktet er mindre end 1 for en passende værdi af m, er det nok at vise, at der findes et tal $\theta < 1$ således, at uligheden $q_{m-1}/q_m < \theta$ gælder for uendelig mange værdier af m. Lad hertil $\theta < 1$ være „det gyldne snit“, bestemt ved $\theta = 1/(1+\theta)$. For alle m gælder vurderingen,

$$\frac{q_m}{q_{m+1}} = \frac{q_m}{x_{m+1}q_m + q_{m-1}} \leqslant \frac{q_m}{q_m + q_{m-1}} = \frac{1}{1 + q_{m-1}/q_m}.$$

Heraf ses, at hvis $q_{m-1}/q_m > \theta$, så er $q_m/q_{m+1} < 1/(1+\theta) = \theta$. Tallet θ har altså specielt den efterlyste egenskab. ∎

(4.11) Bemærkning. At afgøre om to reelle irrationale tal er svagt similære, er ifølge sætningen ækvivalent med at afgøre, om der findes et tal, der forekommer i begge følger af rester. Det sidste problem er i almindelighed ikke endeligt, da følgerne af rester er uendelige følger. Hvis der imidlertid er fundet rester således, at $\tau_m = \hat{\tau}_h$, så giver kædebrøksudviklingen den søgte matrix S med $\hat{\tau} = S(\tau)$. Af

$$\begin{pmatrix} \tau \\ 1 \end{pmatrix} \sim S_m \begin{pmatrix} \tau_m \\ 1 \end{pmatrix} \text{ og } \begin{pmatrix} \hat{\tau} \\ 1 \end{pmatrix} \sim \hat{S}_h \begin{pmatrix} \hat{\tau}_h \\ 1 \end{pmatrix}$$

og $\hat{\tau}_h = \tau_m$, følger nemlig, at

$$\begin{pmatrix} \hat{\tau} \\ 1 \end{pmatrix} \sim \hat{S}_h S_m^{-1} \begin{pmatrix} \tau \\ 1 \end{pmatrix},$$

og så er $\hat{\tau} = \hat{S}_h S_m^{-1}(\tau)$. Matricen $\hat{S}_h S_m^{-1}$ er altså den søgte matrix, og den har determinant 1, hvis h og m har samme paritet.

Som vist i Kapitel 3 er mængden af rester af τ endelig, hvis og kun hvis τ er et kvadratisk tal. For kvadratiske tal er problemet om similaritet altså endeligt. Som vi nu skal se, er problemet altid endeligt, hvis de givne tal er imaginære.

(4.12) Opskrift. Lad τ være et komplekst tal. I det følgende betragtes en variant af opskrift (2.9), som der fås ved at erstatte $\tau \mapsto 1/\tau$ med $\tau \mapsto -1/\tau$ og intervallet $0 \leqslant \tau < 1$ med den *strimmel* i den komplekse plan, hvor $-\frac{1}{2} \leqslant \Re e\, \tau < \frac{1}{2}$. For hvert komplekst tal τ findes et entydigt bestemt helt tal x således, at $\tau - x$ ligger i denne strimmel. Definer altså rekursivt følgen $\tau_0, \tau_1, \ldots$ af tal i strimmelen og en følge $x_0, x_1, x_2, \ldots$ af hele tal således: Skriv

$$\tau = x_0 + \tau_0, \qquad \text{hvor } x_0 \in \mathbb{Z} \text{ og } -\tfrac{1}{2} \leqslant \Re e\, \tau_0 < \tfrac{1}{2}.$$

Stop, hvis $\tau_0 = 0$. I modsat fald skrives

$$\frac{-1}{\tau_0} = x_1 + \tau_1, \qquad \text{hvor } x_1 \in \mathbb{Z} \text{ og } -\tfrac{1}{2} \leqslant \Re e\, \tau_1 < \tfrac{1}{2}.$$

Induktivt, hvis $\tau_{n-1} \neq 0$, skrives

$$\frac{-1}{\tau_{n-1}} = x_n + \tau_n, \qquad \text{hvor } x_n \in \mathbb{Z} \text{ og } -\tfrac{1}{2} \leqslant \Re e\, \tau_n < \tfrac{1}{2}.$$

Tallet τ_n er den n'te *rest* af tallet τ.

Processen stopper i det k'te skridt, hvis $\tau_k = 0$. Er dette tilfældet, får vi to endelige følger, $\tau_0, \tau_1, \ldots, \tau_k = 0$ og $x_0, x_1, \ldots, x_k$.

Som i opskrift (2.9) følger det af ligningerne ovenfor, at

$$\tau = x_0 + \cfrac{-1}{x_1 + \cfrac{-1}{x_2 + \cfrac{-1}{\ddots + \cfrac{-1}{x_n + \tau_n}}}}. \tag{4.12.1}$$

Opskriften fører altså til en *irregulær* kædebrøk, hvor tællerne $t_1, t_2, \ldots$ er lig med -1. Med notationen fra (2.3) er

$$S_n = \begin{pmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{pmatrix} = \begin{pmatrix} 1 & x_0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & x_1 \end{pmatrix} \cdots \begin{pmatrix} 0 & -1 \\ 1 & x_n \end{pmatrix}, \tag{4.12.2}$$

(for $n \geqslant 0$; vi betragter ikke S_{-1}), og som i (2.8) følger det af (4.12.1), at

$$\lambda \begin{pmatrix} \tau \\ 1 \end{pmatrix} = S_n \begin{pmatrix} \tau_n \\ 1 \end{pmatrix}, \text{ med } \lambda = q_n + q_{n-1}\tau_n. \tag{4.12.3}$$

Bemærk, at alle matricerne S_n her har determinant $+1$. (Resterne defineret her må naturligvis ikke forveksles med resterne ved den regulære kædebrøksudvikling, betragtet i Kapitel 2. I resten af dette kapitel betragtes udelukkende den irregulære kædebrøksudvikling, og τ_n er resten defineret ovenfor.)

Øjensynlig vil der for et givet tal τ indtræffe et af følgende tilfælde:

(**a**) Processen stopper ved at $\tau_k = 0$ for et $k \geqslant 0$.

(**b**) For et passende $n \geqslant 1$ er $|x_n| \leqslant 1$ (dvs x_n er lig med -1, 0 eller 1).

(**c**) Processen stopper ikke og for alle $n \geqslant 1$ er $|x_n| \geqslant 2$.

A priori er de tre tilfælde ikke disjunkte, idet det i (**b**) ikke er antaget, at processen ikke stopper. At de faktisk er disjunkte, vil fremgå af det følgende.

Tilfælde (**a**): I dette tilfælde må tallet τ naturligvis være et rationalt tal.

Tilfælde (**b**): I dette tilfælde må tallet τ_{n-1}, og dermed også tallet τ, være imaginært (dvs ikke reelt). Hvis nemlig τ_{n-1} er reelt, og dermed i intervallet $[-\frac{1}{2}, \frac{1}{2}[$, så er $|-1/\tau_{n-1}| \geqslant 2$; tallet x_n, der er det nærmeste hele tal, vil derfor opfylde, at $|x_n| \geqslant 2$. Nedenfor kigger vi nærmere på processen i det imaginære tilfælde.

Tilfælde (**c**): For tallene q_n giver ligning (4.12.2) rekursionsformlen $q_n = x_n q_{n-1} - q_{n-2}$, hvor udgangsværdierne er $q_{-1} = 0$ og $q_0 = 1$. Heraf følger let, da $|x_n| \geqslant 2$, at

$$1 = |q_0| < |q_1| < |q_2| < \cdots .$$

Af Ligning (4.12.3) fås ligningen,

$$\frac{p_{n-1}}{q_{n-1}} - \tau = \frac{1}{q_{n-1}\lambda}, \quad \text{hvor } \lambda = q_n + q_{n-1}\tau_n. \tag{4.12.4}$$

Da τ_n tilhører strimmelen og $|q_n| > |q_{n-1}|$, følger det af ligningen for λ, at $|\Re e\, \lambda| > \frac{1}{2}|q_n|$. Altså er også $|\lambda| > \frac{1}{2}|q_n|$. Af Ligning (4.12.4) fås derfor vurderingen,

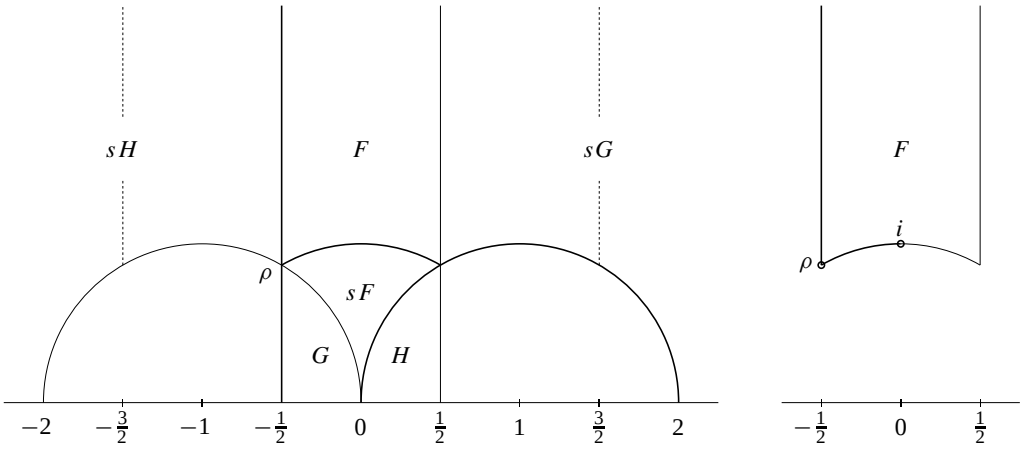
$$\left| \frac{p_{n-1}}{q_{n-1}} - \tau \right| < \frac{2}{|q_n q_{n-1}|} < \frac{2}{q_{n-1}^2} .$$

Da $|q_n|$ går mod uendelig for $n \to \infty$, viser den sidste vurdering, at følgen af konvergenter konvergerer mod τ. Specielt er τ altså et reelt tal i dette tilfælde. Yderligere følger det let af vurderingen, at τ endda må være irrational.

Konklusionerne i de tre tilfælde er åbenbart disjunkte. Vi har således vist følgende:

Tallet τ er rationalt, hvis og kun hvis processen stopper; i dette tilfælde er $|x_n| \geqslant 2$ for alle $n \geqslant 1$. Tallet τ er reelt og irrationalt, hvis og kun hvis processen ikke stopper og $|x_n| \geqslant 2$ for alle $n \geqslant 1$. Tallet τ er imaginært, hvis og kun hvis der findes et $n \geqslant 1$ således, at $|x_n| \leqslant 1$.

(4.13) Opskriften i det imaginære tilfælde. Lad os se nærmere på processen i tilfældet, hvor tallet τ er imaginært. Afbildningen $s(\tau) = -1/\tau$ er en såkaldt Möbius-transformation. Den afbilder den øvre halvplan ind i sig selv; man kan vise, at den er cirkeltro og vinkeltro. Linier medregnes her som „cirkler gennem punktet ∞“, og cirkeltro betyder ikke, at centrum afbildes på centrum. Øjensynlig transformationen s en involution. Vi ser på transformationen i den øvre halvplan. Betragt hertil områderne markeret på den første figur herunder.



Det fælles punkt i områderne F, sF og G er den tredie enhedsrod $\rho := e^{2\pi i/3}$. Da transformationen er cirkeltro, følger det, at den afbilder områderne F, G og H som markeret. De to svagt markerede dele af randen af $F \cup sF$ ombyttes af s; disse to dele af randen medregnes ikke til $F \cup sF$. Lad nu τ være et tal i strimmelen. Det ligger altså i et af områderne F, sF, G eller H (og ikke på linien, hvor realdelen er $\frac{1}{2}$), og $s(\tau) = -1/\tau$ ligger altså i et af områderne sF, F, sG eller sH. Betragt fremstillingen,

$$-1/\tau = x + \tau_0, \qquad \text{hvor } x \in \mathbb{Z} \text{ og } -\tfrac{1}{2} \leqslant \Re e\, \tau_0 < \tfrac{1}{2}.$$

Det er nemt at aflæse følgende: Hvis $x = 1$, så er $\tau_0 \in F$ (og $\tau \in G$). Hvis $x = -1$, så er $\tau_0 \in F$ (og $\tau \in H$). Endelig er $x = 0$, hvis og kun hvis $\tau \in F \cup s(F)$ og $\tau \neq \rho$. For $\tau = \rho$, er $s(\rho) = \rho + 1$, og altså $x = 1$ og $\tau_0 = \rho$.

Betragt nu et vilkårligt tal τ i den øvre halvplan. Lad $n \geqslant 1$ være det første index, for hvilket $|x_n| \leqslant 1$. Den foregående overvejelse, anvendt på τ_{n-1}, viser, at $\tau_n \in F \cup sF$ og, bortset fra undtagelsestilfældet $\tau_n = \rho$, at $x_{n+1} = x_{n+2} = \cdots = 0$; i undtagelsestilfældet er $x_{n+1} = x_{n+2} = \cdots = 1$. Videre er $\tau_n = \tau_{n+2} = \cdots$ og $\tau_{n+1} = \tau_{n+3} = \cdots$; hvis $\tau_n = i$ eller $\tau_n = \rho$ gælder endda $\tau_n = \tau_{n+1} = \cdots$. Følgen af rester er altså særdeles simpel: fra et vist trin er den periodisk med periode to, med rester, der skiftevis ligger i F og sF.

Det er sædvane mere præsist at definere F som det område, der fremkommer af figuren ovenfor ved at fjerne en del af randen, som markeret på den anden figur. Komplekse tal i området F kaldes *reducerede*. Vi har således vist, at der blandt resterne τ_n af et tal τ i den øvre halvplan forekommer præcis én i området F, altså præcis en *reduceret rest*. Hvis τ_m er reduceret, så er $\tau_m = \tau_{m+2} = \ldots$; hvis $\tau_m = i$ eller $\tau_m = \rho$, så er endda $\tau_m = \tau_{m+1} = \cdots$.

(Yderligere er $x_{m+1} = x_{m+2} = \cdots = 0$, bortset fra undtagelsestilfældet $\tau_m = \rho$; her er $x_{m+1} = x_{m+2} = \cdots = 1$.)

(4.15) Sætning. *Lad $\hat{\tau}$ og τ være tal i den øvre halvplan. Da er $\hat{\tau}$ og τ similære, hvis og kun hvis de har den samme reducerede rest.*

Bevis. „hvis" følger ganske som i det reelle tilfælde.

„kun hvis": Antag, at tallene $\hat{\tau}$ og τ er similære. Et tal τ er similært med enhver af sine rester. Følgelig er de reducerede rester af τ og $\hat{\tau}$ similære. Vi kan derfor antage, at både $\hat{\tau}$ og τ er reducerede; det skal så vises, at $\hat{\tau} = \tau$. Ifølge antagelsen findes en ligning,

$$\hat{\tau} = \frac{A\tau + B}{C\tau + D}, \quad \text{hvor } T = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

Af symmetrigrunde kan vi antage, at $\Im m\, \hat{\tau} \geqslant \Im m\, \tau$. Af (4.3.1) får vi derfor den første af følgende vurderinger (de efterfølgende vurderinger udnytter blot, at C og D er reelle tal og at τ ligger i område F):

$$1 \geqslant |C\tau + D|^2 = C^2|\tau|^2 + 2CD\,\Re e\,\tau + D^2 \tag{4.15.1}$$

$$\geqslant C^2 + 2CD\,\Re e\,\tau + D^2 \tag{4.15.2}$$

$$\geqslant C^2 - |CD| + D^2 \tag{4.15.3}$$

$$\geqslant (|C| - |D|)^2. \tag{4.15.4}$$

Da C og D er hele tal, er det sidste udtryk et helt tal, større end eller lig med 0. Vurderingen viser, at udtrykket må være 0 eller 1.

Antag først, at $|C| \neq |D|$. Det følger da, at $|C| - |D| = \pm 1$. Videre følger det, at lighed gælder i alle ulighederne ovenfor. Specielt følger det af lighed i (4.15.4), at $CD = 0$. Altså er enten $C = 0$ eller $D = 0$. I det første tilfælde er $C = 0$, og da $T \in \mathrm{SL}_2(\mathbb{Z})$, følger det, at $A = D = \pm 1$; det kan uden indskrænkning antages, at $A = D = 1$. Altså er $\hat{\tau} = \tau + B$. Da både $\hat{\tau}$ og τ tilhører strimmelen, følger det, at $B = 0$, altså at $\hat{\tau} = \tau$. I det andet tilfælde er $D = 0$, og heraf følger $-B = C = \pm 1$; det kan uden indskrænkning antages, at $-B = C = 1$. Altså er $\hat{\tau} = (A\tau - 1)/\tau = -1/\tau + A$. Da både τ og $\hat{\tau} = -1/\tau + A$ ligger i F, følger det, at enten er $A = 0$ og $\tau = i$ (og så er $\hat{\tau} = -1/\tau = \tau$) eller også er $A = -1$ og $\tau = \rho$ (og så er $\hat{\tau} = -1/\tau - 1 = \tau$). Også i det andet tilfælde er altså $\hat{\tau} = \tau$.

Antag dernæst, at $|C| = |D|$, og dermed, at $C = \pm D = \pm 1$. Det følger, at skarp ulighed gælder i (4.15.4), og dermed at lighed gælder i de øvrige uligheder. Det kan antages, at $C = 1$. Af lighed i (4.15.2) følger, at $|\tau| = 1$, og da $|CD| = 1$ følger det af lighed i (4.15.3), at $D = 1$ og $\Re e\,\tau = -\frac{1}{2}$. Af $C = D = 1$ følger, at $B = A - 1$. Altså er

$$\hat{\tau} = \frac{A\tau + A - 1}{\tau + 1} = \frac{-1}{\tau + 1} + A. \tag{4.15.5}$$

Det er vist, at $|\tau| = 1$ og $\Re e\,\tau = -\frac{1}{2}$. Heraf følger, at τ er den tredie enhedsrod ρ. Brøken $-1/(\tau + 1)$ er derfor lig med τ, så af ligning (4.15.5) følger, at $\hat{\tau} = \tau + A$. Da $\hat{\tau}$ tilhører F, må der her gælde $A = 0$, og altså $\hat{\tau} = \tau$.

Hermed er i alle tilfælde vist, at $\hat{\tau} = \tau$, som ønsket. ∎

I. Index.