Summen von Quadraten
Skript zur ETH Vorlesung HS21

Raphael S. Steiner

Department of Mathematics, ETH Zürich, 8092 Zürich, CH
raphael.steiner@math.ethz.ch

13. Dezember 2021

Inhaltsverzeichnis

1 Motivation und Übersicht

Summen von Quadraten können sich an vielen Orten verstecken. Um dies zu illustrieren, beginnen wir mit einer informalen Besprechung eines Problems. Das Problem ist wiefolgt.

Problem. *Ob und wie kann man ein Quadrat in fünf gleich grosse Teilquadrate zerschneiden mit geraden Schnitten?*

Auf den ersten Blick bemerkt man sofort, dass man das Quadrat in $1, 4, 9, 16, \ldots, n^2, \ldots$ kleinere gleich grosse Quadrate zerschneiden kann und diese scheinen wirklich die einzigen Möglichkeiten zu sein. Dies ist tatsächlich der Fall, denn hätte man N gleich grosse Teilquadrate wäre die Seitenlänge $1/\sqrt{N} \in \mathbb{Q} \Leftrightarrow N = n^2$. Was passiert nun, wenn wir die Bedingung 'gleich gross' weglassen. In diesem Falle kann man beliebiges Teilquadrat nehmen und es in 4 Quadrate zerschneiden. Dies zeigt, dass wenn man ein Quadrat in N kleinere Quadrate zerschneiden kann, so kann man es auch in $N+3$ Quadrate zerschneiden. Nach weiterem überlegen findet man einen Weg ein Quadrat in 6,8,10,... , Teilquadrate zu zerlegen, indem man ein grosses Quadrat mit kleineren Quadraten von zwei Seiten umhüllt. Hier ist eine solche Anordnung für acht Teilquadrate zu sehen (Abbildung 1).

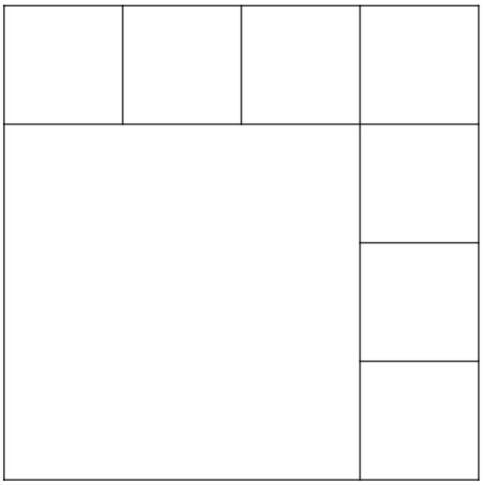

Abbildung 1: Acht Teilquadrate

Insgesamt findet man Wege ein Quadrat in N Teilquadrate zu zerlegen ausser für $N = 2, 3, 5$. Um fünf gleich grosse Quadrate zu erhalten müssen wir also etwas kreativer werden. Versuchen wir uns zuerst mal an zwei gleich grosse Teilquadrate. Durch herum probieren oder motiviert durch die Seitenlänge $1/\sqrt{2}$ und Pythagoras kommt man auf die Diagonalen (Abbildung 2). Hier kann man das linke und rechte Dreieck zu einem Quadrat zusammen kleben und ebenso das obere und untere Dreieck.

Das heisst, erlauben wir (Verschiebungen und) Kleben, so ist das zerschneiden in zwei gleich grosse Quadrate möglich. Wie sieht es aber nun mit fünf aus? Dies ist auch möglich. Dazu schneidet man von einer Ecke zu einer zur einem gegenüberliegenden Mittelpunkt und wiederholt dies für jede Ecke, wie in Abbildung 3, so erhält man fünf gleich grosse Quadrate.

Dass dies, wirklich der Fall ist könnte man nachrechnen, indem man Seitenlängen und Winkel berechnet. Wir möchten aber eine Intuitive Begründung. Der Name der Vorlesung gibt dabei einen kleinen Hinweis. Zwei und Fünf sind Summen von zwei Quadraten: $2 = 1^2 + 1^2$ und $5 = 2^2 + 1^2$. Betrachten wir unsere Schnitte etwas genauer, so sehen wir,

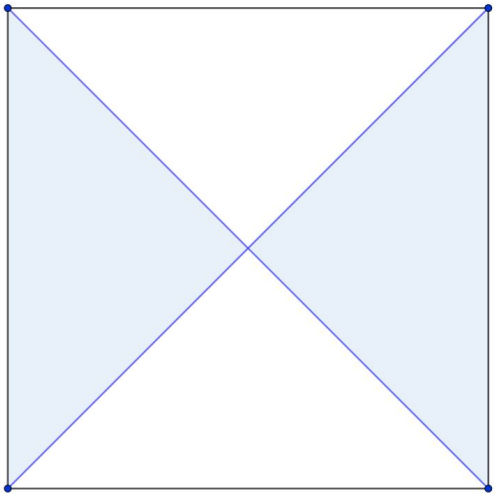

Abbildung 2: Zwei gleich grosse Teilquadrate

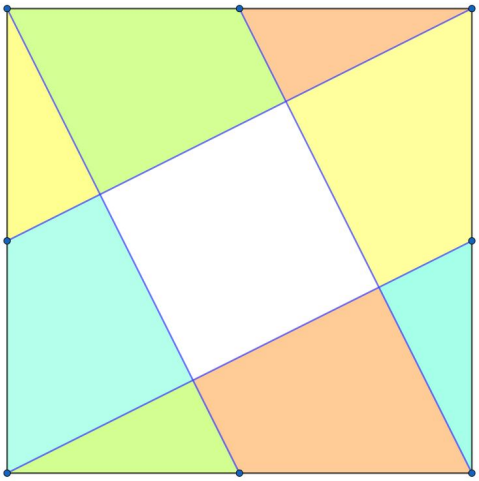

Abbildung 3: Fünf gleich grosse Teilquadrate

dass sie die Hypothenuse eines rechtwinkligen Dreiecks sind mit Kathetenverhältnis $1:1$, respektive $2:1$. Dies scheint kein Zufall zu sein. Tatsächlich ist dies kein Zufall, wie wir jetzt erklären werden.

Betrachte die ganzen Gitterpunkte $\mathbb{Z}^2 \subset \mathbb{R}^2$ in der Ebene. Ferner markiere man alle Punkte rot, welche man vom Ursprung $(0,0)$ durch das Addieren von beliebig vielen Vektoren der Form $(1,2)$, $(2,-1)$, $(-1,-2)$, oder $(-2,1)$ erreichen kann. Zeichnet man nun das quadratische Mesh, welches durch die ganzen Gitterpunkte $\mathbb{Z}^2$ gegeben ist, und ebenso für die rot markierten Punkte[a] so erkennen wir die Abbildung 3 wieder als eines der roten Quadrate (siehe Abbildung 4).

Färbt man nun die kleineren Quadrate so ein wie bei Abbildung 3, so sieht man, dass der Grund warum man fünf Teilquadrate erhält ist, dass es fünf verschiedene Verschiebungen des roten Gitters gibt, dessen gesamthafte Gitterpunkte gerade das Gitter $\mathbb{Z}^2$ ergibt. Letzteres können wir mit etwas Zahlentheorie erklären. Dazu identifizieren wir die

[a] Man überzeuge sich selbst, dass dies wirklich ein quadratisches Mesh darstellt.

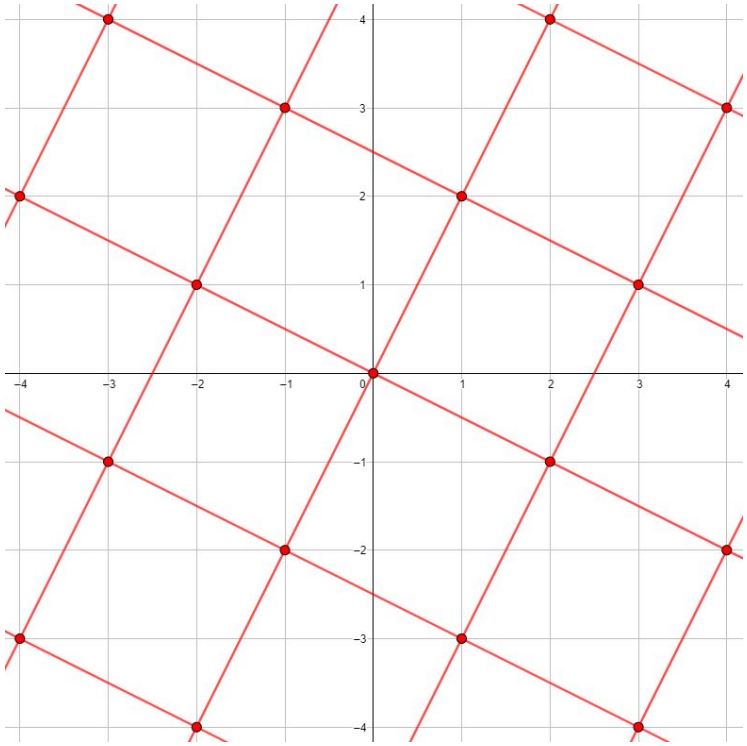

Abbildung 4: Gitter und Mesh in der Ebene

Ebene mit den komplexen Zahlen $\mathbb{C}$ und die ganzen Gitterpunkte mit den Gauss'schen Zahlen $\mathbb{Z}[i] = \mathbb{Z} + \mathbb{Z}i$. Es stellt sich nun heraus, dass die rot markierten Punkte gerade diejenigen Gauss'schen Zahlen sind, welche durch $2 - i$ teilbar sind. So ist zum Beispiel $1 + 2i = i(2 - i)$. In den Gauss'schen Zahlen gilt nun, dass die Division durch $2 - i$ mit Rest genau $|2 - i|^2_{\mathbb{C}} = 2^2 + 1^2 = 5$ verschiede Restklassen besitzt. Jene Restklassen stellen nun genau die verschiedenen Verschiebungen der roten Gitterpunkte dar.

Man kann nun mit dieser Konstruktion etwas Spass haben. Hier ist zum Beispiel, wie man ein Quadrat in $13 = 3^2 + 2^2$ gleich grosse Teilquadrate zerschneiden kann (Abbildung 5).

Wir haben nun motiviert, dass Summen von zwei Quadraten im Zusammenhang mit den Gauss'schen Zahlen $\mathbb{Z}[i]$ stehen. Jene sind ein Analog von den ganzen Zahlen in den komplexen Zahlen $\mathbb{C}$. Nun besitzt $\mathbb{C}$ eine Erweiterung, die Quaternionen $\mathbb{H}$. Sowie die Norm auf $\mathbb{C}$ im Zusammenhang mit Summen von zwei Quadraten steht, so steht die Norm der Quaternionen im Zusammenhang mit Summen von vier Quadraten. Wir werden in der Vorlesung sehen, wie man die Quaternionen konstruiert und dass man auch ein Analog der ganzen Zahlen finden kann. Wir werden jene benutzen, um den vier Quadrate Satz von Lagrange zu beweisen.

Satz (Lagrange)**.** *Jede natürliche Zahl lässt sich als Summe von vier Quadraten ganzer Zahlen schreiben.*

Man stellt sich nun die Frage, wie sieht es mit drei Quadraten aus? Die Theorie für Summen von drei Quadraten ist um einiges komplizierter. Der Grund dafür ist, dass es für Summen von drei Quadraten keine zugrundeliegende multiplikative Struktur gibt wie es im Fall von zwei Quadraten (die komplexen Zahlen) oder vier Quadraten (die Quaternionen) ist. Im Verlaufe der Vorlesung werden wir dies genauer präzisieren und zeigen, dass solche

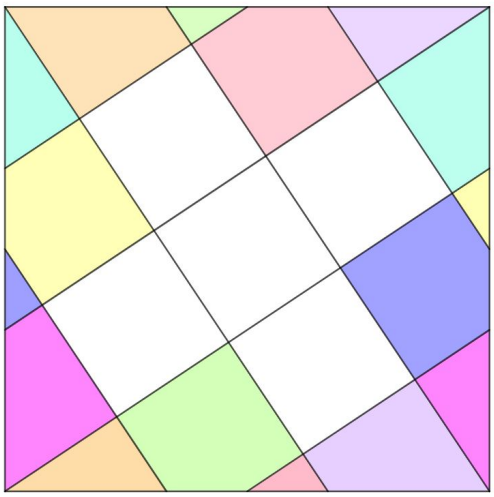

Abbildung 5: 13 gleich grosse Teilquadrate

multiplikativen Strukturen für Summen von N Quadraten gibt genau dann wenn $N = 1, 2, 4, 8$.

Zum Schluss werden wir Quadrate etwas verallgemeinern, nämlich zu positiv definiten binären quadratische Formen, und zeigen, dass man jene auch multiplizieren kann[b]. So hat man zum Beispiel

$$(2r^2 + 2rs + 3s^2)(2u^2 + 2uv + 3v^2) = x^2 + 5y^2,$$

wobei

$$x = 2ru + rv + su + 3sv, \quad y = rv - su.$$

2 Die ganzen Zahlen und ihre Quotienten

Die Theorie in diesem Kapitel lässt sich in den meisten Bücher über algebraische Zahlentheorie oder Algebra auffinden, so zum Beispiel [9], [12] oder [1].

2.1 Teiler und Faktorisierungen

Wir beginnen die Lektion mit einer Repetition der gängigen Eigenschaften der ganzen Zahlen. Wir werden jene Eigenschaften, wie z.B. die eindeutige Primfaktorzerlegung, in einer Art und Weise herleiten, sodass eine Verallgemeinerung auf andere Zahlsystem wie die Gauss'schen Zahlen sofort per Analogie folgt. Dabei erwähnen und verwenden wir die korrekten algebraischen Begriffe. Jedoch sind letztere für uns nicht von grösster Bedeutung, da wir immer mit konkreten algebraischen Strukturen arbeiten.

Zur Auffrischung errinnern wir uns an die Rechenregeln der ganzen Zahlen:

- Die Addition ist kommutativ: $\forall a, b \in \mathbb{Z} : a + b = b + a$.
- Die Addition ist assoziativ: $\forall a, b, c \in \mathbb{Z} : a + (b + c) = (a + b) + c$.
- Die Addition besitzt ein neutrales Element (Null): $\exists 0 \in \mathbb{Z} : \forall a \in \mathbb{Z} : a+0 = 0+a = a$.

[b] Sofern beide Formen dieselbe Diskriminante besitzen

- Es existieren inverse Elemente bezüglich der Addition: $\forall a \in \mathbb{Z} : a+(-a) = (-a)+a = 0$.
- Die Multiplikation ist kommutativ: $\forall a, b \in \mathbb{Z} : a \cdot b = b \cdot a$.
- Die Multiplikation ist assoziativ: $\forall a, b, c \in \mathbb{Z} : a \cdot (b \cdot c) = (a \cdot b) \cdot c$.
- Die Multiplikation besitzt ein neutrales Element (Eins): $\exists 1 \in \mathbb{Z} : \forall a \in \mathbb{Z} : a \cdot 1 = 1 \cdot a = a$.
- Die Multiplikation ist links distributiv über die Addition: $\forall a, b, c \in \mathbb{Z} : a \cdot (b + c) = (a \cdot b) + (a \cdot c)$.
- Die Multiplikation ist rechts distributiv über die Addition: $\forall a, b, c \in \mathbb{Z} : (b + c) \cdot a = (b \cdot a) + (c \cdot a)$.
- Keine von Null verschiedenen Nullteiler: $\forall a, b \in \mathbb{Z} : a \cdot b = 0 \Rightarrow a = 0$ oder $b = 0$.

Bemerkung. Eine solche algebraische Struktur nennt sich *Integritätsbereich*, falls $0 \neq 1$. Nebst $\mathbb{Z}$ sind auch $\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{F}_q, \mathbb{Z}[X], \ldots$ Beispiele eines Intergritätsbereiches.

Definition. *Wir sagen eine Zahl $a \in \mathbb{Z}$* teilt *eine Zahl $b \in \mathbb{Z}$, schreibe $a \mid b$, genau dann wenn eine Zahl $c \in \mathbb{Z}$ existiert, sodass $ac = b$ gilt. Sofern jene Zahl c eindeutig ist, schreiben wir $\frac{b}{a} = c$.*

Bemerkung. Falls $0 \neq a \in \mathbb{Z}$ ein Teiler von $b \in \mathbb{Z}$, so ist c mit $ac = b$ eindeutig bestimmt. Denn für die ganzen Zahlen gilt folgende Eigenschaft $xy = 0$ mit $x, y \in \mathbb{Z}$, so folgt $x = 0$ oder $y = 0$. Wäre also $c' \in \mathbb{Z}$ ein anderes Element mit $ac' = b$ so folgt $a(c - c') = b - b = 0$ also $c - c' = 0 \Leftrightarrow c = c'$, da $a \neq 0$.

Die folgenden Teilereigenschaften sind einfach nachzuweisen.

Lemma 2.1. *Für ganzen Zahlen $a, b, c \in \mathbb{Z}$ gelten folgende Aussagen:*

1. *Falls $a \mid b$ und $b \mid c$, so gilt $a \mid c$,*
2. *Falls $a \mid b$ und $a \mid c$, so gilt $a \mid xb + yc$ für alle $x, y \in \mathbb{Z}$.*

Beweis. Für 1. finden wir $d, e \in \mathbb{Z}$, sodass $b = ad$ und $c = be$. Es folgt $c = be = a(de)$, also a teilt c. Für 2. finden wir $d, e \in \mathbb{Z}$, sodass $b = da$ und $c = ea$. Dann gilt für beliebige $x, y \in \mathbb{Z}$, dass $xb + yc = (xd + ye)a$. Es folgt, dass $a \mid xb + yc$. □

Definition. *Eine ganze Zahl $a \in \mathbb{Z}$ heisst* Einheit *genau dann, wenn es $b \in \mathbb{Z}$ gibt, sodass $ab = 1$. Die Menge der Einheiten von $\mathbb{Z}$ wird mit $\mathbb{Z}^\times$ bezeichnet.*

Bemerkung. Ist a eine Einheit, so ist auch $\frac{1}{a}$ eine Einheit. Anstatten von $\frac{1}{a}$ schreiben wir gängigerweise auch a^{-1}.

Für die ganzen Zahlen $\mathbb{Z}$ bestehen die Einheiten genau aus den Zahlen ± 1.

Definition. *Eine Zahl $c \in \mathbb{Z}$ heisst* grösster gemeinsamer Teiler *von $a, b \in \mathbb{Z}$, schreibe $c = \mathrm{ggT}(a, b)$, falls*

1. *$c \mid a$ und $c \mid b$,*
2. *für alle $d \in \mathbb{Z}$, sodass $d \mid a$ und $d \mid b$, gilt $d \mid c$.*

Der grösste gemeinsame Teiler zweier Zahlen (falls existent), ist bis auf Multiplikation einer Einheit eindeutig bestimmt, (d.h. bis auf Vorzeichen).

Letzteres ist der Fall, denn falls c, d grösste gemeinsame Teiler von $a, b \in \mathbb{Z}$ sind, so gilt nach der zweiten Eigenschaft, dass $c \mid d$ und $d \mid c$. Das heisst es existieren $e, f \in \mathbb{Z}$, sodass $d = ec$ und $c = fd$. Es folgt $d = ec = efd \Rightarrow d(ef - 1) = 0$. Wir unterscheiden zwei Fälle. Falls $d = 0$, so ist auch $c = 0 = d = d \cdot 1$ und wir sind fertig. Falls $d \neq 0$, so gilt $ef = 1$. Also sind e, f Einheiten und es gilt $c = fd$, $d = ec$.

Definition. *Zwei ganzen Zahlen, welche sich nur durch Multiplikation einer Einheit unterscheiden, heissen zueinander* assoziiert. *Dies ist eine Äquivalenzrelation.*

Lesende mögen ein positives Vorzeichen bei der Definition des grössten gemeinsamen Teilers bevorzugen, jedoch ist im Skript von einer beliebigen Wahl auszugehen, d.h. genau genommen müsste man von einer Äquivalenzklasse von Zahlen reden, welche durch Assoziiertheit definiert sind. Da es aber immer einfach ist mit einer Einheit zu multiplizieren, respektive durch eine Einheit zu dividieren, erlauben wir uns diese Ambiguität.

Die wohl wichtigste Eigenschaft, welche die ganzen Zahlen erfüllen ist die Division mit kleinerem Rest.

- Gegeben eine ganze Zahl $n \in \mathbb{Z}$ und eine von null verschiede ganze Zahl $q \in \mathbb{Z}\backslash\{0\}$, so findet man ganze Zahlen $a, b \in \mathbb{Z}$ mit $|b| < |q|$ für welche $n = aq + b$ gilt.

Bemerkung. Man kann sogar verlangen, dass $b \in \mathbb{N}_0$.

Dies ist die letzte Eigenschaft der ganzen Zahlen, welche wir annehmen. Eine wichtige Applikation der Division mit kleinerem Rest ist die effiziente Berechnung eines grössten gemeinsamen Teilers zweier ganzen Zahlen durch den Euklidischen Algorithmus.

Algorithmus 2.2 (Euklidischer Algorithmus)**.** *Gegeben zwei ganze Zahlen $n_0, n_1 \in \mathbb{Z}$ mit $|n_0| \geqslant |n_1|$, so bestimme man sukzessiv ganze Zahlen $n_2, n_3, \ldots$ mittels Division mit kleinerem Rest, welche $n_{i-1} = a_i n_i + n_{i+1}$ und $|n_{i+1}| < |n_i|$ erfüllen, bis man schlussendlich die Zahl $n_m = 0$ erreicht. Dann gilt*

$$\mathrm{ggT}(n_0, n_1) = n_{m-1}.$$

Also, insbesondere existiert der grösste gemeinsame Teiler zweier ganzer Zahlen.

Beweis. Aus $n_{m-2} = a_{m-1}n_{m-1} + n_m = a_{m-1}n_{m-1}$ folgt $n_{m-1} | n_{m-2}$. Weiter folgert man durch $n_{m-3} = a_{m-2}n_{m-2} + n_{m-1}$, dass $n_{m-1} | a_{m-3}$. Mittels Induktion findet man zum Schluss, dass $n_{m-1} | n_0, n_1$. Andersrum, sei b eine Zahl sodass $b | n_0$ und $b | n_1$ gilt, dann folgt $b | n_0 - a_1 n_1 = n_2$ und mittels Induktion schlussendlich $b | n_{m-1}$. Somit haben wir beide Eigenschaften die für einen grössten gemeinsamen Teiler nötig sind. □

Substituiert alle Gleichungen im Algorithmus rückwärts ineinander, d.h.

$$n_{m-1} = n_{m-3} - a_{m-2}n_{m-2} = n_{m-3} - a_{m-2}\left(n_{m-4} - a_{m-3}n_{m-3}\right) = \ldots,$$

so erhält man den Satz von Bézout.

Satz 2.3 (Bézout)**.** *Gegeben zwei ganze Zahlen $a, b \in \mathbb{Z}$, so findet man ganze Zahlen $x, y \in \mathbb{Z}$, sodass*

$$xa + yb = \mathrm{ggT}(a, b)$$

gilt.

Hierzu sollte man noch kurz erwähnen, dass der Euklidische Algorithmus einen grössten gemeinsamen Teiler produziert. Im Satz von Bézout kann aber ein beliebiger grösster gemeinsamer Teiler gewählt werden. Jene unterscheiden sich aber nur durch eine Multiplikation mit einer Einheit. Entsprechend kann man die Gleichung, welche der Euklidische Algorithmus produziert, einfach mit dieser Einheit multiplizieren.

Definition. *Wir sagen, dass zwei ganze Zahlen* $a, b \in \mathbb{Z}$ teilerfremd *sind, genau dann wenn* $\mathrm{ggT}(a, b) = 1$.

Beispiel. *Wendet man den Euklidischen Algorithmus auf das Zahlenpaar* $(17, 7)$*, so findet man*

$$\begin{aligned} 17 &= 2 \cdot 7 + 3, \\ 7 &= 2 \cdot 3 + 1, \\ 3 &= 3 \cdot 1 + 0, \end{aligned}$$

und somit ist $\mathrm{ggT}(17, 7) = 1$*. Mit Rücksubstitution findet man weiter*

$$1 = 7 - 2 \cdot 3 = 7 - 2 \cdot (17 - 2 \cdot 7) = 5 \cdot 7 + (-2) \cdot 17.$$

Das folgende Lemma ist sehr nützlich für das gängige Rechnen mit Teilern.

Lemma 2.4. *Für ganze Zahlen* $a, b, c \in \mathbb{Z}$ *gelten folgende Regeln:*

1. $\mathrm{ggT}(a, b) = 1$ *und* $\mathrm{ggT}(a, c) = 1 \Leftrightarrow \mathrm{ggT}(a, bc) = 1$*;*
2. $a \mid bc$ *und* $\mathrm{ggT}(a, c) = 1 \Rightarrow a \mid b$*;*
3. $a \mid c$ *und* $b \mid c$ *und* $\mathrm{ggT}(a, b) = 1 \Rightarrow ab \mid c$*.*

Beweis. 1. Nach dem Satz von Bézout 2.3 gibt es ganze Zahlen $x, y, z, w \in \mathbb{Z}$, sodass

$$\begin{aligned} xa + yb &= 1, \\ za + wc &= 1, \end{aligned}$$

gilt. Wir folgern aus der zweiten Gleichung, dass $bza + wbc = b$ und setzen dies in der ersten ein und erhalten

$$(x + bzy) \cdot a + yw \cdot bc = 1.$$

Ein gemeinsamer Teiler von a und bc muss folglich auch 1 teilen, d.h. $\mathrm{ggT}(a, bc) = 1$. Umgekehrt gilt, dass ein gemeinsamer Teiler von a und b auch ein gemeinsamer Teiler von a und bc sein muss. Entsprechend teilt jener 1 und es folgt $\mathrm{ggT}(a, b) = 1$. Analog gilt auch $\mathrm{ggT}(a, c) = 1$.

2. Da a und c teilerfremd sind, gibt es nach dem Satz von Bézout 2.3 zwei ganze Zahlen $x, y \in \mathbb{Z}$ mit $xa + yc = 1$. Es folgt

$$a \mid bc \Rightarrow a \mid xba + ybc = b.$$

3. Schreibe $c = ka$ mit $k \in \mathbb{Z}$. Nun gilt $b \mid c = ka$ und da a und b teilerfremd sind, so gilt nach vorherigem Lemma, dass $b \mid k$, d.h. $k = bk'$. Wir folgern, dass $c = k'ab$ und somit gilt $ab \mid c$. □

Definition. *Ein Element* $p \in \mathbb{Z}$ *heisst* irreduzibel *falls* p *nicht Null und keine Einheit ist und ferner gilt:* $p = ab \Rightarrow a$ *oder* b *ist eine Einheit, i.e.* $a \in \mathbb{Z}^\times$ *oder* $b \in \mathbb{Z}^\times$*.*

Definition. *Ein Element $p \in \mathbb{Z}$ heisst* prim *falls p nicht Null und keine Einheit ist und ferner gilt: $p \mid ab \Rightarrow p \mid a$ oder $p \mid b$ für alle $a, b \in \mathbb{Z}$.*

Satz 2.5. *In $\mathbb{Z}$ ist jedes prime Element irreduzibel und jedes irreduzible Element prim.*

Beweis. Sei p prim und $p = ab$ also insbesondere $p \mid ab$, dann nehme o.B.d.A an, dass $p \mid a$, d.h. $a = pc$ für ein $c \in \mathbb{Z}$. Es folgt

$$p = ab = pbc \Leftrightarrow p(1 - bc) = 0 \Leftrightarrow bc = 1$$

und somit ist b eine Einheit und p irreduzibel.

In die andere Richtung nehme an, dass p irreduzibel ist. Ferner sei n ein beliebiges Element von $\mathbb{Z}$. Man wende nun den Euklidischen Algorithmus 2.2 auf n und p an und findet $d = \mathrm{ggT}(n, p)$. Insbesondere hat man $d \mid p$ also $dc = p$ für ein $c \in \mathbb{Z}$. Es muss nun gelten, dass d eine Einheit ist oder dass c eine Einheit ist. Im zweiten Fall folgt $p \mid c^{-1}p = d \mid n$ und im ersten Fall folgt es mittels dem Satz von Bézout 2.3, dass es Zahlen $u, v \in \mathbb{Z}$ gibt, sodass $un + vp = d \Leftrightarrow d^{-1}un + d^{-1}vp = 1$.

Wir nehmen nun an, dass $a, b \in \mathbb{Z}$ existieren, sodass $p \mid ab$ und $p \nmid a, b$. Wir führen obiges Argument für $n = a, b$ durch. Da nun $p \nmid a, b$ kann der erste Fall nicht auftreten. Entsprechend finden wir also ganze Zahlen $x_1, y_1, x_1, y_2 \in \mathbb{Z}$, sodass

$$1 = x_1 a + y_1 p = x_2 b + y_2 p$$

gilt. Nun ist

$$1 = (x_1 a + y_1 p)(x_2 b + y_2 p) = x_1 x_2 \cdot ab + (y_2 x_1 a + y_1 x_2 b + y_1 y_2 p) \cdot p \tag{2.1}$$

durch p teilbar, also $p \mid 1$ und somit ist p eine Einheit - ein Widerspruch! □

Satz 2.6 (Eindeutige Primfaktorzerlegung)**.** *Jede von Null verschiedene ganze Zahl n besitzt eine Faktorisierung*

$$n = \epsilon p_1 p_2 \cdots p_r$$

in eine Einheit ϵ und in prime/irreduzible ganze Zahlen $p_i, i = 1, \ldots, r$. Die primen Faktoren sind eindeutig bis auf die Reihenfolge und Multiplikation mit einer Einheit.

Beweis. Wir beweisen dies mittels Induktion. Die Existenz ist klar für alle ganzen Zahlen mit Absolutbetrag gleich 1, welche genau die Einheiten sind. Nehme nun die Existenz für alle ganzen Zahlen mit Absolutbetrag kleiner gleich N an und sei $n \in \mathbb{Z}$ eine ganze Zahl mit $|n| = N + 1$, also insbesondere nicht Null. Ist n eine Einheit oder ein irreduzibles Element, so ist man fertig. Andernfalls gibt es zwei von Null verschiedene, Nicht-Einheiten $n_1, n_2 \in \mathbb{Z}$ mit $n = n_1 n_2$. Jene Zahlen haben notwendigerweise einen Absolutbetrag (strikt) grösser als Eins. Da der Absolutbetrag multiplikativ ist, folglich haben sie auch einen Absolutbetrag welcher kleiner als $N + 1$ ist, also höchstens N. Aus der Primfaktorzerlegung von n_1, n_2 nach Induktionsannahme folgt durch Multiplikation auch eine Primfaktorzerlegung von n. Dabei beachte man, dass das Produkt von Einheiten wieder eine Einheit ist.

Man nehme nun an, dass eine ganze Zahl n zwei Primfaktorenzerlegungen

$$n = \epsilon p_1 p_2 \cdots p_r = \epsilon' q_1 q_2 \cdots q_s \tag{2.2}$$

besitzt. O.B.d.A sei $r \geqslant s$. Wir zeigen mittels Induktion, dass in der letzteren Gleichung die primen Elemente eindeutig bis auf Reihenfolge und Multiplikation mit einer Einheit. Wir können o.B.d.A annehmen, dass $r \geqslant s \geqslant 0$. Für $r = s = 0$ ist nichts zu zeigen, da dann die Gleichung $\epsilon = \epsilon'$ ist. Sei nun also $r \geqslant 1$. Wir betrachten nun den allgemeinen Fall (2.2)

und möchten das Paar (r, s) (lexigraphisch) verkleinern. Wir betrachten p_r. Da p_r prim ist und prime Elemente keine Einheiten teilen können, so muss zwanghaft $s \geqslant 1$ gelten. Ferner gibt es nach der primen Eigenschaft von p_r ein $1 \leqslant j \leqslant s$, sodass $p_r | q_j$ gilt. Durch Vertauschen der primen Elemente q_i dürfen wir annehmen, dass $j = s$. Da q_s irreduzibel ist und p_r keine Einheit ist, so folgern wir aus $p_r \mid q_s$, dass $q_s = \epsilon_r p_r$. Wir setzen nun in die Gleichung (2.2) ein und sehen, dass beide Seiten durch $p_r \neq 0$ teilbar sind. Folglich dürfen wir durch p_r dividieren und erhalten neu

$$\epsilon p_1 \cdots p_{r-1} = \eta q_1 \cdots q_{s-1},$$

wobei $\eta = \epsilon' \epsilon_r$ eine Einheit ist. Nun sind wir in der Lage die Induktionsvoraussetzung anzuwenden um den Beweis zuvervollständigen. □

2.2 Modulare Arithmetik und Faktorringe

Wenn wir auf die Gleichung (2.1) zurückblicken, können wir uns Fragen ob jene Rechnung nicht ein wenig einfacher ginge indem man die Terme mit p ignorieren würde? Diese Fragestellung führt uns direkt zur modularen Arithmetik. Mehr oder weniger lässt sich diese als 'Rechnen mit Resten' beschreiben. Dazu führen wir die folgende *Kongruenzrelation* auf $\mathbb{Z}$ ein:

$$a \equiv b \bmod(c) \Leftrightarrow c \mid a - b \Leftrightarrow \exists k \in \mathbb{Z} : a = b + kc.$$

Diese Relation ist eine Äquivalenzrelation, d.h. sie ist symmetrisch

$$a \equiv b \bmod(c) \Leftrightarrow c \mid a - b \Leftrightarrow c \mid b - a \Leftrightarrow b \equiv a \bmod(c),$$

reflexiv

$$a \equiv a \bmod(c) \Leftrightarrow c \mid a - a = 0,$$

und transitiv

$$a \equiv b \bmod(c) \text{ und } b \equiv d \bmod(c) \Rightarrow c \mid a - b \text{ und } c \mid b - d$$
$$\Rightarrow c \mid a - b + b - d = a - d \Rightarrow a \equiv d \bmod(c).$$

Falls $a \equiv b \bmod(c)$ sagen wir, dass a und b in derselben Kongruenzklasse modulo c liegen. Öfters sagen wir auch, dass $a \bmod(c)$ und $b \bmod(c)$ die gleiche Klasse beschreiben. Dazu erwähnen wir mehr später.

Man kann überprüfen, dass die Addition, Subtraktion und Multiplikation mit dieser Relation verträglich ist. Wir überprüfen dies kurz für die Multiplikation. Sei $a \equiv a' \bmod(c)$ und $b \equiv b' \bmod(c)$, nun müssen wir zeigen, dass $ab \equiv a'b' \bmod(c)$ gilt. Letzteres ist äquivalent zu

$$c \mid ab - a'b' = a(b - b') + (a - a')b',$$

was wahr ist, da $c \mid b - b'$ und $c \mid a - a'$ nach Voraussetzung. Jene Rechenreduktion nennen wir 'Rechnen modulo c'. Diese Art von Rechnen ist nützlich um zu überprüfen ob z.B. ein komplizierter Ausdruck durch c teilbar ist. Als konkrete Applikation geben wir zwei Beispiele.

Anwendung. *Für ganze Zahlen n gilt $n \equiv s(n) \bmod(9)$, wobei $s(n)$ steht für die Quersumme von n. So kann man kurz überprüfen, ob man doch nicht vielleicht einen kleinen Rechenfehler bei der Multiplikation oder Addition begangen hat.*

$$23 \cdot 43 = 989, \quad 23 \cdot 43 \equiv (2 + 3) \cdot (4 + 3) \equiv 35 \equiv 8 \equiv 9 + 8 + 9 \equiv 989 \bmod(9).$$

Anwendung. *Um Fehler beim Eintippen der IBAN zu verhindern, wird 'mehr oder weniger*[c] *die Kontonummer modulo* 97 *gerechnet und das Resultat sollte kongruent zu* 1 *modulo* 97 *sein.*

Beim Rechnen modulo c können sich nun diverse Ungewohnheiten ereignen. So gilt zum Beispiel $2 \cdot 3 \equiv 0 \bmod(6)$, aber es gilt weder $2 \equiv 0 \bmod(6)$ noch $3 \equiv 0 \bmod(6)$. Das heisst falls ein Produkt $0 \bmod(c)$ ist, muss nicht zwingend eines der Faktoren $0 \bmod(c)$ sein. Eine weitere Ungewohntheit ist, dass es nun öfters der Fall ist, dass man eine ganze Zahl a mit einer ganzen Zahl b multiplizieren kann und das Resultat ist kongruent zu 1 modulo einer Zahl c. So gilt zum Beispiel:

$$5 \cdot 7 \equiv 35 \equiv 1 \bmod(17).$$

Dies ist eine sehr nützliche Eigenschaft die wir schon öfters insgeheim benutzt haben. Entsprechend würdigen wir es mit einer Definition.

Definition. *Seien $a, c \in \mathbb{Z}$ ganz Zahlen. Eine weitere ganze Zahl $b \in \mathbb{Z}$ heisst* multiplikatives Inverses *von a modulo c, falls*

$$a \cdot b \equiv 1 \bmod(c).$$

Bemerkung. In Aufgabe 2.2 werdet Ihr zeigen, dass falls a und c teilerfremd sind, dann besitzt a ein multiplikatives Inverses b modulo c. Ferner ist ein solches Inverses b eindeutig modulo c bestimmt und die Kongruenzklasse $b \bmod(c)$ hängt nur von der Kongruenzklasse $a \bmod(c)$ ab.

Wir schreiben $a^\star$ für ein multiplikatives Inverses modulo c. Von jenem werden wir nur Gebrauch machen, falls wir modulo c rechnen. Entsprechend spielt nach der Bemerkung, die Wahl des multiplikativen Inversen keine Rolle.

Der folgende Satz zeigt, dass wenn wir modulo c rechnen, dann genügt es modulo allen Primzahlpotenzen, welche c exakt teilen, zu rechnen.

Satz 2.7 (Chinesischer Restsatz)**.** *Für paarweise teilerfremde Zahlen $c_1, \ldots, c_n \in \mathbb{Z}$ und ganze Zahlen $a_1, \ldots, a_n \in \mathbb{Z}$ findet sich eine ganze Zahl $x \in \mathbb{Z}$, welche das folgende Kongruenzsystem erfüllt:*

$$\begin{aligned} x &\equiv a_1 \bmod(c_1), \\ x &\equiv a_2 \bmod(c_2), \\ &\vdots \\ x &\equiv a_n \bmod(c_n). \end{aligned}$$

Ferner ist die Kongruenzklasse von einer solchen Zahl x modulo $c_1 c_2 \cdots c_n$ eindeutig bestimmt.

Beweis. Für $n = 1$ ist die Aussage klar. Betrachten wir nun den Fall $n = 2$. Nach dem Satz von Bézout 2.3 existieren zwei ganze Zahlen $y, z \in \mathbb{Z}$, sodass $yc_1 + zc_2 = 1$. Wir bemerken, dass aus dieser Gleichung folgt, dass

$$\begin{aligned} yc_1 &\equiv 1 \bmod(c_2), \\ zc_2 &\equiv 1 \bmod(c_1). \end{aligned}$$

[c] Die ersten vier Zeichen werden hinten angehängt und die Buchstaben werden durch Zahlen ersetzt $A = 10, B = 11, \ldots$.

Ferner gilt auch

$$yc_1 \equiv 0 \bmod(c_1),$$
$$zc_2 \equiv 0 \bmod(c_2).$$

Wenn wir also nun die Zahl $x = a_1 \cdot yc_1 + a_2 \cdot zc_2$ betrachten, so finden wir

$$x \equiv a_1 yc_1 \equiv a_1 \bmod(c_2),$$
$$x \equiv a_2 yc_2 \equiv a_2 \bmod(c_1).$$

Sei nun $x' \in \mathbb{Z}$ eine weitere ganze Zahl mit der selben Eigenschaft. Dann gilt, dass $x \equiv a_i \equiv x' \bmod(c_i)$ für $i = 1, 2$ und somit $c_1 \mid x - x'$ und $c_2 \mid x - x'$. Da nun c_1, c_2 teilerfremd sind gilt nach Lemma 2.4 sogar $c_1 c_2 \mid x - x'$, d.h. x und x' sind in der selben Kongruenzklasse modulo $c_1 c_2$.

Für $n \geqslant 3$ verwenden wir Induktion. Wir wenden zuerst die Induktionsvoraussetzung auf die ersten $n-1$ Kongruenzgleichungen an. Wir finden also eine Zahl $x' \in \mathbb{Z}$, welche $x' \equiv a_i \bmod(c_i)$ für $i = 1, \ldots, n-1$ erfüllt. Nun bemerken wir, dass nach Lemma 2.4 folgt, dass c_n und $c_1 \cdots c_{n-1}$ teilerfremd sind. Wir wenden nun den Satz für $n = 2$ auf das Kongruenzsystem

$$x \equiv x' \bmod(c_1 \cdots c_{n-1}),$$
$$x \equiv a_n \bmod(c_n)$$

an und finden eine Lösung $x \in \mathbb{Z}$. Es gilt nun

$$\begin{aligned} x \equiv x' \bmod(c_1 \cdots c_{n-1}) &\Rightarrow c_1 \cdots c_{n-1} \mid x - x' \\ &\Rightarrow \forall i \in \{1, \ldots, n-1\} : c_i \mid x - x' \Rightarrow \forall i \in \{1, \ldots, n-1\} : x \equiv x' \equiv a_i \bmod(c_i). \end{aligned}$$

Somit erfüllt x das ursprüngliche Kongruenzsystem. Die Eindeutigkeit von x modulo $c_1 \cdots c_n$ nun auch wieder durch die mehrfache Anwendung von Lemma 2.4. □

Zum Schluss gehen wir noch genauers darauf ein was wir genau mit einer Kongruenzklasse meinen. Die Kongruenzklasse $a \bmod(c)$ einer ganzen Zahl a ist nichts anderes als die Menge aller ganzen Zahlen b, welche zu a kongruent modulo c sind, i.e. die Menge

$$\begin{aligned} &\{b \in \mathbb{Z} | b \equiv a \bmod(c)\} \\ =&\{a + kc \in \mathbb{Z} | k \in \mathbb{Z}\} \\ =&a + c\mathbb{Z}. \end{aligned}$$

Wir schreiben kurz $[a]_c := a + c\mathbb{Z}$ für jene Menge. Falls der Modulus c klar ist vom Kontext, lassen wir öfters auch mal das index c weg, also nur $[a]$. Die Zahl a ist dabei ein *Repräsentant* der Menge $[a]_c$ und könnte daher durch jedes andere Element der Menge ersetzt werden. Es gilt zum Beispiel $[2]_7 = [16]_7$ als Mengen.

Alle ganzen Zahlen gehören nun einer Kongruenzklasse modulo c an ($a \in [a]_c$), d.h. die Menge der ganzen Zahlen wird nun partitioniert in Kongruenzklassen: Für $c \neq 0$ hat man

$$\mathbb{Z} = [0]_c \sqcup [1]_c \sqcup [2]_c \sqcup \cdots \sqcup [|c| - 1]_c,$$

wobei $\sqcup$ eine disjunkte[d] Vereinigung von Mengen symbolisiert. Modulo $c \neq 0$ gibt es also genau $|c|$ verschiedene Kongruenzklassen. Die Menge aller Kongruenzklassen modulo c wird mit $\mathbb{Z}/c\mathbb{Z}$ bezeichnet. Zum Beispiel ist

$$\mathbb{Z}/3\mathbb{Z} = \{[0]_3, [1]_3, [2]_3\}.$$

[d] keine gemeinsamen Elemente

Auf $\mathbb{Z}/c\mathbb{Z}$ kann man nun auch eine Addition, Substraktion, und Multiplikation definieren durch

$$[a]_c + [b]_c := [a+b]_c,$$
$$[a]_c - [b]_c := [a-b]_c,$$
$$[a]_c \cdot [b]_c := [a \cdot b]_c.$$

Dass dies wohldefiniert ist folgt gerade von der Verträglichkeit jener Operationen mit der Kongruenzrelation (modulo c). Die Menge $\mathbb{Z}/c\mathbb{Z}$ mit jenen Operationen nennt sich ein *Faktorring*. Unüberaschenderweise sieht man auch, dass bezüglich der Mengenaddition $A +_M B = \{a+b | a \in A, b \in B\}$ folgendes gilt

$$[a]_c +_M [b]_c \subseteq [a+b]_c.$$

Dies stellt eine andere Art und Weise dar wie man sich die Addition auf Kongruenzklassen vorstellen kann. Analoge Aussagen gelten auch für die Subtraktion und Multiplikation. Als Beispiel betrachten wir nochmals $\mathbb{Z}/3\mathbb{Z}$ und stellen die Additions-, Subtraktions-, und Multiplikationstabellen auf.

$+$	$[0]$	$[1]$	$[2]$
$[0]$	$[0]$	$[1]$	$[2]$
$[1]$	$[1]$	$[2]$	$[0]$
$[2]$	$[2]$	$[0]$	$[1]$

$-$	$[0]$	$[1]$	$[2]$
$[0]$	$[0]$	$[2]$	$[1]$
$[1]$	$[1]$	$[0]$	$[2]$
$[2]$	$[2]$	$[1]$	$[0]$

$\cdot$	$[0]$	$[1]$	$[2]$
$[0]$	$[0]$	$[0]$	$[0]$
$[1]$	$[0]$	$[1]$	$[2]$
$[2]$	$[0]$	$[2]$	$[1]$

Wir können nun auch der Chinesischen Restsatz umformulieren.

Satz 2.8 (Chinesischer Restsatz - Faktorring Formulierung)**.** *Sei $n \in \mathbb{N}$ eine natürliche Zahl und $c_1, \ldots, c_n \in \mathbb{Z}$ paarweise teilerfremde ganze Zahlen. Dann ist die Abbildung*

$$\mathbb{Z}/c_1 \cdots c_n\mathbb{Z} \to \mathbb{Z}/c_1\mathbb{Z} \times \mathbb{Z}/c_2\mathbb{Z} \times \cdots \times \mathbb{Z}/c_n\mathbb{Z},$$
$$[x]_{c_1 \cdots c_n} \mapsto ([x]_{c_1}, [x]_{c_2}, \ldots, [x]_{c_n})$$

ein Isomorphismus von Ringen.

2.3 Sätze von Euler-Fermat und Ordnungen

Satz 2.9 (Kleiner Satz von Fermat[e])**.** *Sei $p \in \mathbb{N}$ eine Primzahl und $a \in \mathbb{Z}$ eine zu p teilerfremde ganze Zahl. So gilt $a^{p-1} \equiv 1 \bmod(p)$.*

Beweis. Sowohl $[1], [2], \ldots, [p-1]$ als auch $[a], [2a], \ldots, [(p-1)a]$ bilden ein komplettes System von Null (hier $[0]$) verschiedenen Kongruenzklassen modulo p. Letzteres ist der Fall, da für $i = 1, \ldots, p-1$ die Zahl ia nicht durch p teilbar ist und aus $[ia] = [ja]$ für $i, j \in \{1, \ldots, p-1\}$ folgt $ia \equiv ja \bmod(p) \Leftrightarrow p \mid a(i-j)$. Da nun a und p teilerfremd sind folgt mittels Lemma 2.4, dass $p \mid i-j$, welches für $i, j \in \{1, \ldots, p-1\}$ sogar $i = j$ impliziert. Ensprechend sind $[a], [2a], \ldots, [(p-1)a]$ $(p-1)$ verschiedene von Null verschiedene Kongruenzklassen. Es existieren aber nur $(p-1)$ verschiedene solche Kongruenzklassen. Ensprechend ist es schon ein komplettes System. Es sprechend ist $[1], [2], \ldots, [p-1]$ eine Vertauschung von $[a], [2a], \ldots, [(p-1)a]$ und es muss gelten, dass

$$1 \cdot 2 \cdots (p-1) \equiv a \cdot (2a) \cdots ((p-1)a) \equiv a^{p-1} \cdot 1 \cdot 2 \cdots (p-1) \bmod(p).$$

Nun ist $1 \cdot 2 \cdots (p-1)$ zu p teilerfremd (Lemma 2.4) und somit besitzt es ein mutliplikatives Inverses modulo p. Multiplizieren wir die Gleichung mit jenem erhalten wir $a^{p-1} \equiv 1 \bmod(p)$. □

[e] Für den grossen Satz hat es hier am Rande nicht genügend Platz.

Die Verallgemeinerung des kleinen Satzes von Fermat ist gegeben durch folgenden Satz von Euler.

Satz 2.10 (Euler)**.** *Sei $n \in \mathbb{N}$ eine natürliche Zahl und $a \in \mathbb{Z}$ eine zu n teilerfremde ganze Zahl. So gilt $a^{\varphi(n)} \equiv 1 \bmod(n)$, wobei*

$$\varphi(n) = n \prod_{\substack{p>0,\ prim \\ p|n}} \left(1 - \frac{1}{p}\right)$$

die Euler'sche Phi-Funktion ist.

Beweis. Der Beweis ist analog, jedoch mit einem kompletten System von zu n teilerfremden Kongruenzklassen. Dazu muss man bemerken, dass falls eine ganze Zahl $b \in \mathbb{Z}$, teilerfremd zu zur Zahl n ist, dann ist jede ganze Zahl b' mit $b' \equiv b \bmod(n)$ auch teilerfremd zu n. Schreibe $b = b' + kn$ für ein $k \in \mathbb{Z}$, dann gilt für einen gemeinsamen Teiler d von b' und n, dass d auch ein Teiler von $b' + kn = b$ ist. Es folgt $d \mid \mathrm{ggT}(b, n) = 1$ und somit ist 1 auch ein grösster gemeinsamer Teiler von b' und n, d.h. b' und n sind teilerfremd. Entsprechend können wir sagen, dass die Kongruenzklasse $[b]_n$ und n teilerfremd sind.

Wir erhalten

$$a^{\varphi(n)} \equiv 1 \bmod(n),$$

wobei

$$\varphi(n) = \sharp\{[b]_n \in \mathbb{Z}/n\mathbb{Z} |\ [b]_n \text{ und } n \text{ sind teilerfremd}\}.$$

Es gilt noch zu zeigen, dass $\varphi(n)$ durch die Formel im Satz gegeben ist. Dazu schreiben wir $n = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$ mit $p_i \in \mathbb{N}$ verschiedene Primzahlen.

Es gilt nun für $[b]_n \in \mathbb{Z}/n\mathbb{Z}$, dass

$$\begin{aligned} [b]_n \text{ und } n \text{ teilerfremd} &\Leftrightarrow \mathrm{ggT}(b, n) = 1 \\ &\Leftrightarrow \forall i : \mathrm{ggT}(b, p_i^{\alpha_i}) = 1 \\ &\Leftrightarrow \forall i : [b]_{p_i^{\alpha_i}} \text{ und } p_i^{\alpha_i} \text{ teilerfremd.} \end{aligned}$$

Wobei die zweite Äquivalenz ist gegeben durch Lemma 2.4. Entsprechend gilt nach dem Chinesischen Restsatz 2.7, dass

$$\begin{aligned} &\sharp\{[b]_n \in \mathbb{Z}/n\mathbb{Z} |\ [b]_n \text{ und } n \text{ sind teilerfremd}\} \\ &\qquad = \prod_{i=1}^{r} \sharp\{[b]_{p_i^{\alpha_i}} \in \mathbb{Z}/p_i^{\alpha_i}\mathbb{Z} |\ [b]_{p_i^{\alpha_i}} \text{ und } p_i^{\alpha_i} \text{ sind teilerfremd}\}. \end{aligned}$$

Letzteres ist einfach zu berechnen, denn für eine Primzahl $p \in \mathbb{N}$ und $\alpha \in \mathbb{N}$ ist eine der Kongruenzklassen $[i]_{p^\alpha}$ für $i = 0, 1, \ldots, p^\alpha - 1$ genau dann nicht teilerfremd zu p falls i durch p teilbar ist, d.h.

$$\sharp\{[b]_{p^\alpha} \in \mathbb{Z}/p^\alpha\mathbb{Z} |\ [b]_{p^\alpha} \text{ und } p^\alpha \text{ sind teilerfremd}\} = p^\alpha - p^{\alpha-1}.$$

Die Formel für $\varphi(n)$ folgt nun. □

Definition. *Sei $c \in \mathbb{N}$ und $a \in \mathbb{Z}$ eine zu c teilerfremde Zahl, dann heisst die kleinste natürliche Zahl $r \in \mathbb{N}$ sodass*

$$a^r \equiv 1 \bmod(c)$$

die (multiplikative) Ordnung *von a modulo c.*

Dass überhaupt eine solche natürliche Zahl existiert, garantiert uns der Satz von Euler 2.10.

Satz 2.11. *Sei $c \in \mathbb{N}$, $a \in \mathbb{Z}$ teilerfremd zu c, r die Ordnung von a modulo c und $m \in \mathbb{N}$ eine weitere natürliche Zahl, sodass $a^m \equiv 1 \bmod(c)$ gilt, dann folgt $r \mid m$.*

Beweis. Man wende die Division mit kleinerem Rest bezüglich r auf m an und schreibe $m = br + n$ mit $0 \leqslant n < r$. Dann gilt

$$a^n \equiv a^n \cdot (a^r)^b \equiv a^m \equiv 1 \bmod(r).$$

Falls nun $n > 0$ ergibt dies einen Widerspruch zur Minimalität von r. □

Korollar 2.12. *Sei r wie im Satz. So gilt $r \mid \varphi(c)$.*

Aufgabe 2.1. *Bestimme zwei ganze Zahlen x, y sodass $223x + 157y = 1$ gilt.*

Aufgabe 2.2. *Sei $0 \neq c \in \mathbb{Z}$ eine ganze Zahl und $a \in \mathbb{Z}$ eine weitere ganze Zahl. Zeige, dass a ein multiplikatives Inverses modulo c besitzt genau dann, wenn a und c teilerfremd sind. Ferner zeige, dass jenes multiplikative Inverse eindeutig modulo c bestimmt ist und jene Kongruenzklasse nur von der Kongruenzklasse von a modulo c abhängt.*

Aufgabe 2.3. *Bestimme die Ordnung von 3 modulo 13.*

3 Zwei Quadrate Satz von Fermat

Wir schauen uns die ersten paar natürlichen Zahlen an und ob sie sich als Summe zweier Quadraten ganzer Zahlen schreiben lassen.

$1 = 1^2 + 0^2$,	$11 =$ nicht möglich,	$21 =$ nicht möglich,
$2 = 1^2 + 1^2$,	$12 =$ nicht möglich,	$22 =$ nicht möglich,
$3 =$ nicht möglich,	$13 = 3^2 + 2^2$,	$23 =$ nicht möglich,
$4 = 2^2 + 0^2$,	$14 =$ nicht möglich,	$24 =$ nicht möglich,
$5 = 2^2 + 1^2$,	$15 =$ nicht möglich,	$25 = 5^2 + 0^2 = 4^2 + 3^2$,
$6 =$ nicht möglich,	$16 = 4^2 + 0^2$,	$26 = 5^2 + 1^2$,
$7 =$ nicht möglich,	$17 = 4^2 + 1^2$,	$27 =$ nicht möglich,
$8 = 2^2 + 2^2$,	$18 = 3^2 + 3^2$,	$28 =$ nicht möglich,
$9 = 3^2 + 0^2$,	$19 =$ nicht möglich,	$29 = 5^2 + 2^2$,
$10 = 3^2 + 1^2$	$20 = 4^2 + 2^2$,	$30 =$ nicht möglich.

Aufgrund dieser Tabelle kann man vielleicht vermuten, dass sich eine natürliche Zahl n genau dann als Summe zweier Quadrate schreiben lässt, falls $2n$ sich als Summe zweier Quadrate schreiben lässt. Dies ist tatsächlich der Fall, denn

$$n = a^2 + b^2 \Leftrightarrow 2n = (a + b)^2 + (a - b)^2,$$

und für die Rückrichtung bemerken wir, dass falls $2n = x^2 + y^2$, dann sind entweder x, y beide gerade oder ungerade. Insbesondere sind $a = \frac{x+y}{2}$, $b = \frac{x-y}{2}$ ganze Zahlen.

Schauen wir weiter auf die Tabelle vermuten wir vielleicht weiter, dass das Produkt zweier natürlichen Zahlen, welche sich als Summe zweier ganzer Quadrate schreiben lassen,

wieder als Summe zweier Quadrate schreiben lässt. Dies ist auch wieder der Fall, denn es gilt die Gleichung von Diophantus

$$(a^2 + b^2)(x^2 + y^2) = (ax + by)^2 + (bx - ay)^2, \tag{3.1}$$

welche man schon vielleicht als Multiplikativität des komplexen Absolutbetrags kennt. Diese Identität gibt es auch mit anderem Vorzeichen:

$$(a^2 + b^2)(x^2 + y^2) = (ax - by)^2 + (bx + ay)^2.$$

Dies erklärt auch gleich wieso wir zwei verschiedene Darstellungen für 25 erhalten.

Mit dieser neuer Information im Kopf kommen wir auf die Idee die Primzahlen zu betrachten. Wir vermuten, dass sich eine Primzahl $p \in \mathbb{N}$ sich genau dann als Summe von zwei Quadraten schreiben lässt, falls $p = 2$ oder $p \equiv 1 \bmod(4)$. Dies ist gerade der zwei Quadrate Satz von Fermat, zuerst bewiesen durch Euler.

Satz 3.1 (Fermat's zwei Quadrate Satz (Euler))**.** *Sei $p \equiv 1 \bmod(4)$ eine Primzahl, so ist p die Summe von zwei Quadraten.*

Bevor wir aber jenen Satz beweisen formulieren wir den generellen zwei Quadrate Satz.

Satz 3.2 (Zwei Quadrate Satz)**.** *Eine natürliche Zahl $n \in \mathbb{N}$ lässt sich als Summe von zwei Quadraten (von ganzen Zahlen) schreiben genau dann und nur dann wenn jede Primzahl $p \equiv 3 \bmod(4)$, welche n teilt, eine gerade Anzahl mal vorkommt in der Primfaktorzerlegung von n.*

Wir möchten nun zeigen, dass die Charakterisierung im Satz 3.2 nötig ist. Dazu benötigen wir folgendes Lemma.

Lemma 3.3. *Sei $p \equiv 3 \bmod(4)$ eine Primzahl und a, b ganze Zahlen sodass $p \mid a^2 + b^2$, dann gilt $p \mid a$ und $p \mid b$.*

Beweis. Falls $p \mid b$, dann gilt auch $p \mid a^2$ und somit $p \mid a$. Also dürfen wir $p \nmid b$ annehmen. Insbesondere gilt $\mathrm{ggT}(p, b) = 1$ und somit besitzt b ein multiplikatives inverses $b^\star$ modulo p. Wir rechnen nun

$$p \mid a^2 + b^2 \Leftrightarrow a^2 \equiv -b^2 \bmod(p) \Leftrightarrow (ab^\star)^2 \equiv -1 \bmod(p).$$

Es folgt, dass die Zahl $x = ab^\star$ die (multiplikative) Ordnung 4 modulo p hat, da $x^4 \equiv (-1)^2 \equiv 1 \bmod(p)$, d.h. die Ordnung teilt 4 nach Satz 2.11, und $x^2 \equiv -1 \not\equiv 1 \bmod(p)$ (da $p \neq 2$), d.h. die Ordnung kann kein Teiler von 2 sein. Daraus folgern wir mit Hilfe von Korollar 2.12, dass $4 \mid \varphi(p) = p - 1$ was im Widerspruch zu $p \equiv 3 \bmod(4)$ ist. □

Reduktion von Satz 3.2 auf Satz 3.1. Sei nun $n = x^2 + y^2$ und $p \equiv 3 \bmod(4)$ eine Primzahl welche n teilt. Sei α der exakte Exponent von p welcher n teilt, das heisst $p^\alpha \mid n$ und $p^{\alpha+1} \nmid n$. Da $\alpha > 0$ folgt $p \mid x^2 + y^2$ und mit Hilfe des Lemma also $p \mid x$ und $p \mid y$. Es muss also $\alpha \geqslant 2$ gelten und wir können nun $n/p^2 = (x/p)^2 + (y/p)^2$ schreiben. Falls nun $\alpha - 2 > 0$, so kann man das Argument wiederholen. Da nun α endlich ist muss dieser Prozess enden, aber dies kann nur geschehen falls α gerade ist. Damit haben wir gezeigt, dass die Charakterisation im Satz nötig ist. Um zu zeigen, dass die Charakterisation hinreichend ist, brauchen wir die Identität von Diophantus (3.1). Es genügt also zu zeigen, dass 2, jedes Quadrat p^2 einer Primzahl $p \equiv 3 \bmod(4)$ und jede Primzahl $p \equiv 1 \bmod(4)$ als Summe zweier Quadraten schreiben lässt. Der ersten zwei Fälle sind trivial, da $2 = 1^2 + 1^2$ und $p^2 = p^2 + 0^2$, und der letzte Fall ist ja gerade die Aussage vom Satz 3.1. □

Wir benötigen wieder ein kleines Lemma.

Lemma 3.4. *Sei $p \equiv 1 \bmod(4)$ eine Primzahl. Es existiert eine ganze Zahl $x \in \mathbb{Z}$ sodass $x^2 \equiv -1 \bmod(p)$.*

Beweis. Setze $x = (\frac{p-1}{2})!$, dann gilt

$$x^2 \equiv 1 \cdot 2 \cdots \tfrac{p-1}{2} \cdot \left(\tfrac{p-1}{2} - p\right) \cdot \left(\tfrac{p-3}{2} - p\right) \cdots (1-p) \equiv (-1)^{\frac{p-1}{2}} (p-1)! \bmod(p).$$

Da $p \equiv 1 \bmod(4)$ gilt $2 \mid \frac{p-1}{2}$. Es genügt also zu zeigen, dass $(p-1)! \equiv -1 \bmod(p)$. Dies nennt sich der Satz von Wilson. Wir paaren dafür die Zahl $i \in \{1, \ldots, p-1\}$ mit ihrem multiplikativen Inversen $i^\star$ modulo p, welches existiert, da $\mathrm{ggT}(i,p) = 1$. Wir bemerken dazu, dass $i^\star$ eindeutig in $\{1, \ldots, p-1\}$ gewählt werden kann und dass i ein mutliplikatives Inverses modulo p für $i^\star$ ist, d.h. $i = i^{\star\star}$. Wir erhalten also immer Paare $\{i, i^\star\}$, die sich gegenseitig aufheben, womit wir $i \cdot i^\star \equiv 1$ meinen, ausser falls i zu sich selbst invers ist, d.h. $i = i^\star \Leftrightarrow i^2 \equiv 1 \bmod(p)$. Wir folgern

$$(p-1)! \equiv \prod_{\substack{i \in \{1,\ldots,p-1\} \\ i^2 \equiv 1 \bmod(p)}} i \quad \bmod(p).$$

Aus $i^2 \equiv 1 \bmod(p)$ folgt $p \mid (i-1)(i+1)$. Da p prim ist, sogar $p \mid i-1$ oder $i+1$. Es folgt, dass $i = 1$ oder $i = p-1$. Es gilt also $(p-1)! \equiv 1 \cdot (p-1) \equiv -1 \bmod(p)$. □

Beweis von Satz 3.1. Nach dem Lemma 3.4 gibt es eine natürliche Zahl $m \in \mathbb{N}$ und zwei ganze Zahlen $x, y \in \mathbb{Z}$ teilerfremd zu p sodass $mp = x^2 + y^2$ gilt. Sei nun $m \in \mathbb{N}$ die kleinste natürliche Zahl für welche ganze Zahlen $x, y \in \mathbb{Z}$ existieren, sodass $mp = x^2 + y^2$ und x, y sind teilerfremd zu p. Falls $m = 1$, so ist der Satz bewiesen. Man nehme nun an, dass $m > 1$. Ersetzt man x, y mit den zu ihnen kongruenten Zahlen modulo p im Interval $[-\frac{p-1}{2}, \frac{p-1}{2}]$ so sind jene nicht 0 (da $p \nmid x, y$), entsprechend wird m höchstens kleiner und es folgt, dass $m < p$. Also insbesondere ist m teilerfremd zu p. Ferner folgt von der Minimalität von m, dass x, y teilerfremd sind.

Seien nun $a, b \in [-\frac{m}{2}, \frac{m}{2}]$ zwei ganze Zahlen, sodass $a \equiv x \bmod(m)$ und $b \equiv y \bmod(m)$ gilt. Da x, y teilerfremd sind gilt $a^2 + b^2 > 0$, da ansonsten $m \mid x, y$ mit $m \geqslant 2$ - ein Widerspruch dazu, dass x, y teilerfremd sind. Es gilt nun, dass

$$\begin{aligned} ax + by &\equiv x^2 + y^2 \equiv 0 \bmod(m), \\ ay - bx &\equiv xy - yx \equiv 0 \bmod(m), \\ a^2 + b^2 &\equiv x^2 + y^2 \equiv 0 \bmod(m). \end{aligned}$$

Mithilfe der Identität von Diophantus (3.1) findet man

$$\frac{a^2+b^2}{m} p = \frac{1}{m^2}(a^2+b^2)(x^2+y^2) = \left(\frac{ax+by}{m}\right)^2 + \left(\frac{bx-ay}{m}\right)^2.$$

Nun gilt aber $0 < \frac{a^2+b^2}{m} \leqslant \frac{m}{2} < p$. Insbesondere sind die neuen ganzen(!) Zahlen auf der rechten Seite auch teilerfremd zu p und die Minimalität von m impliziert $m \leqslant \frac{a^2+b^2}{m} \leqslant \frac{m}{2}$, einen Widerspruch! □

Aufgabe 3.1. *Zeige, dass für eine Primzahl $p \equiv 1 \bmod(4)$ die zwei Zahlen x, y für welche $p = x^2 + y^2$ gilt bis auf Reihenfolge und Vorzeichen eindeutig bestimmt sind.*

Hinweis: Betrachte $a^2(x^2+y^2) - x^2(a^2+b^2)$, wobei $p = a^2 + b^2$ eine andere Lösung ist.

Aufgabe 3.2. *Sei $p \in \mathbb{N}$ eine Primzahl und $x, y \in \mathbb{Z}$ zwei ganze Zahlen, welche teilerfremd zu p sind. Nehme an, dass $p \mid x^2 - xy + y^2$ und folgere, dass entweder $p = 3$ oder $p \equiv 1 \bmod(6)$.*

Hinweis: $(x + y)(x^2 - xy + y^2) = x^3 + y^3$.

4 Gauss'schen Zahlen

Eine gute Referenz für dieses Kapitel ist [12, §4]. Es beinhaltet auch eine gute Repetition der Themen, die wir bei den ganzen Zahlen angesprochen haben.

Wie wir bereits erwähnt haben steht die Gleichung von Diophantus (3.1) im Zusammenhang mit der Multiplikativität des komplexen Absolutbetrags:

$$|a + bi|_{\mathbb{C}} \cdot |x - yi|_{\mathbb{C}} = |(a + bi)(x - yi)|_{\mathbb{C}}.$$

Da wir an ganzen Zahlen $a, b, x, y \in \mathbb{Z}$ interessiert sind, kommen wir auf die Idee den Unterring $\mathbb{Z}[i] = \mathbb{Z} + \mathbb{Z}i \subset \mathbb{C}$ zu betrachten. Dass jene Menge tatsächlich bezüglich Addition, Subtraktion, und Multiplikation abgeschlossen ist leicht zu überprüfen und ist dem Leser überlassen. Den Unterring $\mathbb{Z}[i]$ nennt sich der Ring der *Gauss'schen Zahlen*. Wir werden im Verlauf dieses Kapitels sehen, dass die Gauss'schen Zahlen sich in vielen Sinnen wie die ganzen Zahlen verhalten. Dies ist auch der Grund warum wir die ganzen Zahlen von einem generellen Standpunkt betrachtet haben. Wir sehen leicht, dass die Gauss'schen Zahlen alle Eigenschaften der ganzen Zahlen, welche wir im Kapitel 2 angenommen haben, bis auf die Division mit kleinerem Rest erfüllen. Wir wollen nun auch noch diese letzte Eigenschaft zeigen. Doch, dazu müssen wir zuerst definieren was denn genau 'klein' heisst in den Gauss'schen Zahlen. Als Kandidat steht sicherlich der komplexe Absolutbetrag im Raum. Jener ist aber nicht immer ganz für eine Gauss'sche Zahl. Nichtsdestotrotz ist das Quadrat des komplexen Absolutbetrags einer Gauss'schen Zahl immer ganz. Wir bezeichnen jenes als Norm[f] und verwenden die Notation $N(\cdot)$, d.h. für $a + bi \in \mathbb{Z}[i]$ mit $a, b \in \mathbb{Z}$ setzt man

$$N(a + bi) = (a + bi)\overline{(a + bi)} = a^2 + b^2. \tag{4.1}$$

Letzteres bringt uns auch zum Punkt, dass die Gauss'schen Zahlen auch abgeschlossen bezüglich der komplexen Konjugation sind. Die komplexe Konjugation kommutiert mit der Multiplikation, d.h. $\overline{\alpha\beta} = \overline{\alpha}\overline{\beta}$ $\forall \alpha, \beta \in \mathbb{Z}[i]$. Es folgt, dass auch die Norm multiplikativ ist:

$$N(\alpha\beta) = \alpha\beta\overline{\alpha\beta} = \alpha\overline{\alpha}\beta\overline{\beta} = N(\alpha)N(\beta), \quad \forall \alpha, \beta \in \mathbb{Z}[i]. \tag{4.2}$$

Satz 4.1. *Gegeben eine Gauss'sche Zahl $\gamma \in \mathbb{Z}[i]$ und eine von Null verschiedene Gauss'sche Zahl $\rho \in \mathbb{Z}[i]$. So existieren Gauss'sche Zahlen $\alpha, \beta \in \mathbb{Z}[i]$ mit $N(\beta) < N(\rho)$, sodass $\gamma = \alpha\rho + \beta$ gilt.*

Beweis. Man setze $\alpha \in \mathbb{Z}[i]$ diejenige Gauss'sche Zahl welche $\frac{\gamma}{\rho}$ am nächsten liegt. Explizit ist $\alpha = \lfloor \Re(\frac{\gamma}{\rho}) + \frac{1}{2} \rfloor + \lfloor \Im(\frac{\gamma}{\rho}) + \frac{1}{2} \rfloor i$. Die komplexe Distanz von α und $\frac{\gamma}{\rho}$ beträgt also höchstens $\frac{1}{\sqrt{2}}$. Ferner setze man $\beta = \gamma - \alpha\rho \in \mathbb{Z}[i]$. Nun gilt

$$N(\beta) = N(\gamma - \alpha\rho) = |\gamma - \alpha\rho|^2_{\mathbb{C}} = |\rho|^2_{\mathbb{C}} \cdot |\tfrac{\gamma}{\rho} - \alpha|^2_{\mathbb{C}} \leqslant N(\rho) \cdot \tfrac{1}{2} < N(\rho).$$

□

[f] Diese Bezeichnung ist gerechtfertigt, da es gerade die Norm der Körpererweiterung $\mathbb{Q}(i)/\mathbb{Q}$ ist.

Im Generellen heisst ein Ring mit all jenen Eigenschaften ein *Euklidischer Ring*. Ein anderes Beispiel dafür wäre $\mathbb{R}[X]$ (oder allgemeiner der Polynom Ring $K[X]$ über einem Körper K) mit der Euklidischer Norm gegeben durch den Grad des Polynoms.

Das besondere an Euklidischen Ringen ist, das alles was wir in den Sektionen 2.1, 2.2 hergeleitet haben gilt. Dazu muss man einfach überall die ganzen Zahlen $\mathbb{Z}$ mit dem Euklidischen Ring, z.B. den Gauss'schen Zahlen $\mathbb{Z}[i]$, ersetzen. Insbesondere gilt der Satz von Bézout, die eindeutige Primfaktorzerlegung und das Rechnen mit Kongruenzen. Wir erläutern dies nun genauers.

Als erstes möchten wir, dass die Relation 'ein Teiler von' in den Gauss'schen Zahlen eine Erweiterung jener Relation auf den ganzen Zahlen ist. Dies ist im folgenden Lemma präzisiert.

Lemma 4.2. *Für zwei ganze Zahlen $a, b \in \mathbb{Z} \subset \mathbb{Z}[i]$ gelten folgendes:*

1. *Falls a die Zahl b als ganze Zahl teilt, dann teilt a die Zahl b auch als Gauss'sche Zahl;*

2. *Falls a die Zahl b als Gauss'sche Zahlen teilt, dann teilt a die Zahl b auch als ganze Zahlen.*

Beweis. Seien also $a, b \in \mathbb{Z}$ ganze Zahlen.

1. Sei $c \in \mathbb{Z}$ eine ganze Zahl, sodass $ac = b$. Nun kann aber $c \in \mathbb{Z} \subset \mathbb{Z}[i]$ auch als Gauss'sche Zahl aufgefasst werden. Insbesondere teilt a auch b als Gauss'sche Zahl.

2. Falls $a = 0$ gilt, so gilt auch $b = 0$ und die Aussage ist klar. Für $a \neq 0$ sei nun $\gamma \in \mathbb{Z}[i]$ eine Gauss'sche Zahl, sodass $a\gamma = b$ gilt. Dann gilt

$$a\gamma = b = \overline{b} = \overline{a\gamma} = a\overline{\gamma}.$$

   Es folgt $a(\gamma - \overline{\gamma}) = 0$ und da nun $a \neq 0$ ferner $\gamma = \overline{\gamma}$. Insbesondere hat γ Imaginärteil gleich 0 und alle Gauss'schen Zahlen mit Imaginärteil gleich 0 sind gerade die ganzen Zahlen. Folglich gilt a teilt b als ganze Zahl.

□

Als nächstes charakterisieren und berechnen wir alle Einheiten der Gauss'schen Zahlen.

Lemma 4.3. *Die Einheiten der Gauss'schen Zahlen sind genau diejenigen $\epsilon \in \mathbb{Z}[i]$ mit $N(\epsilon) = 1$, d.h. $\epsilon \in \{\pm 1, \pm i\}$.*

Beweis. Für eine Einheit $\epsilon \in \mathbb{Z}[i]$ findet man eine Gauss'sche Zahl $\delta \in \mathbb{Z}[i]$ mit $\epsilon\delta = 1$. Es folgt mit der Multiplikativität der Norm (4.2), dass $N(\epsilon)N(\delta) = N(1) = 1$. Da die Normen $N(\epsilon), N(\delta)$ nicht negative ganze Zahl darstellen, so muss $N(\epsilon) = N(\delta) = 1$ gelten. Umgekehrt gilt $N(\epsilon) = 1$, so ist $\epsilon\overline{\epsilon} = N(\epsilon) = 1$, also ist ϵ eine Einheit. Sei nun $\epsilon = a + bi \in \mathbb{Z}[i]$ mit $a, b \in \mathbb{Z}$. So folgt aus $1 = N(\epsilon) = a^2 + b^2$, dass entweder $a = \pm 1$ und $b = 0$ oder $a = 0$ und $b = \pm 1$, also $\epsilon \in \{\pm 1, \pm i\}$. □

Um den grössten gemeinsamen Teiler zweier Gauss'schen Zahlen zu berechnen können wir genau gleich den Euklidischen Algorithmus anwenden, wobei wir die Division mit kleinerem Rest genau gleich wie im Beweis von Satz 4.1 berechnen.

Beispiel. *Wir berechen den grössten gemeinsamen Teiler von* $7+5i$ *und* $1+3i$*. Wir benutzten den Euklidischen Algorithmus. Wir haben*

$$\frac{7+5i}{1+3i} = \frac{(7+5i)(1-3i)}{10} = \frac{22-16i}{10} \approx 2-2i.$$

Entsprechend finden wir

$$7+5i = (2-2i)(1+3i) + (-1+i).$$

Wir überprüfen zur Sicherheit, dass $N(-1+i) = 2 < 10 = N(1+3i)$*. Weiter finden wir*

$$\frac{1+3i}{-1+i} = \frac{(1+3i)(-1-i)}{2} = \frac{2-4i}{2} = 1-2i.$$

Damit terminiert der Algorithmus und wir finden, dass $-1+i$ *ein grösster gemeinsamer Teiler von* $7+5i$ *und* $1+3i$ *ist. Da wir auch alle Einheiten kennen, können wir auch gleich alle anderen grössten gemeinsamen Teiler von* $7+5i$ *und* $1+3i$ *hinschreiben. Sie sind* $(-1)(-1+i) = 1-i$*,* $i(-1+i) = -1-i$*, und* $(-i)(-1+i) = 1+i$*. Wir können nun auch rückwärts Substituieren und erhalten die Identität von Bézout. In unserem Beispiel ist dies sehr einfach, denn wir erhalten direkt*

$$(1)\cdot(7+5i) + (-2+2i)\cdot(1+3i) = -1+i.$$

Beispiel. *Wir berechnen den grössten gemeinsamen Teiler von* $2+i$ *und* $2-i$*. Wir benutzten den Euklidischen Algorithmus. Wir haben*

$$\frac{2+i}{2-i} = \frac{(2+i)^2}{5} = \frac{3+4i}{5} \approx 1+i.$$

Entsprechend finden wir

$$2+i = (1+i)(2-i) + (-1).$$

Wir überprüfen zur Sicherheit, dass $N(-1) = 1 < 5 = N(2-i)$*. Weiter finden wir*

$$\frac{2-i}{-1} = -2+i.$$

Damit terminiert der Algorithmus und wir finden also $\mathrm{ggT}(2+i, 2-i) = -1$*. Das heisst* $2+i$ *und* $2-i$ *sind teilerfremd und wir haben die Bézout Identität*

$$(-1)\cdot(2+i) + (1+i)\cdot(2-i) = 1.$$

Es folgt, dass -1 *ist ein Multiplikatives Inverses von* $2+i$ *modulo* $2-i$*.*

Letzteres Beispiel zeigte, dass $2+i$ und $2-i$ teilerfremd sind. Das heisst wir sind auch in der Lage den chinesischen Restsatz anzuwenden. Wir bemerken, dass $(2+i)(2-i) = 5$. Wir möchten nun alle möglichen Reste der Gauss'schen Zahlen modulo 5 betrachten. Dazu bemerken wir, dass eine Gauss'sche Zahl $a+bi \in \mathbb{Z}[i]$ mit $a, b \in \mathbb{Z}$ genau dann ein Vielfaches (in den Gauss'schen Zahlen) von 5 ist falls a und b Vielfache (in den ganzen Zahlen) von 5 sind. Es ist nun einfach zu sehen, dass die 25 Kongruenzklassen

$$[a+bi]_5, \quad a, b \in \{0,1,2,3,4\}$$

ein vollständiges System von verschiedenen Kongruenzklassen modulo 5 (in den Gauss'schen Zahlen) sind. Es folgt vom chinesischen Restsatz, dass die Kardinalität von $\mathbb{Z}[i]/5\mathbb{Z}[i]$ gleich

dem Produkt der Kardinalitäten von $\mathbb{Z}[i]/(2+i)\mathbb{Z}[i]$ und $\mathbb{Z}[i]/(2-i)\mathbb{Z}[i]$ ist. Letztere können nicht eine Kardinalität von 1 besitzten, denn sonst wäre ja zum Beispiel $[1]_{2+i} = [0]_{2+i}$ und es wäre also 1 durch $2+i$ teilbar. Also wäre $2+i$ eine Einheit. Dies ist aber im Widerspruch zu Lemma 4.3. Es gibt also 5 verschiede Restklassen modulo $2+i$, beziehungsweise modulo $2-i$. Wenn wir nun auf das Problem in der Motivation 1 zurückblicken, so fehlte uns da nur die Begründung wieso modulo $2-i$ es 5 Restklassen (Kongruenzklassen) gibt.

Wir haben nun gesehen, dass 5 in den Gauss'schen Zahlen nicht mehr prim ist. Es stellt sich also die Frage was sind denn genau die Primzahlen in den Gauss'schen Zahlen. Wir zeigen folgenden Satz mithilfe des zwei Quadrate Satz.

Satz 4.4. *Die primen Elemente der Gauss'schen Zahlen sind bis auf Multiplikation mit einer Einheit gegeben durch*

1. $1+i$,
2. *Ganze Primzahlen* $p \in \mathbb{N}$ *mit* $p \equiv 3 \bmod(4)$,
3. $a \pm bi$, *wobei* $a, b \in \mathbb{N}$, $a > b$ *und* $a^2 + b^2 = p$, *eine ganze Primzahl* $p \equiv 1 \bmod(4)$,

wobei im letzten Punkt, die zwei primen Elemente $a+bi$ *und* $a-bi$ *sich nicht nur durch eine Multiplikation mit einer Einheit unterscheiden, und* a, b *eindeutig bestimmt sind.*

Zuerst zeigen wir, dass jene gelisteten Gauss'sche Zahlen tatsächlich irreduzibel sind. Dazu benötigen wir ein kleines Lemma.

Lemma 4.5. *Sei* $\alpha \in \mathbb{Z}[i]$ *eine Gauss'sche Zahl. Nehme an, dass* $N(\alpha)$ *als ganze Zahl irreduzibel ist, dann ist* α *auch irreduzibel als Gauss'sche Zahl.*

Beweis. Nehme an $\alpha = \beta\gamma$ sei eine Faktorisation in Gauss'sche Zahlen, dann gilt $N(\alpha) = N(\beta)N(\gamma)$. Nun ist aber $N(\alpha)$ irreduzibel als ganze Zahl, somit muss $N(\beta) = 1$ oder $N(\gamma) = 1$ gelten (-1 ist ausgeschlossen da die Norm nicht negativ ist). Aber dann ist entweder β oder γ eine Gauss'sche Einheit nach Lemma 4.3. Es folgt, dass α irreduzibel ist (und somit auch prim). □

Beweis von Satz 4.4. Das vorherige Lemma 4.5 zeigt sofort, dass die Gauss'schen Zahlen, welche in *1.* und *3.* gelistet sind, sind prime Elemente. Ferner bemerken wir auch noch kurz, dass nach dem zwei Quadrate Satz für ganze Primzahlen $p \in \mathbb{N}$ mit $p \equiv 1 \bmod(4)$ immer zwei natürliche (verschiedene) Zahlen $a, b \in \mathbb{N}$ gibt sodass $a^2 + b^2 = p$ gilt, jene sind ferner nach Aufgabe 3.1 eindeutig. Alternativ erhält man die Eindeutigkeit von der eindeutigen Primfaktorzerlegung von p in den Gauss'schen Zahlen. Ferner gilt $(a+bi) \notin \{\pm 1, \pm i\} \cdot (a-bi)$.

Nehme im Fall *2.* an, dass p nicht irreduzibel ist, dann lässt sich also schreiben $\alpha\beta = p$ wobei $\alpha, \beta \in \mathbb{Z}[i]$ keine Einheiten sind. Es folgt $N(\alpha)N(\beta) = p^2$ und somit $N(\alpha) = N(\alpha) = p$. Schreiben wir $\alpha = a+bi$ mit $a, b \in \mathbb{Z}$, dann haben wir $N(\alpha) = a^2 + b^2 = p$. Dies ist aber nach dem zwei Quadrate Satz 3.2 nicht möglich und deshalb muss p als Gauss'sche Zahl irreduzibel sein. Ferner unterscheiden sich die gelisteten primen Elemente nicht durch eine Multiplikation mit einer Einheit, da sie verschiedene Normen haben. Die Ausnahme sind die Primzahlen, welche in *3.* gelistet sind.

Nehme nun an, dass jene Liste von primen Elementen nicht komplett ist und das δ eine weitere prime Gauss'sche Zahl ist. Dann ist $N(\delta) > 1$ und wir finden eine ganze Primzahl $p \in \mathbb{N}$ welche $N(\delta)$ ganz (in $\mathbb{Z}$) teilt. Wir kennen nun die Gauss'sche Primfaktorzerlegung von p. Falls $p = 2$ ist, so hat man $2 = -i(1+i)^2$, falls $p \equiv 3 \bmod(4)$, so ist p schon eine prime Gauss'sche Zahl und falls $p \equiv 1 \bmod(4)$, so ist $p = (a+bi)(a-bi)$ mit $a^2 + b^2 = p$.

Insbesondere sind dies aufgelistete prime Elemente der Gauss'schen Zahlen. Es folgt daher, dass δ teilerfremd zu p ist und wir können nach dem Satz von Bézout zwei Gauss'sche Zahlen $\eta, \nu \in \mathbb{Z}[i]$ finden, sodass $\eta\delta + \nu p = 1$ gilt. Wir finden dann

$$0 \equiv N(\eta)N(\delta) = N(1 - \nu p) = (1 - \nu p)(1 - \overline{\nu} p) \equiv 1 \bmod(p).$$

Es würde folgen, dass $p \mid 1$ (als Gauss'sche Zahl) und somit wäre p eine Einheit im Widerspruch zu Lemma 4.3. Alternativ folgt aus $p \mid 1$ als Gauss'sche Zahl, dass $p \mid 1$ als ganze Zahl nach Lemma 4.2Es folgt wieder ein Widerspruch, da p ein primes Element in $\mathbb{Z}$ ist. Wir folgern, dass unsere Liste von primen Gauss'schen Zahlen schon komplett ist! □

Bemerkung. Letzteres Argument kann man auch gebrauchen, um den Satz 3.1 zu beweisen. Siehe Aufgabe 4.2.

Aufgabe 4.1. *Seien $\alpha, \beta \in \mathbb{Z}[i]$ Gauss'sche Zahlen, sodass ihre Normen $N(\alpha), N(\beta) \in \mathbb{Z}$ teilerfremd sind (als ganze Zahlen). Zeige, dass α und β teilerfremd sind als Gauss'sche Zahlen.*

Aufgabe 4.2. *Sei $\mathbb{N} \ni p \equiv 1 \bmod(4)$ eine ganze Primzahl und sei $x \in \mathbb{Z}$ eine ganze Zahl, sodass $x^2 \equiv -1 \bmod(p)$ (Eine solche Zahl ist durch Lemma 3.4 garantiert). Zeige, dass $x + i$ und p als Gauss'sche Zahlen nicht teilerfremd sind und folgere dass sich p als Summe von zwei ganzen Quadraten schreiben lässt.*

5 Quaternionen

Eine ausführliche Referenz für generelle Quaternion Algberen ist [13].

Ähnlich wie Gleichung von Diophantus (3.1) zeigt, dass das Produkt von zwei Summen zweier Quadrate wieder eine Summe zweier Quadrate ist, zeigt folgende Identität von Euler, dass dasselbe für Summen von vier Quadraten gilt:

$$\begin{aligned}(x_1^2 + x_2^2 + x_3^2 + x_4^2)&(y_1^2 + y_2^2 + y_3^2 + y_4^2) \\ &= (x_1y_1 - x_2y_2 - x_3y_3 - x_4y_4)^2 + (x_1y_2 + x_2y_1 + x_3y_4 - x_4y_3)^2 \\ &\quad + (x_1y_3 - x_2y_4 + x_3y_1 + x_4y_2)^2 + (x_1y_4 + x_2y_3 - x_3y_2 + x_4y_1)^2. \quad (5.1)\end{aligned}$$

Basierend auf dieser Identität gelang es Hamilton ein Quadrupel von Zahlen zu multiplizieren. Als Hamilton die Quaternionenmultiplikation entdeckte, gravierte er jene in die Brücke auf der er sich zur Zeit befand. Die Inschrift ist leider wegerrodiert, aber ein Plaque an dessen Stelle wurde später installiert.

Bevor wir jedoch eine vereinfachte schreibweise und Multiplikationstabelle betrachten, betrachten wir eine allgemeine Konstruktion.

5.1 Cayley–Dickson Konstruktion

Die allgmeine Konstruktion sowie ein alternativer Beweis des Satzes von Hurwitz §5.3 findet man zum Beispiel in [4].

Es hilf uns nochmals die Konstruktion der komplexen Zahlen anzuschauen. Eine komplexe Zahl $z = x + iy \in \mathbb{C}$ besteht aus einem reellen Zahlenpaar $(x, y) \in \mathbb{R}^2$. Wobei die Addition komponentenweise gegeben ist durch

$$(x, y) + (a, b) = (x + a, y + b),$$

Abbildung 6: Quaternion Inschrift auf der Brougham (Broom) Brücke in Dublin. Cone83, Wikipedia, CC BY-SA 4.0 Lizenz.

und die Multiplikation wie folgt

$$(x, y) \cdot (a, b) = (xa - yb, xb + ya).$$

Die Addition erbt alle Eigenschaften von den reellen Zahlen, d.h. assoziativ, kommutativ, Null, und Inverse. Die Multiplikation ist auch assoziativ, kommutativ, besitzt eine Eins und Inverse für alle von Null verschiedenen Elemente. Zudem ist die Multiplikation über die Addition distributiv. Ferner lässt sich zu jeder komplexen Zahl ein konjugiertes Element definieren

$$\overline{(x, y)} = (x, -y)$$

und die reellen Zahlen $\mathbb{R}$ lassen sich als $x \mapsto (x, 0)$ in $\mathbb{C}$ einbetten.

Die Frage stellt sich nun ob und wie man diese Kontruktion verallgemeinern kann um zum Beispiel die komplexen Zahlen zu 'verdoppeln'.

Konstruktion 5.1 (Cayley–Dickson)**.** *Die Quaternionen $\mathbb{H}$ bestehen aus allen Paaren (a, b) von komplexen Zahlen $\mathbb{C} \times \mathbb{C}$ mit komponenterweise Addition (und Subtraktion) und Multiplikation gegeben durch*

$$(a, b) \cdot (c, d) = (ac - \overline{d}b, da + b\overline{c}).$$

Ferner besitzten die Quaternion auch eine Konjugation, welche durch

$$\overline{(a, b)} = (\overline{a}, -b)$$

gegeben ist.

Zuerst bemerken wir, dass die Quaternionen eine Erweiterung der komplexen Zahlen sind, denn man kann sie wiefolgt einbetten:

$$\begin{aligned} \mathbb{C} &\to \mathbb{H} \\ z &\mapsto (z, 0). \end{aligned}$$

Man sieht leicht, dass die Operationen auf beiden Seiten die gleichen sind. Ensprechend können wir die komplexen Zahlen $\mathbb{C}$ und ferner die reellen Zahlen $\mathbb{R}$ als untermengen der

Quaternionen $\mathbb{H}$ auffassen. Eine Quaternion $(a,b) \in \mathbb{H}$ heisst *reell* genau dann falls $a \in \mathbb{R}$ und $b = 0$. Dies ist genau dann der Fall falls $(a,b) = \overline{(a,b)}$.

Wir besprechen nun die weiteren Eigenschaften der Quaternionen.

Addition: Da die Addition komponentenweise definiert ist, über nehmen die Quaternionen alle Eigenschaften der Addition der komplexen Zahlen. Das heisst sie ist assoziativ, kommutativ, besitzt eine Null (=$(0,0)$) und additive Inverse ($-(a,b) = (-a,-b)$).

Multiplikation: Man überprüft leicht, dass die Multiplikation immernoch links- und rechts-distributiv ist, d.h.

$$(a,b)\cdot((c_1,d_1)+(c_2,d_2)) = (a,b)\cdot(c_1,d_1)+(a,b)\cdot(c_2,d_2)$$

und

$$((a_1,b_1)+(a_2,b_2))\cdot(c,d) = (a_1,b_1)\cdot(c,d)+(a_2,b_2)\cdot(c,d).$$

Ferner ist die Multiplikation assoziativ

$$((a,b)\cdot(c,d))\cdot(e,f) = (a,b)\cdot((c,d)\cdot(e,f)),$$

aber nicht mehr kommutativ, da zum Beispiel

$$(i,0)\cdot(0,1) = (0,i) \neq (0,-i) = (0,1)\cdot(i,0).$$

Wir sehen jedoch, dass die reellen Zahlen $\mathbb{R} \subset \mathbb{C} \subset \mathbb{H}$, d.h. $(x,0) \in \mathbb{H}$ mit $x \in \mathbb{R}$ mit allen Quaternionen kommutieren:

$$\begin{aligned}(x,0)\cdot(a,b) &= (xa,bx),\\ (a,b)\cdot(x,0) &= (ax,b\overline{x}) = (xa,bx).\end{aligned}$$

Die Umkehrung ist auch wahr, d.h. falls ein Quaternion mit allen anderen Quaternionen (multiplikativ) kommutiert, so ist sie reell. Dies sieht man in dem man folgende Kommutatoren berechnet:

$$\begin{aligned}(a,b)\cdot(i,0)-(i,0)\cdot(a,b) &= (0,-2bi),\\ (a,b)\cdot(0,1)-(0,1)\cdot(a,b) &= (-b+\overline{b},a-\overline{a}).\end{aligned}$$

Aus der ersten Gleichung folgt $b = 0$ und aus der zweiten $a \in \mathbb{R}$.

Konjugation: Die Konjugation ist eine Involution, da

$$\overline{\overline{(a,b)}} = \overline{(\overline{a},-b)} = (\overline{\overline{a}},b) = (a,b),$$

ist vertauschbar mit der Addition

$$\overline{(a,b)+(c,d)} = (\overline{a+c},-(b+d)) = (\overline{a}+\overline{c},-b-d) = \overline{(a,b)}+\overline{(c,d)},$$

und invertiert die Reihenfolge der Multiplikation

$$\begin{aligned}\overline{(a,b)\cdot(c,d)} &= \overline{(ac-\overline{d}b,da+b\overline{c})} = (\overline{ac}-\overline{\overline{d}b},-da-b\overline{c})\\ &= (\overline{c}\,\overline{a}-\overline{b}d,-b\overline{c}-da) = (\overline{c},-d)\cdot(\overline{a},-b) = \overline{(c,d)}\cdot\overline{(a,b)}.\end{aligned}$$

Spur: Die (reduzierte) Spur einer Quaternion (a,b) ist definiert als

$$\mathrm{tr}((a,b)) = (a,b)+\overline{(a,b)} = (a+\overline{a},0)$$

und stellt eine reelle Zahl dar. Die halbe Spur $\Re((a,b)) = \frac{1}{2}\operatorname{tr}((a,b)) = (\operatorname{Re}(a), 0)$ wird als *Realteil* bezeichnet und der *Imaginärteil* ist gegeben durch $\Im((a,b)) = \frac{1}{2}((a,b) - \overline{(a,b)}) = (\operatorname{Im}(a)i, b)$, sodass

$$(a,b) = \Re((a,b)) + \Im((a,b))$$

gilt.

Norm: Die (reduzierte) Norm einer Quaternion (a,b) ist definiert als

$$N((a,b)) = (a,b) \cdot \overline{(a,b)} = (a\overline{a} + \overline{b}b, 0) = \overline{(a,b)} \cdot (a,b).$$

Auch sie stellt eine nicht-negative reelle Zahl dar, welche genau dann Null ist wenn $(a,b) = (0,0)$. Ferner ist die Norm multiplikativ, da

$$\begin{aligned} N((a,b)\cdot(c,d)) &= (a,b)\cdot(c,d)\cdot\overline{(a,b)\cdot(c,d)} = (a,b)\cdot(c,d)\cdot\overline{(c,d)}\cdot\overline{(a,b)} \\ &= (a,b)\cdot N((c,d))\cdot\overline{(a,b)} = N((a,b))\cdot N((c,d)), \end{aligned} \tag{5.2}$$

wo wir in der letzten Linie benutzt haben, dass die Norm reell ist und daher mit beliebigen Quaternionen kommutiert. Aus jenen Eigenschaften der Norm folgt auch gleich, dass falls das Produkt zweier Quaternionen Null ist, so muss zwingend einer der Faktoren Null sein.

Multiplikative Inverse: Da die reellen Zahlen $\mathbb{R} \subset \mathbb{C} \subset \mathbb{H}$, d.h. $(x,0) \in \mathbb{H}$ mit $x \in \mathbb{R}$ mit allen Quaternionen kommutieren können wir für jede von Null verschiedene Quaternion $(a,b) \neq (0,0)$ ein multiplikatives Inverses definieren. Es ist gegeben durch

$$(a,b)^{-1} = N((a,b))^{-1}\overline{(a,b)},$$

wobei das Inverse $N((a,b))^{-1}$ gegeben ist durch $(r^{-1}, 0)$, falls $N((a,b)) = (r,0)$ mit $r \in \mathbb{R}^+$.

Bemerkung. Die Quaternionen sind ein Beispiel eines *Schiefkörpers* (ein Körper bei welchem die Multiplikation nicht zwingend kommutativ sein muss).

Wie bei den komplexen Zahlen hätten wir gerne eine vereinfachte Schreibweise. Dazu stellen wir fest, dass die Quaternionen einen vier-dimensionalen Vektorraum über $\mathbb{R}$ darstellen. Wir schreiben kurz für die folgende $\mathbb{R}$-Basis

$$1 = (1,0), \quad i = (i,0), \quad j = (0,1), \quad k = (0,i).$$

So haben wir $\mathbb{R} = \mathbb{R}\cdot 1$ und $\mathbb{C} = \mathbb{R}\cdot 1 + \mathbb{R}\cdot i$ wie gewohnt. Wir bemerken nun, dass für eine reelle Zahl $r \in \mathbb{R}$, der Ausdruck $r \cdot j$ im $\mathbb{R}$-Vektorraum $\mathbb{H}$ gerade $(0,r) = (r,0)\cdot(0,1)$ ist. Das heisst, wenn wir $(r,0) \in \mathbb{H}$ mit der reellen Zahl r identifizieren, so macht $r \cdot j$ auch als Quaternion multiplikation Sinn und stimmt überein mit der Vektorraum Definition. Analoge Aussagen gelten auch für die anderen Basiselemente. Wir können also wie gewohnt mit dem Distributivgesetz das Produkt solcher Ausdrücke ausmultiplizieren. Ferner erinnern wir uns, dass die reellen Quaternionen mit allen anderen Quaternionen kommutieren. Das heisst wir können die reellen Zahlen nach vorne ziehen und müssen nur noch die Basiselemente miteinander multiplizieren. Wir berechnen die folgende Multiplikationstabelle.

$\cdot$	1	i	j	k
1	1	i	j	k
i	i	-1	k	$-j$
j	j	$-k$	-1	i
k	k	j	$-i$	-1

Diese Tabelle kann man sich leicht merken indem man sich merkt, dass das Produkt zweier verschiedener 'Buchstaben', i.e. i, j, k, $\pm$ den dritten 'Buchstaben' ergibt und das Vorzeichen ist $+$ falls die alphabetische Reihenfolge korrekt ist, d.h. in $ijki$ vorkommt, ansonsten ist das Vorzeichen -1. Ferner ist das Quadrat eines 'Buchstaben' immer -1.

Beispiel. *Man berechnet mit Hilfe des Distributivgesetzes*

$$\begin{aligned}(1+2j+3k)\cdot(i-j) &= i+2ji+3ki-j-2j^2-3kj\\ &= i-2k+3j-j+2+3i\\ &= 2+4i+2j-2k.\end{aligned}$$

Als Alternative kann man auch mit der $\mathbb{C}$-basis $1=(1,0)$ und $j=(0,1)$ rechnen. Dazu muss man einfach noch bemerken, dass für $\alpha\in\mathbb{C}$ gilt $j\alpha=\overline{\alpha}j$, $j^2=-1$, und $\overline{j}=-j$. Für $\alpha,\beta\in\mathbb{C}$ ist dann $\overline{\alpha+\beta j}=\overline{\alpha}+\overline{j}\,\overline{\beta}=\overline{\alpha}-j\overline{\beta}=\overline{\alpha}-\beta j$ und

$$\begin{aligned}N(\alpha+\beta j) &= (\alpha+\beta j)\overline{(\alpha+\beta j)} = (\alpha+\beta j)(\overline{\alpha}-\beta j)\\ &= \alpha\overline{\alpha}-\alpha\beta j+\beta j\overline{\alpha}-\beta j\beta j\\ &= \alpha\overline{\alpha}-\alpha\beta j+\beta\alpha j-\beta\overline{\beta}j^2\\ &= |\alpha|^2_{\mathbb{C}}+|\beta|^2_{\mathbb{C}}.\end{aligned}$$

Ferner kann man dies auch nutzen um die Quaternionen $\mathbb{H}$ in die 2×2 komplexen Matrizen einzubetten:

$$\begin{aligned}\mathbb{H} &\to \mathrm{Mat}_{2\times 2}(\mathbb{C}),\\ (\alpha,\beta) &\mapsto \begin{pmatrix}\alpha & \beta\\ -\overline{\beta} & \overline{\alpha}\end{pmatrix},\end{aligned}$$

wobei die reduzierte Norm auf der rechten Seite durch die Determinante und die konjugierte Quaternion, $\overline{(\alpha,\beta)}=(\overline{\alpha},-\beta)$, durch die hermitische Transposition gegeben ist.

Anwendung. *Es gilt die Identität*

$$N(a+bi+cj+dk)=a^2+b^2+c^2+d^2$$

und ferner ist die Multiplikativität der Norm (5.2) *äquivalent zur Euler'schen Identität* (5.1). *Entprechend sind die ganzen Zahlen, welche sich als Summe von vier Quadraten schreiben lassen genau diejenigen Zahlen, welche Normen von ganzen Quaternionen* $\mathbb{Z}[1,i,j,k]$ *sind. Wir werden dies uns im nächsten Abschnitt 5.2 zu nutzen machen.*

Anwendung. *Die Quaternionen finden Anwedung in der Computergrafik. Sie bilden die (zur Zeit) schnellste und numerisch stabilste Art Rotationen im Raum zu berechnen und zu verknüpfen. Dazu identifiziert man* S^2 *mit* $\mathbb{H}^{0,1}=\{ai+bj+ck\in\mathbb{H}|a^2+b^2+c^2=1\}$, *dann induziert die (links) Aktion*

$$\begin{aligned}\mathbb{H}^\times\times\mathbb{H}^{0,1} &\to \mathbb{H}^{0,1}\\ (\alpha,\beta) &\mapsto \alpha\beta\alpha^{-1}\end{aligned}$$

einen Isomorphismus

$$\mathbb{H}^\times\Big/\mathbb{R}^\times\cong \mathrm{SO}_3(\mathbb{R}).$$

Die Details dazu findet man zum Beispiel in [13, §2.3].

5.2 Lagrange'scher vier Quadrate Satz und die Hurwitz Quaternionen

Die Arithmetik der Lipschitz Quaternionen findet man in [6] und die einfachere Arithmetik der Hurwitz Quaternionen lässt sich in [7] auffinden.

Das Ziel von diesem Abschnitt ist der vier Quadrate Satz von Lagrange zu beweisen.

Satz 5.2 (Lagrange). *Jede natürliche Zahl n lässt sich als Summe vierer Quadrate ganzer Zahlen schreiben.*

Im Hinblick auf die Identität

$$N(a+bi+cj+dk) = a^2+b^2+c^2+d^2 \tag{5.3}$$

genügt es zu zeigen, dass sich jede natürliche Zahl sich als Norm einer Quaternion mit ganzen Koeffizienten schreiben lässt. Entsprechend würden wir gerne die Arithmetik der 'ganzen' (Lipschitz) Quaternionen $\mathbb{Z}[1,i,j,k]$ studieren. Nun ist es jedoch so, dass die algebraische Struktur von einem leicht grösseren Ring, den Hurwitz Quaternionen, sich besser verhält. Dazu setzen wir

$$\xi = \tfrac{1}{2}(1+i+j+k).$$

Wir sehen, dass

$$\begin{aligned} \overline{\xi} = 1-\xi, \quad i\xi = \xi k = i+k-\xi, \quad j\xi = \xi i = i+j-\xi, \\ k\xi = \xi j = j+k-\xi, \quad \xi^2 = \xi(1-\overline{\xi}) = \xi - N(\xi) = \xi - 1. \end{aligned}$$

Wir folgern, dass der Ring $\mathbb{Z}[i,j,k,\xi]$ additiv, multiplikativ und bezüglich Konjugation abgeschlossen ist. Jener Ring nennt sich die Hurwitz Quaternionen und bemerken

$$\mathbb{Z}[i,j,k,\xi] = \{\tfrac{1}{2}(a+bi+ck+dk) \in \mathbb{H} | \text{ mit } a,b,c,d \in \mathbb{Z} \text{ alle gerade oder alle ungerade}\}$$

und dass die Spur und die Norm all jener Elemente immer ganz ist ($\in \mathbb{Z}$). Wir möchten nun auch eine Theorie der Teilbarkeit für die Hurwitz Quaternionen erarbeiten, jedoch müssen wir dabei viel sorgfältiger sein, da der Ring nicht mehr kommutativ ist. Entsprechend muss man links und rechts teilbar unterscheiden.

Definition. *Seien $\alpha, \beta \in \mathbb{Z}[i,j,k,\xi]$ Hurwitz Quaternionen. Wir sagen, dass β ein* Rechtsteiler *von α ist, falls es ein $\gamma \in \mathbb{Z}[i,j,k,\xi]$ gibt mit*

$$\alpha = \gamma\beta.$$

Analog sagen wir β ist ein Linksteiler *von α, falls es ein $\delta \in \mathbb{Z}[i,j,k,\xi]$ gibt mit*

$$\alpha = \beta\delta.$$

Definition. *Eine Hurwitz Quaternion $\alpha \in \mathbb{Z}[i,j,k,\xi]$ heisst* Einheit*, falls es eine Hurwitz Quaternion $\beta \in \mathbb{Z}[i,j,k,\xi]$ gibt, sodass*

$$\alpha\beta = \beta\alpha = 1$$

gilt.

Lemma 5.3. *Die Einheiten der Hurwitz Quaternionen sind genau diejenigen mit Norm gleich 1 und sind von der Form*

$$\pm 1, \ \pm i, \ \pm j, \ \pm k, \ \pm\tfrac{1}{2} \pm \tfrac{1}{2}i \pm \tfrac{1}{2}j \pm \tfrac{1}{2}k.$$

Beweis. Besitzt ein Hurwitz Quaternion α die Norm 1, so gilt $\alpha\overline{\alpha} = \overline{\alpha}\alpha = 1$. Falls α eine Einheit der Hurwitz Quaternionen ist, so findet man ein $\beta \in \mathbb{Z}[i,j,k,\xi]$, sodass $\alpha\beta = \beta\alpha = 1$ gilt. Nimmt man die Norm so findet man $N(\alpha\beta) = N(\alpha)N(\beta) = 1$. Da nun $N(\alpha), N(\beta) \geqslant 1$ gilt, so folgt $N(\alpha) = N(\beta) = 1$. Durch eine leichte Fallunterscheidung findet man alle Einheiten. □

Lemma 5.4. *Sei $\alpha \in \mathbb{Z}[i,j,k,\xi]$, ein Hurwitz Quaternion, so gibt es immer eine Einheit $\epsilon \in \mathbb{Z}[i,j,k,\xi]$, sodass $\alpha\epsilon \in \mathbb{Z}[1,i,j,k]$.*

Beweis. Falls α schon ein Lipschitz Quaternion ist, dann setzte man $\epsilon = 1$. Ansonsten ist α von der Form

$$\alpha = 2(a + bi + cj + dk) + \tfrac{1}{2}(\epsilon_1 + \epsilon_2 i + \epsilon_3 j + \epsilon_4 k)$$

mit $a, b, c, d \in \mathbb{Z}$ und $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4 \in \{\pm 1\}$. Man setze nun $\epsilon = \frac{1}{2}(\epsilon_1 - \epsilon_2 i - \epsilon_3 k - \epsilon_4 k)$. Dann gilt

$$\begin{aligned}\alpha\epsilon &= (a + bi + cj + dk)(\epsilon_1 - \epsilon_2 i - \epsilon_3 k - \epsilon_4 k) + \overline{\epsilon}\epsilon \\ &= (a + bi + cj + dk)(\epsilon_1 - \epsilon_2 i - \epsilon_3 k - \epsilon_4 k) - 1 \in \mathbb{Z}[1,i,j,k].\end{aligned}$$

□

Es folgt, dass eine natürliche Zahl sich genau dann als Norm eines Lipschitz Quaternionen darstellen lässt, wenn sie sich als Norm eines Hurwitz Quaternion darstellen lässt.

Satz 5.5. *Die Hurwitz Quaternionen $\mathbb{Z}[i,j,k,\xi]$ besitzen eine Rechtsdivision mit kleinerem Rest. Das heisst für alle $\alpha, \beta \in \mathbb{Z}[i,j,k,\xi]$ mit $\beta \neq 0$, existieren $\gamma, \rho \in \mathbb{Z}[i,j,k,\xi]$ mit $N(\rho) < N(\beta)$ und $\alpha = \gamma\beta + \rho$.*

Beweis. Betrachte $\alpha\beta^{-1} \in \mathbb{H}$ und $\gamma \in \mathbb{Z}[i,j,k,\xi]$ diejenige Hurwitz Quaterion, welche $\alpha\beta^{-1}$ am Nächsten ist. Setzte $\delta = \alpha\beta^{-1} - \gamma$. Wir behaupten nun, dass $N(\delta) < 1$. Dazu bemerken wir, dass die nächste Lipschitz quaternion $\widetilde{\gamma}$ höchstens den euklidischen Abstand $\sqrt{(\frac{1}{2})^2 + (\frac{1}{2})^2 + (\frac{1}{2})^2 + (\frac{1}{2})^2} = 1$ hat, d.h. $N(\alpha\beta^{-1} - \widetilde{\gamma}) \leqslant 1$. Dies ist fast was wir benötigen. Nun bemerken wir, dass die Norm genau dann 1 ist, wenn $\alpha\beta^{-1}$ von der Form $\xi + \mathbb{Z}[1,i,j,k]$ ist, aber dann ist $\alpha\beta^{-1}$ eine Hurwitz Quaternion und in diesem Fall gilt $\delta = 0$. Insbesondere haben wir gezeigt, dass $N(\delta) < 1$. Wir finden nun

$$\alpha = \gamma\beta + \delta\beta$$

und setzen $\rho = \delta\beta$. Dann gilt $\rho = \alpha - \gamma\beta \in \mathbb{Z}[i,j,k,\xi]$ und $N(\rho) = N(\delta)N(\beta) < N(\beta)$. □

Wir möchten nun zeigen, dass der euklidische Algorithmus immernoch funktioniert und folgern, dass grösste gemeinsame (Links-/Rechts-) Teiler existieren und auch wieder einen Satz von Bézout haben. Zuerst aber müssen wir unsere Definition von grössten gemeinsamen Teiler anpassen.

Definition. *Eine Hurwitz Quaternion $\gamma \in \mathbb{Z}[i,j,k,\xi]$ heisst* grösster gemeinsamer Rechtsteiler *zweier Hurwitz Quaternionen $\alpha, \beta \in \mathbb{Z}[i,j,k,\xi]$ falls gilt:*

1. *γ ist ein Rechtsteiler von α und β,*
2. *Für jeden gemeinsamen Rechtsteiler δ von α, β ist δ ein Rechtsteiler von γ.*

Bemerkung. Falls ein grösster gemeinsamer Rechtsteiler existiert, so ist jener bis auf Linksmultiplikation mit einer Einheit eindeutig bestimmt.

Satz 5.6. *Ein grösster gemeinsame Rechtsteiler γ zweier Hurwitz Quaternionen $\alpha, \beta \in \mathbb{Z}[i,j,k,\xi]$ existiert immer. Ferner, gibt es zwei Hurwitz Quaternionen τ, σ, sodass*

$$\gamma = \tau\alpha + \sigma\beta.$$

Beweis. Wir benutzen den euklidischen Algorithmus. Ohne Beeinschränkung der Allgemeinheit nehme $N(\alpha) \geqslant N(\beta) \geqslant 0$. Setze $\eta_0 = \alpha$, $\eta_1 = \beta$. Wir bestimmen dann sukzessiv $\delta_{i-1}, \eta_i \in \mathbb{Z}[i,j,k,\xi]$ mit Hilfe der Rechtsdivision mit kleinerem Rest solange $\eta_{i-1} \neq 0$, sodass

$$\eta_{i-2} = \delta_{i-1}\eta_{i-1} + \eta_i \tag{5.4}$$

mit $N(\eta_i) < N(\eta_{i-1})$. Da die Norm eine nicht negative ganze Zahl darstellt und immer kleiner wird, so muss der Algorithmus schlussendlich terminieren und wir erhalten ein $m \in \mathbb{N}$, sodass $\eta_m = 0$. Die Behauptung ist nun, dass η_{m-1} der grösste gemeinsame Rechtsteiler von $\eta_0 = \alpha$ und $\eta_1 = \beta$. Aufgrunde der letzten Gleichung

$$\eta_{m-2} = \delta_{m-1}\eta_{m-1} + \eta_m = \delta_{m-1}\eta_{m-1}$$

sehen wir, dass η_{m-1} ein Rechtsteiler von η_{m-2} ist. Mit Hilfe der Gleichung (5.4) folgern wir induktiv, dass η_{m-1} ein Rechtsteiler von $\eta_{m-2}, \eta_{m-3}, \ldots, \eta_1, \eta_0$. Umgekehrt, falls γ ein gemeinsamer Rechtsteiler von η_0, η_1 ist, so folgt induktiv mit der Gleichung (5.4), dass γ ein Rechtsteiler von $\eta_2, \eta_3, \ldots, \eta_{m-1}$ ist. Es folgt, dass η_{m-1} ein grösster gemeinsamer Rechtsteiler von α und β ist.

Durch Rücksubtitution der Gleichungen (5.4) findet man eine Darstellung

$$\eta_{m-1} = \tau\eta_0 + \sigma\eta_1 = \tau\alpha + \sigma\beta$$

mit $\tau, \sigma \in \mathbb{Z}[i,j,k,\xi]$. □

Wir brauchen noch zwei weitere Lemmata bevor wir den vier Quadrate Satz von Lagrange beweisen können.

Lemma 5.7. *Sei $p \in \mathbb{Z}$ eine Primzahl. Es existieren zwei ganze Zahlen $x, y \in \mathbb{Z}$, sodass*

$$1 + x^2 + y^2 \equiv 0 \bmod(p).$$

Beweis. Für $p = 2$ setze man $x = 1$, $y = 0$. Sei nun p ungerade. Die Menge der Quadratzahlen belegen modulo p genau $\frac{p+1}{2}$ Kongruenzklassen, denn

$$a^2 \equiv b^2 \bmod(p) \Leftrightarrow p \mid (a-b)(a+b) \Leftrightarrow p \mid a-b \text{ oder } p \mid a+b$$

und die Äquivalenzrelation $a \sim b \Leftrightarrow a \equiv \pm b \bmod(p)$ auf $\mathbb{Z}$ induziert die Äquivalenzrelation $[a]_p \sim [b]_p \Leftrightarrow [a]_p = \pm[b]_p$ auf $\mathbb{Z}/p\mathbb{Z}$. Eine Äquivalenzklasse enthält maximal zwei Elemente (Kongruenzklassen), nämlich $[a]_p$ und $-[a]_p$ für ein $[a]_p \in \mathbb{Z}/p\mathbb{Z}$. Jene Elemente (Kongruenzklassen) sind genau dann verschieden, falls $2[a]_p = [2a]_p \neq [0]_p$. Da p ungerade ist und somit teilerfremd zu 2 ist, ist dies genau dann der Fall falls $[a]_p = [0]_p$. Die Äquivalenzrelation $\sim$ teilt also $\mathbb{Z}/p\mathbb{Z}$ in $1 + \frac{p-1}{2} = \frac{p+1}{2}$ Äquivalenzklassen. Es folgt, dass

$$\begin{aligned}
&\left|\{[1]_p + [a]_p^2 \in \mathbb{Z}/p\mathbb{Z} \,|\, [a]_p \in \mathbb{Z}/p\mathbb{Z}\} \cap \{-[b]_p^2 \in \mathbb{Z}/p\mathbb{Z} \,|\, [b]_p \in \mathbb{Z}/p\mathbb{Z}\}\right| \\
&\quad \geqslant \left|\{[1]_p + [a]_p^2 \in \mathbb{Z}/p\mathbb{Z} \,|\, [a]_p \in \mathbb{Z}/p\mathbb{Z}\}\right| + \left|\{-[b]_p^2 \in \mathbb{Z}/p\mathbb{Z} \,|\, [b]_p \in \mathbb{Z}/p\mathbb{Z}\}\right| - |\mathbb{Z}/p\mathbb{Z}| \\
&\quad = \frac{p+1}{2} + \frac{p+1}{2} - p = 1.
\end{aligned}$$

D.h. es gibt Kongruenzklassen $[a]_p, [b]_p \in \mathbb{Z}/p\mathbb{Z}$, sodass

$$[1]_p + [a]_p^2 = -[b]_p^2 \Leftrightarrow 1 + a^2 + b^2 \equiv 0 \,\mathrm{mod}(p)$$

gilt. Nehme nun zwei beliebige Zahlen $x \in [a]_p$ und $y \in [b]_p$ in jenen Kongruenzklassen. □

Lemma 5.8. *Sei $\alpha \in \mathbb{Z}[i,j,k,\xi]$ eine Hurwitz Quaternion und $b \in \mathbb{Z} \subset \mathbb{Z}[i,j,k,\xi]$ eine reelle Hurwitz Quaternion. Dann sind α und b genau dann rechtsteilerfremd in $\mathbb{Z}[i,j,k,\xi]$ wenn $N(\alpha)$ und $N(b) = b^2$ teilerfremd in $\mathbb{Z}$ sind.*

Beweis. Seien α und b rechtsteilerfremd, dann gibt es nach dem Satz 5.6 zwei Hurwitz Quaternionen $\tau, \sigma \in \mathbb{Z}[i,j,k,\xi]$, sodass $1 = \tau\alpha + \sigma b$. Dann gilt

$$\begin{aligned} N(\tau)N(\alpha) = N(\tau\alpha) = N(1-\sigma b) = (1-\sigma b)\overline{(1-\sigma b)} = (1-\sigma b)(1-b\overline{\sigma}) \\ = 1 - \sigma b - b\overline{\sigma} + \sigma b^2 \overline{\sigma} = 1 - \mathrm{tr}(\sigma)b + N(\sigma)b^2, \end{aligned}$$

wobei wir benutzt haben, dass $b\overline{\sigma} = \overline{\sigma}b$, da b reell ist. Dies ist nun eine Gleichung in ganzen Zahlen. Falls nun p eine Primzahl ist, welche $N(\alpha)$ und $N(b) = b^2$ teilt, so folgt zuerst, dass $p \mid b$, und dann, dass $p \mid 1$ anhand der vorherigen Gleichung. Dies ist ein Widerspruch.

Falls α und b einen gemeinsamen Rechtsteiler δ haben, welcher keine Einheit ist. So können wir $\alpha = \alpha'\delta$ und $b = \beta'\delta$ schreiben. Es folgt durch die Multiplikativität der Norm, dass $N(\delta)$ ein Teiler von $N(\alpha)$ und von $N(b)$ ist. Ferner ist $N(\delta) \neq \pm 1$, da δ keine Einheit ist, und somit ist $N(\delta)$ keine Einheit in $\mathbb{Z}$. □

Wir sind nun soweit und können den vier Quadrate Satz beweisen.

Beweis von 5.2. Wir müssen zeigen, dass sich jede Primzahl $p \in \mathbb{N}$ sich als Norm einer Hurwitz Quaternion schreiben lässt. Für $p = 2$ haben wir $N(1+i) = 2$. Sei nun also p ungerade. Nach Lemma 5.7 können wir zwei ganze Zahlen $x, y \in \mathbb{Z}$ finden, sodass $p \mid 1+x^2+y^2$. Wir setzen $\alpha = 1+xi+yj$ und betrachten den grössten gemeinsamen Rechtsteiler δ von α und p. Wir haben $p \mid p^2 = N(p)$ und $p \mid N(\alpha)$. Da nun p reell ist, können wir Lemma 5.8 anwenden. Es folgt, dass α und p nicht rechtteilerfremd sind, d.h. δ ist keine Einheit. Wir schreiben $p = \pi\delta$. Es folgt $p^2 = N(p) = N(\pi)N(\delta)$. Nun ist $N(\delta) \neq 1$ und es gilt den Fall $N(\delta) = p^2$ auszuschliessen. Im letzteren Fall ist aber $N(\pi) = 1$ und somit ist π eine Einheit. Es folgt, dass $\pi^{-1}p = \delta$ ein Rechtsteiler von α ist und somit auch p selbst. Nun hat aber jede Huwitz Quaternion welches ein Vielfaches von p ist jede Koordninate durch p teilbar (d.h. der Zähler in reduzierter Schreibweise ist durch p teilbar). Es muss also $N(\delta) = p$ gelten und damit ist der Beweis abgeschlossen. □

Zum Schluss können wir uns auch noch fragen ob es so etwas wie eine eindeutige Primfaktorisierung in den Hurwitz Quaternionen gibt. Die Antwort ist nicht ganz so simpel, denn es nicht mal klar wie der Begriff prim verallgemeinert werden kann. Andererseits macht der Begriff irreduzibel immernoch Sinn.

Definition. *Eine nicht nulle nicht Einheit Hurwitz Quaternion α heisst irreduzibel, falls für jede Faktorasition $\alpha = \beta\gamma$ in Hurwitz Quaternionen gilt, dass entweder β oder γ eine Einheit ist.*

Lemma 5.9. *Eine Hurwitz Quaternion α ist genau dann irreduzibel wenn $N(\alpha)$ irreduzibel (prim) in $\mathbb{Z}$ ist.*

Beweis. Wir zeigen die schwierigere Richtung. Sei also α irreduzibel und p eine Primzahl, welche $N(\alpha)$ teilt. Es folgt nach Lemma 5.8, dass α und p einen gemeinsamen Rechtsteiler δ haben, welcher keine Einheit ist. Da α irreduzibel ist gilt $\alpha = \epsilon\delta$ für eine Einheit ϵ und es folgt, dass α ein Rechtsteiler von p ist. Aus $p = \alpha'\alpha$ folgt $p^2 = N(p) = N(\alpha')N(\alpha)$, da α keine Einheit ist, ist $N(\alpha) \neq 1$. Falls $N(\alpha') = 1$, so ist α' eine Einheit und es würde folgen, dass $p = \alpha'\alpha$ auch eine irreduzibles Hurwitz Quaternion wäre. Dies ist im Widerspruch zum vorherigen Beweis, wo wir zeigten, dass p einen Rechtsteiler der Norm p besitzt also insbesondere reduzibel ist. Es folgt also $N(\alpha') = N(\alpha) = p$ prim, was zu zeigen war. □

Wir können nun sicherlich eine Hurwitz Quaternion α als ein Produkt von irreduziblen Hurwitz Quaternionen β_i schreiben.

$$\alpha = \beta_1\beta_2\cdots\beta_r$$

Wir können sogar mehr verlangen. Lemma 5.8 erlaubt uns sogar die irreduziblen Elemente β_i so zu wählen, sodass $N(\beta_i) = p_i$ eine beliebig gewählte geordnete Primfaktorisation $N(\alpha) = p_1p_2\cdots p_r$. Zum Beispiel ist $N(-3-i+3j+4k) = 35 = 5\cdot 7 = 7\cdot 5$ und wir haben

$$-3-i+3j+4k = (1+2k)(1+i+j+2k) = (-1+i-j+2k)(1+2i).$$

Dies führt klarerweise zu einem Problem mit einer Eindeutigkeit der Faktorisierung. Weitere Probleme sind, dass man zwischen zwei beliebigen Faktoren immer eine Einheit und ihr Inverses einschieben kann. Aber auch das ist noch nicht genug, denn, zum Beispiel besitzt 5 wirklich verschiedene Faktorisationen in irreduzible Elemente

$$5 = (1+2i)(1-2i) = (1+2j)(1-2j) = (1+2k)(1-2k).$$

Und alle irreduzible Elemente in der vorherigen Gleichungen unterscheiden sich paarweise nicht durch einer Multiplikation mit einer Einheit (vergleiche Aufgabe 5.2).

Aufgabe 5.1. *Zeige, dass im Generellen die Lipschitz Quaternionen $\mathbb{Z}[1,i,j,k]$ kein grössten gemeinsamen Rechtsteiler besitzen. Betrachte dazu alle Rechtsteiler von 2 und $1+i+j+k$.*

Aufgabe 5.2. *Seien $\alpha,\beta \in \mathbb{Z}[i,j,k,\xi]$ zwei Hurwitz Quaternionen, sodass $N(\alpha)$ und $N(\beta)$ nicht teilerfremd in $\mathbb{Z}$ sind. Besitzt dann α,β notwendigerweise einen gemeinsamen Rechtsteiler, welcher keine Einheit ist?*

5.3 Octonionen und der Satz von Hurwitz

Der folgende Abschnitt ist von [10] entnommen.

Der Verdoppelungsschritt von Cayley–Dickson kann man beliebig oft anwenden, jedoch verliert man in jedem Schritt diverse algebraische Gesetze. Im ersten Schritt $\mathbb{R} \to \mathbb{C}$ verloren wir dass alle Zahlen 'reell' sind. Im zweiten Schritt $\mathbb{C} \to \mathbb{H}$ verloren wir die Kommutivität der Multiplikation. Im dritten Schritt $\mathbb{H} \to \mathbb{O}$ von den Quaternionen zu den Octonionen (auch bekannt als Cayley Zahlen) verlieren wir sogar die Assoziativität der Multiplikation! Erstaunlicherweise, bleibt aber die Multiplikativität der Norm vorhanden und die führt zu der grandiosen Identität

$$(x_1^2+\cdots+x_8^2)(y_1^2+\cdots+y_8^2) = z_1^2+\cdots+z_8^2, \tag{5.5}$$

wobei

$$
\begin{aligned}
z_1 &= x_1y_1 - x_2y_2 - x_3y_3 - x_4y_4 - x_5y_5 - x_6y_6 - x_7y_7 - x_8y_8,\\
z_2 &= x_1y_2 + x_2y_1 + x_3y_4 - x_4y_3 + x_5y_6 - x_6y_5 - x_7y_8 + x_8y_7,\\
z_3 &= x_1y_3 + x_3y_1 - x_2y_4 + x_4y_2 + x_5y_7 - x_7y_5 + x_6y_8 - x_8y_6,\\
z_4 &= x_1y_4 + x_4y_1 + x_2y_3 - x_3y_2 + x_5y_8 - x_8y_5 - x_6y_7 + x_7y_6,\\
z_5 &= x_1y_5 + x_5y_1 - x_2y_6 + x_6y_2 - x_3y_7 + x_7y_3 - x_4y_8 + x_8y_4,\\
z_6 &= x_1y_6 + x_6y_1 + x_2y_5 - x_5y_2 - x_3y_8 + x_8y_3 + x_4y_7 - x_7y_4,\\
z_7 &= x_1y_7 + x_7y_1 + x_2y_8 - x_8y_2 + x_3y_5 - x_5y_3 - x_4y_6 + x_6y_4,\\
z_8 &= x_1y_8 + x_8y_1 - x_2y_7 + x_7y_2 + x_3y_6 - x_6y_3 + x_4y_5 - x_5y_4.
\end{aligned}
$$

welche bereits zuvor von C. F. Degan im Jahre 1822[g] entdeckt wurde. Man kann sich nun Fragen, ob weitere Verdoppelungen z.B. von den Octonions zu den Sedonions weitere solche Identitäten liefert und die Antwort ist leider nein. Die Octionen ist das letzte Zahlensystem in dieser Folgen von Verdoppelungen für welche die Multiplikativität der Norm gilt. Dies werden wir nun beweisen.

Satz 5.10. *Die natürlichen Zahlen $n = 1, 2, 4, 8$ sind die einzigen natürlichen Zahlen für welche eine Identität der Form*

$$(x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2) = z_1^2 + \cdots + z_n^2 \tag{5.6}$$

mit z_k eine bilineare Funktion in den Variablen x_i, y_j über $\mathbb{R}$ gilt.

Beweis. Wir geben hier den Beweis von Dickson [5] wieder. Wir haben gesehen, dass solche Identitäten existieren für $n = 1, 2, 4, 8$. Entsprechend nehmen wir an, dass n nicht eine jener Zahlen ist und eine solche Identität existiert. Wir können schreiben

$$\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = A\boldsymbol{y}$$

mit Matrixeinträgen a_{ij} welche lineare reelle Polynome in den Variablen x_i darstellen. Die Gleichung (5.6) ist nun äquivalent zu

$${}^t\boldsymbol{y}\left({}^tAA - (x_1^2 + \cdots + x_n^2)I_{n\times n}\right)\boldsymbol{y} = 0.$$

Da dies wahr ist für alle y_j folgern wir[h], dass

$${}^tAA = (x_1^2 + \cdots + x_n^2)I_{n\times n} \tag{5.7}$$

gilt. Schreiben wir weiter

$$A = A_1x_1 + \cdots + A_nx_n$$

mit $A_i \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, so finden wir, dass (5.7) äquivalent ist zum Gleichungssystem

1. ${}^tA_iA_i = I_{n\times n}$, für alle $1 \leqslant i \leqslant n$,

[g] Cayley entdeckte die Octonions im Jahre 1845 [2].

[h] Man bemerke, dass $\mathrm{char}(\mathbb{R}) \neq 2$.

2. ${}^tA_iA_j + {}^tA_jA_i = 0_{n\times n}$, für alle $1 \leqslant i < j \leqslant n$.

Führen wir nun die Matrizen $B_i = {}^tA_nA_i$ für $i = 1, \ldots, n-1$ ein. So sehen wir, dass jene Matrizen die folgenden Relationen erfüllen

$$
\begin{aligned}
{}^tB_iB_i &= I_{n\times n}, \text{ für } 1 \leqslant i < n, && (5.8)\\
{}^tB_i &= -B_i, \quad \text{für } 1 \leqslant i < n, && (5.9)\\
{}^tB_iB_j + {}^tB_jB_i &= 0_{n\times n}, \text{ für } 1 \leqslant i < j < n. && (5.10)
\end{aligned}
$$

Die ersten zwei Eigenschaft implizieren nun, dass $\det(B_i) \in \{\pm 1\}$ und ferner

$$\det(B_i) = \det({}^tB_i) = \det(-B_i) = (-1)^n \det(B_i).$$

Es folgt, dass n gerade ist.

Wir betrachten nun Matrizen der Form

$$B_{i_1}B_{i_2}\cdots B_{i_r} \text{ mit } 1 \leqslant i_1 < i_2 < \cdots < i_r < n \tag{5.11}$$

für ein $0 \leqslant r < n$ (für $r = 0$ ist das Produkt die Identitätsmatrix).

Lemma 5.11. *Eine Matrix M von der Form* (5.11) *ist symmetrisch falls $r \equiv 0, 3 \bmod(4)$ und schief-symmetrisch falls $r \equiv 1, 2 \bmod(4)$.*

Beweis. Für $r = 0$ ist dies klar und für $r = 1$ ist die Aussage genau (5.9). Kombinieren wir (5.9) und (5.10), so sehen wir, dass die B_i's antikommutieren. Wir folgern, dass

$$
\begin{aligned}
{}^tM = {}^t(B_{i_1}B_{i_2}\cdots B_{i_r}) &= {}^tB_{i_r}\,{}^tB_{i_{r-1}}\cdots{}^tB_{i_1}\\
&= (-1)^r B_{i_r}B_{i_{r-1}}\cdots B_{i_1}\\
&= (-1)^r(-1)^{\frac{r(r-1)}{2}} B_{i_1}B_{i_2}\cdots B_{i_r}
\end{aligned}
$$

nach $(-1)^{\frac{r(r-1)}{2}}$ Vertauschungen. □

Ferner sehen wir, dass die drei Eigenschaften (5.8), (5.9) und (5.10) implizieren, dass Matrizen der Form (5.11) bis auf Vorzeichen multiplikativ abgeschlossen sind (durch Vertauschen und Kürzen gleicher Matrizen). Da jede Matrix jener Form invertierbar ist gilt haben wir sogar folgendes Lemma.

Lemma 5.12. *Die Menge $\mathcal{G}$ aller Matrizen der Form* (5.11) *stellt bis auf Vorzeichen eine Gruppe dar.*

Wir möchten nun zeigen, dass die Hälfte der Matrizen der Form (5.11) linear unabhängig sind. Als Erstes werfen wir alle Matrizen mit negativem Vorzeichen weg, da jene linear abhängig sind zum positiven Vorzeichen. Als Zweites führen wir temporär die Notation einer irreduziblen (linearen) Relation ein. Für eine Menge $\mathcal{M} \subseteq \mathcal{G}$ von Matrizen M der Form (5.11) heisst eine lineare Relation

$$\sum_{M\in\mathcal{M}} \lambda_M M = 0$$

irreduzibel falls sich die Menge $\mathcal{M}$ nicht in zwei nicht-leere disjunkte Mengen $\mathcal{M}_1, \mathcal{M}_2$ aufteilen lässt, sodass

$$\sum_{M\in\mathcal{M}_1} \lambda_M M = 0 \text{ und } \sum_{M\in\mathcal{M}_2} \lambda_M M = 0$$

gilt. Wir bemerken, dass jede solche lineare Relation als Summe von irreduziblen Relationen schreiben lässt. Ferner sind alle Koeffizienten einer irreduziblen Relation nicht null und es können nur symmetrische oder nur schief-symmetrische Matrizen vorkommen. Dies ist leicht zu sehen, indem man die Gleichung/Relation transponiert. Wir können auch eine Relation mit Matrizen aus $\mathbb{R}^{\times}\mathcal{G}$ multiplizieren und erhalten eine neue Relation, welche auch irreduzibel ist falls die ursprüngliche Relation irreduzibel war. Insbesondere können wir jede Relation auf folgende Form bringen:

$$I_{n\times n} = \sum_{\boldsymbol{i}} \lambda_{\boldsymbol{i}} B_{i_1} B_{i_2} \cdots B_{i_r},$$

wobei nur Terme mit $r \equiv 0, 3 \bmod(4)$ vorkommen, da $I_{n\times n}$ symmetrisch ist. Ferner falls ein Term mit $r \equiv 0 \bmod(4)$ und $r > 0$ vorkommt, so können wir die Gleichung mit B_{i_r} multiplizieren und erhalten wieder eine Relation mit symmetrischen sowie schief-symmetrischen Matrizen. Falls nun ein Term vorkommt mit $r \equiv 3 \bmod(4)$ und $r < n-1$, so können wir ein $1 \leqslant j \leqslant n-1$ finden mit $j \notin \{i_1, \ldots, i_r\}$. Multiplizieren wir nun die Gleichung mit B_j, dann erhalten wir eine Relation mit symmetrischen sowohl als schief-symmetrische Matrizen. Ein Widerspruch! Wir folgenern, dass die einzige nicht leere Relation muss von der Form

$$I_{n\times n} = \lambda B_1 B_2 \cdots B_{n-1}$$

mit $r = n-1 \equiv 3 \bmod(4)$ sein. Insbesondere sind die Matrizen der Form (5.11) linear unabhängig falls $n \equiv 2 \bmod(4)$ und die Matrizen der Form (5.11), welche B_{n-1} nicht enthalten, linear unabhängig falls $n \equiv 0 \bmod(4)$. Insbesondere mindestens 2^{n-2} lineare unabhängige Matrizen. Nun sind aber alle Matrizen der Form (5.11) im Raum der $n \times n$ Matrizen enthalten, dessen Dimension aber höchstens n^2 ist. Es folgt $2^{n-2} \leqslant n^2$ also $n \leqslant 8$. Es gilt also noch den Fall $n = 6$ auszuschliessen. Da $6 \equiv 2 \bmod(4)$ gilt sogar, dass wir 2^5 linear unabhängige Matrizen haben, davon sind 16 symmetrisch und 16 schief-symmetrisch. Der Raum der schief-symmetrischen 6×6 Matrizen hat aber nur Dimension $1 + 2 + \cdots + 5 = 15$. Also auch einen Widerspruch! □

Aufgabe 5.3. *Das Produkt PQ eines Polynomes $P \in \mathbb{Q}[X]$, welches sich als Summe von drei Quadraten von Polynomen in $\mathbb{Q}[X]$ schreiben lässt, mit einem Polynom $Q \in \mathbb{Q}[X]$, welches sich als Summe von fünf Quadraten von Polynomen in $\mathbb{Q}[X]$ schreiben lässt, lässt sich als Summe von 15 Quadraten von Polynomen in $\mathbb{Q}[X]$ schreiben. Zeige, dass PQ sich sogar als Summe von sieben Quadraten von Polynomen in $\mathbb{Q}[X]$ schreiben lässt.*

6 Binäre Quadratische Formen

Referenzen für dieses Kapitel mit weiterführendem Stoff sind zum [8] und [3].

Im Kapitel §3 waren wir interessiert an Zahlen $n \in \mathbb{Z}$, welche sich als eine Summe x^2+y^2 von zwei ganzen Zahlen schreiben lassen. Dazu nutzten wir die multiplikative Struktur aus und für Primzahlen $n = p \in \mathbb{N}$ mit $p \equiv 1 \bmod(4)$ nutzten wir ein unendlicher Abstiegs Argument um Lösungen x, y zu generieren. Die Idee in diesem Kapitel ist diese Frage zu verallgemeinern und auf den Kopf zu stellen. Wir motivieren dies nun.

Die Gleichung $n = x^2 + y^2$ lässt sich wiefolgt ausdrücken:

$$n = {}^t\begin{pmatrix} x \\ y \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Nehmen wir nun an, dass x, y teilerfremd sind. Ansonsten, falls d der grösste gemeinsame Teiler ist, können wir $\frac{n}{d^2}$, $\frac{x}{d}$, und $\frac{y}{d}$ betrachten. Nach dem Satz von Bézout 2.3 existieren

nun ganze Zahlen $a, b \in \mathbb{Z}$, sodass $ax - by = 1$. Wir können nun also $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \left(\begin{smallmatrix} x & b \\ y & a \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ mit $\left(\begin{smallmatrix} x & b \\ y & a \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ schreiben und erhalten

$$n = {}^t\begin{pmatrix}1\\0\end{pmatrix}{}^t\begin{pmatrix}x & b\\ y & a\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}x & b\\ y & a\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = {}^t\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}x^2+y^2 & xb+ya\\ xb+ya & a^2+b^2\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}. \tag{6.1}$$

Das heisst, anstatt die Zahlen x, y zu variieren, können wir die (symmetrische) Matrize im Innern variieren. Dies bringt uns zu folgender Definition.

Definition. *Ein homogenes Polynom von Grad zwei in zwei Variablen heisst eine* binäre quadratische Form. *Zu einer binären quadratischen Form, $Q(X,Y) = aX^2 + bXY + cY^2 \in \mathbb{Z}[X,Y]$ können wir die symmetrische Matrix*

$$A_Q = \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix}$$

assozieren, sodass $Q(X,Y) = {}^t\left(\begin{smallmatrix} X \\ Y \end{smallmatrix}\right) A_Q \left(\begin{smallmatrix} X \\ Y \end{smallmatrix}\right)$. Die Zahl $\Delta_Q = b^2 - 4ac$ heisst die Diskriminante *der binären quadratischen Form Q.*

Von unserer Diskussion werden wir die Nullform $Q(X,Y) = 0$ immer ausschliessen.

Bemerkung. Sei $Q(X,Y) = aX^2 + bXY + cY^2$ eine binäre quadratische Form mit ganzen Koeffizienten und Diskriminante Δ_Q. Falls $\Delta_Q > 0$, so ist Q indefinit, d.h. Q nimmt sowohl positive als auch negative Werte an. Falls $\Delta_Q < 0$, so ist Q definit. Genauers gesagt ist Q positiv definit genau dann wenn $a > 0$. Dies sieht kann man anhand der Gleichungen

$$Q(X,Y) = aX^2 + bXY + cY^2 = \frac{1}{4a}\left[(2aX + bY)^2 - \Delta_Q Y^2\right]$$

für $a \neq 0$ und ähnlich falls $c \neq 0$ sehen.

Sei $\mathcal{A}$ die Menge aller binären quadratischen Formen mit ganzen Koeffizienten, dann haben wir das folgende Lemma.

Lemma 6.1. *Die Gruppe $\mathrm{SL}_2(\mathbb{Z})$ von 2×2 Matrizen mit Determinante 1 und ganzen Einträgen agiert auf $\mathcal{A}$ mittels der Rechtsaktion*

$$\begin{aligned} \rho : \mathcal{A} \times \mathrm{SL}_2(\mathbb{Z}) &\to \mathcal{A} \\ (Q, M) &\mapsto Q((X,Y)\,{}^tM) \end{aligned}$$

Die Rechtsaktion kann auch mittels der assoziierten Matrix A_Q ausgedrückt werden:

$$(A_Q, M) \mapsto {}^tMA_QM.$$

Die Diskriminante stellt eine Invariante dieser Aktion dar.

Beweis. Schreiben wir $M = \left(\begin{smallmatrix} u & v \\ w & z \end{smallmatrix}\right)$, dann ist $(X,Y)\,{}^tM = (uX + vY, wX + zY)$ und wir sehen leicht, dass $Q(uX + vY, wX + zY)$ wieder ein homogenes Polynomon von Grad 2 in den Variablen X, Y mit ganzen Koeffizienten ist. Wir sehen auch, dass

$$A_{\rho(Q,M)} = {}^tMA_Q\,{}^t({}^tM) = {}^tMA_QM.$$

Wir verifizieren auch $\rho(A_Q, \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)) = A_Q$ und

$$\rho(\rho(A_Q, M_2), M_1) = {}^tM_1({}^tM_2A_QM_2)M_1 = {}^t(M_2M_1)A_Q(M_2M_1) = \rho(A_Q, M_2M_1).$$

Damit ist ρ eine Rechtsaktion der Gruppe $\mathrm{SL}_2(\mathbb{Z})$. Ferner ist die Diskriminante nichts anderes als $-4\det(A_Q)$ und die Determinante bleibt erhalten:

$$\det({}^tMA_QM) = \det({}^tM)\det(A_Q)\det(M) = \det(A_Q).$$

□

Definition. *Wir sagen, dass zwei binäre quadratische Formen $Q, Q' \in \mathcal{A}$ mit ganzen Koeffizienten* äquivalent *sind, genau dann wenn Q im* Orbit *von Q' unter der Aktion ρ liegt, d.h. es existiert ein $M \in \mathrm{SL}_2(\mathbb{Z})$, sodass $Q = \rho(Q', M)$ gilt. Dies ist eine Äquivalenzrelation und wir schreiben kurz $Q \sim Q'$.*

Bemerkung. Falls Q und Q' äquivalent sind, so haben sie notwendigerweise dieselbe Diskriminante.

Definition. *Wir sagen, dass eine binäre quadratische Formen $Q \in \mathcal{A}$ mit ganzen Koeffizienten eine ganze Zahl $n \in \mathbb{Z}$* darstellt *bzw. n wird von Q* repräsentiert*, falls ganze Zahlen $x, y \in \mathbb{Z}$ existieren, sodass $n = Q(x, y)$. Ferner sagen wir, dass eine ganze Zahl n* primitiv *von Q dargestellt bzw. repräsentiert wird, falls man sogar x, y teilerfremd wählen kann.*

Die Rechnung (6.1) kündigte schon die nächsten Lemmata an.

Lemma 6.2. *Seien $Q, Q' \in \mathcal{A}$ zwei äquivalente binäre quadratische Formen mit ganzen Koeffizient. Dann wird eine ganze Zahl $n \in \mathbb{Z}$ genau dann von Q (primitiv) repräsentiert wenn sie von Q' repräsentiert wird.*

Beweis. Nehme an $n = Q(x, y)$ für zwei ganze Zahlen $x, y \in \mathbb{Z}$ und $Q' = \rho(Q, M)$ für $M \in \mathrm{SL}_2(\mathbb{Z})$. Setzte nun $(u, v) = (x, y)\,{}^tM^{-1}$, dann ist

$$Q'(u, v) = Q((u, v)\,{}^tM) = Q((x, y)\,{}^tM^{-1}\,{}^tM) = Q(x, y) = n.$$

Es bleibt noch zu zeigen, dass falls x, y teilerfremd sind, dann sind auch u, v teilerfremd. Falls u, v nicht teilerfremd wären, dann ist aus der Identität $(x, y) = (u, v)\,{}^tM$ klar, dass x, y auch nicht teilerfremd sind - ein Widerspruch. □

Es folgt, dass die Rechtsaktion ρ auch erhält ob die Form indefinit, positive definit, oder negativ definit ist.

Lemma 6.3. *Eine Zahl $n \in \mathbb{Z}$ verschieden von 0 wird von einer binären quadratischen Form $Q \in \mathcal{A}$ mit ganzen Koeffizienten genau dann repräsentiert, falls eine natürliche Zahl $d \in \mathbb{N}$ existiert, sodass $d^2 \mid n$ und es existiert eine binäre quadratische Form Q' der Form*

$$Q'(X, Y) = \frac{n}{d^2}X^2 + bXY + cY^2,$$

welche zu Q äquivalent ist. Dieselbe Aussage gilt für primitive Darstellungen, jedoch muss dann $d = 1$ sein.

Beweis. Seien $x, y \in \mathbb{Z}$ ganze Zahlen, sodass $n = Q(x, y)$. Ferner sei d ein grösster gemeinsamer Teiler. Da $(x, y) \neq (0, 0)$ (ansonsten wäre $n = 0$) können wir $d \in \mathbb{N}$ wählen. Nach dem Satz von Bézout finden wir zwei ganze Zahlen $u, v \in \mathbb{Z}$ mit $ux - vy = d$. Setze $M = \begin{pmatrix} x/d & v \\ y/d & u \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ und $Q'(X, Y) = \rho(Q, M)(X, Y) = aX^2 + bXY + cY^2$. Nun gilt

$$a = Q'(1, 0) = Q((1, 0)\,{}^tM) = Q(x/d, y/d) = \frac{1}{d^2}Q(x, y) = \frac{n}{d^2}.$$

Die umgekehrte Richtigung ist einfacher, da $Q'(d, 0) = n$ und $Q' \sim Q$, d.h. nach Lemma 6.3 Q stellt auch n dar. □

Die Repräsentation von 0 stellt einen Spezialfall dar, denn jede binäre quadratische form $Q \in \mathcal{A}$ mit ganzen Koeffizienten stellt 0 dar ($Q(0, 0) = 0$). Die bessere Frage wäre also ob 0 primitiv dargestellt wird. Das nächste Lemma zeigt, dass dies nur in einem bestimmten Fall geschieht.

Lemma 6.4. *Eine binäre quadratische Form $Q \in \mathcal{A}$ mit ganzen Koeffizienten stellt 0 primitiv dar genau dann wenn die Diskriminante Δ_Q ein Quadrat ist.*

Beweis. Nehme an $Q \in \mathcal{A}$ stellt 0 primitiv dar. Dann funktioniert das Argument im Beweis des vorherigen Lemmas. Wir finden also eine zu Q äquivalente binäre quadratische Form $Q'(X,Y) = bXY + cY^2$. Es folgt $\Delta_Q = \Delta_{Q'} = b^2$. Sei $Q(X,Y) = aX^2 + bXY + cY^2 \in \mathcal{A}$ mit Diskriminante $\Delta_Q = d^2$. Falls $a = 0$, dann ist $Q(1,0) = 0$. Nehme also $a \neq 0$ and, dann haben wir

$$Q(X,Y) = \frac{1}{4a}\left[(2aX + bY)^2 - \Delta_Q Y^2\right] = \frac{1}{4a}\left(2aX + (b-d)Y\right)\left(2aX + (b+d)Y\right).$$

Es gilt also $Q(b-d, -2a) = 0$. Da $a \neq 0$ ist ein grösster gemeinsamer Teiler e von $b-d$ und $-2a$ nicht 0. Es sind dann $(b-d)/e$ und $-2a/e$ teilerfremd und $Q((b-d)/e, -2a/e) = 0$. □

Lemma 6.5. *Sei $n, \Delta \in \mathbb{Z}$ ganze Zahlen. Die Zahl n wir von einer binären quadratischen Form $Q \in \mathcal{A}$ mit ganzen Koeffizienten und Diskriminante $\Delta_Q = \Delta$ genau dann primitiv dargestellt, falls ein $b \in \mathbb{Z}$ existiert, sodass $\Delta \equiv b^2 \bmod(4n)$.*

Beweis. Der Fall $n = 0$ ist gerade die Aussage des vorherigen Lemmas. Sei nun also $n \neq 0$. Falls $Q \in \mathcal{A}$ mit $\Delta_Q = \Delta$ welche n darstellt. Dann gibt es nach Lemma 6.2 ein $Q' \in \mathcal{A}$ mit $Q \sim Q'$ sodass $Q'(X,Y) = nX^2 + bXY + cY^2$. Nach Lemma 6.1 haben Q und Q' dieselbe Diskriminante $\Delta = b^2 - 4nc$. Es folgt $\Delta \equiv b^2 \bmod(4n)$.

Umgekehrt, falls $\Delta \equiv b^2 \bmod(4n)$, dann gibt es ein $c \in \mathbb{Z}$, sodass $\Delta = b^2 - 4nc$. Dann ist $Q'(X,Y) = nX^2 + bXY + cY^2$ in $\mathcal{A}$ mit Diskriminante Δ. □

Dieses letzte Kriterium, i.e. kongruent zu einem Quadrat modulo $4n$ zu sein, schauen wir uns im Kapitel §7 noch genauers an. Für uns genügt zuerst mal der Fall $\Delta = -4$.

Die quadratische Form $X^2 + Y^2$ hat Diskriminante -4. Für eine ungerade Primzahl $p \in \mathbb{N}$ ist -4 kongruent zu einem Quadrat modulo $4p$ genau dann wenn -1 kongruent zu einem Quadrat modulo p ist und dies ist genau dann der Fall, falls $p \equiv 1 \bmod(4)$ (vgl. Lemmata 3.4 und 3.3). Das heisst, falls wir zeigen möchten, dass $X^2 + Y^2$ p, für $p \equiv 1 \bmod(4)$, (primitiv) repräsentiert, dann genügt es zu zeigen, dass es bis auf Äquivalenz nur eine Klasse an positiven definiten binären quadratischen Formen mit ganzen Koeffizienten und Diskriminante -4 gibt. Dazu zeigen wir, dass jede Klasse einen simplen Repräsentant hat.

Satz 6.6. *Jede binäre quadratische Form $Q \in \mathcal{A}$ mit ganzen Koeffizienten und keine Quadratzahl als Diskriminante besitzt eine dazu äquivalente Form $Q'(X,Y) = aX^2 + bXY + cY^2 \in \mathcal{A}$ mit $|b| \leqslant |a| \leqslant |c|$.*

Beweis. Wir betrachten was mit den Koeffizienten der quadratischen Form geschieht wenn man mit den zwei Matrizzen $\left(\begin{smallmatrix} 1 & n \\ 0 & 1 \end{smallmatrix}\right)$, $n \in \mathbb{Z}$ und $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ agiert. Sei $Q(X,Y) = aX^2 + bXY + cY^2$. Dann sind

$$\rho(Q, \left(\begin{smallmatrix} 1 & n \\ 0 & 1 \end{smallmatrix}\right))(X,Y) = aX^2 + (2na + b)XY + (n^2a + nb + c)Y^2, \tag{6.2}$$

$$\rho(Q, \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right))(X,Y) = cX^2 - bXY + aY^2. \tag{6.3}$$

Wir können nun zuerst (6.2) anwenden um b mit einer beliebigen ganzen Zahl b' mit $b' \equiv b \bmod(2a)$ zu ersetzen. Da $a \neq 0$ (vgl. Lemma 6.4) können wir b' mit minimalem Absolutbetrag wählen, sodass $|b'| \leqslant |a|$. Anschliessend können wir (6.3) benutzen um die Koeffizienten vor X^2 und Y^2 zu vertauschen und den Koeffizienten von XY zu negieren. Nach diesen zwei Schritten erhalten wir also eine neue binäre quadratische Form $Q''(X,Y) = a''X^2 + b''XY + c''Y^2$ welche eine der Ungleichungen, nämlich $|b''| \leqslant |c''|$, erfüllt.

Bei genauer Betrachtung finden wir auch, dass $|b''| < |b|$ falls ursprünglich $|b| > |a| > 0$ galt. Das heisst, solange die Ungleichung $|b''| > |a''| > 0$ gilt können wir die zwei Schritte wiederholen und erhalten eine neue binäre quadratische Form mit Koeffizient von XY kleiner im Absolutbetrag als $|b''|$. Der Koeffizient von X^2 ist nie Null, da die Diskriminante kein Quadrat ist, und der Koeffizient von XY kann man aber nur endlich oft im Absolutbetrag verkleinern, da es sich um eine nicht negative ganze Zahl handelt. Wir müssen daher an eine binäre quadratische Form $Q^\dagger(X,Y) = a^\dagger X^2 + b^\dagger XY + c^\dagger Y^2$ angelangen für welche $|b^\dagger| \leqslant |a^\dagger|$ und $|b^\dagger| \leqslant |c^\dagger|$ gilt. Falls $|a^\dagger| \leqslant |c^\dagger|$, so ist man fertig. Ansonsten, wendet man nocheinmal (6.3) an. □

Definition. *Eine binäre quadratische Form $Q(X,Y) = aX^2 + bXY + cY^2 \in \mathcal{A}$ mit ganzen Koeffizienten und keine Quadratzahl als Diskriminante heisst* reduziert*, falls die Koeffizienten die vorherige Ungleichung $|b| \leqslant |a| \leqslant |c|$ erfüllen.*

Beispiel. *Wir würden gerne die Form $Q_0(X,Y) = 8X^2 + 21XY + 14Y^2$ reduzieren. In einem ersten Schritt agieren wir mit $\left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right)$, sodass $|2\cdot(-1)\cdot 8 + 21| = 5 \leqslant 8$, und erhalten $Q_1(X,Y) = 8X^2 + 5XY + Y^2$. Nun agieren wir mit $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ und erhalten $Q_2(X,Y) = X^2 - 5XY + 8Y^2$. Wir agieren weiter mit $\left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)$ und erhalten $Q_3(X,Y) = X^2 - XY + 2Y^2$. Letztere Form ist reduziert. Wir können nun unsere Schritte auch verfolgen und erhalten*

$$\begin{aligned} Q_3 &= \rho(Q_2, \left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)) = \rho(\rho(Q_1, \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)), \left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)) = \rho(Q_1, \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)) \\ &= \rho(\rho(Q_0, \left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right)), \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)) = \rho(Q_0, \left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)) = \rho(Q_0, \left(\begin{smallmatrix} -1 & -3 \\ 1 & 2 \end{smallmatrix}\right)). \end{aligned}$$

Lemma 6.7. *Sei $Q(X,Y) = aX^2 + bXY + cY^2 \in \mathcal{A}$ eine reduzierte binäre quadratische Form mit ganzen Koeffizienten und Diskriminante Δ_Q, welche kein Quadrat einer ganzen Zahl ist, dann erfüllen die Koeffizienten folgende Ungleichungen:*

$$\begin{aligned} &|b| \leqslant \sqrt{\frac{-\Delta_Q}{3}}, \qquad |a| \leqslant \sqrt{\frac{-\Delta_Q}{3}}, \qquad |c| \leqslant \frac{1-\Delta_Q}{4}, \qquad \textit{falls } \Delta_Q < 0, \\ &|b| \leqslant \sqrt{\frac{\Delta_Q}{5}}, \qquad |a| \leqslant \frac{\sqrt{\Delta_Q}}{2}, \qquad |c| \leqslant \frac{\Delta_Q}{4}, \qquad \textit{falls } \Delta_Q > 0. \end{aligned}$$

Beweis. Nehme zuerst an, dass $\Delta_Q < 0$. Es gilt $b^2 - 4ac = \Delta_Q < 0$, d.h. $4ac > b^2 > 0$ und folglich $-\Delta_Q = 4|ac| - b^2 \geqslant 3|a|^2 \geqslant |b|^2$. Ferner haben wir

$$|c| = \frac{b^2 - \Delta_Q}{4|a|} \leqslant \frac{|a|}{4} - \frac{\Delta_Q}{4|a|} \leqslant \max\left\{\frac{1-\Delta_Q}{4}, \sqrt{\frac{-\Delta_Q}{3}}\right\} = \frac{1-\Delta_Q}{4},$$

da die Funktion $x \mapsto x + C/x$, für $C > 0$, konvex ist auf $\mathbb{R}^+$ und $|a| \in [1, \sqrt{-\Delta_Q/3}]$.

Sei nun $\Delta_Q > 0$. Dann haben wir $|ac| \geqslant b^2 > 4ac$. Es muss also $ac < 0$ gelten. Folglich gilt $\Delta_Q = b^2 + 4|ac| \geqslant 5b^2$. Andererseits gilt auch $\Delta_Q = b^2 + 4|ac| \geqslant 4|a|^2$ und $\Delta_Q = b^2 + 4|ac| \geqslant 4|c|$. □

Es gibt also nur endlich viele reduzierte binäre quadratische Formen mit ganzen Koeffizienten und einer gegebenen Diskriminante, welche kein Quadrat ist.

Wir bestimment nun alle solche reduzierte $Q(X,Y) = aX^2 + bXY + cY^2 \in \mathcal{A}$ mit Diskriminante $\Delta_Q = -4$. Lemma 6.7 besagt nun, dass $|b| \leqslant 1, |a| \leqslant 1$. Ferner bemerken wir, dass aus $b^2 - 4ac = -4$ $2 \mid b$ folgt, also insbesondere $b = 0$. Wir finden zwei Lösungen $b = 0$, $a = c = 1$ und $b = 0$, $a = c = -1$. Letztere Form ist aber negativ definit. Das heisst jede positiv definite binäre quadratische Form mit ganzen Koeffizienten und Diskriminante

-4 ist äquivalent zu $X^2 + Y^2$. Sei nun $p \in \mathbb{N}$ mit $p \equiv 1 \bmod(4)$, sodass -1 ein Quadrat modulo p ist und folglich -4 ein Quadrat modulo $4p$, dann wird nach Lemma 6.5 p durch eine binäre quadratische Form mit ganzen Koeffizienten und Diskriminante -4 dargestellt. Jene kann aber nicht negativ definit sein, insbesondere muss sie äquivalent zu $X^2 + Y^2$ sein. Lemma 6.2 impliziert dann, dass die Form $X^2 + Y^2$ auch p repräsentiert.

Aufgabe 6.1. *Bestimme alle reduzierten binären positiv definiten quadratische Formen mit ganzen Koeffizienten und Diskriminante gleich -7. Sind jene äquivalent? Repräsentiert die Form $11X^2 + 31XY + 22X^2$ die Zahlen 0, 5, 22, und 29?*

7 Quadrate in Restklassen

Im letzten Kapitel ist die Frage aufgetaucht, wann eine ganze Zahl a kongruent zu einer Quadratzahl modulo einer natürlichen Zahl n ist. Für $a = -1$ und $n = 4$ haben wir diese Frage bereits in Lemmata 3.3, 3.4 beantworten. In diesem Kapitel adressieren wir die Frage im Allgemeinen.

Zuerst bemerken wir, dass eine ganze Zahl a kongruent zu einem Quadrat modulo einer natürlichen Zahl n genau dann, wenn im Faktorring $\mathbb{Z}/n\mathbb{Z}$ die Klasse $[a]_n$ ein Quadrat ist, d.h. $[a]_n = [b]_n^2$ für eine Kongruenzklasse $[b]_n \in \mathbb{Z}/n\mathbb{Z}$.

Lemma 7.1. *Sei $a \in \mathbb{Z}$ eine ganze Zahl und $n_1, n_2 \in \mathbb{N}$ zwei teilerfremde natürliche Zahlen, dann ist a (beziehungsweise $[a]_{n_1 n_2}$) genau dann ein Quadrat modulo $n_1 n_2$ falls a (beziehungsweise $[a]_{n_i}$) ein Quadrat modulo n_i für $i = 1, 2$ ist.*

Beweis. Dies ist eine direkte Konsequenz des chinesischen Restsatzes 2.7, denn $\mathbb{Z}/n_1 n_2\mathbb{Z} = \mathbb{Z}/n_1\mathbb{Z} \times \mathbb{Z}/n_2\mathbb{Z}$ als Ringe. Alternativ, falls wir nur die Bijektion, welche wir bewiesen haben, benutzen wollen, kann man man wiefolgt argumentieren. Falls $[a]_{n_1 n_2} = [b]_{n_1 n_2}^2 = [b^2]_{n_1 n_2}$, dann ist

$$[a]_{n_1} = a + n_1\mathbb{Z} = b^2 + n_1 n_2\mathbb{Z} + n_1\mathbb{Z} = b^2 + n_1\mathbb{Z} = [b^2]_{n_1} = [b]_{n_1}^2$$

und analog für n_2.

Umgekehrt, falls $[a]_{n_i} = [x_i^2]_{n_i} = [x_i]_{n_i}^2$ für $i = 1, 2$, dann gibt es nach dem chinesischen Restsatz 2.7 ein $x \in \mathbb{Z}$, sodass $[x]_{n_i} = [x_i]_{n_i}$ für $i = 1, 2$. Es folgt, dass $[x^2]_{n_i} = [a]_{n_i}$ für $i = 1, 2$ und nach der Eindeutigkeit, dass $[x]_{n_1 n_2}^2 = [x^2]_{n_1 n_2} = [a]_{n_1 n_2}$. □

Corollary 7.2. *Sei $p_1^{\alpha_1} \cdot p_r^{\alpha_r}$ die Primfaktorzerlegung einer natürlichen Zahl $n \in \mathbb{N}$ mit paarweise teilerfremden Primzahlen p_i. Ferner sei $a \in \mathbb{Z}$ eine ganze Zahl. Dann ist a ein Quadrat modulo n genau dann falls a ein Quadrat modulo $p_i^{\alpha_i}$ für $i = 1, \ldots, r$ ist.*

Dies erlaubt uns auf den Fall $n = p^\alpha$ gleich einer Primzahlpotenz zu reduzieren. Im nächsten Schritt möchten wir auf den Fall a und p^α teilerfremd reduzieren. Dazu bemerken wir, dass $a = 0$ immer kongruent zu einem Quadrat modulo jeder Zahl ist, denn 0 ist das Quadrat einer ganzen Zahl. Für alle anderen ganzen Zahlen a haben wir folgende Aussage.

Lemma 7.3. *Sei $0 \neq a \in \mathbb{Z}$ eine ganze Zahl, $\alpha \in \mathbb{N}$ eine natürliche Zahl und $p \in \mathbb{N}$ eine Primzahl. Sei $\beta \in \mathbb{N}_0$ die grösste ganze Zahl, sodass $p^\beta \mid a$. Dann ist a kongruent zu einem Quadrat modulo p^α genau dann falls*

- *$\beta \geqslant \alpha$, oder*
- *β ist gerade und $\frac{a}{p^\beta}$ ist kongruent zu einem Quadrat modulo $p^{\alpha-\beta}$.*

Beweis. Falls $\beta \geqslant \alpha$, dann ist $a \equiv 0 \equiv 0^2 \bmod(p^\alpha)$. Sei nun $\beta < \alpha$. Falls β gerade ist und $\frac{a}{p^\beta} \equiv b^2 \bmod(p^{\alpha-\beta})$, dann ist $a \equiv (bp^{\beta/2})^2 \bmod(p^\alpha)$. Umgekehrt, falls $a \equiv c^2 \bmod(p^\alpha)$ gilt, dann ist $c \neq 0$, da sonst $p^\alpha \mid a$. Sei nun $\gamma \in \mathbb{N}_0$ die grösste ganze Zahl, sodass $p^\gamma \mid c$. Aus $p^\alpha \mid a - c^2$ und $\beta < \alpha$ folgt $\beta = 2\gamma$. Ferner gilt dann, dass $p^{\alpha-\beta} \mid \frac{a}{p^\beta} - (\frac{c}{p^\gamma})^2$, folglich ist $\frac{a}{p^\beta}$ kongruent zu einem Quadrat modulo $p^{\alpha-\beta}$. □

Nun dürfen wir also annehmen, dass $n = p^\alpha$ eine Primzahlpotentz ist und a und p teilerfremd sind. Wir beginnen mit dem schwierigsten Fall, nämlich wenn $n = p$ eine Primzahl ist. Für $p = 2$ sind beide Kongruenzklassen $[0]_2, [1]_2$ Quadrate. Für p ungerade führen wir das *Legendre Symbol* ein.

Definition. *Für $p \in \mathbb{N}$ eine ungerade Primzahl und eine ganze Zahl $a \in \mathbb{Z}$ definiere man das Legendre Symbol*

$$\left(\frac{a}{p}\right) = \begin{cases} 0, & p \mid a, \\ 1, & p \nmid a \text{ und } a \text{ ein Quadrat modulo } p, \\ -1, & a \text{ kein Quadrat modulo } p, \end{cases}$$

sodass

$$\sharp\{[b]_p \in \mathbb{Z}/p\mathbb{Z} | [a]_p = [b]_p^2\} = 1 + \left(\frac{a}{p}\right).$$

Als erstes bemerken wir, dass $\left(\frac{a}{p}\right)$ nur von der Kongruenzklasse von a modulo p abhängt. Ferner ist das Legendre Symbol multiplikativ.

Lemma 7.4. *Sei $p \in \mathbb{N}$ eine ungerade Primzahl und $a, b \in \mathbb{Z}$ zwei ganze Zahlen, dann gilt*

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right).$$

Beweis. Falls eine der Seiten Null ist, dann folgt die Gleichheit aus der Eigenschaft, dass p prim ist: $p \mid ab \Leftrightarrow p \mid a$ oder $p \mid b$. Sei nun also $p \nmid ab$. Falls a und b Quadrate modulo p sind, so ist klarerweise auch ab ein Quadrat modulo p. Falls a ein Quadrat modulo p ist und b nicht, so ist ab auch kein Quadrat modulo p, denn falls ab ein Quadrat modulo p ist, dann ist a^2b modulo p ein Quadrat sowie $(a^\star)^2$, wobei $a^\star$ ein multiplikatives Inverses von a modulo p. Es folgt, dass die Kongruenzklasse von b modulo p auch ein Quadrat ist, denn $[b]_p = [a^\star]_p^2 \cdot [a^2 b]_p$. Ein Widerspruch!

Seien nun a, b keine Quadrate modulo p. Da $p \nmid a$ ist die Abbildung $\mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$ gegeben durch $[x]_p \mapsto [ax]_p = [a]_p \cdot [x]_p$ bijektiv. Sie sendet $[0]_p$ auf $[0]_p$ und Quadrate (modulo p) auf nicht-Quadrate (modulo p). Nun gibt es in $\mathbb{Z}/p\mathbb{Z}$ aber genau $\frac{p+1}{2}$ Quadrate (modulo p inklusive $[0]_p$). Es müssen also die nicht-Quadrate auf Quadrate abgebildet werden, d.h. ab ist ein Quadrat modulo p. □

Wir benötigen nun einen Satz aus der Körpertheorie.

Satz 7.5 (Primitives Element)**.** *Sei $p \in \mathbb{Z}$ eine Primzahl, dann gibt es eine ganze Zahl g, sodass g modulo p die multiplikative Ordnung $p - 1$ hat, d.h. $(\mathbb{Z}/p\mathbb{Z})^\times$ ist zyklisch. Eine solche Zahl g nennt sich ein* primitives Element *modulo p und jede Kongruenzklasse $[a]_p$ lässt sich als $[g]_p^n$ schreiben für ein $n \in \mathbb{Z}$, welches eindeutig modulo $p - 1$ ist.*

Beweis. Siehe [1, §3.6 Satz 14]. Skizze: $\mathbb{Z}/p\mathbb{Z}$ ist ein Körper mit p Elementen. Die Einheitengruppe $(\mathbb{Z}/p\mathbb{Z})^\times$ ist abelsch und nach dem Elementarteilersatz isomorph zur einer (additiven) Gruppe $\mathbb{Z}/e_1\mathbb{Z} \times \ldots \mathbb{Z}/e_r\mathbb{Z}$ mit $e_1 \mid \cdots \mid e_r$ und $e_1 > 1$. Dann sind alle Elemente von $(\mathbb{Z}/p\mathbb{Z})^\times$ Nullstellen vom Polynom $X^{e_r} - 1$, welches aber höchstens e_r Nullstellen besitzt. Es folgt $r = 1$ und $e_1 = p - 1$, d.h. $(\mathbb{Z}/p\mathbb{Z})^\times$ ist zyklisch. □

Lemma 7.6. *Sei $p \in \mathbb{N}$ eine ungerade Primzahl und $a \in \mathbb{Z}$ eine ganze Zahl teilerfremd zu p. Ferner sei g ein primitives Element modulo p, dann ist a ein Quadrat modulo p genau dann wenn $[a]_p = [g]_p^n$ für ein gerades n. Ferner gilt*

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

Beweis. Falls n gerade ist, so ist a klar ein Quadrat modulo p. Falls a ein Quadrat modulo p ist, d.h. $[a]_p = [b]_p^2$. Schreibe nun $[b]_p = [g]_p^m$, dann ist $[a]_p = [g]_p^{2m}$. Wir bemerken noch, dass n eindeutig modulo $p-1$ (gerade) ist, d.h. die Parität vom Exponenten ist eindeutig bestimmt.

Wir haben $(a^{\frac{p-1}{2}})^2 = a^{p-1} \equiv 1 \bmod(p)$ nach dem kleinen Satz von Fermat 2.9. Es folgt, dass $a^{\frac{p-1}{2}} \equiv \pm 1 \bmod(p)$. Wenn wir nun $a \equiv g^n \bmod(p)$ schreiben, so sehen wir, dass $a^{\frac{p-1}{2}} \equiv (g^{\frac{p-1}{2}})^n \bmod(p)$ genau dann kongruent zu 1 ist falls n gerade ist. □

Der nächste Satz ist einer der zentralen Sätze der Zahlentheorie. Über hundert Beweise sind heutzutage bekannt. Einige von ihnen gaben Anlass zu komplett neuen Gebieten in der Zahlentheorie. Als solches kann jener Satz als Geburt der modernen Zahlentheorie angesehen werden. Der Satz wurde zuerst von Gauss im Jahre 1796 bewiesen. Hier reprodizieren wir den Beweis von Rousseau [11].

Satz 7.7 (Quadratische Reziprozität)**.** *Seien $p, q \in \mathbb{N}$ zwei verschiedene ungerade Primzahlen, dann gilt*

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

Beweis. Wir betrachten die Menge aller Einheiten (invertierbare Kongruenzklassen) in $\mathbb{Z}/pq\mathbb{Z}$. Nach dem chinesischen Restsatz 2.7 haben wir

$$(\mathbb{Z}/pq\mathbb{Z})^\times \cong (\mathbb{Z}/p\mathbb{Z})^\times \times (\mathbb{Z}/q\mathbb{Z})^\times. \tag{7.1}$$

Wir betrachten nun die Äquivalenzrelation $[a]_{pq} \sim [b]_{pq} \Leftrightarrow [a]_{pq} = \pm[b]_{pq} \Leftrightarrow a \equiv \pm b \bmod(pq)$. Auf der rechten Seite des Isomorphismus (7.1) is die Äquivalenzrelation gegeben durch

$$([a_1]_p, [a_2]_q) \sim ([b_1]_p, [b_2]_q) \Leftrightarrow [a_1]_p = \pm[b_1]_p \text{ und } [a_2]_q = \pm[b_2]_q,$$

wobei das Vorzeichen gleich gewählt werden muss. Wir wollen nun aus jeder Äquivalenzklasse $\{[a]_{pq}, [-a]_{pq}\}$ genau ein Element auswählen und das Produkt all jener Elemente betrachten. Ein solches Produkt ist dann auch bis auf Äquivalenz bestimmt. Es gibt nun drei natürliche Arten ein Element aus jeder Äquivalenzklasse zu wählen.

- $\{[a]_p \in \mathbb{Z}/p\mathbb{Z} | a = 1, \ldots, \frac{p-1}{2}\} \times (\mathbb{Z}/q\mathbb{Z})^\times$,

- $(\mathbb{Z}/p\mathbb{Z})^\times \times \{[a]_q \in \mathbb{Z}/q\mathbb{Z} | a = 1, \ldots, \frac{q-1}{2}\}$,

- $\{[a]_{pq} \in \mathbb{Z}/pq\mathbb{Z} | 1 \leqslant a \leqslant \frac{pq-1}{2} \text{ mit } \mathrm{ggT}(a, pq) = 1\}$.

Wir berechnen das Produkt auf der rechten Seite von (7.1) aus. Im ersten Fall erhalten wir

$$\left(\left[\left(\left(\tfrac{p-1}{2} \right)! \right)^{q-1} \right]_p, \left[((q-1)!)^{\frac{p-1}{2}} \right]_q \right).$$

Wir erinnern uns nun an den Beweis von 3.4 und finden

$$\left(\tfrac{p-1}{2} \right)^2 \equiv (-1)^{\frac{p-1}{2}} (p-1)! \mod(p).$$

Wir erhalten also im ersten Fall, dass das Produkt

$$\left(\left[(-1)^{\frac{p-1}{2}\frac{q-1}{2}} ((p-1)!)^{\frac{q-1}{2}} \right]_p, \left[((q-1)!)^{\frac{p-1}{2}} \right]_q \right) \tag{7.2}$$

ist. Im zweiten Fall erhalten wir analogue

$$\left(\left[((p-1)!)^{\frac{q-1}{2}} \right]_p, \left[(-1)^{\frac{p-1}{2}\frac{q-1}{2}} ((q-1)!)^{\frac{p-1}{2}} \right]_q \right). \tag{7.3}$$

Im letzten Fall berechnen wir zuerst modulo p. Das Produkt ist gleich

$$\frac{1 \cdot 2 \cdots (p-1) \times (p+1) \cdot (p+2) \cdots (p+p-1) \times \cdots \times (\frac{q-1}{2}p + 1) \cdots (\frac{q-1}{2}p + p - 1)}{q \cdot (2q) \cdots (\frac{p-1}{2}q) \times (\frac{q-1}{2}p + \frac{p+1}{2}) \cdots (\frac{q-1}{2}p + p - 1)}$$
$$\equiv \frac{((p-1)!)^{\frac{q+1}{2}}}{q^{\frac{p-1}{2}} \cdot (p-1)!} \equiv \left(\frac{q}{p} \right) ((p-1)!)^{\frac{q-1}{2}} \mod(p).$$

Wir bemerken hier, das wir eigentlich einen Bruch $\frac{a}{b}$ als $ab^\star$ schreiben müssten. Die Rechnung und das Resultat sind aber gleich.[i]

Die Rechnung modulo q ist analog. Wir erhalten also im dritten Fall, dass das Produkt gleich

$$\left(\left[\left(\frac{q}{p} \right) ((p-1)!)^{\frac{q-1}{2}} \right]_p, \left[\left(\frac{p}{q} \right) ((q-1)!)^{\frac{p-1}{2}} \right]_q \right). \tag{7.4}$$

Wir vergleichen nun (7.2) mit (7.4). Da die Produkte bis auf Äquivalenz gleich sein müssen muss also

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

gelten. □

Das erstaunliche am Satz ist, dass die Lösbarkeit von $p \equiv b^2 \mod(q)$ sagt etwas aus über die Lösbarkeit von $q \equiv c^2 \mod(p)$. Dies ist *a priori* überhaupt nicht klar und sogar überaschend. Nebst dieser Reziprozität gibt es auch noch zwei Hilfsidentitäten, welche uns erlaubt die Frage ob a kongruent zu einem Quadrat modulo p rekursiv zu beantworten.

[i] Die Lokalisation von $\mathbb{Z}/p\mathbb{Z}$ an $(\mathbb{Z}/p\mathbb{Z})^\times$ ist isomorph zu $\mathbb{Z}/p\mathbb{Z}$.

Satz 7.8. *Sei $p \in \mathbb{N}$ eine ungerade Primzahl, dann gilt*

$$\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}} \text{ und } \left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}} = \begin{cases} 1, & p \equiv 1,7 \mod(8), \\ -1, & p \equiv 3,5 \mod(8). \end{cases}$$

Beweis. Der erste Teil ist eine Konsequenz von Lemmata 3.3, 3.4 beziehungsweise ein Spezialfall von Lemma 7.6. Für den zweiten Teil bemerken wir zuerst, dass die rechte Seite nur von $p \mod(8)$ abhängt. Um die Identität zu zeigen werden wir die Gauss'schen Zahlen zuhilfe nehmen. In den Gauss'schen Zahlen haben wir $2 = -i(1+i)^2$ und p und 2 sind teilerfremd (da p ungerade ist). Wir erhalten

$$\begin{aligned} \left(\frac{2}{p}\right) &\equiv 2^{\frac{p-1}{2}} \equiv (-i)^{\frac{p-1}{2}}(1+i)^{p-1} \mod(p) \\ &\equiv (-i)^{\frac{p-1}{2}}(1+i)^{p}(1+i)^{\star} \mod(p) \\ &\equiv (-i)^{\frac{p-1}{2}}(1+i^p)(1+i)^{\star} \mod(p). \end{aligned} \tag{7.5}$$

Wobei wir zuerst Lemma 7.6 benutzt haben. Die Kongruenz gilt in den ganzen Zahlen und somit sicherlich auch in den Gauss'schen Zahlen $\mathbb{Z}[i]$. Danach multiplizieren wir mit $1+i$ und dem multiplikativen Inversen von $1+i$ modulo p. Dies ist möglich, da $1+i$ und p teilerfremd sind (als Gauss'sche Zahlen). Zum Schluss gilt auch noch $(1+i)^p \equiv 1+i^p \mod(p)$ nach der ersten binomischer Formel, da die Koeffizienten $\binom{p}{i}$ für $i = 1, \ldots, p-1$ durch p teilbar sind für eine Primzahl p. Nun ist die rechte Seite von (7.5) auch nur abhängig von $p \mod(8)$. Wir finden

$$1 + i^p = \begin{cases} 1+i, & p \equiv 1,5 \mod(5), \\ (-i)(1+i), & p \equiv 3,7 \mod(8). \end{cases}$$

Einsetzen in (7.5) zeigt dann

$$\left(\frac{2}{p}\right) \equiv (-1)^{\frac{p^2-1}{8}} \mod(p)$$

und da $p \neq 2$ muss sogar Gleichheit gelten. □

Beispiel. *Wir wollen bestimmen, ob* 198 *ein Quadrat modulo* 59 *(Primzahl) ist. Wir finden, dass* $198 \equiv 21 \mod(59)$. *Es gilt also*

$$\left(\frac{198}{59}\right) = \left(\frac{21}{59}\right) = \left(\frac{3}{59}\right)\left(\frac{7}{59}\right).$$

Quadratische Reziprozität besagt nun

$$\left(\frac{3}{59}\right) = -\left(\frac{59}{3}\right) = -\left(\frac{2}{3}\right) = 1,$$

da 2 *kongruent zu keinem Quadrat modulo* 3 *ist, und*

$$\left(\frac{7}{59}\right) = -\left(\frac{59}{7}\right) = -\left(\frac{3}{7}\right).$$

Entweder kann man auch wieder kurz testen, ob 3 *kongruent zu einem Quadrat ist modulo* 7 *oder die quadratische Reziprozität nochmals benützen und* $\left(\frac{3}{7}\right) = -\left(\frac{7}{3}\right) = -\left(\frac{1}{3}\right) = -1$ *finden. Wir folgern* $\left(\frac{198}{59}\right) = 1$, *also ist* 198 *kongruent zu einem Quadrat modulo* 59. *Tatsächlich ist* $198 \equiv 27^2 \mod(59)$. *Letzteres hätte aber ein wenig länger gedauert um es direkt zu finden.*

Lemma 7.9 (Hensel's Lemma)**.** *Sei $a \in \mathbb{Z}$ eine ganze Zahl, $\alpha \in \mathbb{N}$ eine natürliche Zahl, und $p \in \mathbb{N}$ eine ungerade Primzahl. Nehme an a und p sind teilerfremd. Dann ist a kongruent zu einem Quadrat modulo p^α genau dann wenn a kongruent zu einem Quadrat modulo p ist.*

Beweis. Falls a kongruent zu einem Quadrat modulo p^α ist, dann ist sicherlich a kongruent zu einem Quadrat modulo p, da $p \mid p^\alpha$. Für die andere Richtung verwenden wir Induktion nach α. Für $\alpha = 1$ ist nichts zu zeigen. Falls a kongruent zu einem Quadrat modulo p ist, so ist nach Induktionsannahme a kongruent zu einem Quadrat modulo $p^{\alpha-1}$. Wir betrachten nun a modulo p^α. Sei $b \in \mathbb{Z}$ eine ganze Zahl, sodass $a \equiv b^2 \bmod(p^{\alpha-1})$. Wir finden also ein k, sodass $a = b^2 + kp^{\alpha-1}$. Sei $(2b)^\star$ ein multiplikatives Inverses von $2b$ modulo p ($2b$ ist teilerfremd zu p da $p \neq 2$ und a teilerfremd zu p ist). Dann gilt

$$(b + (2b)^\star kp^{\alpha-1})^2 = b^2 + 2b(2b)^\star kp^{\alpha-1} + ((2b)^\star kp^{\alpha-1})^2 \equiv b^2 + kp^{\alpha-1} \equiv a \bmod(p^\alpha),$$

da $2(\alpha - 1) \geqslant \alpha$ und $2b \cdot (2b)^\star \equiv 1 \bmod(p)$. □

Für die Primzahl 2 klappt dieser Beweis nicht ganz in der selben Art und Weise, da 2 und $p = 2$ nicht teilerfremd sind und es daher kein multiplikatives Inverse gibt. Wir haben aber folgendes.

Lemma 7.10. *Sei a eine ungerade ganze Zahl und $\alpha \in \mathbb{N}$ eine natürliche Zahl, dann ist a ein Quadrat modulo 2^α genau dann wenn*

$$\begin{cases} \textit{immer}, & \alpha = 1, \\ a \equiv 1 \bmod(4), & \alpha = 2, \\ a \equiv 1 \bmod(8), & \alpha \geqslant 3. \end{cases}$$

Beweis. Für $\alpha = 1, 2, 3$ ist dies schnell nachgerechnet. Für $\alpha \geqslant 4$ wollen wir argumentieren wie im vorherigen Lemma durchführen. Also auch wieder mittels Induktion. Sei $a = b^2 + k2^{\alpha-1}$. Falls k gerade ist, so ist $a \equiv b^2 \bmod(2^\alpha)$ und wir sind fertig. Sei also k ungerade. Dann ist

$$(b + 2^{\alpha-2})^2 = b^2 + b2^{\alpha-1} + 2^{2\alpha-4} \equiv b^2 + k2^{\alpha-1} \equiv a \bmod(2^\alpha),$$

da $\alpha \geqslant 4$ und $b \equiv k \bmod(2)$. □

Aufgabe 7.1. *Bestimme ob 173 und 177 kongruent zu einem Quadrat modulo $8 \cdot 37$ sind.*

Aufgabe 7.2. *Sei $p > 5$ eine Primzahl. Zeige, dass -5 genau dann ein Quadrat modulo p ist, wenn $p \equiv \pm 1 \bmod(5)$.*

Literatur

[1] Siegfried Bosch. *Algebra.* Berlin: Springer Spektrum, 2020.

[2] A. Cayley. On Jacobi's elliptic functions, in reply to the Rev. Brice Bronwin and on quaternions. *Philosophical Magazine and Journal of Science*, 3:210–213, 1845.

[3] David A. Cox. *Primes of the form $x^2 + ny^2$. Fermat, class field theory, and complex multiplication.* Hoboken, NJ: John Wiley & Sons, 2013.

[4] C. W. Curtis. The four and eight square problem and division algebras. Stud. in Math. 2, 100-125 (1963)., 1963.

[5] L. E. Dickson. On quaternions and their generalization and the history of the eight square theorem. *Ann. of Math. (2)*, 20(3):155–171, 1919.

[6] L. E. Dickson. Arithmetic of Quaternions. *Proc. London Math. Soc. (2)*, 20(3):225–232, 1921.

[7] Adolf Hurwitz. Vorlesungen über die Zahlentheorie der Quaternionen. Berlin: J. Springer, IV u. 74 S. gr. 8° (1919)., 1919.

[8] Franz Lemmermeyer. Binary Quadratic Forms. Available at http://www.rzuser.uni-heidelberg.de/∼hb3/publ/bf.pdf, 2010.

[9] Stefan Müller-Stach and Jens Piontkowski. *Elementare und algebraische Zahlentheorie. Ein moderner Zugang zu klassischen Themen.* Wiesbaden: Vieweg+Teubner, 2011.

[10] A. R. Rajwade. *Squares*, volume 171 of *London Mathematical Society Lecture Note Series.* Cambridge University Press, Cambridge, 1993.

[11] G. Rousseau. On the quadratic reciprocity law. *J. Austral. Math. Soc. Ser. A*, 51(3):423–425, 1991.

[12] Alexander Schmidt. *Einführung in die algebraische Zahlentheorie.* Berlin: Springer, 2007.

[13] John Voight. *Quaternion algebras*, volume 288. Cham: Springer, 2021.