Elementare Zahlentheorie

Goethe–Universität Frankfurt — Sommersemester 2015 und 2017
für Bachelor und L3

JAKOB STIX

Zusammenfassung. — Die Vorlesung Elementare Zahlentheorie behandelt Primzahlen, Arithmetik in Restklassen, das quadratische Reziprozitätsgesetz, quadratische Formen und quadratische Zahlkörper.

Inhaltsverzeichnis

1. Einführung 3
1.1. Unendlich viele Primzahlen 3
1.2. Primzahllücken 5
1.3. Primzahlverteilung 6
1.4. Diophantische Gleichungen 7
1.5. Zahlentheorie über endlichen Körpern 10
1.6. Literatur 10

Teil 1. Arithmetik in $\mathbb{Z}$ 11
2. Die Peanoaxiome 11
3. Arithmetik der Teilbarkeit 12
3.1. Teilbarkeit in ganzen Zahlen 12
3.2. Der Euklidische Algorithmus 13
3.3. Kleinstes gemeinsames Vielfaches 16
3.4. Teilerfremde Zahlen 17
3.5. Lineare Gleichungen 18
4. Kongruenzen 19
4.1. Arithmetik modulo m 19
4.2. Teilbarkeitsregeln 21
5. Primzahlen 23
5.1. Das Sieb des Eratosthenes 23
5.2. Eindeutige Primfaktorzerlegung 24
5.3. Elementare Sätze über Primzahlen 27
5.4. Quadrate und Primzahlen 31
5.5. Die Sätze von Tschebyscheff zur Primzahlverteilung 35
6. Das abc der Arithmetik 41
6.1. Die abc-Vermutung 41
6.2. *abc* für Polynome 43
7. Zahlentheoretische Funktionen 45
7.1. Multiplikative Funktionen 45
7.2. Vollkommene Zahlen 46
7.3. Formale Potenzreihen 47
7.4. Möbiusinversion 49
8. Kettenbrüche 50
8.1. Der Wert eines Kettenbruchs 51
8.2. Der Kettenbruchalgorithmus 53
8.3. Diophantische Approximation mit Kettenbrüchen 57

Teil 2. Arithmetik in Restklassenringen 62
9. Der Chinesische Restsatz 62
9.1. Algebraisch: Teilbarkeit und Kongruenzrechnung 62
9.2. Algebraisch: Faktorisierung in Primpotenzen — der Chinesische Restsatz 62
10. Einheiten in Restklassenringen 64
10.1. Der Chinesische Restsatz für Einheiten 64
10.2. Primitivwurzeln 67
10.3. Struktur der Einheitengruppe modulo Primpotenzen 70
11. Arithmetik in der Kryptographie 73
11.1. RSA 73
11.2. Diffie-Hellman Schlüsselaustausch und der diskrete Logarithmus 75
11.3. Primzahltests 76
12. Quadratische Reste 81
12.1. Das quadratische Reziprozitätsgesetz 81
12.2. Ein erster Beweis mittels Gauß-Lemma 85
12.3. Ein zweiter Beweis mittels Gauß-Summen 88
12.4. Das Jacobi-Symbol 91
12.5. Das Kronecker-Symbol 92
13. Primzahlen in arithmetischen Folgen 94
13.1. Algebraische Beweise in Spezialfällen 94
13.2. Charaktere und Dualität für endliche abelsche Gruppen 96
13.3. Dirichlet-Reihen 100
13.4. Eulerprodukte 101
13.5. Dirichlet-Dichte 104
14. Diophantische Gleichungen in endlichen Körpern 106
14.1. Endliche Körper 106
14.2. Chevalley–Warning 108
14.3. Eine Anwendung 110

Teil 3. Arithmetik in quadratischen Zahlkörpern 115
15. Gaußsche ganze Zahlen 115
15.1. Arithmetik in $\mathbb{Z}[i]$ 115
15.2. Die Primzahlen der Gaußschen Zahlen 117
15.3. Anwendungen 119
16. Quadratische Zahlkörper 123
16.1. Quadratische Zahlkörper 123
16.2. Der Ganzzahlring 126
16.3. Ordnungen 128
16.4. Spurform und Diskriminante 129
17. Ganzzahlige binäre quadratische Formen 133
17.1. Binäre quadratische Formen 133
17.2. Äquivalenzklassen quadratischer Formen 134
17.3. Werte quadratischer Formen 136
18. Die Pell-Gleichung und Kettenbrüche 138
18.1. Die Pell-Gleichung 138
18.2. Periodische Kettenbrüche 138
19. Die Einheitengruppe quadratischer Zahlkörper 143
19.1. Imaginär-quadratische Zahlkörper 143
19.2. Reell-quadratische Zahlkörper 144
19.3. Die Fundamentaleinheit eines reell-quadratischen Zahlkörpers 146

1. Einführung

Elementare Zahlentheorie kommt ohne Primzahlen nicht aus. Sie sind gewissermaßen die Atome, die unteilbaren Bestandteile der multiplikativen Arithmetik.

Definition 1.1. Eine **Primzahl** ist eine ganze Zahl $p > 1$, die nur durch 1 und p teilbar[1].

Mathematische Laien fragen oft, warum 1 keine Primzahl ist, hat doch auch 1 nur die Teiler 1 und sich selbst. Die 1 bildet zusammen mit -1 die Einheitengruppe $\mathbb{Z}^\times = \{1, -1\}$ im Ring $\mathbb{Z}$ der ganzen Zahlen. Aussagen über eindeutige Faktorisierung wie der noch zu besprechende Fundamentalsatz der Arithmetik sind stets nur eindeutig bis auf Multiplikation mit Einheiten. Daher sind 1 und -1 keine Primzahlen.

1.1. Unendlich viele Primzahlen.

Seit mehr als 2000 Jahren weiß man, daß es unendlich viele Primzahlen gibt. Der Beweis von Euklid[2] hierfür gehört unbedingt zur Allgemeinbildung.

Theorem 1.2 (Euklid)**.** *Es gibt unendlich viele Primzahlen.*

Erster Beweis nach Euklid. Man mache sich zuerst klar, daß jede ganze Zahl $n > 1$ durch eine Primzahl teilbar ist. Entweder ist n selbst Primzahl oder aber $n = ab$ mit ganzen Zahlen $a, b > 1$. Also ist $1 < a = n/b < n$ und man verfährt weiter mit a. Die Teiler von a sind nämlich auch Teiler von n. Nach endlich vielen Schritten findet man einen Primteiler. (Eigentlich verwendet man hier schon das subtile Induktionsaxiom!)

Jetzt nehmen wir an, daß es eine endliche Liste aller Primzahlen gibt. Diese Liste sei $p_1, \ldots, p_r$. Dann ist

$$N = p_1 \cdot \ldots \cdot p_r + 1 \tag{1.1}$$

eine ganze Zahl > 1 und hat daher einen Primteiler ℓ. Wir behaupten nun, daß ℓ nicht in der Liste vorkommen kann. In der Tat geht die Division N/ℓ ohne Rest auf. Aber für alle $i = 1, \ldots, r$ gilt

$$N = 1 + p_i \cdot \prod_{j \neq i} p_j,$$

so daß die Division von N durch p_i den Rest 1 läßt. Die Primzahl ℓ war also noch nicht auf unserer Liste. Widerspruch. □

Euklids Beweis ist äußerst elegant, weil er durch den Trick (1.1) geschickt die definierende Eigenschaft von Primzahlen gegen die multiplikative Struktur (das Produkt) und die additive Struktur (plus 1) ausspielt.

Aber das Ergebnis ist zu wichtig, um nur einmal bewiesen zu werden. In der Hoffnung, daß uns ein weiterer Beweis mehr über Primzahlen erzählt, führen wir einen zweiten. Ein guter Beweis zeigt nämlich nicht nur, daß etwas wahr ist, sondern hebt wichtige Eigenschaften hervor und wirft neue Fragen auf.

Zweiter Beweis nach Euler.[3] Nach dem Fundamentalsatz der Arithmetik hat jede natürliche Zahl n eine eindeutige Faktorisierung in Primzahlen. Wir fassen die Primfaktoren zur gleichen Primzahl zusammen und erhalten

$$n = p_1^{m_1} \cdot \ldots \cdot p_r^{m_r}$$

mit eindeutigen Primzahlen p_i und eindeutigen Exponenten $m_i \geq 1$. Die Zahl $n = 1$ gehört zum leeren Produkt mit $r = 0$ Faktoren.

---

[1] Für die Belange der Einführung gehen wir davon aus, daß einige Begriffe, die später genauer studiert werden, bereits bekannt sind ist. Teilbarkeit ist einer davon.

[2] Euklid von Alexandria, griechischer Mathematiker, wahrscheinlich aus dem 3. Jahrhundert v. Chr.

[3] Leonhard Euler (1707–1783), Schweizer Mathematiker.

Dann erinnern wir uns noch an die geometrische Reihe für $|t| < 1$

$$\frac{1}{1-t} = 1 + t + t^2 + \ldots = \sum_{m \geq 0} t^m.$$

Nehmen wir wieder an, wir hätten eine vollständige endliche Liste $p_1, \ldots, p_r$ aller Primzahlen. Dann können wir rechnen

$$\begin{aligned} \prod_{i=1}^{r} \frac{1}{1 - 1/p_i} &= \prod_{i=1}^{r} \left(1 + \frac{1}{p_i} + \frac{1}{p_i^2} + \ldots\right) && \text{(ausmultiplizieren)} \\ &= \sum_{m_1 \geq 0} \cdots \sum_{m_r \geq 0} \frac{1}{p_1^{m_1} \cdot \ldots \cdot p_r^{m_r}} = \sum_{n \geq 1} \frac{1}{n}, && \text{(Fundamentalsatz)} \end{aligned}$$

und die letzte Summe, die harmonische Reihe, divergiert (Analysis 1) im Widerspruch zur Endlichkeit des ersten Produkts. □

Eulers Beweis beruht auf der feineren Aussage des Fundamentalsatzes der Arithmetik im Gegensatz zu Euklids Beweis, der nur die Existenz von Primfaktoren benötigt. Außerdem betont Eulers Ansatz die Analysis. Zahlentheorie steht als mathematische Dispziplin ganz oben in der Nahrungspyramide und bedient sich aller zur Verfügung stehender Methoden. Interessanterweise können auch viele Gebiete zur Zahlentheorie beisteuern. In Wahrheit ist diese Beziehung natürlich wechselseitig. Die angewandten Methoden beeinflussen auch die zahlentheoretischen Fragen, die gestellt und beantwortet werden können.

Um zu illustrieren, daß man in der Zahlentheorie nicht vor überraschenden Anwendungen aus den verschiedensten Gebieten der Mathematik sicher ist, beweisen wir die Unendlichkeit der Menge der Primzahlen mit (unanschaulichen aber formal korrekten) topologischen Methoden ein drittes Mal.

Dritter Beweis nach Fürstenberg, 1955. Auf der Menge der ganzen Zahlen $\mathbb{Z}$ definieren wir eine Topologie, und zwar die gröbste Topologie, in der alle Kongruenzklassen offen sind. Eine Teilmenge $U \subseteq \mathbb{Z}$ ist demnach offen, wenn für jedes $a \in U$ ein $m \in \mathbb{Z}$, $m \neq 0$ existiert mit

$$B_m(a) := a + m\mathbb{Z} \subseteq U.$$

Die Mengen $B_m(a)$ spielen die Rolle von kleinen (je größer m) Bällen um die Punkte a. Man überlege sich, daß diese Definition eine Topologie auf $\mathbb{Z}$ beschreibt. Es gilt nun:

- Jede nichtleere offene Menge in $\mathbb{Z}$ ist unendlich.
- Jede Menge der Form $B_a(m)$ ist offen und abgeschlossen.

Den zweiten Punkt sieht man wegen

$$B_a(m) = \mathbb{Z} \setminus \bigcup_{b=a+1}^{a+m-1} B_b(m).$$

Nehmen wir nun an, wir hätten eine vollständige endliche Liste $p_1, \ldots, p_r$ aller Primzahlen. Weil jede ganze Zahl außer ± 1 durch mindestens eine Primzahl teilbar ist, erhalten wir

$$\mathbb{Z} \setminus \bigcup_{i=1}^{r} B_0(p_i) = \{1, -1\}.$$

Diese Menge ist offen als Komplement von endlich vielen abgeschlossenen Mengen. Andererseits kann sie nicht offen sein, denn sie hat nur zwei Elemente. Widerspruch. □

1.2. **Primzahllücken.** Nachdem man nun weiß, daß es unendlich viele Primzahlen gibt, möchte man etwas über die Regelmäßigkeit oder Zufälligkeit ihres Auftretens in der Folge aller natürlichen Zahlen wissen. Beginnen wir elementar. Sei dazu

$$p_n$$

die n-te Primzahl. Also

$$p_1 = 2,\ p_2 = 3,\ p_3 = 5,\ p_4 = 7,\ p_5 = 11,\ p_6 = 13, \ldots, p_{25} = 97, \ldots$$

Es fällt auf, daß die Primzahlen anteilsmäßig immer weniger werden. Die Wahrscheinlichkeit für n, nicht durch eine der Primzahlen $p \leq n$ teilbar zu sein, ist heuristisch — wir nehmen an, daß *durch p teilbar sein* für verschiedene Primzahlen unabhängige Zufallsvariablen sind — gegeben durch

$$P(n \text{ ist Primzahl}) = \prod_{p \leq n} \left(1 - \frac{1}{p}\right),$$

und das konvergiert gegen 0, wie wir durch eine Variante zur Rechnung aus obigem Beweis Eulers sehen. Vorsicht: dies ist nur eine Heuristik!

Proposition 1.3. *Es gibt beliebig lange Lücken in der Folge der Primzahlen $(p_n)_{n \in \mathbb{N}}$. Mit anderen Worten*

$$\limsup_i\ (p_{i+1} - p_i) = \infty.$$

Beweis. Sei $n \geq 2$ eine ganze Zahl. Das Intervall $n! + 2, \ldots, n! + n$ ist garantiert frei von Primzahlen, denn für $2 \leq k \leq n$ gilt

$$n! + k = k(n \cdot \ldots \cdot (k+1) \cdot (k-1) \cdot \ldots \cdot 1 + 1),$$

somit ist k ein Teiler, der von 1 und $n!+k$ verschieden ist. Wenn p_m die größte Primzahl $\leq n!+1$ ist, dann gilt

$$p_{m+1} - p_m \geq (n! + n + 1) - (n! + 1) = n.$$

Mit einer divergenten Teilfolge haben die Differenzen $p_{i+1} - p_i$ den $\limsup$ unendlich. □

Es ist eine große Überraschung aus dem Jahr 2013, daß man auch etwas über kleine Lücken sagen kann. Das ursprüngliche Problem ist das der Zwillingsprimzahlen. Die Primzahlen 2 und 3 sind die einzigen Primzahlen, die sich nur um 1 unterscheiden. Klar, von aufeinanderfolgenden Zahlen ist stets eine gerade. Gerade Zahlen sind durch 2 teilbar, also nur dann Primzahl, wenn es sich bereits um 2 handelt.

Definition 1.4. Ein Primzahlzwilling ist ein Paar aufeinanderfolgender Primzahlen, die sich um genau 2 unterscheiden.

Die Primzahlzwillingsvermutung besagt, daß es unendlich viele Primzahlzwillinge

$$(3,5),\ (5,7),\ (11,13),\ (17,19),\ (29,31),\ (41,43),\ (71,73), \ldots$$

geben soll. Das weiß nach meinem Wissen bis heute keiner. Aber wir haben das folgende Theorem nach Vorarbeiten von Zhang.

Theorem 1.5 (Maynard, Tao 2013)**.** *Es gibt unendlich viele i mit $p_{i+1} - p_i \leq 600$. Mit andern Worten*

$$\liminf_i\ (p_{i+1} - p_i) \leq 600.$$

Das ist nun aber wirklich zu schwer für elementare Zahlentheorie. Der qualitative Sprung der durch Zhang und das Theorem von Maynard und Tao gemacht wurde, besteht in der Endlichkeit des $\liminf$. Vorher war nicht bekannt, daß es eine feste Schranke C und unendlich oft eine Primzahllücke höchstens der Länge C gibt (egal wie groß man C angesetzt hätte). Nun ist $C \leq 600$, nach einem Polymath-Projekt sogar $C \leq 246$. Allerdings ist es bis zu $C = 2$, also der Primzahlzwillingsvermutung noch ein weiter Weg, der vermutlich mit den bekannten Methoden nicht gegangen werden kann.

1.3. **Primzahlverteilung.** Der Primzahlsatz bestimmt die Asymptotik der Primzahlzählfunktion

$$\pi(x) := \#\{p \leq x \ ; \ p \text{ Primzahl}\}.$$

Dies ist eine Treppenfunktion, die jedesmal, wenn $x = p$ eine Primzahl ist, spontan um 1 ansteigt.

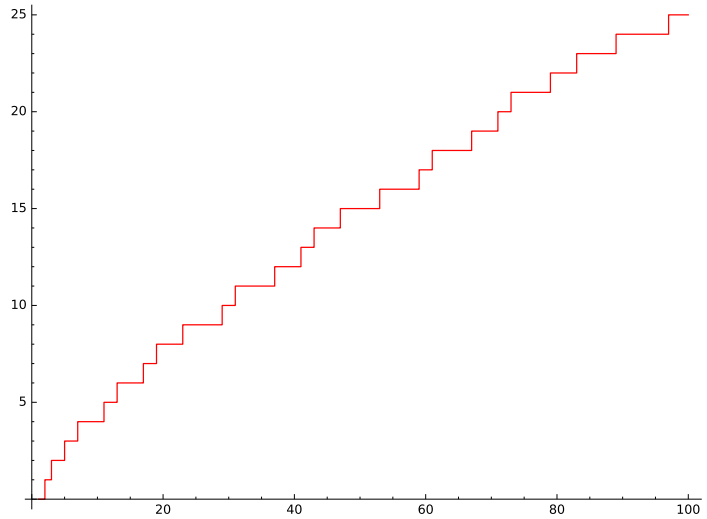


Abbildung 1. Die Primzahlzählfunktion $\pi(x)$

Schon 1793 hat der 15-jährige Gauß[4] (und unabhängig davon Legendre[5]) die Vermutung geäußert, die Primzahlzählfunktion verhielte sich für $x \to \infty$ wie $x/\ln(x)$ beziehungsweise genauer wie der Integrallogarithmus

$$\mathrm{Li}(x) = \int_2^x \frac{1}{\ln(t)} dt.$$

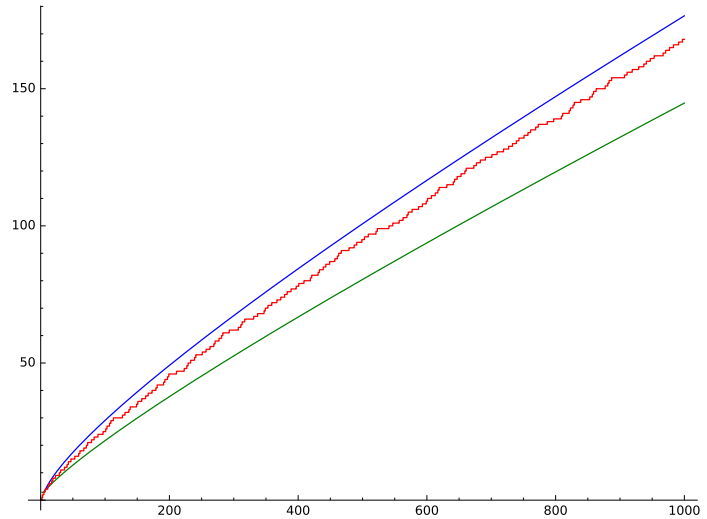


Abbildung 2. Die vorgeschlagenen Approximationen der Primzahlzählfunktion $\pi(x)$, in grün durch $x/\ln(x)$ und in blau die bessere Approximation durch $\mathrm{Li}(x)$.

[4] Johann Carl Friedrich Gauß (1777–1855), deutscher Mathematiker, Astronom, Geodät und Physiker.
[5] Adrien-Marie Legendre (1752–1833), französischer Mathematiker.

Wir schreiben $f(x) \sim g(x)$, wenn sich die Funktion $f(x)$ asymptoisch für $x \to \infty$ wie die Funktion $g(x)$ verhält. Das bedeutet, daß der Quotient den folgenden Limes hat:

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1.$$

Der Primzahlsatz bestätigt dieVermutung zur Asymptotik der Primzahlzählfunktion (wegen $\mathrm{Li}(x) \sim x/\ln(x)$ ist es egal, welche Asymptotik man beweist).

Theorem 1.6 (Primzahlsatz, Hadamard[6] und de La Vallée Poussin[7] 1896)**.** *Es gilt asymptotisch für $x \to \infty$*

$$\pi(x) \sim \frac{x}{\ln x}.$$

Aus dem Primzahlsatz folgt eine weitere Heuristik für die „Wahrscheinlichkeit“ einer natürlichen Zahl n, eine Primzahl zu sein. Dazu sein $x < n \leq 2x$ eine beliebige natürliche Zahl. Als Wahrscheinlichkeit schätzen wir den Anteil der Primzahlen im Intervall $(x, 2x]$, also

$$\begin{aligned} P(n \in (x, 2x] \text{ ist Primzahl}) = \frac{\pi(2x) - \pi(x)}{2x - x} &\approx \frac{2x/\ln(2x) - x/\ln(x)}{x} \\ &= \frac{1}{\ln(x)} \cdot \frac{\ln(x) - \ln(2)}{\ln(x) + \ln(2)} \approx \frac{1}{\ln(x)} \approx \frac{1}{\ln(n)} \end{aligned}$$

für $x \to \infty$ und $x < n \leq 2x$.

Interessanterweise weist Eulers Beweis für unendlich viele Primzahlen den Weg zum Primzahlsatz, auf welchem unabhängig voneinander Hadamard und de La Vallée Poussin den Beweis im Jahre 1896 führen konnten. Die Riemannsche Zetafunktion ist definiert als

$$\zeta(s) = \sum_{n\geq 1} \frac{1}{n^s},$$

zumindest als Funktion einer komplexen Variablen s mit $\Re(s) > 1$, denn dort konvergiert die angegebene Summe absolut auf kompakten Teilmengen (Vergleich mit geometrischer Reihe). Der Fundamentalsatz der Arithmetik führt (genau wie im obigen Beweis) zum Eulerprodukt

$$\zeta(s) = \prod_p \left(\frac{1}{1 - p^{-s}} \right).$$

Die Unendlichkeit der Menge der Primzahlen ist über Eulers Beweis direkt mit einer funktionentheoretischen Eigenschaft von $\zeta(s)$ verbunden, nämlich einem Pol bei $s = 1$. Dies zeigt bereits die enge Verzahnung von $\zeta(s)$ und der Asymptotik der Primzahlen. Wie genau die Riemannsche Zetafunktion es anstellt, mit seinen Nullstellen die Primzahlen zu zählen, werden wir hier nicht besprechen.

Ein elementarer Beweis für den Primzahlsatz wurde 1949 von Atle Selberg[8] und Paul Erdős[9] erdacht. Dabei bezieht sich *elementar* darauf, daß der Beweis ohne Methoden der Funktionentheorie auskommt, kompliziert ist er trotzdem.

1.4. **Diophantische Gleichungen.** Diophantische Gleichungen sind Gleichungen polynomialer Art, die man in ganzen Zahlen zu lösen versucht. Die bekannteste darunter ist sicher für eine feste natürliche Zahl n die Gleichung

$$X^n + Y^n = Z^n,$$

[6] Jacques Hadamard (1865–1963), französischer Mathematiker.
[7] Charles-Jean de La Vallée Poussin (1866–1962), belgischer Mathematiker.
[8] Atle Selberg (1917–2007), norwegisch-US-amerikanischer Mathematiker.
[9] Paul Erdős (1913–1996), ungarischer Mathematiker.

von der der große Amateurmathematiker Fermat[10] behauptete, daß es für $n \geq 3$ nur die trivialen ganzzahligen Lösungen (x, y, z) mit $xyz = 0$ geben kann. Seit etwa 1995 wissen wir dank Andrew Wiles (und Richard Taylor und vielen anderen Mathematikern), daß der Jurist Fermat recht hatte. Dieser Fermat'sche Große Satz spielt in der Entwicklung der modernen Algebra und Zahlentheorie eine bedeutende Rolle, nicht so sehr aus Interesse an der Aussage, sondern vielmehr als Testfrage, die über die Stärke der entwickelten Methoden Auskunft geben kann.

Fermat selbst bewies nachweislich den Fall $n = 4$ mit der von ihm entwickelten Methode des unendlichen Abstiegs. Offensichtlich reicht dazu der stärkere Satz 1.7. Die grundsätzliche Idee des unendlichen Abstiegs ist die folgende. Wir nehmen eine Lösung an und konstruieren daraus eine kleinere Lösung. Was genau kleiner bedeutet, muß man erklären, jedenfalls muß dies so beschaffen sein, daß es nur endlich viele kleinere Lösungen geben kann und man daher nicht unendlich oft absteigen kann.

Satz 1.7. *Die Gleichung $X^4 + Y^4 = Z^2$ hat keine ganzzahligen Lösungen $x, y, z \in \mathbb{Z}$ außer den trivialen mit $x = 0$ oder $y = 0$ oder $z = 0$.*

Beweis. Sei x, y, z eine nichttriviale Lösung. Dann behaupten wir, daß es eine nichttriviale Lösung x', y', z' gibt mit

$$|z'| < |z|.$$

Das geht offensichtlich (hier geht implizit wieder das Induktionsaxiom der natürlichen Zahlen ein) nur endlich oft.

Schritt 0: Indem wir notfalls z durch $-z$ ersetzen, dürfen wir $z > 0$ annehmen. Dann gilt auch

$$z = \sqrt{x^4 + y^4} \geq \max\{x^2, y^2\}.$$

Schritt 1: Wir dürfen annehmen, daß je zwei der drei Zahlen x, y und z teilerfremd sind.

Sei nämlich $d = \mathrm{ggT}(x, y)$, dann gilt

$$d^4 \mid x^4 + y^4 = z^2,$$

also $d^2 \mid z$. Wir setzen $x' = x/d$, $y' = y/d$ und $z' = z/d^2$, was wegen

$$z'^2 = \left(\frac{z}{d^2}\right)^2 = \frac{z^2}{d^4} = \frac{x^4 + y^4}{d^4} = \left(\frac{x}{d}\right)^4 + \left(\frac{y}{d}\right)^4 = x'^4 + y'^4$$

zu einer weiteren nichttrivialen ganzzahligen Lösung führt. Wenn $d > 1$, dann ist die neue Lösung kleiner. Daher nehmen wir an, daß $d = 1$ gilt. Somit sind x und y schon einmal teilerfremd.

Angenomen z und (wegen Symmetrie ohne Einschränkung) y haben einen gemeinsamen Teiler > 1, somit auch einen gemeinsamen Primteiler p. Dann teilt p auch $x^4 = z^2 - y^4$ und somit ist p Teiler von x. Dann sind aber x und y nicht teilerfremd. Widerspruch.

Schritt 2: Es sind nicht beide x und y ungerade.

Quadrate gerader Zahlen $n = 2m$ sind durch 4 teilbar:

$$n^2 = 4m^2,$$

während Quadrate ungerader Zahlen $n = 2m + 1$ den Rest 1 bei Division durch 4 lassen:

$$n^2 = (2m + 1)^2 = 4m(m + 1) + 1.$$

Wären x und y ungerade, so auch $x^2 = 2n + 1$ und $y^2 = 2m + 1$, und somit läßt

$$z^2 = x^4 + y^4 = (2n + 1)^2 + (2m + 1)^2 = 4(n(n + 1) + m(m + 1)) + 2$$

den Rest 2 bei Division durch 4. Widerspruch.

Wir nehmen daher an, daß x gerade, y ungerade und folglich z auch ungerade ist.

Schritt 3: Faktorisieren. Wir schreiben

$$x^4 = z^2 - y^4 = (z - y^2)(z + y^2)$$

[10] Pierre de Fermat (1607–1665), französischer Mathematiker und Jurist.

und studieren die Faktoren. Aus Schritt 0 folgt, daß beide Faktoren $z \pm y^2 > 0$ sind. Wir berechnen nun den ggT der Faktoren. Beide sind gerade, daher (nach dem Euklidischen Algorithmus)

$$2 \mid \mathrm{ggT}(z - y^2, z + y^2) = \mathrm{ggT}(2z, z + y^2) \mid \mathrm{ggT}(2z, 2z + 2y^2) = \mathrm{ggT}(2z, 2y^2) = 2.$$

Da in der Primfaktorzerlegung von x^4 alle Primzahlen in Vielfachen von 4 vorkommen, diese aber jeweils nur einen der beiden Faktoren $z \pm y^2$ teilen können (bis auf die 2), muß es ganze Zahlen A (ungerade) und B geben mit

$$\text{I:} \left\{ \begin{array}{rcl} z - y^2 & = & 2A^4 \\ z + y^2 & = & 8B^4 \end{array} \right. \qquad \text{oder} \qquad \text{II:} \left\{ \begin{array}{rcl} z + y^2 & = & 2A^4 \\ z - y^2 & = & 8B^4 \end{array} \right.$$

Hierbei ist $z \pm y^2 > 0$ wichtig, denn sonst könnte jeweils noch ein Minuszeichen in den Gleichungen auftreten.

Im Fall I bekommen wir

$$y^2 = (8B^4 - 2A^4)/2 = 4B^4 - A^4,$$

was bei Division durch 4 den Rest 3 läßt. Widerspruch. Es gilt also der Fall II.

Schritt 4: Nochmals faktorisieren! Aus II bekommen wir $8B^4 + y^2 = z = 2A^4 - y^2$ oder

$$4B^4 = (A^2 - y)(A^2 + y).$$

Wieder sind beide Faktoren $A^2 \pm y \geq 0$ wegen

$$A^4 = y^2 + 4B^4 \geq y^2.$$

Und wieder sind beide Faktoren gerade und

$$2 \mid \mathrm{ggT}(A^2 - y, A^2 + y) \mid 2\,\mathrm{ggT}(A^2, y) = 2,$$

denn jeder Primteiler, der in y und A^2 aufgeht, geht auch in $z = 2A^4 - y^2$ auf. Wir schließen wie vorher auf ganze Zahlen C und D mit

$$\left\{ \begin{array}{rcl} A^2 - y & = & 2C^4 \\ A^2 + y & = & 2D^4 \end{array} \right.$$

(wieder sind keine Vorzeichen nötig!), woraus wegen

$$A^2 = C^4 + D^4$$

eine neue Lösung $x' = C$, $y' = D$ und $z' = A$ der ursprünglichen Gleichung wird. Diese Lösung ist nichttrivial, denn wegen $B^4 = (CD)^4$ und $x^4 = 16(AB)^4$ folgt

$$ACD = 0 \implies AB = 0 \implies x = 0,$$

Widerspruch. Außerdem ist die neue Lösung kleiner, denn

$$|z| = |A^4 + 4B^4| \geq |A^4| \geq |A|$$

mit Gleichheit nur bei $B = 0$, was wir bereits ausgeschlossen haben. Damit ist die kleinere Lösung gefunden und der Beweis erbracht. □

Der Beweis von Satz 1.7 sieht aus wie eine Ansammlung von Tricks (und die Feinheiten der Teilbarkeitsargumente werden wir erst noch begründen müssen), die man allerdings systematisieren kann (nicht in dieser Vorlesung). Das systematische Studium Diophantischer Gleichungen ist ein weites Feld der Zahlentheorie mit Ausrichtung auf arithmetische algebraische Geometrie.

1.5. **Zahlentheorie über endlichen Körpern.** Aufgrund struktureller Ähnlichkeiten kann man Zahlentheorie nicht nur mit ganzen Zahlen sondern auch mit Polynomen betreiben. Den ganzen Zahlen am ähnlichsten sind dabei Polynome mit Koeffizienten in einem endlichen Körper, etwa in $\mathbb{F}_p$ für eine Primzahl p. Man spricht vom *geometrischen Fall* und kann in der Tat erstaunlich viel geometrische Intuition benutzten.

Die strukturelle Ähnlichkeit beginnt damit, daß $\mathbb{Z}$ und $\mathbb{F}_p[T]$ beides Hauptidealringe sind. Die „Primzahlen“ unter den Polynomen sind die irreduziblen Polynome, und diese kann man über einem endlichen Körper exakt zählen; ein Indiz, daß im *geometrischen Fall* vieles leichter ist als im *arithmetischen Fall* der ganzen Zahlen $\mathbb{Z}$.

Der geometrische Fall findet über die Theorie der Funktionenkörper Anwendungen in der Kryptographie und Codierungstheorie. Diese Anwendungen sind jahrzehnte[11] nach den ersten Schritten auf diesem Gebiet entstanden, ein starkes Plädoyer für freie Grundlagenforschung.

Literatur

[MP11] Stefan Müller-Stach, Jens Piontkowski, *Elementare und algebraische Zahlentheorie: ein moderner Zugang zu klassischen Themen*, zweite Auflage, Viehweg+Teubner, 2011, 261 Seiten.

[Leu96] Armin Leutbecher, *Zahlentheorie: eine Einführung in die Algebra*, Grundwissen Mathematik, Springer, 1996, xi+354 Seiten.

[Sch07] Alexander Schmidt, *Einführung in die algebraische Zahlentheorie*, Springer, 2007, xi+215 Seiten.

[Ser73] Jean-Pierre Serre, *A course in Arithmetic*, Springer, Graduate Texts in Mathematics **7**, Original 1973, 6. Auflage 2001, viii+115 Seiten.

[Wol11] Jürgen Wolfart, *Einführung in die Zahlentheorie und Algebra*, zweite Auflage: Viehweg+Teubner, 2011, xiii+308 Seiten.

[Za90] Don Zagier, *A One-Sentence Proof that That Every Prime* $p \equiv 1 \pmod 4$ *Is a Sum of Two Squares*, American Mathematical Monthly **97** (1990), no. 2, 144.

[11] Die Anfänge reichen bis zu Évariste Galois (1811–1832), französischer Mathematiker.

Teil 1. Arithmetik in $\mathbb{Z}$

2. Die Peanoaxiome

Jeder weiß hoffentlich, was die natürlichen und die ganzen Zahlen sind. Schließlich gilt nach einem Ausspruch Kroneckers[12] „Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk“. Wir skizzieren kurz die axiomatische Einführung.

Axiom 1 (Peanoaxiome)**.** Die natürlichen Zahlen sind ein Modell für das folgende Axiomensystem.

(1) Es gibt eine Menge $\mathbb{N}$, genannt die **natürlichen Zahlen**, und ein Element $1 \in \mathbb{N}$, **Eins** genannt.

(2) Jedes $n \in \mathbb{N}$ hat einen **Nachfolger** in $\mathbb{N}$, den wir mit $N(n)$ bezeichnen. Die Funktion

$$N : \mathbb{N} \to \mathbb{N}$$

hat die folgenden Eigenschaften:

(i) Die 1 ist kein Nachfolger: für alle $n \in \mathbb{N}$ gilt $N(n) \neq 1$.

(ii) Für alle $n, m \in \mathbb{N}$ folgt aus $N(n) = N(m)$ bereits $n = m$.

(3) **Vollständige Induktion**: Jede Teilmenge $M \subseteq \mathbb{N}$ mit

(i) $1 \in M$ und

(ii) wenn $m \in M$, dann auch $N(m) \in M$

ist bereits ganz $\mathbb{N}$.

Aus den Peanoaxiomen lassen sich Addition und Multiplikation rekursiv (nutzt das Induktionsaxiom!) definieren:

$$\begin{aligned} n + 1 &:= N(n), \\ n + N(m) &:= N(n + m), \\ n \cdot 1 &:= n, \\ n \cdot N(m) &:= n \cdot m + n. \end{aligned}$$

Die **ganzen Zahlen** $\mathbb{Z}$ sind definiert als Äquivalenzklassen von Paaren $(n, m) \in \mathbb{N} \times \mathbb{N}$ mit

$$(n, m) \sim (n', m') \ :\Longleftrightarrow \ n + m' = m + n'.$$

Die Klasse von (n, m) sei mit $[(n, m)]$ bezeichnet und der Gewohnheit folgend sei abkürzend

$$\begin{aligned} n &= [(n + 1, 1)], \\ 0 &= [(n, n)], \\ -n &= [(1, n + 1)]. \end{aligned}$$

Dann gilt $\mathbb{N} \hookrightarrow \mathbb{Z}$ mit $n \mapsto n$ und Addition und Multiplikation lassen sich wie gewohnt auf $\mathbb{Z}$ ausdehnen. Daß die so definierten Strukturen assoziativ, distributiv, kommutativ etc. sind, hat zum Glück schon einmal jemand überprüft. Wir gewinnen nichts daraus und beenden damit den axiomatischen Exkurs. Nur noch eins: es gibt eine Anordnung auf $\mathbb{Z}$

$$n < m \ :\Longleftrightarrow \ m + (-n) \in \mathbb{N}.$$

Dann gilt der wichtige folgende Satz.

Satz 2.1. *Jede nichtleere Teilmenge von $\mathbb{N}$ hat ein kleinstes Element.*

Beweis. Sei $A \subseteq \mathbb{N}$ eine Teilmenge ohne kleinstes Element. Sei M die Menge der $m \in \mathbb{N}$, so daß

$$\{n \in \mathbb{N} \ ; \ n \leq m\} \subseteq \mathbb{N} \setminus A.$$

[12] Leopold Kronecker (1823–1891), deutscher Mathematiker.

Dann gilt $1 \in M$, denn sonst wäre $1 \in A$ ein kleinstes Element. Weiter gilt mit $m \in M$ auch $m+1 \in M$, denn ansonsten wäre $m+1 \in A$ ein kleinstes Element. Nach dem Prinzip der vollständigen Induktion gilt dann $M = \mathbb{N}$ und $A = \emptyset$. Widerspruch. $\square$

Wir haben Satz 2.1 bereits im Kapitel 1 versteckt benutzt: zum Beispiel in Satz 1.7 in der Methode des unendlichen Abstiegs. Sei nämlich M die Menge der $|z|$ der nichttrivialen Lösungen. Dann hat diese Menge ein kleinstes Element, was im Beweis ad absurdum geführt wird, indem ein noch kleineres $|z|$ erzeugt wird.

Notation 2.2. Wir bezeichnen natürlich weiter die natürlichen Zahlen mit

$$\mathbb{N} = \{1, 2, 3, 4, 5, \ldots\}.$$

Es bezeichne weiter $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ und $\mathbb{Z}$ die ganzen Zahlen. Ob die 0 zu den natürlichen Zahlen gehört oder nicht, ist reine Konvention. Es tut nichts zur Sache.

3. Arithmetik der Teilbarkeit

3.1. **Teilbarkeit in ganzen Zahlen.** Nicht jede Division in $\mathbb{Z}$ geht auf, wie man in der Grundschule sagt. Dies führt zur Definition der Teilbarkeit.

Definition 3.1. Die ganze Zahl $a \in \mathbb{Z}$ **teilt** $b \in \mathbb{Z}$, wenn es ein $x \in \mathbb{Z}$ gibt mit $b = ax$. Wir schreiben dann

$$a \mid b,$$

nennen a einen **Teiler** von b und b ein **Vielfaches** von a.

Hier sind die einfachen formalen Eigenschaften der Teilbarkeitsrelation.

Proposition 3.2. *Die Teilerrelation hat die folgenden Eigenschaften. Seien a, b, c, t, b_i und t_i ganze Zahlen (für $i = 1, \ldots, n$), und $t \neq 0$. Dann gilt:*

(1) $0 \mid a$, *dann* $a = 0$,
(2) $a \mid 0$ *und* $1 \mid a$ *für jedes* a,
(3) $a \mid b$ *und* $b \mid c$, *dann* $a \mid c$,
(4) $a \mid b$, *dann* $a \mid bc$,
(5) $at \mid bt \iff a \mid b$,
(6) $a \mid b \iff a \mid -b \iff -a \mid b$,
(7) $a \mid b_i$ *für* $i = 1, \ldots, n$, *dann* $a \mid \sum_{i=1}^n t_i b_i$,
(8) $a \mid b$, *dann* $b = 0$ *oder* $|a| \leq |b|$,
(9) $a \mid b$ *und* $b \mid a$, *dann* $a = \pm b$.

Beweis. Der Beweis ist eine einfache Übung. Sie sollten die Beweise selbst ausarbeiten und nur zur Not hier nachschauen.

(1) Wenn $0 \mid a$, dann gibt es $x \in \mathbb{Z}$ mit $a = 0 \cdot x = 0$. (2) Es gilt $a \cdot 0 = 0$ und $1 \cdot a = a$ für jedes $a \in \mathbb{Z}$. (3) Wenn $a \mid b$ und $b \mid c$, dann gibt es $x, y \in \mathbb{Z}$ mit $b = ax$ und $c = by$, also gilt auch $c = by = a(xy)$ und somit $a \mid c$. (4) Weil offensichtlich $b \mid bc$ folgt die Ausssage sofort aus Aussage (3). Für (5) überlegen wir

$$at \mid bt \iff \exists x \in \mathbb{Z}: \ bt = atx \iff \exists x \in \mathbb{Z}: \ b = ax \iff a \mid b.$$

(6) ist eine formale Folgerung aus (4) und (5) jeweils für $t = -1$.

(7) Nach Voraussetzung gibt es $y_i \in \mathbb{Z}$ mit $b_i = ax_i$. Dann ist mit

$$x = \sum_{i=1}^n t_i x_i$$

wegen des Distributivgesetzes auch a ein Teiler von $ax = \sum_{i=1}^n t_i(ax_i) = \sum_{i=1}^n t_i b_i$.

(8) Sei $b = ax$ mit $x \in \mathbb{Z}$. Wenn $x = 0$, dann ist $b = 0$. Andernfalls ist $|x| \geq 1$ und daher $|b| = |a| \cdot |x| \geq |a|$.

(9) Wenn weder a noch b gleich 0 ist, dann folgt aus (8) $|a| \leq |b| \leq |a|$ und daher $a = \pm b$. Andernfalls, wenn oBdA $a = 0$, dann ist auch $b = 0$ wegen (1). Wieder gilt $a = \pm b$. □

Korollar 3.3. *Sei $n \neq 0$ eine ganze Zahl. Dann hat n nur endlich viele Teiler.*

Beweis. Nach Proposition 3.2 (8) sind die Teiler von n in $\{d;\ -n \leq d \leq n\}$ enthalten. □

Dieses offensichtliche Korollar macht die folgende Definition wohldefiniert, denn sie garantiert die Existenz des Maximums.

Definition 3.4. Der **größte gemeinsame Teiler (ggT)** von $a, b \in \mathbb{Z}$ ist die natürliche Zahl

$$d = \max\{t\ ;\ t \mid a \text{ und } t \mid b\},$$

wenn $a \neq 0$ oder $b \neq 0$ gilt, oder 0, wenn $a = b = 0$. Der größte gemeinsame Teiler ist wohldefiniert, denn $\{t\ ;\ t \mid a \text{ und } t \mid b\}$ ist bei $a \neq 0$ oder $b \neq 0$ eine endliche nicht-leere Menge (sie enthält 1) nach Korollar 3.3 und besitzt daher ein Maximum. Wir schreiben

$$(a, b) := \mathrm{ggT}(a, b) := d$$

für den größten gemeinsamen Teiler d von a und b.

Allgemeiner ist für $a_1, \ldots, a_n \in \mathbb{Z}$ der größte gemeinsame Teiler 0, wenn alle $a_i = 0$, und sonst

$$(a_1, \ldots, a_n) := \mathrm{ggT}(a_1, \ldots, a_n) := \max\{t \in \mathbb{Z}\ ;\ t \mid a_i \text{ für alle } i = 1, \ldots, n\}.$$

Proposition 3.5. *Der ggT erfüllt die folgenden formalen Eigenschaften. Für alle $a, b, n \in \mathbb{Z}$ gilt:*

(1) $(a, b) = (b, a)$,
(2) $(a, b) = (a, -b)$,
(3) $(0, b) = |b|$,
(4) $(a, b) = (a, b - na)$.

Beweis. (1) Die Definition des ggT ist symmetrisch, (2) folgt aus Proposition 3.2 (6), und (3) aus Proposition 3.2 (2), weil $|b|$ nach Proposition 3.2 (8) der größte Teiler von b ist.

(4) Wenn $d \mid a$ und $d \mid b$, dann gilt auch $d \mid b - na$ nach Proposition 3.2 (7). Und wenn $d \mid a$ und $d \mid b - na$, dann gilt mit demselben Argument auch

$$d \mid (b - na) - (-n)a = b.$$

Damit sind (a, b) und $(a, b - na)$ gleich als Maximum derselben Menge gemeinsamer Teiler. □

3.2. **Der Euklidische Algorithmus.** Eine fundamentale arithmetische Eigenschaft der ganzen Zahlen ist die Division mit Rest.

Satz 3.6. *Seien $a, b \in \mathbb{Z}$ und $a \neq 0$. Dann gibt es eindeutig $q, r \in \mathbb{Z}$ mit*

(i) $b = qa + r$,
(ii) $0 \leq r < |a|$.

*Wir bezeichnen q als das Ergebnis der **Ganzzahldivision** von b durch a und r als den **Rest** bei dieser Ganzzahldivision.*

Beweis. Die Teilmenge $\mathscr{R} \subseteq \mathbb{Z}$ definiert durch

$$\mathscr{R} = \{b - na\ ;\ n \in \mathbb{Z} \text{ und } b - na \geq 0\}$$

ist nicht leer[13]. In der Tat ist $b - ba \in \mathscr{R}$ oder $b + ba = b - (-b)a \in \mathscr{R}$, denn andernfalls wäre $b - ba < 0$ und $b + ba < 0$, also wäre das Produkt echt positiv:

$$0 < (b - ba)(b + ba) = b^2(1 - a)(1 + a) = b^2(1 - a^2).$$

[13]Ein klarerer Beweis benutzt Analysis: mittels des Archimedischen Prinzips zeigt man, daß für $a > 0$ bei $n \ll 0$ (und für $a < 0$ bei $n \gg 0$) die Zahlen $b - na \geq 0$ werden. Der im Text angegebene Beweis bleibt innerhalb der ganzen Zahlen.

Weil b^2 stets positiv ist, folgt $1 - a^2 > 0$, und damit $-1 < a < 1$, ein Widerspruch zu $a \neq 0$.

Wir setzen $r = \min \mathscr{R}$ und $q = (b - r)/a$, was wegen $r \in \mathscr{R}$ eine ganze Zahl ist. Es gilt dann $b = qa + r$ und $r \geq 0$. Angenommen $r \geq |a|$, dann wäre $0 \leq r - |a| = b - (q \pm 1)a$ (mit Vorzeichen je nach Vorzeichen von a), somit $r - |a| \in \mathscr{R}$ ein Widerspruch zur Minimalität von r. Dies zeigt die Existenz.

Zur Eindeutigkeit nehmen wir an, es gäbe auch noch $q', r' \in \mathbb{Z}$ mit den geforderten Eigenschaften. Dann ist oBdA $r' \geq r$ und dann

$$r' - r = (b - q'a) - (b - qa) = (q - q')a,$$

also a ein Teiler von $r' - r$. Wenn $r' = r$, so folgt auch $q' = q$. Sei also $r' \neq r$, dann gilt nach Proposition 3.2 (8)

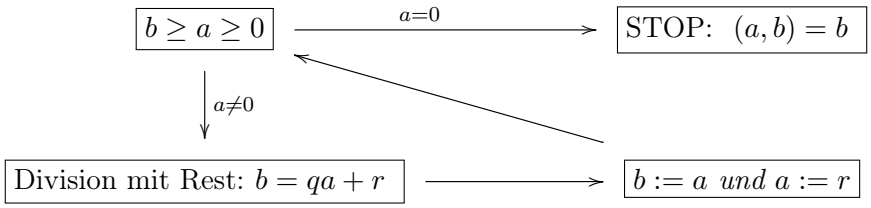
$$|a| \leq |r' - r| = r' - r \leq r' < |a|,$$

Widerspruch. □

Bemerkung 3.7. Division mit Rest etabliert auf dem Ring $\mathbb{Z}$ die Struktur eines euklidischen Rings. Solche sind Hauptidealringe. Jedes Ideal von $\mathbb{Z}$ ist also ein Hauptideal. Diese wichtige arithmetische Eigenschaft ist vom algebraischen Standpunkt äquivalent zum Fundamentalsatz der Arithmetik, dem wir in Theorem 5.8 begegnen.

Die euklidische Struktur bringt zusätzliche algorithmische Vorteile, wie wir gleich sehen: auf Satz 3.6 und Proposition 3.5 (4) beruht die effektive schnelle Berechnung des ggT. Nach Proposition 3.5 (2) dürfen wir uns auf den ggT von natürlichen Zahlen beschränken.

Satz 3.8 (Euklidischer Algorithmus)**.** *Seien $b \geq a \geq 0$ und $b \neq 0$ ganze Zahlen. Dann berechnet der folgende Algorithmus den größten gemeinsamen Teiler (a, b):*



Beweis. Der Algorithmus ist wohldefiniert, denn Division mit Rest führt zu $0 \leq r < a$, so daß die neuen Werte für a, b wieder die oBdA-Annahme $b \geq a \geq 0$ erfüllen.

Der Algorithmus **terminiert**, denn im Iterationsschritt gilt $b + a > a + r \geq 0$. Die Summe wird also stets kleiner. Da es nur endlich viele natürliche Zahlen unterhalb einer gegebenen gibt, muß nach endlich vielen Schritten die Abbruchbedingung des Algorithmus erreicht werden.

Der Algorithmus ist **korrekt**, das heißt, er berechnet, was er zu berechnen vorgibt. Das folgt sofort aus Proposition 3.5. □

Bemerkung 3.9. In [Wol11] findet man eine Laufzeitabschätzung für den Euklidischen Algorithmus mit einer Laufzeit, die logarithmisch in den Eingabedaten ist. Das ist recht gut. Man vergleiche mit dem naiven Durchprobieren aller Zahlen $d \leq \min\{|a|, |b|\}$ als Teiler von a und b.

Lemma 3.10 (Lemma von Bézout[14])**.** *Sei $a, b \in \mathbb{Z}$ und $d = (a, b)$. Dann gibt es $x, y \in \mathbb{Z}$ mit*

$$d = xa + yb.$$

Beweis. Indem wir notfalls a oder b durch sein negatives $-a$ oder $-b$ ersetzen und eventuell a mit b tauschen, dürfen wir ohne Beschränkung der Allgemeinheit annehmen, daß $0 \leq a \leq b$ gilt.

[14] Étienne Bézout (1730–1783), französischer Mathematiker.

Sei (a_i, b_i) die Folge der Paare aus dem Euklidischen Algorithmus zu den Startwerten $a_1 = a$ und $b_1 = b$. Seien weiter

$$\begin{aligned} x_0 &= 0 \text{ und } y_0 = 1, \\ x_1 &= 1 \text{ und } y_1 = 0. \end{aligned}$$

Wir berechnen dann rekursiv aus der Division mit Rest $b_i = q_i a_i + r_i$ (solange $a_i > 0$)

$$\begin{aligned} b_{i+1} &= a_i \\ a_{i+1} &= r_i \\ x_{i+1} &= x_{i-1} - q_i x_i \\ y_{i+1} &= y_{i-1} - q_i y_i. \end{aligned} \qquad (i \geq 1)$$

Wir zeigen dann per Induktion, daß für alle $i \geq 0$

$$b_{i+1} = x_i a + y_i b.$$

Für $i = 0, 1$ folgt dies aus den Startwerten und $b_2 = a_1 = a$. Für $i + 1 \geq 2$ gilt dann

$$\begin{aligned} x_{i+1}a + y_{i+1}b &= (x_{i-1} - q_i x_i)a + (y_{i-1} - q_i y_i)b \\ &= (x_{i-1}a + y_{i-1}b) - q_i(x_i a + y_i b) \\ &= b_i - q_i b_{i+1} = b_i - q_i a_i = r_i = a_{i+1} = b_{i+2}. \end{aligned}$$

Wenn der Algorithmus nach n Schritten bei (a_{n+1}, b_{n+1}) abbricht, dann gilt $a_{n+1} = 0$ und

$$d = (a, b) = b_{n+1} = x_n a + y_n b.$$

Dies beweist das Lemma von Bézout konstruktiv: mögliche Koeffizienten x, y der $\mathbb{Z}$-Linearkombination $d = xa + yb$, werden vom Euklidischen Algorithmus in der obigen Form mitberechnet. □

Bemerkung 3.11. Wir können die Formeln aus dem Beweis des Lemma 3.10 kompakter in der folgenden Weise in Matrixform schreiben. Wir fassen die Zwischenschritte a_i, b_i, x_i, y_i, q_i zu Vektoren und Matrizen in $\mathbb{R}^2$ bzw. $\mathrm{M}_2(\mathbb{R})$ zusammen (genauer in $\mathbb{Z}^2$ und $\mathrm{M}_2(\mathbb{Z})$), indem wir setzen:

$$\begin{aligned} v_i &= \begin{pmatrix} a_i \\ b_i \end{pmatrix} && 1 \leq i \leq n+1, \\ A_i &= \begin{pmatrix} -q_i & 1 \\ 1 & 0 \end{pmatrix} && 1 \leq i \leq n, \\ M_i &= \begin{pmatrix} x_{i+1} & y_{i+1} \\ x_i & y_i \end{pmatrix} && 0 \leq i \leq n. \end{aligned}$$

Dann gilt $v_1 = \begin{pmatrix} a \\ b \end{pmatrix}$ und $M_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, und für alle $1 \leq i \geq n$ gilt

$$\begin{aligned} v_{i+1} &= \begin{pmatrix} a_{i+1} \\ b_{i+1} \end{pmatrix} = \begin{pmatrix} b_i - q_i a_i \\ a_i \end{pmatrix} = \begin{pmatrix} -q_i & 1 \\ 1 & 0 \end{pmatrix} v_i = A_i v_i, \\ M_i &= \begin{pmatrix} x_{i-1} - q_i x_i & y_{i-1} - q_i y_i \\ x_i & y_i \end{pmatrix} = \begin{pmatrix} -q_i & 1 \\ 1 & 0 \end{pmatrix} M_{i-1} = A_i M_{i-1}. \end{aligned}$$

Per Induktion nach i folgt für alle $0 \leq i \leq n$

$$M_i = A_i \cdot \ldots \cdot A_1 \quad \text{und} \quad v_{i+1} = M_i v_1.$$

Mit $d = \mathrm{ggT}(a, b)$ und Abbruch des euklidischen Algorithmus nach n Schritten folgt

$$\begin{pmatrix} 0 \\ d \end{pmatrix} = v_{n+1} = M_n v_1 = M_n \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} * & * \\ x_n & y_n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} * \\ x_n a + y_n b \end{pmatrix},$$

also wieder die den ggT darstellende Linearkombination $d = x_n a + y_n b$.

Das Lemma von Bézout hat eine Reihe von Korollaren. Zuerst stellen wir fest, daß für eine Menge von Zahlen der vom Absolutbetrag her größte gemeinsame Teiler auch der bezüglich Teilbarkeit größte gemeinsame Teiler ist.

Korollar 3.12. *Seien $a, b, t \in \mathbb{Z}$ und $d = (a, b)$. Dann gilt*

$$t \mid a \text{ und } t \mid b \iff t \mid d.$$

Beweis. Wenn $t \mid d$, dann gilt wegen $d \mid a$ und $d \mid b$ aus Proposition 3.2 (3) schon $t \mid a$ und $t \mid b$.

Sei umgekehrt $t \mid a$ und $t \mid b$. Nach Lemma 3.10 wählen wir $x, y \in \mathbb{Z}$ mit $d = ax + by$. Dann folgt aus Proposition 3.2 (7) schon $t \mid ax + by = d$. □

Bemerkung 3.13. Die Charakterisierung des ggT als bezüglich Teilbarkeit größter gemeinsamer Teiler nach Korollar 3.12 führt zusammen mit der Transitivität des Teilens nach Proposition 3.2 (3) für ganze Zahlen $a_1, \dots, a_n \in \mathbb{Z}$ zu

$$(a_1, \dots, a_n) = (a_1, \dots, a_{n-2}, (a_{n-1}, a_n)) = \Big(a_1, \big(a_2, \dots, a_{n-2}, (a_{n-1}, a_n) \dots\big)\Big).$$

Die zweite Gleichung folgt per Induktion aus der ersten. Die erste Gleichung ist bei $a_n = 0$ trivial und folgt bei $a_n \neq 0$ aus Korollar 3.12, weil damit

$$\begin{aligned}(a_1, \dots, a_n) &= \max\{t \in \mathbb{Z}\ ;\ t \mid a_i \text{ für alle } i = 1, \dots, n\}\\ &= \max\{t \in \mathbb{Z}\ ;\ t \mid a_i \text{ für alle } i = 1, \dots, n-2 \text{ und } t \mid (a_{n-1}, a_n)\}\\ &= (a_1, \dots, a_{n-2}, (a_{n-1}, a_n)).\end{aligned}$$

Iteriertes Anwenden des Euklidischen Algorithmus führt somit zu einer algorithmischen Bestimmung des ggTs von mehr als zwei ganzen Zahlen.

Genauso zeigt man, daß gilt

$$t \mid (a_1, \dots, a_n) \iff \text{für alle } i = 1, \dots, n:\ t \mid a_i.$$

Korollar 3.14. *Seien $a, b, t \in \mathbb{Z}$ und $t \neq 0$. Dann gilt*

$$(ta, tb) = |t| \cdot (a, b).$$

Beweis. Wegen $t \mid ta$ und $t \mid tb$ folgt $t \mid \mathrm{ggT}(ta, tb)$ aus Korollar 3.12. Daher gibt es $e \in \mathbb{Z}$ mit

$$\mathrm{ggT}(ta, tb) = |t| \cdot e.$$

Aus $te \mid ta$ und $te \mid tb$ folgt nach Proposition 3.2 (5) bereits $e \mid a$ und $e \mid b$, also aus Korollar 3.12 auch $e \mid \mathrm{ggT}(a, b) =: d$. Damit gilt $e \leq d$, und es reicht zu zeigen, daß td ein gemeinsamer Teiler von ta und tb ist. Das ist offensichtlich. □

Bemerkung 3.15. Alternativ kann man Korollar 3.14 auch dadurch beweisen, daß man den euklidischen Algorithmus mit den Startwerten a, b und den Startwerten ta, tb vergleicht.

3.3. **Kleinstes gemeinsames Vielfaches.** In die im Vergleich zum ggT entgegengesetzte Richtung bewegt sich das kleinste gemeinsame Vielfache.

Definition 3.16. Das **kleinste gemeinsame Vielfache (kgV)** ganzer Zahlen $a, b \neq 0$ ist die natürliche Zahl

$$D = \min\{T \in \mathbb{Z}\ ;\ T > 0 \text{ und } a \mid T \text{ und } b \mid T\}.$$

Wir schreiben $[a, b] = \mathrm{kgV}(a, b)$ für das kleinste gemeinsame Vielfache von a und b.

Allgemeiner ist für $a_1, \dots, a_n \in \mathbb{Z}$, alle $a_i \neq 0$, das kleinste gemeinsame Vielfache

$$[a_1, \dots, a_n] = \mathrm{kgV}(a_1, \dots, a_n) = \min\{T \in \mathbb{Z}\ ;\ T > 0 \text{ und } a_i \mid T \text{ für alle } i = 1, \dots, n\}.$$

Das kgV ist wohldefiniert, weil die Menge der positiven gemeinsamen Vielfachen nicht leer ist: sie enthält $|ab|$, oder bei mehreren Zahlen

$$|a_1 \cdot \ldots \cdot a_n|,$$

und weil nach Satz 2.1 jede nichtleere Menge natürlicher Zahlen ein Minimum hat.

Wie beim ggT hat das kgV zunächst eine Definition (als Minimum) über die Anordnung. Aber auch hier gilt, daß das kgV auch minimal ist bezüglich Teilbarkeit.

Proposition 3.17. *Seien $a, b \neq 0$ ganze Zahlen. Dann ist jedes gemeinsame Vielfache von a und b ein Vielfaches des* $\mathrm{kgV}(a, b)$.

Beweis. Sei $D = [a, b]$ und sei $a \mid T$ und $b \mid T$. Wir teilen T durch D mit Rest und erhalten $T = qD + r$ mit $0 \leq r < D$ und $q \in \mathbb{Z}$. Dann ist $a \mid r = T - qD$ und genauso $b \mid r$. Damit kann D nur dann minimal unter den positiven gemeinsamen Vielfachen sein, wenn $r = 0$. Dann ist aber $T = qD$ ein Vielfaches von D wie behauptet. □

Korollar 3.18. *Seien $a, b, t \neq 0$ ganze Zahlen. Dann gilt*

$$[at, bt] = [a, b] \cdot |t|.$$

Beweis. Ohne Einschränkung ist $t > 0$. Ansonsten ersetzen wir t durch $-t$.

Sei $D = [a, b]$. Aus $a \mid D$ und $b \mid D$ folgt wegen Proposition 3.2 (5) auch $at \mid Dt$ und $bt \mid Dt$. Daher ist $Dt = [a, b]t$ ein gemeinsames Vielfaches von at und bt.

Sei nun $T > 0$ ein gemeinsames Vielfaches von at und bt. Da $t \mid at \mid T$, gibt es $x \in \mathbb{Z}$ mit $T = xt$. Aus $at \mid T = tx$ folgt $a \mid x$ wegen Proposition 3.2 (5), und genauso folgt $b \mid x$. Daher ist x ein gemeinsames Vielfaches von a und b, also $D \leq x$. Folglich ist auch

$$Dt \leq xt = T,$$

und somit ist Dt das kleinste gemeinsame Vielfache von at und bt. Das war zu zeigen. □

3.4. **Teilerfremde Zahlen.** Wenn ganze Zahlen nur die offensichtlichen gemeinsamen Teiler ± 1 haben, dann nennen wir sie teilerfremd.

Definition 3.19. Zwei ganze Zahlen $a, b \in \mathbb{Z}$ heißen **teilerfremd**, wenn $(a, b) = 1$.

Lemma 3.20. *Seien $a, b \in \mathbb{Z}$, nicht beide 0, und sei $d = \mathrm{ggT}(a, b)$. Dann gibt es $\alpha, \beta \in \mathbb{Z}$ mit*

$$a = d\alpha, \quad b = d\beta$$

und α und β sind teilerfremd: $(\alpha, \beta) = 1$.

Beweis. Der ggT d teilt a und b, also gibt es α und β mit $a = d\alpha$ und $b = d\beta$. Aus Korollar 3.14 folgt sofort $(\alpha, \beta) = \frac{1}{d}(a, b) = 1$, somit sind α und β teilerfremd. □

Proposition 3.21. *Seien a, b, c ganze Zahlen und a und b teilerfremd. Dann gilt*

$$a \mid bc \Longrightarrow a \mid c.$$

Beweis. Nach Lemma 3.10 wählen wir $x, y \in \mathbb{Z}$ mit $1 = ax + by$. Weiter sei $bc = za$. Dann gilt

$$a \mid a(cx + zy) = c(ax + by) = c.$$ □

Lemma 3.22. *Seien $a, b \neq 0$ teilerfremde ganze Zahlen und sei $m \in \mathbb{Z}$ ein gemeinsames Vielfaches: $a \mid m$ und $b \mid m$. Dann gilt*

$$ab \mid m.$$

Beweis. Nach Lemma 3.10 wählen wir $x, y \in \mathbb{Z}$ mit $1 = ax + by$. Außerdem seien $u, v \in \mathbb{Z}$ mit $m = au = bv$. Dann folgt

$$m = m \cdot (ax + by) = m(ax) + m(by) = (bv)(ax) + (au)(by) = ab(vx + uy),$$

somit wird m von ab geteilt. □

Korollar 3.23. *Seien $a, b \neq 0$ teilerfremde ganze Zahlen. Dann gilt*

$$\mathrm{kgV}(a, b) = |ab|.$$

Beweis. Das folgt aus Lemma 3.22, der Definition des kgV sowie Proposition 3.2 (8). □

Satz 3.24. *Seien $a, b \neq 0$ ganze Zahlen. Dann gilt*

$$\mathrm{ggT}(a,b) \cdot \mathrm{kgV}(a,b) = |ab|.$$

Beweis. Sei $t \neq 0$. Dann gilt der Satz wegen Korollar 3.14 und Korollar 3.18 für a, b genau dann, wenn er für ta, tb gilt: beide Seiten skalieren sich mit $|t|^2$.

Sei $d = (a, b)$ und seien $\alpha, \beta \in \mathbb{Z}$ mit $a = d\alpha$ und $b = d\beta$ wie in Lemma 3.20 also mit $(\alpha, \beta) = 1$. Es reicht demnach, den Satz für die teilerfremden α und β zu beweisen. Dieser Fall ist nichts anderes als Korollar 3.23. □

3.5. **Lineare Gleichungen.** Wir sind nun in der Lage, eine lineare diophantische Gleichung zu diskutieren. Bei diophantischen Gleichungen interessiert man sich für (in einem gewissen Sinne) ganzzahlige Lösungen der Gleichungen.

Satz 3.25. *Seien $a, b, c \in \mathbb{Z}$.*

(1) Die Gleichung in den Variablen X, Y

$$c = aX + bY \tag{3.1}$$

hat genau dann eine Lösung $x, y \in \mathbb{Z}$, wenn c ein Vielfaches von $d = \mathrm{ggT}(a, b)$ ist.

(2) Wenn $c = d\gamma$ gilt und $d = ua + vb$ eine $\mathbb{Z}$-Linearkombination wie in Lemma 3.10 ist, dann erhält man mit $x_0 = u\gamma$ und $y_0 = v\gamma$ eine ganzzahlige Lösung von (3.1).

(3) Seien $a = d\alpha$ und $b = d\beta$ wie in Lemma 3.20, und sei $x_0, y_0 \in \mathbb{Z}$ eine Lösung von (3.1). Dann ist die Menge aller Lösungen gegeben durch

$$\{(x, y)\ ;\ x = x_0 + n\beta, y = y_0 - n\alpha \text{ für ein } n \in \mathbb{Z}\}.$$

Beweis. (1) Wenn $x, y \in \mathbb{Z}$ eine Lösung ist, dann gilt $d \mid ax + by = c$. Die Umkehrung folgt aus Behauptung (2) und damit der trivialen Rechnung

$$c = d\gamma = (ua + vb)\gamma = a(u\gamma) + b(v\gamma) = ax_0 + by_0.$$

(3) Sei x, y eine weitere ganzzahlige Lösung neben x_0, y_0. Dann gilt

$$\alpha(x - x_0) = \frac{1}{d}a(x - x_0) = \frac{1}{d}b(y_0 - y) = \beta(y_0 - y).$$

Aus Korollar 3.14 folgt $(\alpha, \beta) = 1$. Nach Proposition 3.21 folgt daher

$$\alpha \mid y_0 - y,$$

somit gibt es $n \in \mathbb{Z}$ mit $y_0 - y = n\alpha$. Daraus folgt sofort $x - x_0 = n\beta$. Dies zeigt, daß alle Lösungen die angegebene Form haben. Umgekehrt, wenn wir durch $x = x_0 + n\beta$ und $y = y_0 - n\alpha$ ganze Zahlen x, y definieren, dann lösen auch diese die Gleichung $c = aX + bY$, wie man durch leichtes Nachrechnen verifiziert. □

Bemerkung 3.26. Die ganzen Zahlen der Form $ax + by$ mit $x, y \in \mathbb{Z}$ sind das von a und b erzeugte Ideal $(a, b) \subseteq \mathbb{Z}$. Dies sind genau die Zahlen c, für die $c = aX + bY$ Lösungen aus $\mathbb{Z}$ besitzt. Wir haben also gezeigt, daß

$$(a, b) = (d)$$

als Ideale, wenn $d = (a, b)$ der ggT ist. Weil $\mathbb{Z}$ ein Hauptidealring ist, gibt es a priori ein $t \in \mathbb{Z}$ mit $(a, b) = (t)$ als Ideale. Es folgt sofort aus Proposition 3.2 (7), daß $d \mid t$, und aus Lemma 3.10, daß $d \in (a, b) = (t)$, also $t \mid d$. Zusammengenommen folgt $d = \pm t$ oder eben $(d) = (t) = (a, b)$.

4. Kongruenzen

4.1. **Arithmetik modulo m.** Die arithmetischen Eigenschaften der Teilbarkeit lassen sich einfacher mit dem Kalkül der Kongruenzrechnung gebrauchen.

Definition 4.1. Sei $m \geq 1$ eine ganze Zahl. Die Zahlen $a, b \in \mathbb{Z}$ sind **kongruent modulo** m, wenn

$$m \mid a - b.$$

Wir schreiben dann

$$a \equiv b \pmod{m}$$

oder auch einfacher

$$a \equiv b \ (m).$$

Die Zahl m heißt manchmal der **Modulus** der **Kongruenz** $a \equiv b$. Teilbarkeit $m \mid a$ wird zu

$$a \equiv 0 \pmod{m}.$$

Bemerkung 4.2. Sei $m \geq 1$ eine ganze Zahl. Die Kongruenzrelation $\equiv \pmod{m}$ ist eine Äquivalenzrelation, denn es gilt für alle $a, b, c \in \mathbb{Z}$

(i) $a \equiv a$,
(ii) $a \equiv b \implies b \equiv a$,
(iii) $a \equiv b$ und $b \equiv c \implies a \equiv c$.

Das folgt aus $m \mid (a-a)$, und $m \mid (a-b) \iff m \mid (b-a)$, und wenn $m \mid (a-b)$ und $m \mid (b-c)$, dann auch

$$m \mid (a-b) + (b-c) = (a-c).$$

Die Kongruenzrelation $\equiv \pmod{m}$ hat die folgenden arithmetischen Eigenschaften. Diese führen zum Kalkül der Kongruenzrechnung.

Proposition 4.3. *Sei $m \geq 1$ und seien $a, a', b, b' \in \mathbb{Z}$ mit $a \equiv a'$ und $b \equiv b' \pmod{m}$. Dann gilt auch*

(1) $a + b \equiv a' + b' \pmod{m}$,
(2) $a - b \equiv a' - b' \pmod{m}$,
(3) $a \cdot b \equiv a' \cdot b' \pmod{m}$.
(4) *Für alle $n \in \mathbb{N}_0$ gilt $a^n \equiv a'^n \pmod{m}$.*

Beweis. Nach Voraussetzung teilt m sowohl $a - a'$ als auch $b - b'$. Damit gilt auch

$$m \mid (a - a') \pm (b - b') = (a \pm b) - (a' \pm b'),$$

und

$$m \mid a(b - b') + b'(a - a') = a \cdot b - a' \cdot b'.$$

Eigenschaft (4) folgt per Induktion nach n aus (3). □

Bemerkung 4.4. Die Division ist nicht so unproblematisch: es gilt $2 \cdot 3 \equiv 2 \cdot 8 \pmod{10}$, aber nicht $3 \equiv 8 \pmod{10}$.

Proposition 4.5. *Sei $m \geq 1$ und $a, b, t \in \mathbb{Z}$.*

(1) *$at \equiv bt \pmod{mt}$ ist äquivalent zu $a \equiv b \pmod{m}$.*
(2) *Wenn $(t, m) = 1$, dann ist $at \equiv bt \pmod{m}$ äquivalent zu $a \equiv b \pmod{m}$.*

Beweis. Aussage (1) folgt aus

$$mt \mid at - bt = t(a - b) \iff m \mid a - b.$$

Aussage (2) folgt wegen $(m, t) = 1$ mit Proposition 3.21 aus

$$m \mid t(a - b) \iff m \mid (a - b).$$ □

Wir lösen nun eine lineare Kongruenz-Gleichung.

Satz 4.6. *Sei $m \geq 1$ und $a, c \in \mathbb{Z}$.*

(1) Die Gleichung in der Variablen X

$$aX \equiv c \pmod{m} \tag{4.1}$$

hat genau dann eine ganzzahlige Lösung $x \in \mathbb{Z}$, wenn $d := (a, m) \mid c$.

(2) Wenn $d \mid c$ gilt und $u, v, \gamma \in \mathbb{Z}$ mit $d = ua + vm$ und $c = d\gamma$ sind, dann ist $x_0 = u\gamma$ eine ganzzahlige Lösung von (4.1).

(3) Sei $m = d\mu$ mit $\mu \in \mathbb{Z}$. Dann gibt es ein bis auf Kongruenz modulo μ eindeutiges $\xi \in \mathbb{Z}$ mit $a\xi \equiv c \pmod{m}$.

Beweis. Ganzzahlige Lösungen von (4.1) entsprechen eineindeutig den ganzzahligen Lösungen der Gleichung $aX + mY = c$. Damit folgen die Aussagen sofort aus Satz 3.25. □

Korollar 4.7. *Sei $m \geq 1$ und $a, c \in \mathbb{Z}$. Wenn a und m teilerfremd sind, dann gibt es bis auf Kongruenz modulo m ein eindeutiges $x \in \mathbb{Z}$ mit*

$$ax \equiv c \pmod{m}.$$

Beweis. Das folgt sofort aus Satz 4.6. □

Bemerkung 4.8. Sei $m \geq 1$ fixiert. Zu $(b, m) = 1$ gibt es nach Korollar 4.7 ein b' mit $bb' \equiv 1 \pmod{m}$, und dieses b' ist eindeutig bis auf Kongruenz modulo m. Wir setzen

$$\frac{a}{b} := ab' \pmod{m}.$$

Dies ist wohldefiniert als Äquivalenzklasse modulo m und verträglich mit den üblichen arithmetischen Operationen Addition und Multiplikation. Wenn $(c, m) = 1$ und $cc' \equiv 1 \pmod{m}$, dann ist modulo m

- wohldefiniert (a/b modulo m hängt nur ab von $a/b \in \mathbb{Q}$ als Bruch):
$$\frac{ac}{bc} \equiv ac(b'c') \equiv (ab')(cc') \equiv ab' \equiv \frac{a}{b} \pmod{m}.$$
- Addition:
$$\frac{a}{b} + \frac{d}{c} \equiv ab' + dc' \equiv (ac + db)(c'b') \equiv \frac{ac + db}{bc} \pmod{m},$$
was der Addition der Brüche entspricht.
- Multiplikation:
$$\frac{a}{b} \cdot \frac{d}{c} \equiv (ab') \cdot (dc') \equiv (ad)(b'c') \equiv \frac{ad}{bc} \pmod{m},$$
was der Multiplikation der Brüche entspricht.

Wir bemerken, daß sowohl Addition als auch Multiplikation den Bereich der Brüche, deren Nenner teilerfremd zu m sind, nicht verlassen. Von nun an erlauben wir also auch solche Brüche in Kongruenzgleichungen.

Bemerkung 4.9. Sei $m \geq 1$. Division mit Rest zeigt, daß jedes $a \in \mathbb{Z}$ zu genau einem $r \in \mathbb{Z}$ mit $0 \leq r \leq m - 1$ kongruent ist modulo m. Es gibt nämlich $q, r \in \mathbb{Z}$ mit $a = qm + r$, wobei $0 \leq r \leq m-1$, und daher $a \equiv r \pmod{m}$. Angenommen, es gäbe r und s mit dieser Eigenschaft, oBdA $r \geq s$. Dann folgt aus $r \equiv a \equiv s \pmod{m}$, daß $m \mid r - s$. Aber $0 \leq r - s \leq r \leq m - 1$ erzwingt $r = s$.

Definition 4.10. Sei $m > 0$ eine natürliche Zahl. Ein **vollständiges Restsystem** modulo m ist eine Menge ganzer Zahlen $a_1, \ldots, a_r$, so daß jede ganze Zahl $x \in \mathbb{Z}$ zu genau einer der Zahlen $a_1, \ldots, a_r$ modulo m kongruent ist.

Ein vollständiges Restsystem modulo m besteht stets aus genau m Elementen.

Beispiel 4.11. Die (alte) ISBN-10-Codierung der Bücher hat 10 Stellen, wobei die 10. Stelle eine Prüfziffer ist, die abweichend vom Rest auch den Wert X annehmen kann. Ein solcher Code ist also eine Ziffernfolge

$$a = (a_1, a_2, \ldots, a_{10})$$

mit Ziffern $a_i \in \{0, 1, \ldots, 9\}$ für alle $i = 1, \ldots, 9$ und $a_{10} \in \{0, 1, \ldots, 9\} \cup \{X\}$. Es werden nun nur solche ISBN-10 Codes verwendet, bei denen als Kontrollsumme

$$\sum_{i=1}^{10} i \cdot a_i \equiv 0 \pmod{11}$$

herauskommt. Dabei ist X als 10 zu werten (und das erklärt über das römische Zahlsymbol für 10 auch die Wahl von X). Weil $10 \cdot a_{10} \equiv -a_{10} \pmod{11}$ kann mit

$$a_{10} \equiv \sum_{i=1}^{9} i \cdot a_i \pmod{11}$$

durch Wahl des Repräsentanten modulo 11 im Intervall $[0, 10]$ die Prüfziffer a_{10} aus den restlichen Ziffern berechnet werden.

Der ISBN-10-Code erkennt nun die zwei häufigsten menschlichen Fehler. Sei zuerst nur eine Ziffer fehlerhaft, etwa an der Stelle j, also statt a eine Ziffernfolge

$$b = (b_1, \ldots, b_{10})$$

mit $b_i = a_i$ für alle $i \neq j$ und $b_j \neq a_j$. Wenn b als ISBN-10-Code geprüft wird, dann ist

$$\sum_{i=1}^{10} i \cdot b_i \equiv j(b_j - a_j) + \sum_{i=1}^{10} i \cdot a_i \equiv j(b_j - a_j) \not\equiv 0 \pmod{11},$$

denn als Primzahl teilt 11 das Produkt $j(b_j - a_j)$ nicht, da die Faktoren $1 \leq j \leq 10$ und $-10 \leq b_j - a_j \leq 10$ mit $b_j - a_j \neq 0$ nicht von 11 geteilt werden.

Als zweites nehmen wir an, daß der fehlerhafte Code b an genau zwei Stellen $j \neq k$ die Ziffern vertauscht hat. Wenn dann b als ISBN-10-Code geprüft wird, dann ist

$$\begin{aligned}\sum_{i=1}^{10} i \cdot b_i &\equiv j(b_j - a_j) + k(b_k - a_k) + \sum_{i=1}^{10} i \cdot a_i \\ &\equiv j(a_k - a_j) + k(a_j - a_k) \equiv (j - k)(a_k - a_j) \not\equiv 0 \pmod{11}\end{aligned}$$

mit im Wesentlichen dem gleichen Argument.

4.2. **Teilbarkeitsregeln.** Die traditionelle Darstellung einer natürlichen Zahl n im Zehnersystem als Ziffernfolge

$$n = (a_d, a_{d-1}, \ldots, a_1, a_0)_{10}$$

mit Ziffern $a_i \in \{0, \ldots, 9\}$ für alle $0 \leq i \leq d$ bedeutet

$$n = \sum_{i=0}^{d} a_i 10^i.$$

Die Quersumme von n ist definiert als

$$Q(n) = \sum_{i=0}^{d} a_i$$

und die alternierende Quersumme als

$$A(n) = \sum_{i=0}^{d} (-1)^i a_i.$$

Bemerkung 4.12. Da $10 \equiv 1 \pmod 9$, folgt

$$n \equiv \sum_{i=0}^{d} a_i 10^i \equiv \sum_{i=0}^{d} a_i 1^i \equiv \sum_{i=0}^{d} a_i \equiv Q(n) \pmod 9,$$

und damit gilt $9 \mid n \iff 9 \mid Q(n)$. Da $10 \equiv -1 \pmod{11}$, folgt

$$n \equiv \sum_{i=0}^{d} a_i 10^i \equiv \sum_{i=0}^{d} a_i (-1)^i \equiv A(n) \pmod{11},$$

und damit gilt $11 \mid n \iff 11 \mid A(n)$.

Sei $b \geq 2$ eine natürliche Zahl. Dann können wir $n \in \mathbb{N}$ auch eindeutig im **Zahlsystem zur Basis** b als Ziffernfolge

$$n = (a_d, a_{d-1}, \ldots, a_1, a_0)_b \tag{4.2}$$

mit Ziffern $a_i \in \{0, \ldots, b-1\}$ für alle $0 \leq i \leq d$ darstellen, wobei (4.2) bedeutet

$$n = \sum_{i=0}^{d} a_i b^i.$$

Die b-Quersumme von $n = (a_d, a_{d-1}, \ldots, a_1, a_0)_b$ ist definiert als

$$Q_b(n) = \sum_{i=0}^{d} a_i$$

und die alternierende b-Quersumme als

$$A_b(n) = \sum_{i=0}^{d} (-1)^i a_i.$$

Proposition 4.13. *Es gelten die folgenden Teilbarkeitsregeln.*

(1) $b-1 \mid n \iff b-1 \mid Q_b(n)$.
(2) $b+1 \mid n \iff b+1 \mid A_b(n)$.

Beweis. Da $b \equiv 1 \pmod{b-1}$, folgt

$$n \equiv \sum_{i=0}^{d} a_i b^i \equiv \sum_{i=0}^{d} a_i 1^i \equiv \sum_{i=0}^{d} a_i \equiv Q_b(n) \pmod{b-1}.$$

Da $b \equiv -1 \pmod{b+1}$, folgt

$$n \equiv \sum_{i=0}^{d} a_i b^i \equiv \sum_{i=0}^{d} a_i (-1)^i \equiv A_b(n) \pmod{b+1}.$$ □

Beispiel 4.14. Als Anwendung erhält man eine Teilbarkeitsregel zu gleichzeitig 7, 11 und 13. Es gilt nämlich

$$1001 = 7 \cdot 11 \cdot 13$$

und daher gilt für $p \in \{7, 11, 13\}$ und $n \in \mathbb{N}$

$$p \mid n \iff p \mid A_{1000}(n).$$

Die Darstellung von n im 1000-er System erfordert keine Rechnung. Dies bedeutet nur, daß im Zehnersystem von rechts beginnend jeweils drei Ziffern zu einem Block zusammengefaßt werden. Diese muß man nun alternierend aufsummieren und erhält so eine Prüfzahl für Teilbarkeit durch $1000 + 1$ oder eben alle Teiler von 1001.

5. Primzahlen

5.1. **Das Sieb des Eratosthenes.** Wir haben bereits definiert, was eine Primzahl ist. Mit Teilbarkeit in ganzen Zahlen müssen wir genauer Folgendes definieren.

Definition 5.1. Eine **Primzahl** ist eine ganze Zahl $p > 1$, die nur durch ± 1 und $\pm p$ teilbar ist.

Wie findet man heraus, ob ein $n \in \mathbb{N}$ Primzahl ist? Der naive Primzahltest nach Definition (und Proposition 3.2 (8)) probiert alle $1 < d < n$ als Teiler von n. Wenn kein solches d teilt, dann ist n Primzahl. Etwas besser ist das folgende semi-naive Kriterium, nach dem man nur die Primzahlen $1 < d \leq \sqrt{n}$ testen muß.

Proposition 5.2. *Ist $n > 1$ keine Primzahl, dann gibt es einen Teiler $p \mid n$ mit $1 < p \leq \sqrt{n}$ und p ist Primzahl.*

Beweis. Sei n keine Primzahl. Weil eine ganze Zahl nur endlich viele Teiler hat, gibt es einen kleinsten von 1 verschiedenen Teiler $p \mid n$. Weil Teiler von p auch Teiler von n sind, folgt aus der Minimalität, daß p eine Primzahl ist.

Es gibt nun eine ganze Zahl $m > 1$ mit $n = pm$. Aufgrund der Minimalität von p folgt $p \leq m$ und daher $p^2 \leq pm = n$, woraus die gesuchte Abschätzung $p \leq \sqrt{n}$ folgt. □

Bemerkung 5.3. Es reicht somit aus, nur die Primzahlen $1 < p \leq \sqrt{n}$ auf $p \mid n$ zu testen. Wenn keine dieser Primzahlen n teilt, dann ist n selbst Primzahl. Darauf beruht das **Sieb des Eratosthenes**[15], mittels dessen man rekursiv die Primzahlen aus den natürlichen Zahlen aussiebt. Wir konstruieren eine Folge $(S_i)_{i\geq 0}$ von Teilmengen $S_i \subseteq \mathbb{N}$ und die Folge $(p_i)_{i\geq 1}$ der Größe nach aufsteigend angeordneten Primzahlen rekursiv wie folgt. Als Startwert setzen wir

$$S_0 = \mathbb{N} \setminus \{1\}.$$

Seien S_i und $p_1, \ldots, p_i$ bereits konstruiert, dann ist

$$\begin{aligned} p_{i+1} &:= \min S_i, \\ S_{i+1} &:= S_i \setminus p_{i+1}\mathbb{N}. \end{aligned}$$

Dabei haben wir mit der Menge S_i bereits die Vielfachen der Primzahlen $p_1 = 2, \ldots, p_i$ ausgesiebt. Es gilt genauer:

(a) p_i ist Primzahl und $\{p_1, \ldots, p_i\}$ enthält genau die Primzahlen $\leq p_i$ der Größe nach geordnet:
$$p_1 = 2 < p_2 < \ldots < p_i.$$
(b) S_i besteht genau aus den $n \in \mathbb{N}$, deren Primteiler sämtlich $> p_i$ sind.
(c) Sei $n \in \mathbb{N}$ und sei i groß genug, so daß $p_{i+1} > \sqrt{n}$. Dann ist n Primzahl genau dann, wenn $n \in \{p_1, \ldots, p_i\}$ oder $n \in S_i$.

Beweis. Wir zeigen dies per Induktion nach i. Für $i = 1$ gilt $p_1 = 2$ und $P_1 = \{p_1\}$ und S_1 sind genau die ungeraden Zahlen.

Gelten die Behauptungen bis i, so enthält S_i genau die Zahlen mit Primteilern $> p_i$ und $p_1, \ldots, p_i$ sind die ersten i Primzahlen der Größe nach. Damit ist die kleinste Zahl in S_i die nächste Primzahl p_{i+1}. Diese ist sicher in S_i und da alle $n \in S_i$ nur durch Primteiler $> p_i$ teilbar sind, ist jedenfalls $p_{i+1} \leq n$ für alle diese Zahlen. Dies zeigt (a) für $i+1$. (b) für $i+1$ folgt sofort, da nun aus S_i alle Vielfache der nächsten Primzahl p_{i+1} entfernt werden.

Beweisen wir nun (c). Wenn n Primzahl ist, dann ist $n \in \{p_1, \ldots, p_i\}$, wenn $n \leq p_i$ nach (a), und $n \in S_i$, wenn $n > p_i$ nach (b).

Andersherum, wenn $n \in \{p_1, \ldots, p_i\}$, dann ist n Primzahl nach (a), und wenn $n \in S_i$, dann hat n nur Primfaktoren $> p_i$, also $\geq p_{i+1} > \sqrt{n}$. Nach Proposition 5.2 geht das nur, wenn n Primzahl ist. □

[15] Eratosthenes, 3. Jahrhundert v. Chr., griechischer Gelehrter.

5.2. **Eindeutige Primfaktorzerlegung.** Wir beginnen mit einer Dichotomie.

Lemma 5.4. *Sei p eine Primzahl und $n \in \mathbb{Z}$. Dann gilt genau eine der beiden folgenden Aussagen:*

(i) p *teilt* n,
(ii) p *und* n *sind teilerfremd.*

Beweis. Der $\mathrm{ggT}(p, n)$ ist als positiver Teiler von p entweder 1, dann sind p und n teilerfremd, oder p, dann wird n von p geteilt. □

Satz 5.5. *Eine ganze Zahl $p > 1$ ist eine Primzahl genau dann, wenn für alle $a, b \in \mathbb{Z}$ gilt*

$$p \mid ab \Longrightarrow p \mid a \text{ oder } p \mid b.$$

Beweis. Sei zuerst p eine Primzahl und $p \mid ab$ für ganze Zahlen a, b. Wir wenden Lemma 5.4 auf p und a an. Wenn $p \mid a$ gilt, sind wir fertig. Andernfalls ist $(p, a) = 1$, und aus Proposition 3.21 und $p \mid ab$ folgt schon $p \mid b$. Dies zeigt die eine Richtung.

Für die umgekehrte Richtung nehmen wir an $p = ab$ mit $a, b \in \mathbb{Z}$. Dann gilt $p \mid ab$, also nach Voraussetzung oBdA $p \mid a$, d.h. es gibt $u \in \mathbb{Z}$ mit $a = pu$. Dann ist

$$p = ab = pub,$$

also $ub = 1$. Als Teiler von 1 gilt für b dann $0 < |b| \leq 1$, d.h. $b = \pm 1$ und folglich $a = \pm p$. Somit ist p eine Primzahl. □

Bemerkung 5.6. Der Begriff der Teilbarkeit ist auch in einem beliebigen Integritätsring R sinnvoll. Ein Element $\pi \in R$, das selbst keine Einheit und bis auf Einheiten nur durch 1 und sich selbst teilbar ist, nennt man **irreduzibel**. Per Definition sind also die Primzahlen die positiven irreduziblen Elemente von $\mathbb{Z}$.

Ein Element $\pi \in R$ wird Primelement genannt, wenn es keine Einheit ist und die Eigenschaft aus Satz 5.5 besitzt: für alle $a, b \in R$ mit $\pi \mid ab$ folgt $\pi \mid a$ oder $\pi \mid b$. Wir haben damit gerade bewiesen, daß Primzahlen genau die Primelemente aus $\mathbb{Z}$ sind. Hier geht bereits das Lemma von Bézout ein und nicht nur unmittelbar die Definition einer Primzahl.

Proposition 5.7. *Seien p, q Primzahlen. Dann sind äquivalent*

(a) $p = q$,
(b) $p \mid q$,
(c) $(p, q) \neq 1$.

Zwei Primzahlen sind also entweder gleich oder teilerfremd.

Beweis. Trivial ist (a) $\Longrightarrow$ (b) $\Longrightarrow$ (c). Nehmen wir daher (c) an und zeigen (a). Wenn $(p, q) = d \neq 1$, dann ist d ein Teiler > 1 der Primzahlen p und q. Dies kann nur $p = d = q$ selbst sein, und das zeigt (a). □

Wir zerlegen nun jede ganze Zahl $\neq 0$ auf eindeutige Weise in ein Vorzeichen und ein Produkt von Primfaktoren. Das mit dem Vorzeichen ist trivial, daher beschränken wir uns auf natürliche Zahlen.

Theorem 5.8 (Fundamentalsatz der Arithmetik)**.** *Sei $n \in \mathbb{N}$ eine natürliche Zahl.*

*(1) Die Zahl n hat eine **Primfaktorzerlegung**: es gibt ein $r \geq 0$ und Primzahlen $p_1, \ldots, p_r$ mit*

$$n = p_1 \cdot \ldots \cdot p_r.$$

(2) Die Primfaktorzerlegung ist eindeutig bis auf Permutation: sei $s \geq 0$ und seien $q_1, \ldots, q_s$ Primzahlen mit

$$n = q_1 \cdot \ldots \cdot q_s,$$

dann ist $r = s$ und es gibt eine Permutation σ mit $q_i = p_{\sigma(i)}$ für alle $i = 1, \ldots, r$.

Beweis. Wir beweisen zunächst die Existenz einer Primfaktorzerlegung. Sei dazu

$$A = \{n \in \mathbb{N}\ ;\ n \text{ hat keine Primfaktorzerlegung}\}$$

die Menge der Gegenbeispiele. Wenn $A \neq \emptyset$, dann gibt es nach Satz 2.1 ein minimales Element $a \in A$, also ein minimales Gegenbeispiel.

Dieses a kann keine Primzahl sein, denn Primzahlen sind ihre eigene Primfaktorzerlegung. Also gibt es $x, y \in \mathbb{Z}$ mit $a = xy$ und $x, y > 1$. Also folgt $x = a/y < a$ und genauso $y < a$. Aufgrund der Minimalität folgt $x, y \notin A$. Daher haben x und y eine Primfaktorzerlegung. Man bekommt dann eine für $a = xy$, indem man die für x mit der für y multipliziert.

Es bleibt die Eindeutigkeit der Primfaktorzerlegung nachzuweisen. Dazu machen wir Induktion nach der Länge der minimalen Primfaktorzerlegung:

$$\ell(n) := \min\{r\ ;\ \text{ es gibt Primfaktorzerlegung } n = p_1 \cdot \ldots \cdot p_r\}.$$

Für $\ell(n) = 0$ bleibt nur $n = 1$. Da $n = 1$ keinen Primteiler hat (die Teiler sind ≤ 1 und Primzahlen ≥ 2), hat 1 auch nur das leere Produkt als Primfaktorzerlegung.

Sei nun $\ell(n) \geq 1$ und für alle natürlichen Zahlen kleinerer Länge die Eindeutigkeit der Primfaktorzerlegung bereits gezeigt. Sei $n = p_1 \cdot \ldots \cdot p_r$ eine Primfaktorzerlegung minimaler Länge und sei

$$n = q_1 \cdot \ldots \cdot q_s$$

eine weitere Primfaktorzerlegung. Weil $\ell(n) = r \geq 1$, gibt es ein $p_1 \mid n$. Aus $p_1 \mid n = q_1 \cdot \ldots \cdot q_s$ folgt durch iteriertes Anwenden von Satz 5.5, daß p_1 einen Faktor der alternativen Primfaktorzerlegung teilt. Es gibt somit ein j mit $p_1 \mid q_j$ und wegen Proposition 5.7 dann sogar $p_1 = q_j$.

Teilen wir n durch p_1, so erhalten wir zwei Primfaktorzerlegungen

$$p_2 \cdot \ldots \cdot p_r = \frac{n}{p_1} = \prod_{i=1,\ i \neq j}^{s} q_i.$$

Da offensichtlich die Länge mindestens (da definiert als die minimal mögliche Länge einer Faktorisierung) um 1 runtergeht

$$\ell(n/p_1) \leq \ell(n) - 1,$$

haben wir nun per Induktionsvoraussetzung Eindeutigkeit der Faktorisierung bis auf Permutation der Faktoren für n/p_1. Es folgt $r - 1 = s - 1$ sowie eine Permutation σ mit $q_i = p_{\sigma(i)}$ für alle $i \neq j$. Wir setzen noch $\sigma(j) = 1$ und haben damit die Eindeutigkeit für die Faktorisierung von n gezeigt. □

Bemerkung 5.9. Eine ganze Zahl $n \neq 0$ hat ein Vorzeichen

$$\operatorname{sign}(n) = \frac{n}{|n|}.$$

Für ganze Zahlen n, m ungleich 0 gilt

$$\operatorname{sign}(nm) = \operatorname{sign}(n)\operatorname{sign}(m).$$

Definition 5.10. Für eine ganze Zahl $n \in \mathbb{Z}$, $n \neq 0$, und eine Primzahl p definieren wir

$$v_p(n) = \text{ Anzahl der Faktoren } p \text{ in der Primfaktorzerlegung von } |n|.$$

Proposition 5.11. *Die Abbildung*

$$v_p : \mathbb{Z} \setminus \{0\} \to \mathbb{N}_0$$

ist multiplikativ im logarithmischen Sinn: für ganze Zahlen n, m verschieden von 0 gilt

$$v_p(nm) = v_p(n) + v_p(m).$$

Beweis. Das folgt sofort aus der Eindeutigkeit der Primfaktorzerlegung, weil daher die Primfaktorzerlegung von nm das Produkt der Primfaktorzerlegungen von n und m ist. □

Proposition 5.12. *Die Funktionen v_p haben die folgenden Eigenschaften. Sei $n \in \mathbb{Z}$, $n \neq 0$.*

(1) $v_p(n) > 0 \iff p \mid n$.

(2) $p^m \mid n \iff m \leq v_p(n)$.

(3) *Die Primfaktorzerlegung von n schreibt sich kurz als*

$$n = \operatorname{sign}(n) \cdot \prod_{p \mid n} p^{v_p(n)} = \operatorname{sign}(n) \cdot \prod_p p^{v_p(n)}$$

(4) *Sei $m \in \mathbb{Z}$, $m \neq 0$ und auch $n + m \neq 0$. Dann*

$$v_p(n+m) \geq \min\{v_p(n), v_p(m)\}.$$

Wenn $v_p(n) \neq v_p(m)$, dann gilt Gleichheit.

Beweis. (1) und (3) folgen sofort aus der Eindeutigkeit der Primfaktorzerlegung.

(2) Wenn $p^m \mid n$, dann ist $n = p^m x$ für ein $x \in \mathbb{Z}$ und $v_p(n) = v_p(p^m) + v_p(x) = m + v_p(x) \geq m$. Umgekehrt, wenn $v_p(n) \geq m$, dann kommt der Faktor p^m in der Primfaktorzerlegung von n vor.

Für (4) folgt mit $e = \min\{v_p(n), v_p(m)\}$, daß $p^e \mid n$ und $p^e \mid m$. Daher gilt auch $p^e \mid n + m$. Wegen (2) haben wir dann $v_p(n+m) \geq e$.

Nehmen wir nun an, daß oBdA $e = v_p(n) < v_p(m)$. Wir führen $v_p(n+m) > e$ zu einem Widerspruch, indem

$$p^{e+1} \mid (n+m) - m = n,$$

was (2) widerspricht. □

Korollar 5.13. *Sei p eine Primzahl und $n \in \mathbb{Z}$, $n \neq 0$. Dann gilt*

$$v_p(n) = \max\{e \in \mathbb{N}_0 \ ; \ p^e \mid n\}.$$

Beweis. Das ist eine Umformulierung von Proposition 5.12 (2). □

Definition 5.14. Die Abbildung $v_p : \mathbb{Q}^\times = \mathbb{Q} \setminus 0 \to \mathbb{Z}$ definiert durch

$$v_p(\frac{a}{b}) = v_p(a) - v_p(b)$$

ist wohldefiniert und multiplikativ: für alle $x, y \in \mathbb{Q}^\times$ gilt

$$v_p(xy) = v_p(x) + v_p(y).$$

Beweis. Wenn $a/b = c/d$, dann gilt $ad = bc$ und aus Proposition 5.11 folgt

$$v_p(a) + v_p(d) = v_p(ad) = v_p(bc) = v_p(b) + v_p(c).$$

Umstellen liefert $v_p(a) - v_p(b) = v_p(c) - v_p(d)$ und damit die Wohldefiniertheit von v_p auf rationalen Zahlen $\neq 0$. Sei nun $x = a/b$ und $y = c/d$. Dann ist

$$v_p(xy) = v_p(\frac{ac}{bd}) = v_p(ac) - v_p(bd) = \big(v_p(a) - v_p(b)\big) + \big(v_p(c) - v_p(d)\big) = v_p(x) + v_p(y). \quad \square$$

Bemerkung 5.15. Für eine rationale Zahl x gibt $v_p(x)$ den Exponenten von p im gekürzten Bruch, der x darstellt, an. Wenn $v_p(x) > 0$, dann tritt $p^{v_p(x)}$ im gekürzten Zähler auf, wenn $v_p(x) < 0$ ist, dann tritt $p^{-v_p(x)}$ im gekürzten Nenner auf, und wenn $v_p(x) = 0$, dann tritt p weder im gekürzten Zähler noch Nenner auf.

Satz 5.16. *Sei $n \in \mathbb{N}$ und $a \in \mathbb{Z}$, $a \neq 0$. Wenn es eine Primzahl p gibt, so daß*

$$n \nmid v_p(a),$$

dann ist $\sqrt[n]{a}$ keine rationale Zahl.

Beweis. Angenommen $\alpha = \sqrt[n]{a}$ wäre rational. Dann ist $\alpha \neq 0$ und

$$v_p(a) = v_p(\alpha^n) = n v_p(\alpha)$$

ist durch n teilbar, ein Widerspruch. □

Korollar 5.17. *Die Zahl $\sqrt{2}$ ist irrational.*

Beweis. Das folgt aus Satz 5.16 mit $n = a = p = 2$, weil $v_2(2) = 1$ keine gerade Zahl ist. □

Satz 5.18. *Es gilt die folgende Beschreibung der ganzen Zahlen innerhalb der rationalen Zahlen:*

$$\mathbb{Z} = \{x \in \mathbb{Q}\ ;\ x = 0 \text{ oder } x \neq 0 \text{ und für alle Primzahlen } p \text{ gilt } v_p(x) \geq 0\}.$$

Beweis. Nach Bemerkung 5.15 sind die $x \in \mathbb{Q}^\times$ mit $v_p(x) \geq 0$ für alle Primzahlen p genau diejenigen rationalen Zahlen, bei denen im gekürzeten Bruch in der Primfaktorzerlegung des Nenners keine Primzahl auftritt. Der Nenner muß daher $= 1$ sein und x demnach eine ganze Zahl. □

Korollar 5.19. *Sei $n \in \mathbb{N}$ und $a \in \mathbb{Z}$ keine n-te Potenz in $\mathbb{Z}$, d.h., die Gleichung $a = X^n$ hat keine Lösung $x \in \mathbb{Z}$. Dann ist $\sqrt[n]{a}$ keine rationale Zahl, d.h., die Gleichung $a = X^n$ hat auch keine Lösung $x \in \mathbb{Q}$.*

Beweis. Angenommen $a = x^n$ mit $x \in \mathbb{Q}$. Wir müssen zeigen, daß x eine ganze Zahl ist. Sei p eine beliebige Primzahl. Wegen

$$0 \leq v_p(a) = nv_p(x)$$

gilt $v_p(x) \geq 0$, und nach Satz 5.18 folgt $x \in \mathbb{Z}$. □

5.3. **Elementare Sätze über Primzahlen.** Für beliebiges $n \in \mathbb{N}$ und Variablen X, Y gilt:

$$(X - Y)(X^{n-1} + X^{n-2}Y + \ldots + X^{n-1-i}Y^i + \ldots + Y^{n-1}) = X^n - Y^n. \tag{5.1}$$

Es gilt daher für alle $a, b \in \mathbb{Z}$ und alle $n \in \mathbb{N}$

$$a - b \mid a^n - b^n.$$

Polynome $P(X) = a_0 + a_1X + \ldots + a_dX^d$ mit ganzen Koeffizienten $a_i \in \mathbb{Z}$ nehmen für $x \in \mathbb{Z}$ ganzzahlige Werte $P(x) \in \mathbb{Z}$ an. Diese Werte studieren wir nun.

Proposition 5.20. *Sei $P(X) = a_0 + a_1X + \ldots + a_dX^d$ mit $a_i \in \mathbb{Z}$ ein Polynom mit ganzen Koeffizienten. Dann gilt für alle $a, b \in \mathbb{Z}$*

$$a - b \mid P(a) - P(b).$$

Beweis. Wir haben gerade gesehen, daß $a-b$ ein Teiler von $a^i - b^i$ ist für alle $i \geq 0$. Multiplizieren mit a_i und Aufsummieren ergibt die Behauptung.

Variante: wir rechnen modulo $m = a - b$. Weil $a \equiv b \pmod{m}$, folgt dann

$$P(a) \equiv \sum_i a_i a^i \equiv \sum_i a_i b^i \equiv P(b) \pmod{m}$$

und das ist die Behauptung. □

Satz 5.21. *Ein nichtkonstantes Polynom $P(X) \in \mathbb{Z}[X]$ mit ganzen Koeffizienten nimmt für $n \in \mathbb{N}$ nicht nur Primzahlen als Werte $P(n)$ an.*

Beweis. Sei $P(X)$ vom Grad $d > 0$. Als Polynom vom Grad d nimmt P auf $\mathbb{R}$ jeden Wert höchstens d-mal an. Damit ist für alle bis auf endlich viele $a \in \mathbb{Z}$ der Wert $P(a)$ weder 0 noch ± 1. Wir können daher ein $a \in \mathbb{N}$ und dann eine Primzahl p wählen, so daß $p \mid P(a)$. Dann gilt für alle $b \equiv a \pmod{p}$

$$P(b) \equiv P(a) \equiv 0 \pmod{p}$$

nach Proposition 5.20. Da von den Werten $P(a + n\mathbb{Z})$ mit $n \in \mathbb{Z}$ nur endlich viele $\pm p$ sein können, sind alle bis auf endlich viele davon keine Primzahlen. □

Beispiel 5.22. Das Polynom $P(X) = X^2 - X + 41$ nimmt für $x = 1, \ldots, 40$ nur Primzahlen als Werte an. Aufgrund der Symmetrie $X \mapsto 1 - X$ gilt das sogar für $x = -39, \ldots, 40$, aber die dadurch gewonnen Primzahlwerte sind nicht neu.

Proposition 5.23. *Seien $a \in \mathbb{N}$ und $n \in \mathbb{N}$, $n \geq 2$. Wenn $p = a^n - 1$ eine Primzahl ist, dann muß $a = 2$ und n eine Primzahl sein.*

Beweis. Für $a = 1$ ist $p = 0$ keine Primzahl. Sei daher $a > 1$. Nach Proposition 5.20 ist $a - 1$ ein Teiler von $a^n - 1^n = p$, und wegen $n \geq 2$ auch ein Teiler $< p$. Damit muß $a - 1 = 1$ somit $a = 2$ sein.

Wenn n keine Primzahl ist, dann gibt es $n = rs$ mit $r, s > 1$ und genau wie eben ist $2^r - 1 \mid (2^r)^s - 1^s = p$ ein echter Teiler:

$$1 = 2^1 - 1 < 2^r - 1 < 2^n - 1 = p.$$

Das geht nicht. □

Definition 5.24. Eine **Mersenne-Primzahl**[16] ist eine Primzahl der Form $2^n - 1$ mit $n \in \mathbb{N}$.

Bemerkung 5.25. Nach Proposition 5.23 sind Kandidaten für Mersenne-Primzahlen von der Form

$$M_p = 2^p - 1$$

mit einer Primzahl p. Die ersten Mersenne-Primzahlen sind

p	2	3	5	7	13	...
M_p	3	7	31	127	8191	...

aber

$$M_{11} = 2047 = 23 \cdot 89.$$

Es ist nicht bekannt, ob unter den Zahlen der Form $2^p - 1$ mit p Primzahl unendlich viele Primzahlen vorkommen. Es ist aber auch nicht bekannt, ob von diesen unendlich viele zusammengesetzte Zahlen (keine Primzahlen) sind. Man kennt wohl aktuell 45 Mersenne-Primzahlen (Stand April 2015), und die aktuell größte bekannte Primzahl ist eine Mersenne-Primzahl.

Für beliebiges **ungerades** $n \in \mathbb{N}$ und Variablen X, Y gilt:

$$(X + Y)(X^{n-1} - X^{n-2}Y \pm \ldots + X^{n-1-i}(-Y)^i + \ldots + Y^{n-1}) = X^n + Y^n. \tag{5.2}$$

Es gilt daher für alle $a, b \in \mathbb{Z}$ und alle ungeraden $n \in \mathbb{N}$

$$a + b \mid a^n + b^n.$$

Definition 5.26. Eine **Fermat-Primzahl** ist eine Primzahl der Form $2^n + 1$ mit $n \in \mathbb{N}$.

Bemerkung 5.27. Es folgt aus (5.2), daß zu einem echten ungeraden Teiler $a \mid n$ mit $b = n/a$ der echte Faktor $2^b + 1 \mid 2^n + 1$ gehört. Kandidaten für Fermat-Primzahlen haben also die Form

$$F_n = 2^{2^n} + 1$$

mit $n \in \mathbb{N}_0$. Die einzigen bekannten Fermat-Primzahlen sind

n	0	1	2	3	4
F_n	3	5	17	257	65537.

Euler fand den Faktor

$$641 \mid F_5 = 2^{32} + 1.$$

Es ist nicht bekannt, ob unter den Zahlen der Form $2^{2^n} + 1$ unendlich viele Primzahlen vorkommen. Es ist aber auch nicht bekannt, ob von diesen unendlich viele zusammengesetzte Zahlen (keine Primzahlen) sind.

Fermat-Primzahlen spielen eine Rolle in der Antwort auf die antike Frage, für welches $n \in \mathbb{N}$, $n \geq 3$ das regelmäßige n-Eck allein mit Zirkel und Lineal konstruierbar ist. Die Antwort lautet, eine solche Konstruktion mit Zirkel und Lineal gibt es genau für die n von der Form

$$n = 2^e \cdot p_1 \cdot \ldots \cdot p_r,$$

[16] Marin Mersenne (1588–1648), französischer Theologe, Mathematiker und Musiktheoretiker.

wobei $p_1, \ldots, p_r$ paarweise verschiedene Fermat-Primzahlen sind. Der Beweis dieser Aussage benötigt Galois-Theorie und gehört in die Vorlesung *Algebra*.

Satz 5.28 (Satz von Wilson[17]). *Sei p eine Primzahl. Dann gilt*

$$(p-1)! \equiv -1 \pmod{p}.$$

Beweis. Jedes $1 \leq a \leq p-1$ ist teilerfremd zu p. Es gibt daher nach Satz 4.6 genau eine Lösung $x =: \iota(a)$ für die Kongruenzgleichung $aX \equiv 1 \pmod{p}$ mit der Nebenbedingung $0 \leq \iota(a) \leq p-1$ (Division mit Rest!). Der Fall $\iota(a) = 0$ scheidet aus. Wir haben somit eine Abbildung

$$\iota : \{1, \ldots, p-1\} \to \{1, \ldots, p-1\}$$

definiert. Diese Abbildung ist eine Involution, denn aus $a\iota(a) \equiv 1$ folgt auch $\iota(a)a \equiv 1$, und aus der Eindeutigkeit folgt $\iota(\iota(a)) = a$.

Wir bestimmen nun die Fixpunkte der Involution ι. Es gilt $\iota(a) = a$ genau dann, wenn

$$p \mid a^2 - 1 = (a-1)(a+1),$$

also für die zwei Fälle $a = 1$ und $a = p-1$. Damit gibt es $a_1, \ldots, a_{\frac{p-3}{2}}$ mit

$$\{1, \ldots, p-1\} = \{1, p-1, a_1, \ldots, a_{\frac{p-3}{2}}, \iota(a_1), \ldots, \iota(a_{\frac{p-3}{2}})\}.$$

Wir multiplizieren nun auf und bilden Paare $(a, \iota(a))$, wann immer das geht:

$$(p-1)! = \prod_{i=1}^{p-1} i = 1 \cdot (p-1) \cdot \prod_{i=1}^{\frac{p-3}{2}} a_i \iota(a_i) \equiv 1 \cdot (p-1) \prod_{i=1}^{\frac{p-3}{2}} 1 \equiv -1 \pmod{p}. \qquad \square$$

Bemerkung 5.29. (1) Die Involution ι auf der Menge $\{1, \ldots, p-1\}$ im Beweis von Satz 5.28 kann man mittels Kongruenzen von Brüchen mit zum Modulus teilerfremden Nenner wie in Bemerkung 4.8 kompakt und geschlossen als

$$\iota(a) \equiv \frac{1}{a} \pmod{p}$$

eindeutig beschreiben. Auch im Brüchekalkül ist $\iota(\iota(a)) = a$ offensichtlich.

(2) Der Satz von Wilson liefert ein Primzahlkriterium. Es gilt auch die Umkehrung von Satz 5.28: wenn für eine natürliche Zahl $n \geq 2$ gilt $(n-1)! \equiv -1 \pmod{n}$, dann ist n Primzahl. Sei nämlich $n = ab$ eine echte Faktorisierung, dann ist $a \leq n-1$ und daher a einTeiler von $(n-1)!$. Es folgt

$$-1 \equiv (n-1)! \equiv 0 \pmod{a},$$

und das bedeutet $a = 1$, Widerspruch.

Satz 5.30 (Kleiner Satz von Fermat). *Sei p eine Primzahl. Dann ist für alle $a \in \mathbb{Z}$ mit $p \nmid a$*

$$a^{p-1} \equiv 1 \pmod{p}.$$

Beweis. Für alle $x \in \{1, \ldots, p-1\}$ gibt es nach Division mit Rest ein eindeutiges $0 \leq r \leq p-1$ mit $ax \equiv r \pmod{p}$. Da $p \nmid ax$ folgt $r > 0$. Multiplikation mit a und anschließende Division mit Rest durch p definiert also eine Abbildung

$$[a] : \{1, \ldots, p-1\} \to \{1, \ldots, p-1\},$$

die durch $[a](x) \equiv ax \pmod{p}$ eindeutig bestimmt ist. Die Abbildung $[a]$ ist injektiv, denn aus $ax \equiv ay \pmod{p}$ folgt wegen $(a, p) = 1$ bereits $x \equiv y \pmod{p}$. Als Selbstabbildung einer endlichen Menge muß $[a]$ als injektive Abbildung automatisch surjektiv sein (alternativ folgt surjektiv daraus, daß die Kongruenzgleichung $aX \equiv r \pmod{p}$ stets eine Lösung hat).

[17] John Wilson (1741–1793), britischer Mathematiker und Jurist

Nun gilt

$$a^{p-1} \cdot \prod_{x=1}^{p-1} x = \prod_{x=1}^{p-1} ax \equiv \prod_{x=1}^{p-1} [a](x) = \prod_{x=1}^{p-1} x \pmod{p}.$$

Weil $p \nmid \prod_{x=1}^{p-1} x$ und weil p eine Primzahl ist[18] darf man kürzen und erhält

$$a^{p-1} \equiv 1 \pmod{p}. \qquad \square$$

Bemerkung 5.31. Mittels Kongruenzen von Brüchen mit zum Modulus teilerfremden Nenner wie in Bemerkung 4.8 kann man die zu $[a]$ inverse Abbildung einfach als $[1/a]$ angeben. Man sieht nämlich leicht ein, daß für zu p teilerfremde a und b gilt

$$[ab] = [a] \circ [b],$$

weiter hängt $[a]$ nur von a modulo p ab, und außerdem ist $[1]$ die Identität.

Wenn man in Satz 5.30 den Fall $p \mid a$ nicht ausschließen will, muß man die Aussage wie folgt formulieren.

Korollar 5.32. *Sei p eine Primzahl. Dann ist für alle $a \in \mathbb{Z}$*

$$a^p \equiv a \pmod{p}.$$

Beweis. Wegen $a^p - a = a(a^{p-1} - 1)$ folgt dies für $(a, p) = 1$ aus Satz 5.30, und für $a \equiv 0 \pmod{p}$ ist nichts weiter zu zeigen. $\square$

Satz 5.33 (Satz von Thue[19])**.** *Sei p eine Primzahl und $a, b \in \mathbb{N}$ mit $ab > p$. Dann gibt es für jedes $n \in \mathbb{Z}$ ganze Zahlen $x, y \in \mathbb{Z}$ mit*

$$0 \leq x < \max\{2, a\}$$
$$1 \leq y < \max\{2, b\} \text{ und } (p, y) = 1,$$

so daß für die richtige Wahl des Vorzeichens $n \equiv \pm\frac{x}{y} \pmod{p}$.

Beweis. Wenn $n \equiv 0 \pmod{p}$, dann reicht $x = 0$ und $y = 1$. Sei daher nun $n \not\equiv 0 \pmod{p}$ und $1 \leq m \leq p - 1$ mit $nm \equiv 1 \pmod{p}$.

Wenn $a \geq p$, dann reicht $y = 1$, denn jedes $n \in \mathbb{Z}$ ist wegen Division mit Rest zu einer der Zahlen $0, 1, \ldots, p-1$ kongruent modulo p. Wenn $b \geq p$, dann reicht $x = 1$ und $y = m$. Sei daher von nun an $a < p$ und $b < p$.

Die Menge S der Paare (α, β) mit $0 \leq \alpha < a$ und $0 \leq \beta < b$ besteht aus $ab > p$ Paaren. Nach dem **Schubfachprinzip** sind daher unter den Werten $\beta n - \alpha$ mit $(\alpha, \beta) \in S$ mindestens zwei kongruent modulo p. Sei also $(\alpha_1, \beta_1) \neq (\alpha_2, \beta_2)$ aus S mit

$$\beta_1 n - \alpha_1 \equiv \beta_2 n - \alpha_2 \pmod{p}.$$

Wenn $\beta_1 \equiv \beta_2 \pmod{p}$, dann folgt $\alpha_1 \equiv \alpha_2 \pmod{p}$. Weil beide Paare aus S sind, folgt sogar schon Gleichheit, ein Widerspruch. Wir setzen

$$x = |\alpha_1 - \alpha_2|$$
$$y = |\beta_1 - \beta_2|,$$

so daß $0 \leq x \leq \max\{\alpha_1, \alpha_2\} < a$ und $1 \leq y \leq \max\{\beta_1, \beta_2\} < b$ gilt und außerdem $(p, y) = 1$. Die Rechnung

$$n \equiv \frac{\alpha_1 - \alpha_2}{\beta_1 - \beta_2} = \pm\frac{x}{y} \pmod{p}$$

beweist dann den Satz. $\square$

[18] Nach Satz 5.28 weiß man sogar, daß diese Produkt $\equiv -1 \pmod{p}$ ist, und könnte damit zum Ziel kommen. Aber der genaue Wert ist hier nicht wesentlich, und daher soll der Satz von Wilson im Sinne der Unterscheidung zwischen wesentlichen und unwesentlichen Argumenten hier nicht verwendet werden.

[19] Axel Thue (1863–1922), norwegischer Mathematiker.

Bemerkung 5.34. Seien $a, b \geq 2$. Für $(x, y) \in [0, a) \times [1, b)$ mit $(p, y) = 1$ und $ab > p$ nimmt der Bruch $\pm\frac{x}{y}$ demnach alle Restklassen modulo p als Wert an. Wenn man bedenkt, daß man dafür nur ungefähr $2ab$-viele Ausdrücke zur Verfügung hat, was in der Größenordnung von $2p$ liegen kann, dann ist der Satz von Thue schon eine recht knappe (und damit starke) Aussage.

5.4. **Quadrate und Primzahlen.**

Satz 5.35. *Sei p eine Primzahl $p \equiv -1 \pmod 4$. Dann hat $X^2 \equiv -1 \pmod p$ keine Lösung.*

Beweis. Angenommen $a \in \mathbb{Z}$ mit $a^2 \equiv -1 \pmod p$. Dann gilt $a^4 \equiv 1 \pmod p$. Wir schreiben nun nach Voraussetzung $p - 1 = 4n + 2$ mit $n \in \mathbb{N}$ und finden nach Satz 5.30

$$1 \equiv a^{p-1} \equiv a^{4n+2} \equiv (a^4)^n \cdot a^2 \equiv a^2 \equiv -1 \pmod p.$$

Dies bedeutet $p \mid 2$, ein Widerspruch. □

Als Korollar zu Satz 5.35 erhalten wir einen nicht-zwei-Quadrate-Satz

Korollar 5.36. *Sei p eine Primzahl $p \equiv -1 \pmod 4$. Dann gibt es keine $a, b \in \mathbb{Z}$ mit $p = a^2+b^2$.*

Beweis. Wenn $p = a^2 + b^2$, dann ist $p \nmid a, b$ und dann gibt es

$$(\frac{a}{b})^2 \equiv \frac{p - b^2}{b^2} \equiv -1 \pmod p$$

im Widerspruch zu Satz 5.35. □

Satz 5.37. *Sei p eine Primzahl $p \equiv 1 \pmod 4$. Dann hat $X^2 \equiv -1 \pmod p$ genau zwei Lösungen (gemeint ist bis auf Kongruenz modulo p).*

Beweis. Nach Voraussetzung gibt es ein $n \in \mathbb{N}$ mit $p = 4n+1$. Wir setzen $x = (2n)!$ und rechnen

$$\begin{aligned} x^2 &\equiv (-1)^{2n}(2n)! \cdot (2n)! \equiv (2n)! \cdot \prod_{i=1}^{2n}(-i) \equiv (2n)! \cdot \prod_{i=1}^{2n}(p - i) \\ &\equiv (2n)! \cdot \prod_{i=2n+1}^{p-1} i \equiv (p - 1)! \equiv -1 \pmod p \end{aligned}$$

nach dem Satz von Wilson 5.28.

Wenn x und y zwei Lösungen sind, dann gilt $x^2 \equiv -1 \equiv y^2 \pmod p$, also

$$p \mid x^2 - y^2 = (x - y)(x + y).$$

Da p Primzahl ist, teilt p einen der beiden Faktoren und $x \equiv \pm y \pmod p$. Da $p \nmid x$ ist $x \not\equiv -x \pmod p$ und es gibt genau 2 Lösungen. □

Bemerkung 5.38. Im Vorgriff auf das Rechnen mit Restklassen in den Kapiteln §9 und §10 führen einen weiteren Beweis für die Sätze 5.35 und 5.37. Sei p eine ungerade Primzahl. Auf $\mathbb{F}_p^\times$ definieren

$$\sigma(x) = -x \qquad \text{und} \qquad \tau(x) = x^{-1}$$

zwei Involutionen σ und τ. Wegen $(-x)^{-1} = -(x^{-1})$ kommutieren diese Involutionen und definieren so eine Operation der Kleinschen Vierergruppe $V_4 = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ auf $\mathbb{F}_p^\times$. Die Fixpunkte der Elemente der Ordnung 2, das sind die Bahnen mit nichttrivialem Stabilisator, sind wie folgt:

$$\begin{aligned} \operatorname{Fix}(\sigma; \mathbb{F}_p^\times) &= \{x \in \mathbb{F}_p^\times \ ; \ \sigma(x) = x\} = \emptyset, \\ \operatorname{Fix}(\tau; \mathbb{F}_p^\times) &= \{x \in \mathbb{F}_p^\times \ ; \ \tau(x) = x\} = \{1, -1\}, \\ \operatorname{Fix}(\sigma\tau; \mathbb{F}_p^\times) &= \{x \in \mathbb{F}_p^\times \ ; \ \sigma\tau(x) = x\} = \{x \in \mathbb{F}_p^\times \ ; \ x^2 = -1\}, \end{aligned}$$

also

$$\#\operatorname{Fix}(\sigma\tau; \mathbb{F}_p^\times) = \begin{cases} 2 & \text{die Gleichung } X^2 \equiv -1 \pmod p \text{ hat Lösungen,} \\ 0 & \text{die Gleichung } X^2 \equiv -1 \pmod p \text{ hat keine Lösung.} \end{cases} \tag{5.3}$$

Die Orbits mit trivialem Stabilisator haben die Länge $|V_4| = 4$. Die Bilanzgleichung der Bahnenformel besagt daher moduo 4, daß

$$p - 1 \equiv \#\operatorname{Fix}(\sigma; \mathbb{F}_p^\times) + \#\operatorname{Fix}(\tau; \mathbb{F}_p^\times) + \#\operatorname{Fix}(\sigma\tau; \mathbb{F}_p^\times) \equiv 2 + \#\operatorname{Fix}(\sigma\tau; \mathbb{F}_p^\times) \pmod 4,$$

und damit $\#\operatorname{Fix}(\sigma\tau; \mathbb{F}_p^\times) \equiv p+1 \pmod 4$. Durch den Vergleich mit (5.3) folgt die Behauptung:

$$\#\{x \in \mathbb{F}_p^\times \ ; \ x^2 = -1\} = \#\operatorname{Fix}(\sigma\tau; \mathbb{F}_p^\times) = \begin{cases} 2 & p \equiv 1 \pmod 4, \\ 0 & p \equiv -1 \pmod 4. \end{cases}$$

Wir brauchen eine Notation für unteres und oberes Runden, die **untere und obere Gauß-Klammer**. Für $x \in \mathbb{R}$ setzen wir

$$\lfloor x \rfloor := \max\{n \in \mathbb{Z} \ ; \ n \leq x\},$$
$$\lceil x \rceil := \min\{n \in \mathbb{Z} \ ; \ n \geq x\}.$$

Theorem 5.39 (Fermat, Zwei-Quadrate-Satz)**.** *Jede Primzahl* $p \equiv 1 \pmod 4$ *ist Summe zweier Quadratzahlen.*

Beweis. Wir wählen nach Satz 5.37 ein $n \in \mathbb{Z}$ mit $n^2 \equiv -1 \pmod p$. Sei weiter $a = b = \lceil \sqrt{p} \rceil$. Weil $p \geq 5$ ist, haben wir $a = b \geq 2$. Weiter, weil $\sqrt{p} \notin \mathbb{Z}$ gilt $ab = \lceil \sqrt{p} \rceil^2 > p$ und wir können den Satz von Thue, Satz 5.33, anwenden. Wir erhalten $0 \leq x < \lceil \sqrt{p} \rceil$ und $1 \leq y < \lceil \sqrt{p} \rceil$, also

$$0 \leq x < \sqrt{p} \text{ und } 1 \leq y < \sqrt{p}$$

mit $n \equiv \pm \frac{x}{y} \pmod p$. Dann gilt

$$x^2 + y^2 = y^2(1 + \frac{x^2}{y^2}) \equiv y^2(1 + n^2) \equiv 0 \pmod p$$

und somit $p \mid x^2 + y^2$. Andererseits gilt

$$0 < 0^2 + 1^2 \leq x^2 + y^2 < \sqrt{p}^2 + \sqrt{p}^2 = 2p.$$

Das einzige Vielfache von p echt zwischen 0 und $2p$ ist p selbst. Daher gilt $p = x^2 + y^2$. □

Bemerkung 5.40. Theorem 5.39 erlaubt einen zweiten Blick auf Satz 5.37, denn nach Theorem 5.39 ist ein p wie im Satz darstellbar als $p = a^2 + b^2$, eine Summe zweier Quadrate. Wenn $p \mid a$, dann folgt $p \mid p - a^2 = b^2$, also $p \mid b$. Dann gilt aber $p^2 \mid a^2 + b^2 = p$, ein Widerspruch. Eine Lösung erhalten wir nun durch die Rechnung

$$(\frac{a}{b})^2 \equiv \frac{p - b^2}{b^2} \equiv -1 \pmod p.$$

Das ist kein neuer Beweis von Satz 5.37, weil wir die Aussage im Beweis von Theorem 5.39 benutzt haben.

Alternativer Beweis von Theorem 5.39. Wir erklären den Beweis nach Heath-Brown (1984) in der Version von Zagier [Za90], der aus einem einzigen englischen Satz besteht (wir brauchen mehr). Dazu betrachten wir die Menge

$$S = \{(x, y, z) \in \mathbb{N}^3 \ ; \ x^2 + 4yz = p\}.$$

Gesucht wird eine Lösung der Gleichung $x^2 + 4yz = p$, die ein Fixpunkt unter der Involution von S

$$\beta : (x, y, z) \mapsto (x, z, y)$$

ist. Dann ist nämlich $y = z$ und somit

$$p = x^2 + 4yz = x^2 + (2y)^2.$$

(Nebenbei: weil p ungerade ist, muß eines der gesuchten Quadrate gerade sein, und so ist die Gleichung $p = x^2 + (2y)^2$ keine Einschränkung gegenüber des Ziels des Theorems.)

Schritt 1: Die Menge S ist endlich. Weil $x, y, z \geq 1$ sein sollen, gilt $x \leq \sqrt{p}$ und $y, z \leq p/4$. Damit ist S endlich.

Schritt 2: Partitionierung. Es gilt $y - z < 2y$ für alle $(x, y, z) \in S$. Daher zerlegt sich die Menge S disjunkt in die folgenden drei Mengen

$$\begin{aligned} A &= \{(x, y, z) \in S \ ; \ x < y - z\} \\ B &= \{(x, y, z) \in S \ ; \ y - z < x < 2y\} \\ C &= \{(x, y, z) \in S \ ; \ 2y < x\}, \end{aligned}$$

je nachdem wo x im durch $y - z$ und $2y$ unterteilten Zahlenstrahl positioniert ist. Dabei ist noch zu prüfen, warum für $(x, y, z) \in S$ keine Gleichheit in den verwendeten Abschätzungen gelten kann. Aus $x = 2y$ folgt $4 \mid x^2 + 4yz = p$, Widerspruch. Und aus $x = y - z$ folgt $p = x^2 + 4yz = (y + z)^2$, was für eine Primzahl auch nicht geht.

Schritt 3: Eine neue Involution. Wir definieren eine neue Involution $\alpha : S \to S$, und zwar abschnittsweise linear durch

$$\alpha(x, y, z) = \begin{cases} (x + 2z, z, y - x - z) & \text{falls } (x, y, z) \in A, \\ (2y - x, y, x - y + z) & \text{falls } (x, y, z) \in B, \\ (x - 2y, x - y + z, y) & \text{falls } (x, y, z) \in C. \end{cases}$$

Hier ist einiges zu zeigen. Zunächst gilt wieder $\alpha(x, y, z) \in S$, denn

- Wenn $(x, y, z) \in A$, dann ist

  $$(x + 2z)^2 + 4z(y - x - z) = x^2 + 4xz + 4z^2 + 4zy - 4zx - 4z^2 = x^2 + 4yz = p,$$

  und $x + 2z > 0$, $z > 0$ und $y - x - z = (y - z) - x > 0$. Weiter gilt $x + 2z > 2z$, also ist

  $$\alpha(A) \subseteq C.$$

- Wenn $(x, y, z) \in B$, dann ist

  $$(2y - x)^2 + 4y(x - y + z) = x^2 - 4xy + 4y^2 + 4yx - 4y^2 + 4yz = x^2 + 4yz = p,$$

  und $2y - x > 0$, $y > 0$ und $x - y + z = x - (y - z) > 0$. Weiter gilt

  $$y - (x - y + z) = 2y - x - z < 2y - x < 2y,$$

  also ist

  $$\alpha(B) \subseteq B.$$

- Wenn $(x, y, z) \in C$, dann ist

  $$(x - 2y)^2 + 4(x - y + z)y = x^2 - 4xy + 4y^2 + 4yx - 4y^2 + 4yz = x^2 + 4yz = p,$$

  und $x - 2y > 0$, $x - y + z = x - (y - z) > 0$ und $y > 0$. Weiter gilt

  $$x - 2y < x - 2y + z = (x - y + z) - y$$

  also ist

  $$\alpha(C) \subseteq A.$$

Die Abbildung α ist involutiv, weil

- für $(x, y, z) \in A$

  $$\begin{aligned} \alpha(\alpha(x, y, z)) &= \alpha(x + 2z, z, y - x - z) \\ &= ((x + 2z) - 2z, (x + 2z) - z + (y - x - z), z) = (x, y, z), \end{aligned}$$

- für $(x, y, z) \in B$

  $$\begin{aligned} \alpha(\alpha(x, y, z)) &= \alpha(2y - x, y, x - y + z) \\ &= (2y - (2y - x), y, (2y - x) - y + (x - y + z)) = (x, y, z), \end{aligned}$$

- für $(x,y,z) \in C$
$$\begin{aligned}\alpha(\alpha(x,y,z)) &= \alpha(x-2y, x-y+z, y)\\ &= ((x-2y)+2y, y, (x-y+z)-(x-2y)-y) = (x,y,z).\end{aligned}$$

Schritt 4: Fixpunkte von α. Ein Fixpunkt der Involution α muß ein $(x,y,z) \in B$ sein mit demnach
$$(x,y,z) = \alpha(x,y,z) = (2y-x, y, x-y+z).$$
Das ist äquivalent zu $x = 2y - x$, also $x = y$. Da $(x,y,z) \in S$ folgt
$$p = x^2 + 4xz = x(x+4z).$$
Jetzt erst kommt so richtig zum Tragen, daß p eine Primzahl ist. Da $x + 4z > x > 0$ der größere Faktor ist, gilt
$$x = 1 \quad \text{und} \quad x + 4z = p.$$
Der einzige Fixpunkt von α ist somit
$$(1, 1, \frac{p-1}{4}) \in B,$$
womit auch $p \equiv 1 \pmod 4$ eingeht (das muß so sein, weil der Satz für $p \equiv 3 \pmod 4$ falsch ist).

Schritt 5: Ein Paritätsargument. Sei ι eine Involution einer endlichen Menge X und $F = \{x \in X \ ; \ \iota(x) = x\}$ die Menge der Fixpunkte. Dann ist
$$\#F \equiv \#X \pmod 2,$$
denn alle Punkte in $X \setminus F$ treten ja paarweise als $\{x, \iota(x)\}$ auf.

Die Menge S ist endlich und hat bezüglich der Involution α genau einen Fixpunkt. Die Fixpunkte bezüglich der Involution β sind genau die gesuchten Lösungen der Darstellung als Summe zweier Quadrate. Es gilt somit
$$1 = \#\{\text{Fixpunkte von } \alpha\} \equiv \#S \equiv \#\{\text{Fixpunkte von } \beta\} \pmod 2.$$
Insbesondere hat β mindestens einen Fixpunkt. □

Satz 5.41 (Jacobi[20]). *Eine natürliche Zahl n ist genau dann von der Form*
$$n = a^2 + b^2$$
mit $a, b \in \mathbb{Z}$, wenn alle Primteiler $p \mid n$ der Form $p \equiv 3 \pmod 4$ mit geradem Exponenten in der Primfaktorzerlegung von n vorkommen (d.h. für solche p ist $v_p(n) \equiv 0 \pmod 2$).

Beweis. Für $u, v, x, y \in \mathbb{Z}$ gilt
$$(u^2 + v^2) \cdot (x^2 + y^2) = (ux - vy)^2 + (uy + vx)^2.$$
Daher sind die Zahlen der Form $a^2 + b^2$ mit ganzen Zahlen a, b abgeschlossen unter Multiplikation.

Wenn $v_p(n)$ gerade ist für alle Primzahlen $p \equiv 3 \pmod 4$, dann kann man n in die folgenden Arten von Faktoren zerlegen

(a) $2 = 1^2 + 1^2$,
(b) eine Primzahl $p \equiv 1 \pmod 4$,
(c) ein Primzahlquadrat $p^2 = p^2 + 0^2$ für ein $p \equiv 3 \pmod 4$.

Nach dem Zwei-Quadrate-Satz, Theorem 5.39, kann jeder Faktor dieser Art als Summe zweier Quadrate geschrieben werden. Das gilt dann auch für n.

Wir zeigen nun die Umkehrung und zwar per Induktion über n. Der Induktionsanfang sind alle Fälle n ohne Primteiler p der Form $p \equiv 3 \pmod 4$.

Sei also $n = a^2 + b^2$ und $p \mid n$ mit $p \equiv 3 \pmod 4$. Wenn $p \nmid a$, dann auch $p \nmid b$ und mit $x = \frac{a}{b}$ wäre $x^2 \equiv -1 \pmod p$. Das ist ein Widerspruch zu Satz 5.35. Also gilt $p \mid a$ und dann auch

[20] Carl Gustav Jacob Jacobi (1804–1851), deutscher Mathematiker.

$p \mid b$. Zusammen folgt $p^2 \mid n$, und es gibt $\alpha, \beta, \mu \in \mathbb{Z}$ mit $n = p^2\mu$, $a = p\alpha$ und $b = p\beta$. Dann folgt

$$\mu = \alpha^2 + \beta^2$$

und der Beweis schließt per Induktion. □

Ohne Beweis erwähnen wir hier den 4-Quadrate-Satz.

Theorem 5.42 (Lagrange[21]). *Sein $n \in \mathbb{N}$ beliebig. Dann gibt es $a, b, c, d \in \mathbb{Z}$ mit*

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

Bemerkung 5.43. Die Aussage von Satz 5.41 findet eine algebraische Erklärung mittels der Gaußschen ganzen Zahlen

$$\mathbb{Z}[i] = \{a + bi \ ; \ a, b \in \mathbb{Z}\} \subseteq \mathbb{C}$$

und dem Zerlegungsverhalten der Primzahlen $p \in \mathbb{Z}$ darin. Dies werden wir später in Kapitel §15 behandeln. Eine Ahnung bekommt man vom Zusammenhang durch die Formel

$$a^2 + b^2 = (a + ib)(a - ib).$$

Der 4-Quadrate-Satz benutzt analog dazu die ganzzahligen Quaternionen (nach Hurwitz), also

$$\{a + bi + cj + dk \ ; \ a, b, c, d \in \mathbb{Z}\} \subseteq \mathbb{H}.$$

Der Zusammenhang hier bedient sich der Formel

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

5.5. **Die Sätze von Tschebyscheff zur Primzahlverteilung.** In diesem Abschnitt erarbeiten wir den Satz von Tschebyscheff, einen Vorläufer des Primzahlsatzes, der bis auf Konstanten bereits die richtige Asymptotik $x/\log(x)$ für die Primzahlzählfunktion

$$\pi(x) := \#\{p \leq x \ ; \ p \text{ Primzahl}\} = \sum_{p \leq x} 1$$

angibt. Der wesentliche Trick besteht in zwei Beobachtungen:

(1) Der Binomialkoeffizient

$$\binom{2n}{n}$$

wird von allen Primzahlen $n < p \leq 2n$ geteilt und man kann ihn gut nach oben und unten abschätzen.

(2) Anstelle von $\pi(x)$ betrachtet man die Primzahlen gewichtet mit ihrem Logarithmus

$$\vartheta(x) = \sum_{p \leq x} \log(p).$$

Die behauptete Asymptotik für $\pi(x)$ folgt aus (und ist in der Tat äquivalent zu) der linearen Asymptotik für $\vartheta(x)$. Eine solche Asymptotik ist leichter zu behandeln.

Das Gewicht $\log(p)$ für eine Primzahl p findet eine tiefere Begründung in Arakelov-Theorie. Für Physiker: die Zeta-Funktion ist die Zustandsfunktion eines (fiktiven?) Mehrteilchensystems aus Primzahlen p, deren Energie proportional zu $\log(p)$ ist.

Zuerst kümmern wir uns um die benötigte Abschätzung für den Binomialkoeffizienten.

Proposition 5.44. *Für alle $n \in \mathbb{N}$ gilt*

$$\frac{1}{\sqrt{4n}} 4^n \leq \binom{2n}{n} \leq \begin{cases} 4^{n-1} & n \geq 5 \\ 4^n & n \geq 1. \end{cases}$$

[21] Joseph-Louis Lagrange (1736–1813), französischer Mathematiker und Astronom italienischer Abstammung.

Beweis. Aus der binomischen Formel folgt zunächst

$$\binom{2n}{n} \leq \sum_{i=0}^{2n} \binom{2n}{i} = (1+1)^{2n} = 4^n.$$

Als nächstes wollen wir einen Faktor 4 sparen. Das gilt für $n = 5$ wegen

$$\binom{10}{5} = 252 < 256 = 4^4.$$

Weiter benutzen wir Induktion

$$\binom{2(n+1)}{n+1} = \frac{(2n+2)(2n+1)}{(n+1)^2} \cdot \binom{2n}{n} \leq \frac{4n+2}{n+1} \cdot 4^{n-1} \leq \frac{4n+4}{n+1} \cdot 4^{n-1} \leq 4^n.$$

Die untere Abschätzung folgt ebenfalls per Induktion und benutzt die Ungleichung vom arithmetischen versus geometrischen Mittel:

$$\begin{aligned}\binom{2(n+1)}{n+1} &= \frac{(2n+2)(2n+1)}{(n+1)^2} \cdot \binom{2n}{n} = \frac{4}{(n+1)} \cdot \frac{n+(n+1)}{2} \cdot \binom{2n}{n} \\ &\geq \frac{4}{(n+1)} \cdot \frac{n+(n+1)}{2} \cdot \frac{1}{\sqrt{4n}} 4^n \\ &\geq \frac{4}{(n+1)} \cdot \sqrt{n(n+1)} \cdot \frac{1}{\sqrt{4n}} 4^n = \frac{1}{\sqrt{4(n+1)}} 4^{n+1}.\end{aligned}$$

Der Induktionsanfang bei $n = 1$ gilt wegen $\binom{2}{1} = 2 = \frac{1}{\sqrt{4}}4$. □

Proposition 5.45. *Sei $n \in \mathbb{N}$ und p eine Primzahl. Dann gilt*

$$v_p(n!) = \sum_{m=1}^{\infty} \lfloor \frac{n}{p^m} \rfloor,$$

wobei die Summe nur bis $\lceil \log n / \log p \rceil$ summiert werden muß.

Beweis. Der Summand $\lfloor \frac{n}{p^m} \rfloor$ gibt an, wieviele der Zahlen von 1 bis n durch p^m geteilt werden. Durch die Summation wird die Zahl x mit $1 \leq x \leq n$ genau $v = v_p(x)$-mal gezählt: durch p teilbar, durch p^2 teilbar, ..., durch p^v teilbar und dann nicht mehr. Daher ist

$$\sum_{m=1}^{\infty} \lfloor \frac{n}{p^m} \rfloor = \sum_{x=1}^{n} v_p(x) = v_p(n!).$$ □

Korollar 5.46. *Sei $n \in \mathbb{N}$ und p eine Primzahl. Dann gilt*

$$v_p(\binom{2n}{n}) = \sum_{m=1}^{\infty} \lfloor \frac{2n}{p^m} \rfloor - 2\lfloor \frac{n}{p^m} \rfloor.$$

Es gilt speziell

$$v_p(\binom{2n}{n}) = \begin{cases} 1 & n < p \leq 2n, \\ 0 & \frac{2}{3}n < p \leq n \text{ und } n \geq 3, \\ \leq 1 & \sqrt{2n} < p. \end{cases}$$

Beweis. Die erste Aussage folgt sofort aus Proposition 5.45

$$v_p(\binom{2n}{n}) = v_p(\frac{(2n)!}{n!n!}) = v_p((2n)!) - 2v_p(n!) = \sum_{m=1}^{\infty} \lfloor \frac{2n}{p^m} \rfloor - 2\lfloor \frac{n}{p^m} \rfloor.$$

Wir bemerken nun, daß die Summanden für $x = n/p^m$ von der Form

$$f(x) = \lfloor 2x \rfloor - 2\lfloor x \rfloor$$

sind, das nur die Werte 0 und 1 annimmt. Wegen $\lfloor x+1 \rfloor = \lfloor x \rfloor + 1$, folgt $f(x) = f(x+1)$ für alle x, und es reicht aus, die Abschätzung $0 \leq f(x) \leq 1$ für $0 \leq x < 1$ nachzurechnen, wo es offensichtlich ist.

In allen speziell zu betrachtenden Fällen ist $p^2 > 2n$, und daher sind die Summanden 0 ab $m \geq 2$. Es bleibt also dann

$$v_p(\binom{2n}{n}) = \lfloor \frac{2n}{p} \rfloor - 2\lfloor \frac{n}{p} \rfloor \leq 1.$$

Falls $n < p \leq 2n$, gilt genauer $\lfloor \frac{2n}{p} \rfloor - 2\lfloor \frac{n}{p} \rfloor = 1 - 0 = 1$. Und falls $\frac{2}{3}n < p \leq n$ und $n \geq 3$, dann gilt genauer $\lfloor \frac{2n}{p} \rfloor - 2\lfloor \frac{n}{p} \rfloor = 2 - 2 \cdot 1 = 0$. □

Korollar 5.47. *Sei $n \in \mathbb{N}$ und p eine Primzahl. Dann gilt*

$$p^{v_p(\binom{2n}{n})} \leq 2n.$$

Beweis. Wenn $p^m > 2n$, dann ist der Summand $\lfloor \frac{2n}{p^m} \rfloor - 2\lfloor \frac{n}{p^m} \rfloor = 0$. Es sind genau $v_p(\binom{2n}{n})$ der Summanden $= 1$. Daraus folgt sofort die Aussage durch einen Widerspruch. Wäre $p^{v_p(\binom{2n}{n})} > 2n$, so könnten höchstens die Summanden zu $m = 1, \ldots, v_p(\binom{2n}{n}) - 1$ von 0 verschieden sein. Das ist mindestens einer zu wenig. Widerspruch. □

Satz 5.48. *Für alle $x \in \mathbb{R}$, $x \geq 1$ gilt*

$$\vartheta(x) \leq (x - \frac{1}{2}) \cdot \log 4.$$

Beweis. Aus Korollar 5.46 und Proposition 5.44 folgt für alle $n \geq 5$

$$\vartheta(2n) - \vartheta(n) \leq \log \binom{2n}{n} \leq (n-1) \cdot \log 4. \tag{5.4}$$

Explizites Nachrechnen liefert

n	1	2	3	4
$\vartheta(2n) - \vartheta(n)$	$\log 2$	$\log 3$	$\log 5$	$\log 35$
$(n-1) \cdot \log 4$	0	$\log 4$	$\log 16$	$\log 64$

und (5.4) gilt auch für $n \geq 2$.

Jetzt machen wir einen Teleskopsummenansatz. Dazu setzen wir

$$x_i := \lceil \frac{x}{2^i} \rceil.$$

Es gilt

$$x_i - 1 < \frac{x}{2^i} \leq x_i. \tag{5.5}$$

Für alle $y \in \mathbb{R}$ gilt $2\lceil \frac{y}{2} \rceil \geq \lceil y \rceil$. Speziell für $y = \frac{x}{2^i}$ folgt daraus für alle i

$$2x_{i+1} \geq x_i.$$

Daher überdecken für alle $r \geq 1$ die Intervalle $(x_i, 2x_i]$ mit $1 \leq i \leq r$ das Intervall $(x_r, x]$.

Im Fall $x \leq 2$ gilt die Aussage des Satzes trivialerweise. Sei daher von nun an $x > 2$. Dann gibt es ein $r \geq 1$ mit $x_r = 2$. Dies bedeutet nämlich $1 < \frac{x}{2^r} \leq 2$, was durch Logarithmieren und Umstellen äquivalent zu

$$r < \frac{\log(x)}{\log(2)} \leq r + 1$$

ist. Wir wählen somit speziell $r = \lceil \log(x)/\log(2) \rceil - 1$. Es gilt nun wegen (5.4) und (5.5) und der Überdeckung durch die Intervalle $(x_i, 2x_i]$

$$\begin{aligned}\vartheta(x) &\leq \vartheta(x_r) + \sum_{i=1}^{r} (\vartheta(2x_i) - \vartheta(x_i)) \leq \log 2 + \sum_{i=1}^{r} (x_i - 1) \cdot \log 4 \\ &\leq \log 2 + \sum_{i=1}^{r} \frac{x}{2^i} \cdot \log 4 = \log 2 + x \cdot (1 - \frac{1}{2^r}) \log 4 = x \cdot \log 4 + \log 2 - \frac{x}{2^r} \cdot \log 4 \\ &< x \cdot \log 4 + \log 2 - \log 4 = (x - 1/2) \cdot \log 4,\end{aligned}$$

weil $\frac{x}{2^r} > 1$. □

Lemma 5.49. *Es gilt für alle $x \in \mathbb{R}$, $x \geq 1$ die naive Abschätzung*

$$\pi(x) \leq \frac{x+5}{3}.$$

Beweis. Wenn $n \geq 3$, dann sind von den 6 Zahlen $n+1, \ldots, n+6$ höchstens die zu 6 teilerfremden Zahlen Primzahlen, denn die anderen sind durch 2 oder 3 teilbar und selbst größer als 3. Von diesen 6 Zahlen sind genau zwei zu 6 teilerfremd, nämlich die $\equiv \pm 1 \pmod 6$.

Wenn die Abschätzung für $x \geq 3$ gilt, dann gilt sie auch für $x + 6$: wir setzen $n = \lfloor x \rfloor$ und finden

$$\pi(x+6) = \pi(n+6) \leq \pi(n) + 2 = \pi(x) + 2 \leq \frac{x+5}{3} + 2 = \frac{(x+6)+5}{3}.$$

Daher reicht es, die Abschätzung für $1 \leq x < 9$ nachzuweisen. Das ist elementar, indem man die Abschätzung an den Sprungstellen auswertet (die rechte Seite ist monoton wachsend):

x	2	3	5	7
$\pi(x)$	1	2	3	4
$\frac{x+5}{3}$	7/3	8/3	10/3	4

□

Theorem 5.50 (Bertrands Postulat; Tschebyscheff)**.** *Zu jeder natürlichen Zahl n gibt es im Intervall $n < x \leq 2n$ eine Primzahl.*

Beweis. Wir argumentieren durch Widerspruch und nehmen an, das Intervall $(n, 2n]$ sei frei von Primzahlen. Dann hat $\binom{2n}{n}$ nach Korollar 5.46 nur Primfaktoren $p \leq \frac{2n}{3}$. Der Beitrag derjenigen $p \leq \sqrt{2n}$ in der Primfaktorzerlegung ist überdies $\leq 2n$, und für $p > \sqrt{2n}$ ist der Beitrag höchstens ein Faktor p. Aus Korollar 5.46 und Proposition 5.44 folgt somit

$$\begin{aligned}n \cdot \log 4 - \log(\sqrt{4n}) \leq \log \binom{2n}{n} &\leq \vartheta(\frac{2n}{3}) - \vartheta(\sqrt{2n}) + \sum_{p \leq \sqrt{2n}} \log(2n) \\ &= \vartheta(\frac{2n}{3}) + \sum_{p \leq \sqrt{2n}} \big(\log(2n) - \log(p)\big) \leq \vartheta(\frac{2n}{3}) + \sum_{p \leq \sqrt{2n}} \log(n).\end{aligned}$$

Die Anzahl der Primzahlen $p \leq \sqrt{2n}$ wird brutal nach Lemma 5.49 durch $(\sqrt{2n} + 5)/3$ abgeschätzt, so daß mit Satz 5.48 folgt

$$n \cdot \log 4 - \log \sqrt{4n} \leq (\frac{2n}{3} - \frac{1}{2}) \cdot \log 4 + \frac{\sqrt{2n} + 5}{3} \log(n).$$

Umsortieren und Multiplikation mit 3 liefert

$$n \cdot \log 4 \leq (\sqrt{2n} + \frac{13}{2}) \log(n).$$

Für $n \geq 25$ ist $\sqrt{2n} \geq \sqrt{49} = 7 > 13/2$. Das nehmen wir nun an und erhalten

$$n \cdot \log 4 \leq 2\sqrt{2n} \cdot \log(n),$$

oder äquivalent dazu

$$f(n) = \frac{\log(n)}{\log(2)} - \sqrt{\frac{n}{2}} \geq 0.$$

Die Ableitung von $f(x)$ ist

$$f'(x) = \frac{1}{x\log 2} - \frac{1}{\sqrt{8x}},$$

und bei $x > 0$

$$0 \geq f'(x) \iff \frac{1}{\sqrt{8x}} \geq \frac{1}{x\log 2} \iff x \geq \frac{8}{(\log(2))^2} \approx 16.6509.$$

Damit ist $f(x)$ ab $x = 128$ sicher monoton fallend, und wegen $f(128) = 7 - \sqrt{64} = -1 < 0$ dann auch negativ für $x \geq 128$. Das ist der gesuchte Widerspruch.

Für die kleinen Werte $n \leq 128$ kontrolliert man die Gültigkeit der Aussage des Theorems direkt durch eine Folge von Primzahlen p_i mit $p_{i+1} \leq 2p_i$, konkret etwa

$$2, 3, 5, 7, 13, 23, 43, 83, 163, \ldots. \qquad \square$$

Wir beweisen nun das historisch wichtige Resultat, daß nämlich der Primzahlsatz zumindest die Größenordnung der Primzahlzählfunktion trifft.

Satz 5.51. *Es gibt ein $x_0 \in \mathbb{R}$, so daß für alle $x \in \mathbb{R}$, $x \geq x_0$ gilt*

$$\vartheta(x) \geq \frac{x}{2}.$$

Zusatz: Es reicht $x_0 = 11$.

Beweis. Der Beweis beginnt parallel zum Beweis von Theorem 5.50. Für alle $n \in \mathbb{N}$ gilt nach Korollar 5.46, daß Primfaktoren $p > \sqrt{2n}$ in $\binom{2n}{n}$ höchstens einfach auftreten und nach Korollar 5.47 ist der Beitrag der kleinen Primzahlen $p \leq \sqrt{2n}$ höchstens $2n$. Daraus folgt

$$\begin{aligned}\log\binom{2n}{n} &\leq \vartheta(2n) - \vartheta(\sqrt{2n}) + \sum_{p\leq\sqrt{2n}} \log(2n) \\ &= \vartheta(2n) + \sum_{p\leq\sqrt{2n}} \big(\log(2n) - \log(p)\big) \leq \vartheta(2n) + \sum_{p\leq\sqrt{2n}} \log(n).\end{aligned}$$

Die Anzahl der Primzahlen $p \leq \sqrt{2n}$ wird brutal nach Lemma 5.49 durch $(\sqrt{2n}+5)/3$ abgeschätzt, so daß mittels Proposition 5.44 und Umstellen folgt

$$\begin{aligned}\vartheta(2n) &\geq n \cdot \log(4) - \log(\sqrt{4n}) - \frac{\sqrt{2n}+5}{3} \cdot \log(n) \\ &= (n - \frac{1}{2}) \cdot \log(4) - \big(\frac{\sqrt{2n}+5}{3} + \frac{1}{2}\big) \cdot \log(n)\end{aligned}$$

Aus dieser Formel für gerades $2n$ schließen wir mit $x \geq 2\lfloor\frac{x}{2}\rfloor = 2n \geq x - 2$

$$\begin{aligned}\vartheta(x) \geq \vartheta(2\lfloor\frac{x}{2}\rfloor) &\geq (\lfloor\frac{x}{2}\rfloor - \frac{1}{2}) \cdot \log(4) - \big(\frac{\sqrt{2\lfloor\frac{x}{2}\rfloor}+5}{3} + \frac{1}{2}\big) \cdot \log(\lfloor\frac{x}{2}\rfloor) \\ &\geq (x-3) \cdot \log(2) - \frac{2\sqrt{x}+13}{6} \cdot \log(x/2) \\ &= x \cdot (\log(2) - \varepsilon(x))\end{aligned}$$

mit dem Fehlerterm

$$\begin{aligned}\varepsilon(x) &= \frac{3\log(2)}{x} + \frac{2\sqrt{x}+13}{6x} \cdot \log(x/2) \\ &= \frac{3\log(2)}{x} + \big(\frac{2}{3\sqrt{2}} + \frac{13}{3\sqrt{2x}}\big) \cdot \frac{\log(\sqrt{x/2})}{\sqrt{x/2}}.\end{aligned}$$

Da $\lim_{t\to\infty} \frac{\log(t)}{t} = 0$, geht der Fehler $\varepsilon(x) \to 0$ für $x \to \infty$. Wir wählen x_0 derart, daß für alle $x \geq x_0$ gilt

$$\varepsilon(x) < \log(2) - \frac{1}{2} \approx 0.193147180559945\ldots.$$

Den Zusatz $x_0 = 11$ und die damit verbundene Kurvendiskussion von $\varepsilon(x)$ überlassen wir zur Übung. □

Bemerkung 5.52. Mit der Methode des Beweises von Satz 5.51 kann man für jedes $\varepsilon > 0$ und $x \geq x_\varepsilon$ die Abschätzung

$$(\log(2) - \varepsilon) \cdot x \leq \vartheta(x) \leq \log(4) \cdot x$$

zeigen (nach Erdős und Erdős–Kalmár). Es gilt numerisch

$$\log(2) = 0.693147180559945\ldots < 1 < 1.38629436111989\ldots = \log(4).$$

Die erreichte Abschätzung ist also noch ein wenig von der wahren Asymptotik

$$\vartheta(x) \sim x$$

für $x \to \infty$ entfernt.

Lemma 5.53. *Für alle $x \in \mathbb{R}$ mit $x > 1$ gilt*

$$\sqrt{x} \leq \frac{x}{\log x}.$$

Beweis. Substitutieren wir $x = y^2$, dann wird die Abschätzung äquivalent zu

$$\frac{\log(y)}{y} \leq \frac{1}{2}.$$

Das Maximum der Funktion $\frac{\log y}{y}$ für $y > 1$ bestimmt man leicht mit Differentialrechnung als den Wert bei der Nullstelle der Ableitung

$$\frac{d}{dy}\frac{\log y}{y} = \frac{1}{y^2} - \frac{\log y}{y^2}.$$

Die Nullstelle liegt bei $y = e$ und führt zum Wert $1/e$, und das ist kleiner als $1/2$. □

Theorem 5.54 (Tschebyscheff, 1851 (mit anderen Konstanten))**.** *Für alle $n \geq 3$ gilt*

$$\frac{1}{2} \cdot \frac{n}{\log n} \leq \pi(n) \leq (1 + 2\log 4) \cdot \frac{n}{\log n},$$

wobei numerisch $1 + 2\log 4 \approx 3.7725887\ldots$ gilt.

Beweis. Es gilt nach Satz 5.51 für $n \geq 11$

$$\pi(n) = \sum_{p \leq n} 1 \geq \frac{1}{\log(n)} \sum_{p \leq n} \log(p) = \frac{\vartheta(n)}{\log(n)} \geq \frac{1}{2} \cdot \frac{n}{\log n}.$$

Explizites Nachrechnen liefert

n	2	3	4	5	6	7	8	9	10
$\pi(n)$	1	2	2	3	3	4	4	4	4
$\frac{1}{2} \cdot \frac{n}{\log n}$	1.44	1.37	1.44	1.55	1.67	1.80	1.92	2.05	2.17

und damit die untere Abschätzung für $n \geq 3$.

Für die Abschätzung nach oben arbeiten wir mit der großzügigen Abschätzung $\vartheta(\sqrt{n}) \geq 0$ für alle $n \in \mathbb{N}$, sowie Lemma 5.53 und Satz 5.48:

$$\begin{aligned}\pi(n) &\leq \sqrt{n} + (\pi(n) - \pi(\sqrt{n})) \leq \sqrt{n} + \frac{\vartheta(n) - \vartheta(\sqrt{n})}{\log(\sqrt{n})} \\ &\leq \sqrt{n} + \frac{\vartheta(n)}{\log(\sqrt{n})} \leq \frac{n}{\log n} + \frac{n \cdot \log(4)}{1/2 \cdot \log(n)} = (1 + 2\log(4)) \cdot \frac{n}{\log n}.\end{aligned}$$

□

Bemerkung 5.55. Mit ähnlichen Techniken wie denen aus dem Beweis von Theorem 5.54 haben Paul Erdős und Atle Selberg einen „elementaren“ Beweis des Primzahlsatzes führen können.

6. Das abc der Arithmetik

6.1. **Die abc-Vermutung.** Ein bemerkenswerter Zusammenhang zwischen Addition und Primfaktorzerlegung wird durch die *abc*-**Vermutung** ausgedrückt.

Definition 6.1. Das **Radikal** einer natürlichen Zahl n ist definiert als

$$\mathrm{Rad}(n) = \prod_{p|n} p.$$

Vermutung 6.2 (Masser und Oesterlé 1985: die starke *abc*-Vermutung)**.** *Für jedes $\varepsilon > 0$ gibt es eine Konstante $C_\varepsilon > 0$, so daß für jedes Tripel $(a, b, c) \in \mathbb{Z}^3$ teilerfremder ganzer Zahlen mit*

$$a + b + c = 0$$

die folgende Abschätzung gilt:

$$\max\{|a|, |b|, |c|\} < C_\varepsilon \cdot \mathrm{Rad}(abc)^{1+\varepsilon}.$$

Bemerkung 6.3. (1) Die Vermutung beschreibt ein Spannungsverhältnis zwischen Addition und Multiplikation und quantifiziert die Beobachtung, daß wenn teilerfremde a und b durch hohe Potenzen kleiner Primzahlen teilbar sind, hingegen die Summe $-c = a + b$ dazu tendiert, durch kleine Potenzen großer Primzahlen teilbar zu sein.

(2) Typischerweise ist das Radikal größer als das Maximum:

$$2^4 + 3^2 = 5^2$$
$$1 + 5^4 = 2 \cdot 313$$

haben (313 ist Primzahl)

$$\max\{9, 16, 25\} = 25 < 30 = \mathrm{Rad}(2^4 \cdot 3^2 \cdot 5^2)$$
$$\max\{1, 625, 626\} = 626 < 3130 = \mathrm{Rad}(1 \cdot 5^4 \cdot 626).$$

In diesen Fällen reicht für alle $\varepsilon > 0$ bereits $C_\varepsilon = 1$.

(3) Sei $p > 2$ eine Primzahl. Wir lernen später in Satz 10.8, daß $p^2 \mid 2^{p(p-1)} - 1$. Dann führt

$$1 + (2^{p(p-1)} - 1) = 2^{p(p-1)}$$

zu

$$\begin{aligned} 2^{p(p-1)} = \max\{1, 2^{p(p-1)} - 1, 2^{p(p-1)}\} &< C_\varepsilon \cdot \mathrm{Rad}\left(1 \cdot (2^{p(p-1)} - 1) \cdot 2^{p(p-1)}\right)^{1+\varepsilon} \\ &\leq C_\varepsilon \cdot \left(2 \cdot \frac{2^{p(p-1)} - 1}{p}\right)^{1+\varepsilon} < C_\varepsilon \cdot \left(\frac{2^{p(p-1)+1}}{p}\right)^{1+\varepsilon} \end{aligned}$$

Dies zeigt, daß Vermutung 6.2 in der Variante für $\varepsilon = 0$ falsch ist, denn dann führt dies zu

$$2^{p(p-1)} < C_0 \cdot \frac{2^{p(p-1)+1}}{p}$$

und nach Umstellung zu $p < 2C_0$ im Widerspruch zur Tatsache, daß die Menge der Primzahlen nicht nach oben beschränkt ist.

Definition 6.4. Die **Qualität** $q(a, b, c) \in \mathbb{R}$ eines *abc*-Tripels, d.h. teilerfremder ganzer Zahlen $a, b, c \in \mathbb{Z}$ mit $a + b + c = 0$, ist definiert durch

$$\max\{|a|, |b|, |c|\} = \mathrm{Rad}(abc)^{q(a,b,c)}.$$

Bemerkung 6.5. Jedes *abc*-Tripel a, b, c stellt für jedes $\varepsilon > 0$ eine Bedingung an die Konstante C_ε. Endlich vielen Ausreißern, bei denen das Radikal klein gegen das Maximum ist, kann man durch Vergrößern von C_ε begegnen. Hätte man unendlich viele *abc*-Tripel a, b, c der Qualität $q(a, b, c) \geq q_0 > 1$, dann erfordern diese insbesondere ein C_ε mit

$$\mathrm{Rad}(abc)^{q_0} \leq \max\{|a|, |b|, |c|\} < C_\varepsilon \cdot \mathrm{Rad}(abc)^{1+\varepsilon}$$

also

$$\mathrm{Rad}(abc)^{q_0-1-\varepsilon} < C_\varepsilon.$$

Ferner ist unter diesen *abc*-Tripeln a, b, c der Qualität $q(a, b, c) \geq q_0$ das Radikal $\mathrm{Rad}(abc)$ unbeschränkt, weil sonst das Maximum ebenfalls beschränkt wäre und dies nur endlich viele Tripel erlaubt. Da für $0 < \varepsilon < q_0 - 1$ der Exponent $q_0 - 1 - \varepsilon > 0$ wird, kann kein C_ε allen diesen Bedingungen genügen: die linke Seite wird beliebig groß.

Die Vermutung 6.2 verlangt also insbesondere, daß für jedes $q_0 > 1$ die Qualität $q(a, b, c)$ nur endlich oft $\geq q_0$ ist. Die Qualität mißt also, vor welche „Probleme“ ein *abc*-Tripel die Vermutung stellt.

Beispiel 6.6. Hier ist eine Liste der Rekorde in Bezug auf die Qualität eines *abc*-Tripels: nach dem Maximum sortiert ist die Qualität eines Rekords größer als die Qualität aller kleineren Tripel:

a	b	$-c$	$q(a, b, c)$
1	2^3	3^2	1.2263
1	$2^4 \cdot 5$	3^4	1.2920
3	5^3	2^7	1.4266
1	$2^5 \cdot 3 \cdot 5^2$	7^4	1.4557
1	$2 \cdot 3^7$	$5^4 \cdot 7$	1.5679
2	$3^{10} \cdot 109$	23^5	1.6299

Das letzte Beispiel

$$2 + 3^{10} \cdot 109 = 23^5$$

wurde von Eric Reyssat gefunden und ist gegenwärtig das *abc*-Tripel mit der größten bekannten Qualität. Es beruht auf der erstaunlich guten Approximation

$$\sqrt[5]{109} = 2,55555539\ldots \approx \frac{23}{9}.$$

Gemäß http://www.math.leidenuniv.nl/∼desmit/abc/ gibt es (Stand 2016) mit $\max < 10^{20}$ nur **239** bekannte *abc*-Tripel mit Qualität $> 1,4$.

Wir illustrieren nun die Faszination, die von der *abc*-Vermutung ausgeht, indem wir den großen Satz von Fermat für hinreichend große Exponenten folgern.

Satz 6.7 (Asymptotischer Fermat)**.** *Die abc-Vermutung mit Konstanten* $0 < \varepsilon < 1$ *und* $C = C_\varepsilon \geq 1$ *zeigt, daß für*

$$n \geq 3(1 + \varepsilon) + \log_3(C)$$

die Gleichung

$$X^n + Y^n = Z^n$$

keine ganzzahligen Lösungen $x, y, z \in \mathbb{Z}$ *mit* $xyz \neq 0$ *hat.*

Beweis. Sei x, y, z eine ganzzahlige Lösung $x^n + y^n = z^n$ mit $xyz \neq 0$. Durch Kürzen mit dem ggT dürfen wir annehmen, daß x, y und z teilerfremd sind. Außerdem können wir durch Permutieren der Variablen und geeignete Multiplikation mit -1 erreichen, daß $0 < x < y < z$ sind. Insbesondere ist $z \geq 3$.

Wir wenden die Ungleichung der abc-Vermutung auf $a = x^n$, $b = y^n$ und $c = z^n$ an, die dann auch teilerfremd sind. Das ergibt

$$z^n = \max\{x^n, y^n, z^n\} < C \cdot \mathrm{Rad}((xyz)^n)^{1+\varepsilon} = C \cdot \mathrm{Rad}(xyz)^{1+\varepsilon} \leq C \cdot (xyz)^{1+\varepsilon} < C \cdot z^{3(1+\varepsilon)}.$$

Da nach Voraussetzung $n - 3(1+\varepsilon) \geq \log_3(C) \geq 0$ und $z \geq 3$ folgt der Widerspruch

$$C = 3^{\log_3(C)} \leq z^{n-3(1+\varepsilon)} < C. \qquad \square$$

Korollar 6.8. *Die abc-Vermutung mit Konstanten $\varepsilon = 1$ und $C_\varepsilon = 1$:*

$$\max\{|a|, |b|, |c|\} < \mathrm{Rad}(abc)^2$$

impliziert den großen Fermat'schen Satz.

Beweis. Nach Satz 6.7 führen die explizit gewählten Konstanten zu einem Beweis für $n \geq 6$. Die Fälle $n = 3$ (Euler), $n = 4$ (Fermat) and $n = 5$ (Dirichlet, Legendre) sind als klassische Resultate bereits bewiesen. $\square$

6.2. ***abc* für Polynome.** Als Literatur zu diesem Abschnitt sei auf den berühmten und sehr lesenswerten Vortrag/Text von Serge Lang [Lan93] verwiesen.

Die Arithmetik in $\mathbb{Z}$ und die in Polynomringen $k[T]$ über einem Körper k haben große Ähnlichkeiten (und gewisse Unterschiede). Oft lassen sich zahlentheoretische Aussagen zu ganzen Zahlen in Aussagen über Polynome übersetzen, und umgekehrt, sofern man das richtige Wörterbuch findet. Dies ist eine Maschine zum Beweis neuer Sätze, aber auch Inspiration für Beweise und Intuition.

Definition 6.9. Für ein Polynom $0 \neq f \in \mathbb{C}[T]$ ist das **logarithmische Radikal** gegeben als

$$r(f) = \#\{\alpha \in \mathbb{C}\ ;\ f(\alpha) = 0\}.$$

Bemerkung 6.10. (1) Es gilt stets $r(f) \leq \deg(f)$: ein Polynom hat höchstens so viele Nullstellen wie sein Grad.

(2) Es kann $r(f)$ sehr klein gegenüber $\deg(f)$ sein, z.B. für $f(T) = (T-\alpha)^m$ gilt

$$r(f) = 1 \leq m = \deg(f).$$

(3) Zu einer Nullstelle α von f gehört ein Linearfaktor $T-\alpha$ von $f(T)$, ein Primteiler von $f(T)$. Während das Radikal einer ganzen Zahl die Primfaktoren multipliziert, zählen (additiv) wir im Fall von Polynomen. Daher hat $r(f)$ einen logarithmischen Charakter im Vergleich zum arithmetischen Fall beim Radikal von ganzen Zahlen.

Nach dem Radikal benötigen wir auch noch ein Analogon für den Absolutbetrag einer ganzen Zahl. Dafür nehmen wir als logarithmisches Maß für die Größe eines Polynoms den Grad

$$\deg(f),$$

der sich bei Produkten additiv verhält.

Theorem 6.11 (Mason [Mas84] and Stothers [Sto81], *abc* für Polynome)**.** *Seien $f, g, h \in \mathbb{C}[T]$ ohne gemeinsame Nullstelle und nicht alle konstant, so daß*

$$f + g + h = 0.$$

Dann gilt

$$\max\{\deg(f), \deg(g), \deg(h)\} \leq r(fgh) - 1.$$

Bemerkung 6.12. (1) Bemerkenswerterweise gilt das Analogon der *abc*-Vermutung für Polynome bereits mit $\varepsilon = 0$. (Das ε würde hier aufgrund der im Vergleich zu *abc* logarithmierten Form der Abschätzung als Faktor $1 + \varepsilon$ auftreten.)

(2) Es ist wahrscheinlich, daß Masser und Oesterlé bei der Formulierung der Vermutung 6.2 vom analogen Fall in Polynomringen beeinflußt waren.

(3) Was hat der Polynomring $\mathbb{C}[T]$ den ganzen Zahlen $\mathbb{Z}$ voraus? Man kann Differenzieren!

Zum Beweis benötigen wir ein Lemma. Auch für Polynome existiert ein größter gemeinsamer Teiler, weil wie $\mathbb{Z}$ der Ring $\mathbb{C}[T]$ ein euklidischer Ring und damit Hauptidealring ist.

Lemma 6.13. *Seien $f, g \in \mathbb{C}[T]$ ohne gemeinsame Nullstelle. Dann gilt:*

(1) $r(fg) = r(f) + r(g)$,
(2) $\deg(\mathrm{ggT}(f, f')) = \deg(f) - r(f)$.

Beweis. Behauptung (1) folgt sofort aus der Definition. Für Behauptung (2) nutzen wir den Fundamentalsatz der Algebra und faktorisieren f vollständig in Linearfaktoren:

$$f(T) = \prod_{i=1}^{r} (T - \alpha_i)^{m_i}$$

mit paarweise verschiedenen $\alpha_i \in \mathbb{C}$ und $m_i \geq 1$. Dann gilt für jedes i

$$\begin{aligned} f'(T) &= (T - \alpha_i)^{m_i - 1} \cdot (m_i \cdot \prod_{j \neq i} (T - \alpha_j)^{m_j} + (T - \alpha_i) \cdot \frac{d}{dT} \prod_{j \neq i} (T - \alpha_j)^{m_j}) \\ &= (T - \alpha_i)^{m_i - 1} \cdot Q_i(T) \end{aligned}$$

mit $Q_i(\alpha_i) \neq 0$. Daher ist

$$\mathrm{ggT}(f, f') = \prod_{i=1}^{r} (T - \alpha_i)^{m_i - 1}$$

und die Behauptung (2) folgt aus $r = r(f)$ durch Vergleich der Grade beider Seiten. □

Beweis von Theorem 6.11 nach Snyder [Sny00]. Aus $f + g + h = 0$ folgt

$$\frac{g}{h} = -1 - \frac{f}{h}$$

und nach Ableiten mittels Quotientenregel

$$\frac{hg' - gh'}{h^2} = \frac{fh' - hf'}{h^2}.$$

Der Zähler $F := fh' - hf' = hg' - gh'$ hat die folgenden Eigenschften:

- Es gilt $F \neq 0$, weil sonst $(g/h)' = 0$ und somit $g(T) = c \cdot h(T)$ mit $c \in \mathbb{C}$ der Voraussetzung an die Nullstellen widerspricht, es sei denn, g und h sind konstant, was auch nicht erlaubt ist.
- Die Polynome $\mathrm{ggT}(f, f')$ und $\mathrm{ggT}(g, g')$ und $\mathrm{ggT}(h, h')$ teilen F und haben als Teiler jeweils von f, g und h keine gemeinsamen Nullstellen.
- Beim Ableiten geht der Grad um eins runter:

$$\deg(F) \leq \deg(g) + \deg(h) - 1.$$

Daher folgt

$$\mathrm{ggT}(f, f') \cdot \mathrm{ggT}(g, g') \cdot \mathrm{ggT}(h, h') \mid F$$

und mit Lemma 6.13

$$\deg(f) - r(f) + \deg(g) - r(g) + \deg(h) - r(h) \leq \deg(F) \leq \deg(g) + \deg(h) - 1.$$

Ergo

$$\deg(f) \leq r(f) + r(g) + r(h) - 1 = r(fgh) - 1$$

und das beweist das Theorem aufgrund der Symmetrie durch Permutationen von f, g, h. □

Korollar 6.14. *Sei $n \geq 3$. Dann hat die Gleichung*

$$f^n + g^n = h^n$$

keine Lösung in Polynomen $f, g, h \in \mathbb{C}[T]$ mit $fgh \neq 0$ und nicht alle konstant.

Beweis. Sei $d = \max\{\deg(f), \deg(g), \deg(h)\} > 0$. Dann folgt aus Theorem 6.11 ein Widerspruch durch

$$3d \leq n \cdot d \leq \max\{\deg(f^n), \deg(g^n), \deg(h^n)\} \leq r((fgh)^n) - 1 = r(fgh) - 1 \leq 3d - 1. \qquad \square$$

Literatur

[Lan93] Serge Lang. Die *abc*-Vermutung. *Elem. Math.*, 48(3):89–99, 1993.
[Mas84] R. C Mason. Equations over function fields. 1068:149–157, 1984.
[Sny00] Noah Snyder. An alternate proof of Mason's theorem. *Elem. Math.*, 55(3):93–94, 2000.
[Sto81] W. W Stothers. Polynomial identities and Hauptmoduln. *Quart. J. Math. Oxford Ser. (2)*, 32(127):349–370, 1981.

7. Zahlentheoretische Funktionen

Wir betrachten Funktionen, die nur auf der Menge der natürlichen Zahlen $\mathbb{N}$ definiert sind, und zwar durch spezielle Beispiele und bezüglich der Operation der Faltung.

7.1. Multiplikative Funktionen.

Definition 7.1. Eine **arithmetische Funktion** ist eine Abbildung $f : \mathbb{N} \to \mathbb{C}$.

(1) Eine arithmetische Funktion f ist **multiplikativ**, wenn $f(1) = 1$ und
$$f(nm) = f(n)f(m) \tag{7.1}$$
für alle teilerfremden $n, m \in \mathbb{N}$ gilt.

(2) Eine arithmetische Funktion f ist **vollständig multiplikativ**, wenn $f(1) = 1$ und (7.1) für alle $n, m \in \mathbb{N}$ gilt.

Beispiel 7.2. (1) Die **Teileranzahl**
$$\tau(n) = \#\{d \ ; \ d \mid n\}$$
ist multiplikativ. In der Tat ist für teilerfremde $n, m \in \mathbb{Z}$ die Multiplikation
$$\begin{aligned} \{d \ ; \ d \mid n\} \times \{e \ ; \ e \mid m\} &\to \{f \ ; \ f \mid nm\} \\ (d, e) &\mapsto de \end{aligned} \tag{7.2}$$
eine Bijektion. Die Umkehrabbildung bildet $f \mid nm$ auf $d = (f, n) \mid n$ und $e = (f, m) \mid m$ ab. Übung! Durch Vergleich der Kardinalität beider Seiten erhält man
$$\tau(nm) = \tau(n)\tau(m).$$

(2) Die Funktion $\mathbf{1} : \mathbb{N} \to \mathbb{C}$ mit $\mathbf{1}(n) = 1$ konstant 1 ist vollständig multiplikativ.

(3) Die Funktion $\delta : \mathbb{N} \to \mathbb{C}$ mit
$$\delta(n) = \begin{cases} 1 & n = 1, \\ 0 & n > 1, \end{cases}$$
ist vollständig multiplikativ.

(4) Für jedes $k \in \mathbb{N}_0$ ist die Funktion n^k multiplikativ. Der Fall $k = 0$ ist gerade $\mathbf{1}$, und $k = 1$ ist die Inklusion $\mathrm{id} : \mathbb{N} \hookrightarrow \mathbb{C}$.

Bemerkung 7.3. (1) Die Forderung $f(1) = 1$ an eine multiplikative Funktion f schließt nur die Funktion $f(n) = 0$, für alle $n \in \mathbb{N}$, aus. Denn aus $n = m = 1$ folgt $f(1) = f(1)^2$, also $f(1) = 1$ oder $f(1) = 0$. In letzterem Fall folgt dann $f(n) = f(n)f(1) = 0$ für alle n.

(2) Die Menge der multiplikativen Funktionen kein $\mathbb{C}$-Vektorraum mit der üblichen werteweisen Addition und $\mathbb{C}$-Skalarmultiplikation. Im Gegensatz dazu bilden die Menge der arithmetischen Funktionen einen $\mathbb{C}$-Vektorraum, den Vektorraum der mit $\mathbb{N}$ indizierten Folgen komplexer Zahlen.

(3) Anstelle von $\mathbb{C}$ als Wertebereich für arithmetische oder (vollständig) multiplikative Funktionen kann man jeden anderen beliebigen Ring nehmen.

Definition 7.4. Die **Faltung** zweier arithmetischen Funktionen f, g ist die arithmetische Funktion $f * g : \mathbb{N} \to \mathbb{C}$ gegeben für alle $n \in \mathbb{N}$ durch
$$f * g(n) = \sum_{d \mid n} f(d) g(\frac{n}{d}).$$

Proposition 7.5. *Die Faltung hat die folgenden Eigenschaften. Seien f, g und h arithmetische Funktionen.*

(1) *Wenn f und g multiplikativ sind, dann ist $f * g$ multiplikativ.*
(2) $f * g = g * f$.
(3) $f * \delta = f = \delta * f$.
(4) $f * (g * h) = (f * g) * h$.

Beweis. Wir zeigen (1), der Rest folgt sofort oder durch entsprechendes triviales Nachrechnen und wird zur Übung überlassen. Seien n, m teilerfremde natürliche Zahlen. Dann ist

$$\begin{aligned} f * g(nm) &= \sum_{d|nm} f(d)g(\frac{nm}{d}) = \sum_{x|n,\ y|m} f(xy)g(\frac{nm}{xy}) \\ &= \sum_{x|n,\ y|m} f(x)f(y)g(\frac{n}{x})g(\frac{m}{y}) \\ &= \sum_{x|n} f(x)g(\frac{n}{x}) \cdot \sum_{y|m} f(y)g(\frac{m}{y}) \\ &= f * g(n) \cdot f * g(m). \end{aligned}$$

Der Übergang von $d \mid nm$ zu $x \mid n$ und $y \mid m$ mit $d = xy$ beruht auf der Bijektion (7.2). □

Beispiel 7.6. Die folgenden Funktionen sind als Faltung multiplikativer Funktionen multiplikativ.

(1) $\tau = \mathbf{1} * \mathbf{1}$, also
$$\tau(n) = \sum_{d|n} 1$$
die Teileranzahlfunktion (das wissen wir schon!).

(2) für jedes $k \in \mathbb{N}_0$ die Funktion $\sigma_k = \mathbf{1} * (n^k)$, also
$$\sigma_k(n) = \sum_{d|n} d^k$$
die Summe der k-ten Potenzen der Teiler.

(3) Speziell $\sigma = \sigma_1 = \mathbf{1} * \mathrm{id}$, die Teilersumme
$$\sigma(n) = \sum_{d|n} d.$$

7.2. **Vollkommene Zahlen.** Wir behandeln nun ein weiteres antikes Thema, das bis heute ungelöste Fragen aufwirft.

Definition 7.7. Eine **vollkommene Zahl** ist eine natürliche Zahl $n > 1$, die gleich der Summe ihrer Teiler $< n$ ist, d.h.,
$$2n = \sigma(n) = \sum_{d|n} d.$$

Beispiel 7.8. Hier sind ein paar vollkommene Zahlen und ein paar, die es nicht sind.

$$\begin{aligned} 6 &= 1 + 2 + 3 \\ 28 &= 1 + 2 + 4 + 7 + 14 \\ 496 &= 1 + 2 + 4 + 8 + 16 + 31 + 62 + 124 + 248 \\ 24 &\neq 1 + 2 + 3 + 4 + 6 + 8 + 12 = 35 \\ 100 &\neq 1 + 2 + 4 + 5 + 10 + 20 + 25 + 50 = 117 \end{aligned}$$

Satz 7.9 (Euklid–Euler)**.** *Eine gerade Zahl n ist genau dann eine vollkommene Zahl, wenn es eine Mersenne-Primzahl $2^p - 1$ gibt und*

$$n = 2^{p-1} \cdot (2^p - 1).$$

Beweis. Sei $n = 2^{p-1}(2^p - 1)$ mit $q = 2^p - 1$ Primzahl. Dann gilt für die mulitplikative Teilerfunktion

$$\sigma(n) = \sigma(q)\sigma(2^{p-1}) = (q+1) \cdot \sum_{i=0}^{p-1} 2^i = 2^p \cdot (2^p - 1) = 2n.$$

Dies war Euklids Anteil des Beweises. Wir müssen nun zeigen, daß eine natürliche Zahl

$$n = 2^a \cdot b$$

mit $a \geq 1$ und b ungerade nur dann eine vollkommene Zahl sein kann, wenn $b = 2^{a+1} - 1$ und b Primzahl ist. Sei also n vollkommen, so daß gilt:

$$2^{a+1} \cdot b = 2n = \sigma(n) = (2^{a+1} - 1)\sigma(b).$$

Da $2^{a+1} - 1 > 1$ ein ungerader Teiler von $2n$ ist, gibt es $c \in \mathbb{N}$ mit $b = (2^{a+1} - 1) \cdot c > c$. Dann haben wir

$$2^{a+1} \cdot c = \sigma(b) \geq b + c = (2^{a+1} - 1) \cdot c + c = 2^{a+1} \cdot c.$$

Die Ungleichung ist somit eine Gleichung und b, c sind die einzigen zwei Teiler von b. Deshalb ist b Primzahl, $c = 1$ und $b = 2^{a+1} - 1$ wie verlangt. □

Bemerkung 7.10. Man weiß nicht, ob es ungerade vollkommene Zahlen gibt. Jedenfalls ist keine bekannt und man kann auch nicht beweisen, daß es keine solchen gibt.

7.3. **Formale Potenzreihen.** Wir erinnern an den Ring der formalen Potenzreihen mit komplexen Koeffizienten

$$\mathbb{C}[[T]] = \{\sum_{k=0}^{\infty} a_k T^k \ ; \ a_k \in \mathbb{C}\}.$$

Bei den formalen Potenzreihen macht man sich das Leben leichter, indem man auf Konvergenzüberlegungen verzichtet. Allerdings handelt man sich damit ein, daß formale Potenzreihen nur in $T = 0$ ausgewertet werden können: dies ergibt den Koeffizienten von T^0, den konstanten Koeffizienten.

Definition 7.11. Sei $f : \mathbb{N} \to \mathbb{C}$ eine multiplikative Funktion. Zu jeder Primzahl p definieren wir die komplexe formale Potenzreihe

$$\lambda_p(f, T) = \sum_{k=0}^{\infty} f(p^k) T^k \quad \in 1 + T\mathbb{C}[[T]]$$

mit konstantem Term $\lambda_p(f, 0) = 1$.

Beispiel 7.12. Wenn f vollständig multiplikativ ist und $f(p) = a_p$ für eine Primzahl p, dann gilt

$$\lambda_p(f, T) = \sum_{k=0}^{\infty} f(p^k) T^k = \sum_{k=0}^{\infty} a_p^k T^k = \sum_{k=0}^{\infty} (a_p T)^k = \frac{1}{1 - a_k T},$$

wobei der gebrochen rationale Ausdruck in $\mathbb{C}[[T]]$ nur durch die Potenzreihe definiert ist.

Somit folgt speziell für die Funktion $\mathbf{1}$

$$\lambda_p(\mathbf{1}, T) = \sum_{k=0}^{\infty} T^k = \frac{1}{1 - T},$$

und für δ berechnen wir

$$\lambda_p(\delta, T) = 1.$$

Proposition 7.13. *(1) Eine multiplikative Funktion $f : \mathbb{N} \to \mathbb{C}$ ist eindeutig durch die Werte $f(p^k)$ für alle Primzahlen p und alle $k \in \mathbb{N}$ festgelegt.*

(2) Umgekehrt gibt es zu jedem Tupel von Werten $a_{p,k} \in \mathbb{C}$ mit p Primzahl und $k \in \mathbb{N}$ genau eine multiplikative Funktion f mit $f(p^k) = a_{p,k}$ für alle p und $k \geq 1$.

(3) Die Abbildung $f \mapsto (\lambda_p(f,T))_p$, die einer multiplikativen Funktion das Tupel aller Potenzreihen $\lambda_p(f,T)$ zuordnet, ist eine Bijektion

$$\{\textit{multipl. Funktion}\} \leftrightarrow \{(\lambda_p)_{p\ prim}\ ;\ \lambda_p \in 1 + T\mathbb{C}[[T]]\}$$

Beweis. (1) Sei f multiplikativ. Wenn $n = \prod_{i=1}^r p_i k_i$, dann gilt

$$f(n) = \prod_{i=1}^r f(p_i^{k_i}).$$

(2) Umgekehrt können wir nach Vorgabe der $a_{p,k}$ durch

$$f(n) = \prod_{p|n} a_{p,v_p(n)}$$

eine arithmetische Funktion $f : \mathbb{N} \to \mathbb{C}$ definieren. Die Definition ist wohldefiniert aufgrund des Fundamentalsatzes der Arithmetik. Damit beweist man auch, daß die so definierte Funktion multiplikativ ist.

(3) folgt sofort aus (1) und (2). □

Proposition 7.14. *Seien f, g multiplikative Funktionen und p eine Primzahl. Dann gilt*

$$\lambda_p(f * g, T) = \lambda_p(f,T)\lambda_p(g,T).$$

Beweis. Das ist eine einfache formale Rechnung:

$$\begin{aligned}
\lambda_p(f * g, T) &= \sum_{k\geq 0} (f * g)(p^k)T^k = \sum_{k\geq 0} \big(\sum_{d|p^k} f(d)g(p^k/d)\big)T^k \\
&= \sum_{k\geq 0} \big(\sum_{i=0}^{k} f(p^i)g(p^{k-i})\big)T^k = \big(\sum_{k=0}^{\infty} f(p^k)T^k\big) \cdot \big(\sum_{k=0}^{\infty} g(p^k)T^k\big) \\
&= \lambda_p(f,T)\lambda_p(g,T).
\end{aligned}$$

□

Proposition 7.15. *Eine arithmetische Funktion $f : \mathbb{N} \to \mathbb{C}$ ist von der Form*

$$f = g_1 * \ldots * g_r$$

für vollständig multiplikative Funktionen $g_1, \ldots, g_r$ genau dann, wenn die folgenden Bedingungen erfüllt sind:

(i) f ist multiplikativ,

(ii) für alle Primzahlen p ist $\lambda_p(f,T)$ das multiplikative Inverse eines Polynoms aus $\mathbb{C}[T] \subseteq \mathbb{C}[[T]]$ vom Grad $\leq r$.

Beweis. Wenn $f = g_1 * \ldots * g_r$ gilt mit vollständig multiplikative Funktionen $g_1, \ldots, g_r$, dann ist f multiplikativ nach Proposition 7.5. Außerdem ist wegen Proposition 7.14

$$\lambda_p(f,T)^{-1} = \prod_{i=1}^r \lambda_p(g_i,T)^{-1} = \prod_{i=1}^r (1 - g_i(p)T)$$

ein Polynom vom Grad $\leq r$.

Für die Umkehrung zerlegen wir nach dem Fundamentalsatz der Algebra für alle Primzahlen p die Polynome $\lambda_p(f,T)^{-1}$ in Linearfaktoren, und zwar geht das wegen $\lambda_p(f,0) = 1$ in der Form

$$\lambda_p(f,T)^{-1} = \prod_{i=1}^r (1 - \alpha_{i,p}T)$$

mit $\alpha_{1,p}, \ldots, \alpha_{r,p} \in \mathbb{C}$. Wenn der Grad $< r$ ist, sind entsprechend viele der $\alpha_{i,p} = 0$.

Dann definieren wir vollständig multiplikative Funktionen g_i, $i = 1, \ldots, r$ eindeutig dadurch, daß wir fordern

$$g_i(p) = \alpha_{i,p}.$$

Aus Proposition 7.14 und Beispiel 7.12 folgt

$$\lambda_p(f, T) = \lambda_p(g_1 * \ldots * g_r, T),$$

woraus die gewünschte Form $f = g_1 * \ldots * g_r$ mittels Proposition 7.13 folgt. □

7.4. **Möbiusinversion.**

Definition 7.16. Die **Möbiusfunktion** ist die multiplikative Funktion $\mu : \mathbb{N} \to \mathbb{C}$ mit

$$\mu(p^k) = \begin{cases} -1 & k = 1, \\ 0 & k > 1. \end{cases}$$

Definition 7.17. Eine ganze Zahl n ist **quadratfrei**, wenn es kein $a \in \mathbb{N}$, $a > 1$ gibt mit $a^2 \mid n$. Äquivalent dazu ist die Bedingung, daß $v_p(n) \leq 1$ für alle Primzahlen p gilt.

Bemerkung 7.18. Für eine quadratfreie natürliche Zahl n gilt

$$\mu(n) = (-1)^{\#\{p \mid n \ ; \ p \text{ Primzahl}\}},$$

und $\mu(n) = 0$ sonst. Außerdem berechnet man leicht die zu einer Primzahl p gehörende Potenzreihe zu

$$\lambda_p(\mu, T) = 1 - T.$$

Es folgt

$$\lambda_p(\mu * \mathbf{1}, T) = \lambda_p(\mu, T)\lambda_p(\mathbf{1}, T) = 1 = \lambda_p(\delta, T),$$

was eigentlich bereits das folgende Lemma beweist. Der Beweis, den wir gleich führen, enthält die wesentliche Rechnung aus der Relation mit den Potenzreihen, ohne diese explizit zu erwähnen.

Lemma 7.19. *Es gilt* $\mu * \mathbf{1} = \delta$.

Beweis. Als multiplikative Funktion müssen wir $\mu * \mathbf{1}$ nur für $n = p^k$ eine Primpotenz (p Primzahl und $k \geq 1$) mit δ vergleichen. Es gilt

$$\mu * \mathbf{1}(p^k) = \sum_{d \mid p^k} \mu(d) = \mu(1) + \mu(p) = 1 + (-1) = 0 = \delta(p^k).$$ □

Korollar 7.20 (Möbiussche Umkehrformel)**.** *Für eine arithmetische Funktion* $F = f * \mathbf{1}$ *gilt* $f = F * \mu$, *also aus*

$$F(n) = \sum_{d \mid n} f(d) \qquad \forall n \in \mathbb{N}$$

folgt für alle $n \in \mathbb{N}$

$$f(n) = \sum_{d \mid n} F(d)\mu(\frac{n}{d}).$$

Beweis. Es gilt $F * \mu = (f * \mathbf{1}) * \mu = f * (\mathbf{1} * \mu) = f * \delta = f$. □

Definition 7.21. Die **Eulersche φ-Funktion** ist die Funktion $\varphi : \mathbb{N} \to \mathbb{C}$ definiert durch

$$\varphi(n) = \#\{a \ ; \ 1 \leq a \leq n \text{ und } (a, n) = 1\}$$

die Anzahl der zu n teilerfremden natürlichen Zahlen $\leq n$.

Proposition 7.22. *Es gilt* $\varphi * \mathbf{1} = \mathrm{id}$.

Beweis. Die Behauptung besagt für alle $n \in \mathbb{N}$:

$$\sum_{d|n} \varphi(d) = n.$$

Unter den n Brüchen $\frac{a}{n}$ mit $1 \leq a \leq n$ treten in gekürzter Form nur Nenner d mit $d \mid n$ auf, und von diesen genau $\varphi(d)$-viele (bei denen man nicht weiter kürzen kann, also der Zähler teilerfremd zu d ist!). Da jeder Bruch eine eindeutige gekürzte Form hat, folgt durch Abzählen die Behauptung. □

Korollar 7.23. *Die Eulersche φ-Funktion ist multiplikativ und es gilt*

$$\varphi(n) = \sum_{d|n} d \cdot \mu(\frac{n}{d}).$$

Beweis. Die Formel folgt sofort aus Korollar 7.20 für $f(n) = \varphi(n)$ mit der Summenformel aus dem Beweis von Proposition 7.22, also mit $F(n) = n$. Oder wir rechnen nochmlas direkt: wegen

$$\varphi = \varphi * \delta = \varphi * (\mathbf{1} * \mu) = (\varphi * \mathbf{1}) * \mu = \mathrm{id} * \mu$$

ist φ als Faltung zweier multiplikativer Funktionen selbst multiplikativ nach Proposition 7.5. Die angegebene Formel drückt nur die Faltungsgleichung aus. □

Beispiel 7.24. Zum Schluß des Kapitels berechnen wir die zu φ gehörigen Potenzreihen. Nach Korollar 7.23 gilt für eine Primzahlpotenz p^k

$$\varphi(p^k) = \sum_{d|p^k} d \cdot \mu(p^k/d) = p^k - p^{k-1} = p^k(1 - \frac{1}{p})$$

und damit

$$\begin{aligned}\lambda_p(\varphi, T) &= \sum_{k\geq 0} \varphi(p^k)T^k = 1 + \sum_{k\geq 1} \varphi(p^k)T^k \\ &= 1 + (1 - \frac{1}{p}) \cdot \sum_{k\geq 1} p^k T^k = 1 + (1 - \frac{1}{p})\frac{pT}{1-pT} = 1 + \frac{(p-1)T}{1-pT} = \frac{1-T}{1-pT}.\end{aligned}$$

Diese Potenzreihe belegt erneut die Gleichung $\varphi * \mathbf{1} = \mathrm{id}$ wegen

$$\frac{1-T}{1-pT} \cdot \frac{1}{1-T} = \frac{1}{1-pT}.$$

Aus der Berechnung von $\varphi(p^k)$ folgt weiter für alle $n \in \mathbb{N}$ die Formel

$$\varphi(n) = n \cdot \prod_{p|n} (1 - \frac{1}{p}).$$

Übungsaufgaben zu §7

Übungsaufgabe 7.1. Sei $f : \mathbb{N} \to \mathbb{C}$ eine multiplikative Funktion. Zeigen Sie, daß f bezüglich Faltung invertierbar ist genau dann, wenn $f \neq 0$ nicht die Nullfunktion ist. Dabei hießt f invertierbar bezüglich Faltung, wenn ein $g : \mathbb{N} \to \mathbb{C}$ existiert, so daß $f * g = \delta$. Ist das Inverse von f multiplikativ?

8. Kettenbrüche

Ein Kettenbruch ist ein Ausdruck der Form

$$a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \dots}}},$$

nachdem wir sinnvoll die ... erklärt haben.

8.1. **Der Wert eines Kettenbruchs.** Vergleichbar zum Wert einer unendlichen Reihe wird der Wert eines unendlichen Kettenbruchs als Limes der Näherungsbrüche, die durch Abschneiden des Kettenbruchs entstehen, definiert. wir benötigen zur Begriffsbildung zunächst ein paar Eigenschaften endlicher Kettenbrüche.

Definition 8.1. Der **endliche Kettenbruch** $[x_0, x_1, \ldots, x_n]$ mit **Teilnennern** $x_i \in \mathbb{R}$ zu $i = 0, \ldots, n$ mit $x_i > 0$ für alle $0 < i \leq n$ ist rekursiv (über n) definiert als die reelle Zahl

$$[x_0] := x_0,$$

$$[x_0, x_1, \ldots, x_{n-1}, x_n] := [x_0, x_1, \ldots, x_{n-1} + \frac{1}{x_n}].$$

Bemerkung 8.2. Endliche Kettenbrüche sind wohldefiniert, denn mit $x_{n-1} \geq 0$ und $x_n > 0$ ist der neue $(n-1)$-te Teilnenner $x_{n-1} + \frac{1}{x_n} > 0$. Es gilt dann

$$[x_0, x_1, \ldots, x_n] = x_0 + \frac{1}{x_1 + \frac{1}{x_2 + \ldots \frac{1}{x_n}}},$$

wie man sofort per Induktion nach n beweist.

Um den Limesprozess bei unendlichen Kettenbrüchen zu verstehen, benötigen wir ein paar Lemmata.

Lemma 8.3. *Es seien $x_0, \ldots, x_{n+m} \in \mathbb{R}$ mit $x_i > 0$ für alle $0 < i \leq n+m$. Dann gilt*

$$[x_0, x_1, \ldots, x_{n-1}, x_n, x_{n+1} \ldots, x_{n+m}] = [x_0, x_1, \ldots, x_{n-1}, [x_n, x_{n+1}, \ldots, x_{n+m}]].$$

Beweis. Dies folgt sofort per Induktion nach m. Für $m = 0$ ist nichts zu tun, und für den Schritt von $m-1$ auf m gelten die gleichen rekursiven Formeln:

$$\begin{aligned}
[x_0, x_1, \ldots, x_{n-1}, x_n, x_{n+1} \ldots, x_{n+m}] &= [x_0, x_1, \ldots, x_{n-1}, x_n, x_{n+1} \ldots, x_{n+m-1} + \frac{1}{x_{n+m}}] \\
&= [x_0, x_1, \ldots, x_{n-1}, [x_n, x_{n+1}, \ldots, x_{n+m-1} + \frac{1}{x_{n+m}}]] \\
&= [x_0, x_1, \ldots, x_{n-1}, [x_n, x_{n+1}, \ldots, x_{n+m}]].
\end{aligned}$$

□

Bemerkung 8.4. Wir erinnern an die Wirkung durch Möbiustransformationen auf $\mathbb{P}^1(\mathbb{R})$. Für $\begin{pmatrix} p & r \\ q & s \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R})$ ist

$$\begin{pmatrix} p & r \\ q & s \end{pmatrix} x := \begin{cases} \frac{px+r}{qx+s} & \text{sofern } qx + s \neq 0, \\ \infty & qx + s = 0, \\ \frac{p}{q} & x = \infty. \end{cases}$$

Für Matrizen $M_1, M_2 \in \mathrm{GL}_2(\mathbb{R})$ und $x \in \mathbb{P}^1(\mathbb{R}) = \mathbb{R} \cup \{\infty\}$ gilt dann

$$M_1(M_2 x) = (M_1 M_2)x,$$

wie man sofort zur Übung selbst nachrechnet.

Das folgende Lemma hat Ähnlichkeiten mit Matrizenrechnungen aus dem Kontext des euklidischen Algorithmus, so wie dieser in Bemerkung 3.11 formuliert wird.

Lemma 8.5. *Es seien $a_0, \ldots, a_n, x \in \mathbb{R}$ mit $x > 0$ und $a_i > 0$ für alle $0 < i \leq n$. Sei weiter*

$$\begin{pmatrix} p & r \\ q & s \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R}).$$

Dann gilt

(1) $\quad [a_0, a_1, \ldots, a_n, x] = \begin{pmatrix} p & r \\ q & s \end{pmatrix} x.$

(2) $[a_0, a_1, \ldots, a_n] = \frac{p}{q}$.

Beweis. Man zeigt durch direktes Nachrechnen

$$[a, x] = a + \frac{1}{x} = \frac{ax+1}{x} = \begin{pmatrix} a & 1 \\ 1 & 0 \end{pmatrix} x.$$

Die Behauptung (1) folgt nun per Induktion nach n mit Lemma 8.3 aus

$$\begin{aligned}[a_0, a_1, \ldots, a_n, x] &= [a_0, [a_1, \ldots, a_n, x]] = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} [a_1, \ldots, a_n, x] \\ &= \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix} x = \begin{pmatrix} p & r \\ q & s \end{pmatrix} x.\end{aligned}$$

Behauptung (2) folgt aus (1) im Limes $x \to \infty$. □

Satz 8.6. *Sei $(x_i)_{i \in \mathbb{N}_0}$ eine Folge reeller Zahlen mit $x_i \geq 1$ für $i > 0$. Dann existiert der Grenzwert der endlichen Kettenbrüche*

$$[x_0, x_1, \ldots] := \lim_{n \to \infty} [x_0, x_1, \ldots, x_n].$$

Beweis. Wir setzen $\vartheta_n = [x_0, x_1, \ldots, x_n]$ und beschreiben dies vermöge der Matrizen

$$M_n = \begin{pmatrix} p_n & r_n \\ q_n & s_n \end{pmatrix} = \begin{pmatrix} x_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} x_n & 1 \\ 1 & 0 \end{pmatrix}$$

mittels Lemma 8.5 als

$$\vartheta_n = \frac{p_n}{q_n}.$$

Aus der Gleichung

$$\begin{pmatrix} p_n & r_n \\ q_n & s_n \end{pmatrix} = \begin{pmatrix} p_{n-1} & r_{n-1} \\ q_{n-1} & s_{n-1} \end{pmatrix} \begin{pmatrix} x_n & 1 \\ 1 & 0 \end{pmatrix}$$

folgt für $n \geq 1$ die Rekursion

$$\begin{aligned} p_n &= x_n p_{n-1} + r_{n-1}, \\ r_n &= p_{n-1}, \\ q_n &= x_n q_{n-1} + s_{n-1}, \\ s_n &= q_{n-1}, \end{aligned}$$

bei Startwerten $p_0 = x_0$, $r_0 = 1$, $q_0 = 1$, $s_0 = 0$. Da wir nur an $p_n = \vartheta_n q_n$ interessiert sind, machen wir daraus die zweischrittige Rekursion für $n \geq 1$

$$\begin{aligned} p_n &= x_n p_{n-1} + p_{n-2}, \\ q_n &= x_n q_{n-1} + q_{n-2}, \end{aligned}$$

bei Startwerten $p_0 = x_0$, $p_{-1} = 1$, $q_0 = 1$ und $q_{-1} = 0$. Dies entspricht $M_{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Wir zeigen nun per Induktion nach $n \geq 1$ die Abschätzung

$$q_n > 0.$$

Als Induktionsanfang dient $q_1 = x_1 q_0 + q_{-1} = x_1 > 0$. Da $x_i \geq 1$ für alle $i \geq 1$ folgt der Induktionsschritt für $n \geq 2$ aus

$$q_n = x_n q_{n-1} + q_{n-2} \geq q_{n-1} + q_{n-2} \geq 1 + q_{n-2} > 0,$$

denn es gilt auch $q_0 = 1 > 0$ (Fall $n = 2$). Weiter folgt nun für $n \geq 1$

$$q_n = x_n q_{n-1} + q_{n-2} \geq q_{n-1} + q_{n-2} \geq q_{n-1}$$

und schließlich für alle $n \geq 2$:

$$0 < \frac{q_{n-2}}{x_n q_{n-1} + q_{n-2}} \leq \frac{q_{n-2}}{q_{n-1} + q_{n-2}} \leq \frac{q_{n-2}}{q_{n-2} + q_{n-2}} = \frac{1}{2}.$$

Damit können wir die Werte der endlichen Kettenbrüche kontrollieren als ($n \geq 1$)

$$\begin{aligned}\vartheta_n &= \frac{p_n}{q_n} = \frac{x_n p_{n-1} + p_{n-2}}{x_n q_{n-1} + q_{n-2}} \\ &= \frac{x_n \vartheta_{n-1} q_{n-1} + \vartheta_{n-2} q_{n-2}}{x_n q_{n-1} + q_{n-2}} = \vartheta_{n-1} + (\vartheta_{n-2} - \vartheta_{n-1}) \frac{q_{n-2}}{x_n q_{n-1} + q_{n-2}},\end{aligned}$$

woraus für alle $n \geq 2$ folgt

$$\vartheta_n - \vartheta_{n-1} = (\vartheta_{n-2} - \vartheta_{n-1}) \frac{q_{n-2}}{x_n q_{n-1} + q_{n-2}}.$$

Die Differenz $\vartheta_n - \vartheta_{n-1}$ hat also alternierendes Vorzeichen und ist wegen

$$|\vartheta_n - \vartheta_{n-1}| = |\vartheta_{n-2} - \vartheta_{n-1}| \frac{q_{n-2}}{x_n q_{n-1} + q_{n-2}} \leq \frac{1}{2} |\vartheta_{n-2} - \vartheta_{n-1}| \leq \frac{1}{2^{n-1}} \cdot |\vartheta_1 - \vartheta_0|$$

eine Nullfolge. Somit konvergiert die Folge (ϑ_n). □

Lemma 8.7. *Sei $(x_i)_{i \in \mathbb{N}_0}$ eine Folge reeller Zahlen mit $x_i \geq 1$ für $i > 0$. Dann gilt für alle $n \geq 0$*

$$[x_0, x_1, \ldots, x_n, x_{n+1} \ldots] = [x_0, x_1, \ldots, x_n, [x_{n+1}, \ldots]].$$

Beweis. Die Funktion in $x \in \mathbb{R}_{>0}$

$$f(x) = [x_0, x_1, \ldots, x_n, x]$$

ist nach Lemma 8.5 eine gebrochenrationale Funktion und daher stetig bis auf höchstens einen Pol. Aus dem Beweis von Satz 8.6 und in der Notation desselben folgt, daß der Pol bei der Nullstelle von $q_n x + q_{n-1}$, also bei einer negativen reellen Zahl liegt ($q_n > 0$ für alle n). Damit ist f stetig und vertauscht mit existierenden Grenzwerten:

$$\begin{aligned}[x_0, x_1, \ldots, x_n, [x_{n+1}, \ldots]] &= f([x_{n+1}, \ldots]) = f\big(\lim_{m \to \infty} [x_{n+1}, \ldots, x_{n+m}]\big) \\ &= \lim_{m \to \infty} f([x_{n+1}, \ldots, x_{n+m}]) \\ &= \lim_{m \to \infty} [x_0, x_1, \ldots, x_n, [x_{n+1}, \ldots, x_{n+m}]] \qquad \text{(Lemma 8.3)} \\ &= \lim_{m \to \infty} [x_0, x_1, \ldots, x_n, x_{n+1}, \ldots, x_{n+m}] = [x_0, x_1, \ldots, x_n, x_{n+1}, \ldots].\end{aligned}$$ □

8.2. **Der Kettenbruchalgorithmus.** Wir gehen nun der Frage nach, wie man eine reelle Zahl als Kettenbruch darstellen kann. Dabei fordern wir nun Ganzzahligkeit der Teilnenner.

Definition 8.8. Ein **unendlicher Kettenbruch** besteht aus einer Folge $(a_n)_{n \geq 0}$ ganzer Zahlen mit $a_i \geq 1$ für alle $i > 0$ und beschreibt die reelle Zahl

$$[a_0, a_1, \ldots, a_n, \ldots] := \lim_{n \to \infty} [a_0, a_1, \ldots, a_n].$$

Der endliche Kettenbruch aus den Anfangsgliedern $[a_0, a_1, \ldots, a_n]$ heißt n-ter Näherungsbruch.

Wir definieren für das Folgende die Funktion $T : \mathbb{R} \setminus \mathbb{Z} \to \mathbb{R}$ durch

$$T(x) = \frac{1}{x - \lfloor x \rfloor}.$$

Lemma 8.9. *Es gilt für alle $x \in \mathbb{R} \setminus \mathbb{Z}$:*

(1) $T(x)$ *ist wohldefiniert.*
(2) $T(x) > 1$.
(3) *Mit* $a = \lfloor x \rfloor$ *haben wir* $x = [a, T(x)]$.
(4) $x \in \mathbb{Q} \iff T(x) \in \mathbb{Q}$.

Beweis. (1) Weil $x = \lfloor x \rfloor \iff x \in \mathbb{Z}$ ist $T(x) \in \mathbb{R}$.

(2) Das folgt aus $0 < x - \lfloor x \rfloor < 1$ für $x \in \mathbb{R} \setminus \mathbb{Z}$.

(3) Das ist eine einfache Rechnung

$$x = a + (x - \lfloor x \rfloor) = a + \frac{1}{T(x)} = [a, T(x)]. \qquad \square$$

(4) Das folgt sofort aus $x - T(x)^{-1} \in \mathbb{Z}$.

Bemerkung 8.10. Im Versuch, die reelle Zahl x in einen Kettenbruch zu entwickeln, stellt

$$x = [a, T(x)]$$

mit $a = \lfloor x \rfloor$ den Ansatz nach dem „greedy–Prinzip“ dar. Man ist so gierig, wie es geht, und hofft, daß alles gut geht. Man spaltet die größte ganze Zahl, die möglich ist, ab und verfährt mit dem Rest rekursiv genauso. Das ist die Idee, aber keine korrekte rekursive Definition, weil $T(x)$ ja noch keine Kettenbruchentwicklung zugewiesen bekommen hat. Die korrekte Rekursion kommt jetzt.

Beispiel 8.11. Wir probieren den Algorithmus nach dem „greedy–Prinzip“ aus. Sei $x = \frac{97}{67}$. Dann gilt

$$x = \frac{97}{67} = 1 + \frac{1}{\frac{67}{30}} = 1 + \frac{1}{2 + \frac{1}{\frac{30}{7}}} = 1 + \frac{1}{2 + \frac{1}{4 + \frac{1}{\frac{7}{2}}}} = 1 + \frac{1}{2 + \frac{1}{4 + \frac{1}{3 + \frac{1}{2}}}} = [1, 2, 4, 3, 2].$$

Satz 8.12 (Kettenbruchalgorithmus)**.** *Sei $x \in \mathbb{R}$. Wir definieren rekursiv $x_n \in \mathbb{R}$ durch $x_0 = x$ und*

$$x_{n+1} = \begin{cases} T(x_n) & \text{falls } x_n \notin \mathbb{Z} \\ \text{nicht definiert} & \text{falls } x_n \in \mathbb{Z}. \end{cases}$$

Sodann setzen wir $a_n = \lfloor x_n \rfloor \in \mathbb{Z}$.

(1) Die Folge (a_n) hat endlich viele Folgenglieder genau dann, wenn $x \in \mathbb{Q}$.

(2) $a_n \geq 1$ für alle $n \geq 1$ (sofern definiert).

(3) Für alle $n \geq 0$ gilt

$$x = [a_0, a_1, \ldots, a_{n-1}, x_n].$$

(4) Sei $x \in \mathbb{Q}$ und $x_n = a_n \in \mathbb{Z}$ das letzte definierte Folgenglied. Dann gilt:

$$x = [a_0, a_1, \ldots, a_n].$$

(5) Sei $x \in \mathbb{R} \setminus \mathbb{Q}$. Dann kovergiert der Kettenbruch $[a_0, a_1, \ldots]$, und es gilt:

$$x = [a_0, a_1, \ldots].$$

Beweis. (1) Wenn (a_n) nur endlich viele Folgenglieder hat, dann gibt es ein $n \geq 0$ mit $x_n \in \mathbb{Z}$, also $x_n \in \mathbb{Q}$. Nach Lemma 8.9 (4) schließen wir fallend induktiv auf $x_m \in \mathbb{Q}$ für alle $m \leq n$. Dann ist auch $x = x_0 \in \mathbb{Q}$.

Sei umgekehrt $x_0 = \frac{p}{q} \in \mathbb{Q}$ mit $p, q \in \mathbb{Z}$ teilerfremd und $q > 0$. Sei

$$p = qd + r$$

die Division mit Rest: $d \in \mathbb{Z}$ und $0 \leq r < q$. Dann ist $d = \lfloor x \rfloor$ und

$$x_1 = T(\frac{p}{q}) = \frac{1}{\frac{p}{q} - d} = \frac{q}{r}.$$

Da $1 = (p, q) = (q, r)$ ist dies wieder ein gekürzter Bruch. Wir sehen, daß $T(x)$ mit Zähler und Nenner die Zahlen des Euklidischen Algorithmus zu (p, q) reproduziert. Insbesondere wird irgendwann der $\mathrm{ggT} = 1$ erreicht und dann ist $x_n \in \mathbb{Z}$. Die Folge bricht dann ab.

(2) Das folgt aus $T(x) > 1$, Lemma 8.9 (2).

(3) Es gilt $x_n = [a_n, x_{n+1}]$ nach Lemma 8.9 (3), so daß (3) aus Lemma 8.3 per Induktion folgt:

$$x = [a_0, a_1, \ldots, a_{n-1}, x_n] = [a_0, a_1, \ldots, a_{n-1}, [a_n, x_{n+1}]] = [a_0, a_1, \ldots, a_{n-1}, a_n, x_{n+1}].$$

(4) Folgt offensichtlich aus (3).

(5) Der Kettenbruch konvergiert, weil die Voraussetzungen von Satz 8.6 nach (2) erfüllt sind. Wir berechnen nach (3) mit der Notation von Satz 8.6

$$\begin{aligned} x - [a_0, a_1, \ldots, a_n] &= [a_0, a_1, \ldots, a_n, x_{n+1}] - [a_0, a_1, \ldots, a_n] \\ &= \frac{p_n x_{n+1} + r_n}{q_n x_{n+1} + s_n} - \frac{p_n}{q_n} \\ &= \frac{-\det(M_n)}{q_n \cdot (q_n x_{n+1} + q_{n-1})} = \frac{(-1)^n}{q_n \cdot (q_n x_{n+1} + q_{n-1})}. \end{aligned} \tag{8.1}$$

Wir zeigen nun per Induktion nach $n \geq 1$ die Abschätzung

$$q_n \geq n.$$

Als Induktionsanfang dient $q_1 = x_1 q_0 + q_{-1} = x_1 \geq 1$. Da $x_i \geq 1$ für alle $i \geq 1$, folgt der Induktionsschritt für $n \geq 2$ aus

$$q_n = x_n q_{n-1} + q_{n-2} \geq q_{n-1} + q_{n-2} \geq n - 1 + q_{n-2} \geq n,$$

denn $q_0 = 1$ (Fall $n = 2$) und per Induktionsvoraussetzung $q_{n-2} \geq n - 2 \geq 1$, sobald $n \geq 3$ ist.

Damit gilt

$$|x - [a_0, a_1, \ldots, a_n]| = \frac{1}{q_n \cdot (q_n x_{n+1} + q_{n-1})} \leq \frac{1}{n(2n-1)}$$

und so die behauptete Konvergenz

$$x = [a_0, a_1, \ldots]. \qquad \square$$

Der Kettenbruchalgorithmus funktioniert also tatsächlich nach dem „greedy–Prinzip“. Jetzt wollen wir wissen, ob wir überhaupt eine Wahl haben, also wie eindeutig die Darstellung einer reellen Zahl als Kettenbruch ist.

Satz 8.13. *Die Kettenbruchentwicklung ist*

(1) eindeutig, wenn $x \in \mathbb{R} \setminus \mathbb{Q}$,

(2) mehrdeutig in der folgenden Form, wenn $x \in \mathbb{Q}$. Sei $x = [a_0, a_1, \ldots, a_n]$ die Darstellung aus dem Kettenbruchalgorithmus. Dann ist $a_n \geq 2$, sofern $n > 0$ und

$$[a_0, a_1, \ldots, a_n] = [a_0, a_1, \ldots, a_n - 1, 1]$$

ist die einzige andere Kettenbruchdarstellung von x.

Beweis. (1) Sei $x \in \mathbb{R} \setminus \mathbb{Q}$ und $x = [a_0, a_1, \ldots]$ die Darstellung aus dem Kettenbruchalgorithmus. Angenommen es gibt einen weiteren Kettenbruch, der den Wert x hat. Dann muß dieser Kettenbruch ein unendlicher Kettenbruch $x = [b_0, b_1, \ldots]$ sein, weil sonst $x \in \mathbb{Q}$.

Für jedes $a \in \mathbb{Z}$ ist $[a, x] = a + \frac{1}{x}$ als Funktion von $x > 0$ streng monoton fallend. Sind alle Teilnenner > 0, so ist jeder Näherungsbruch > 0 und auch der Limes (alternierend!) > 0. Daraus folgt

$$[a_0, a_1, \ldots] = [a_0, [a_1, \ldots]] > \lim_{x \to \infty} [a_0, x] = a_0$$

und

$$[a_0, a_1, \ldots] = [a_0, [a_1, \ldots]] < [a_0, [a_1]] = a_0 + \frac{1}{a_1} \leq a_0 + 1.$$

Folglich ist bei $[a_0, a_1, \ldots] = x = [b_0, b_1, \ldots]$

$$a_0 = \lfloor x \rfloor = b_0.$$

Aus $[a, x] = [a, y]$ folgt $x = y$, so daß wir weiter auf $[a_1, \ldots] = T(x) = [b_1, \ldots]$ schließen können. Wenn die Kettenbruchentwicklungen verschieden sind, so gibt es ein kleinstes n mit $a_n \neq b_n$. Dies widerspricht der iteriert ausgeführten obigen Abschätzung (per Induktion nach n zeigen wir $a_n = b_n$).

(2) Weil $[a,1] = [a+1]$ folgt die Gleichheit der angegebenen endlichen Kettenbrüche aus Lemma 8.3. Solange der Kettenbruch ≥ 2 Teilnenner hat, argumentiert man wie in (1). Es bleibt also der Fall $x = a \in \mathbb{Z}$. Dann ist die Darstellung aus dem Kettenbruchalgorithmus $x = [a]$. Sei $[b_0, \ldots, b_m]$ oder $[b_0, \ldots]$ eine weitere Darstellung von a. Wenn $m = 0$, dann ist $a = b_0$ und wir haben die gleiche Darstellung. Wenn $m \geq 1$ oder der Kettenbruch unendlich ist, dann ist mit den Abschätzungen aus (1)

$$b_0 < [b_0, \ldots, b_m] = a \leq b_0 + \frac{1}{b_1}$$

und das geht nur für $a = b_0$ (und wieder $m = 0$) oder eben $b_0 = a - 1$ und $b_1 = 1$. Aber dann ist bereits $a = [b_0, b_1]$ und somit muß $m = 1$ sein. Der Teilnenner $a_n - 1$ ist nur erlaubt, wenn $a_n \geq 2$ ist, sofern $n > 0$. □

Korollar 8.14. *Ein endlicher Kettenbruch ist eine Zahl in $\mathbb{Q}$, ein unendlicher Kettenbruch konvergiert gegen eine Zahl in $\mathbb{R} \setminus \mathbb{Q}$.*

Beweis. Die erste Teilaussage ist trivial, und die zweite Teilaussage folgt aus der Eindeutigkeit und Satz 8.12 (1). □

Beispiel 8.15. Wir bestimmen ein paar Kettenbrüche mit dem Kettenbruchalgorithmus.

(1) Der Goldene Schnitt φ gilt als besonders ästhetisches Teilungsverhältnis einer Strecke. Und zwar habe das große Teilstück zum kleinen Teilstück dasselbe Verhältnis wie das ganze zum größeren Teilstück. Sei das kleine Teilstück der Länge normiert auf 1. Dann hat das größere die Länge φ und das Ganze die Länge $1 + \varphi$. Gefordert ist nun

$$\varphi = \frac{1+\varphi}{\varphi}$$

oder $\varphi^2 = \varphi + 1$. Aus der Nebenbedingung $\varphi > 1$ folgt sofort

$$\varphi = \frac{1+\sqrt{5}}{2}.$$

Dann ist $a_0 = \lfloor \varphi \rfloor = 1$ und

$$T(\varphi) = \frac{1}{\varphi - 1} = \frac{\varphi}{\varphi^2 - \varphi} = \varphi.$$

Wir finden somit

$$\varphi = [1, \varphi]$$

und finden die Kettenbruchentwicklung durch iteriertes in sich selbst Einsetzen als

$$\varphi = [1, 1, 1, 1, 1, 1, 1, \ldots].$$

Der Goldene Schnitt hat demnach die Kettenbruchentwicklung mit den kleinsten möglichen Teilnennern. Damit konvergiert die Folge der Näherungsbrüche für den Goldenen Schnitt am langsamsten unter allen reellen Zahlen.

(2) Es gilt

$$e = [2, 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, 1, 1, \ldots]$$

und es gilt genau das, wonach es aussieht! Die Kettenbruchentwicklung der Eulerschen Zahl e ist also sehr regelmäßig!

(3) Dasselbe kann man von der Kreiszahl π nicht behaupten:

$$\pi = [3, 7, 15, 1, 292, 1, 1, 1, 2, 1, 3, 1, 14, 2, 1, 1, \ldots].$$

Hier sind die Näherungsbrüche interessant:

$$3, \frac{22}{7}, \frac{333}{106}, \frac{355}{113} \approx 3,14159292\ldots.$$

Das wird also schnell sehr genau und beinhaltet die bereits aus dem Altertum bekannten Approximationen! Das ist kein Zufall.

8.3. **Diophantische Approximation mit Kettenbrüchen.** Die Kettenbruchentwicklung einer reellen Zahl x führt mit den Näherungsbrüchen zu einer ausgezeichneten Folge rationaler Zahlen $\frac{p_n}{q_n}$ mit

$$x = \lim_{n\to\infty} \frac{p_n}{q_n}.$$

Die Näherungsbrüche optimieren die Approximation von x mit möglichst kleinen Nennern.

Beispiel 8.16. Das tropische Jahr, das die Position der Jahreszeiten im Kalender konstant hält und daher in der originären Absicht einer Kalenderkonstruktion möglichst gut approximiert werden soll, hat laut Wikipedia

$$a = 365,24219052$$

Tage. Die Schaltjahrregelung des Gregorianischen Kalenders führt zu einer durchschnittlichen Jahreslänge von

$$365 + \frac{1}{4} - \frac{1}{100} + \frac{1}{400} = 365\frac{97}{400} = 365,2425.$$

Damit geht der Gregorianische Kalender in

$$\frac{1}{365,2425 - 365,24219052} \approx 3231$$

Jahren um einen Tag falsch.

Vergleichen wir dies mit der Kettenbruchentwicklung

$$a = [365, 4, 7, 1, 3, 19, 1, 20, 18, 1, 3, \ldots]$$

und den Näherungsbrüchen

$$\begin{aligned}
[365] &= 365\\
[365,4] &= 365\frac{1}{4} = 365,25\\
[365,4,7] &= 365\frac{7}{29} \approx 365,2413\ldots\\
[365,4,7,1] &= 365\frac{8}{33} \approx 365,2424\ldots\\
[365,4,7,1,3] &= 365\frac{31}{128} \approx 365,2421875.
\end{aligned}$$

Schlagen wir daher als alternative Schalttagsregelung

$$365\frac{31}{128} = 365 + \frac{1}{4} - \frac{1}{128} \approx 365,2421875$$

vor, also alle 4 Jahre ein Schalttag, aber alle 128 Jahre doch kein Schalttag. Diese Schalttagsregelung geht in

$$\frac{1}{365,24219052 - 365,2421875} \approx 333125$$

Jahren um einen Tag falsch. Die Periodenlänge der Regelung ist mit 128 kürzer als die offiziellen 400, aber trotzdem dauert es 100-Mal länger, bis der Kalender um einen Tag falsch läuft.

Bemerkung 8.17. Natürlich kann man jede reelle Zahl beliebig genau durch rationale Zahlen approximieren: $\mathbb{Q}$ liegt dicht in $\mathbb{R}$. Wir betrachten eine Approximation von $x \in \mathbb{R}$ durch eine rational Zahl $\frac{p}{q}$ als besonders gut, wenn bei möglichst kleinem Nenner q ein möglichst kleiner Fehler $|x - \frac{p}{q}|$ erreicht wird.

Definition 8.18. Eine reelle Zahl $x \in \mathbb{R}$ heißt durch rationale Zahlen **approximierbar von Ordnung** r, wenn es eine Konstante $c = c(x,k)$ gibt, so daß unendlich viele (gekürzte) $\frac{p}{q} \in \mathbb{Q}$ existieren mit

$$|x - \frac{p}{q}| < \frac{c}{q^r}.$$

Bemerkung 8.19. (1) Sei $r > 0$. Wenn man auf die Bedingung „gekürzt“ verzichtet, dann macht man sich das Leben nur schwerer: dieselbe rationale Zahl hat dann ein größeres q und muß dann x genauer approximieren (der erlaubte Fehler $\frac{c}{q^r}$ wird kleiner).

Ein Sonderfall ist $x = \frac{p}{q} \in \mathbb{Q}$, bei dem jedes $\frac{pd}{qd}$ eine gute Approximation darstellt. Diese wollen wir nur als eine Approximation ansehen, also eben doch auf gekürzten Brüchen bestehen (aber eben nur zu diesem Zweck).

(2) Wenn $0 < r < s$, dann ist
$$\frac{c}{q^s} < \frac{c}{q^r}$$
und $x \in \mathbb{R}$ ist durch rationale Zahlen approximierbar von Ordnung r, wenn x sogar von Ordnung s approximierbar ist. Je größer der Exponent, desto stärker die Aussage.

Proposition 8.20. *Jedes $x \in \mathbb{R}$ ist von Ordnung 1 approximierbar durch rationale Zahlen.*

Beweis. Zu $q \in \mathbb{N}$ betrachten wir dasjenige p mit $p \leq qx < p+1$. Dann ist x im Intervall $\{y \ ; \ \frac{p}{q} \leq y < \frac{p+1}{q}\}$ der Länge $\frac{1}{q}$ und
$$|x - \frac{p}{q}| < \frac{1}{q}. \qquad \square$$

Proposition 8.21. *Sei $\varepsilon > 0$ und sei $x \in \mathbb{R}$ von Ordnung $1+\varepsilon$ approximierbar durch rationale Zahlen. Dann ist $x \notin \mathbb{Q}$.*

Beweis. Sei x rational, $x = \frac{a}{b}$ als gekürzter Bruch und von Ordnung $1+\varepsilon$ approximierbar. Dann gibt es $c > 0$ und unendlich viele gekürzte $\frac{p}{q}$ mit
$$|x - \frac{p}{q}| < \frac{c}{q^{1+\varepsilon}}.$$
Wir dürfen also annehmen, daß $\frac{a}{b} \neq \frac{p}{q}$ gilt. Demnach ist $aq - bp \in \mathbb{Z}$ von 0 verschieden, also
$$|aq - bp| \geq 1.$$
Daraus folgt nun
$$\frac{c}{q^{1+\varepsilon}} > |x - \frac{p}{q}| = |\frac{aq-bp}{bq}| \geq \frac{1}{bq}$$
oder äquivalent
$$bc > q^{\varepsilon}.$$
Da b und c fest sind, gibt es nur endlich viele q, die in Frage kommen, ein Widerspruch. $\square$

Bemerkung 8.22. Der Beweis von Proposition 8.21 enthält die Hauptidee der Diophantischen Approximation: zwischen 0 und 1 liegt keine weitere ganze Zahl.

Proposition 8.23 (Dirichlet)**.** *Jedes $x \in \mathbb{R}\backslash\mathbb{Q}$ ist von Ordnung 2 approximierbar durch rationale Zahlen.*

Beweis. Sei $Q \in \mathbb{N}$ gegeben. Wir betrachten von den Q-vielen Zahlen
$$x, 2x, \ldots, qx, \ldots, Qx$$
die gebrochenen Anteile $\langle y \rangle := y - \lfloor y \rfloor$ im (halboffenen) Einheitsintervall $[0,1)$, oder äquivalent in $\mathbb{R}/\mathbb{Z}$. Dazu unterteilen wir das Intervall in die (halboffenen) Intervalle
$$\left[\frac{i}{Q}, \frac{i+1}{Q}\right)$$
für $i = 0, \ldots, Q-1$. Wenn einer der $\langle qx \rangle$ im Intervall $[0, \frac{1}{Q})$ liegt, dann gilt hier für $p = \lfloor qx \rfloor \in \mathbb{Z}$
$$|x - \frac{p}{q}| = |\frac{qx-p}{q}| = |\frac{\langle qx \rangle}{q}| \leq \frac{1/Q}{q} = \frac{1}{qQ} \leq \frac{1}{q^2}.$$

Wenn das Intervall $[0, \frac{1}{Q})$ keinen der Werte $\langle qx \rangle$ abbekommt, dann verteilen sich die Q-vielen Werte auf $Q-1$-viele restliche Intervalle (Schubfächer), so daß nach dem Dirichlet'schen Schubfachprinzip ein Intervall existiert, in dem mindestens 2 Werte auftreten. Sei daher $1 \leq a < b \leq Q$ und $1 \leq i \leq Q-1$ mit

$$\langle ax \rangle, \langle bx \rangle \in \left[\frac{i}{Q}, \frac{i+1}{Q}\right).$$

Dann ist mit $q = b - a$, erstens $1 \leq q < Q$ und zweitens $qx = \langle bx \rangle - \langle ax \rangle + p$ für ein $p \in \mathbb{Z}$. Daraus folgt

$$|x - \frac{p}{q}| = |\frac{qx-p}{q}| = |\frac{\langle bx \rangle - \langle ax \rangle}{q}| \leq \frac{1/Q}{q} = \frac{1}{qQ} \leq \frac{1}{q^2}.$$

Insgesamt haben wir nun bewiesen, daß es zu jedem Q ein $1 \leq q \leq Q$ und $p \in \mathbb{Z}$ gibt, mit

$$|x - \frac{p}{q}| \leq \frac{1}{qQ}.$$

Sei $\frac{p}{q}$ fixiert. Für $Q \to \infty$ geht die rechte Seite nach 0, während die linke Seite konstant und > 0 ist, weil $x \notin \mathbb{Q}$. Daher kommt jedes feste $\frac{p}{q}$ nur für endlich viele Q in Frage. Es muß daher unendlich viele verschiedene $\frac{p}{q}$ geben mit

$$|x - \frac{p}{q}| \leq \frac{1}{q^2}. \qquad \square$$

Eine algebraische Zahl in $\mathbb{C}$ ist die Nullstelle eines nichtkonstanten Polynoms mit Koeffizienten aus $\mathbb{Q}$. Das folgende Ergebnis können wir im Rahmen dieser Vorlesung nicht beweisen. Thue[22] und Siegel[23] haben Vorarbeiten mit größeren Werten als $2+\varepsilon$ geleistet, aber bereits die wesentliche Strategie zum Beweis benutzt. Roths[24] Beitrag wurde 1958 mit der Fields-Medaille gewürdigt.

Theorem 8.24 (Thue–Siegel–Roth)**.** *Sei $x \in \mathbb{R}$ eine algebraische Zahl und $\varepsilon > 0$. Dann ist x nicht von Ordnung $2+\varepsilon$ approximierbar durch rationale Zahlen. Für jedes $c > 0$ hat die Ungleichung*

$$|x - \frac{p}{q}| \leq \frac{c}{q^{2+\varepsilon}}$$

nur endlich viele gekürzte Lösungen $\frac{p}{q} \in \mathbb{Q}$.

Die besten Approximationen entstammen der Kettenbruchentwicklung:

Satz 8.25. *Sei $x \in \mathbb{R} \setminus \mathbb{Q}$.*

(1) Sei $x = [a_0, a_1, \ldots]$ die Kettenbruchentwicklung von x, und sei

$$\begin{pmatrix} p_n & r_n \\ q_n & s_n \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix} \in \mathrm{GL}_2(\mathbb{Z}).$$

Dann ist für alle $n \geq 0$ der Näherungsbruch $\frac{p_n}{q_n} = [a_0, a_1, \ldots, a_n]$ bereits gekürzt, d.h., $(p_n, q_n) = 1$, und für $n \geq 1$ gilt

$$|x - \frac{p_n}{q_n}| < \frac{1}{q_n^2}.$$

(2) Für alle $n \geq 1$ ist mindestens eine der beiden Abschätzungen erfüllt:

$$|x - \frac{p_n}{q_n}| < \frac{1}{2q_n^2} \quad \textit{oder} \quad |x - \frac{p_{n+1}}{q_{n+1}}| < \frac{1}{2q_{n+1}^2}.$$

[22] Axel Thue (1863–1922), norwegischer Mathematiker.
[23] Carl Ludwig Siegel (1896–1981), deutscher Mathematiker, Professor in Frankfurt 1922–1935.
[24] Klaus Friedrich Roth (1925–2015), aus Deutschland stammender britischer Mathematiker.

(3) Wenn $p, q \in \mathbb{Z}$ teilerfremd mit $q > 0$ sind und

$$|x - \frac{p}{q}| < \frac{1}{2q^2}$$

gilt, dann ist $\frac{p}{q}$ ein Näherungsbruch aus der Kettenbruchentwicklung von x.

Beweis. (1) Wir wissen bereits aus Lemma 8.5, daß

$$\frac{p_n}{q_n} = [a_0, a_1, \ldots, a_n]$$

gilt. Ein gemeinsamer Teiler von p_n und q_n teilt

$$\det(\begin{pmatrix} p_n & r_n \\ q_n & s_n \end{pmatrix}) = \prod_{i=0}^{n} \det(\begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix}) = \pm 1.$$

Daher sind p_n und q_n teilerfremd. Weiter gilt nach der Rechnung (8.1) aus Satz 8.12 mit dem Restglied $x_{n+1} > a_{n+1} \geq 1$

$$x - \frac{p_n}{q_n} = \frac{(-1)^{n+1}}{q_n(q_n x_{n+1} + q_{n-1})}, \tag{8.2}$$

daher

$$|x - \frac{p_n}{q_n}| = \frac{1}{q_n(q_n x_{n+1} + q_{n-1})} \leq \frac{1}{q_n(q_n a_{n+1} + q_{n-1})} = \frac{1}{q_n q_{n+1}} < \frac{1}{q_n^2}.$$

(2) Beide Abschätzungen gehen schief nach (8.2), wenn gleichzeitig gilt (Gleichheit ist nämlich ausgeschlossen, weil mit $x \notin \mathbb{Q}$ auch x_{n+1} und x_{n+2} nicht rational sind):

$$\begin{aligned} 2q_n &> q_n x_{n+1} + q_{n-1}, \\ 2q_{n+1} &> q_{n+1} x_{n+2} + q_n. \end{aligned} \tag{8.3}$$

Dann gilt $1 < x_{n+1} < 2$ und somit $a_{n+1} = \lfloor x_{n+1} \rfloor = 1$. Daraus folgt

$$\begin{aligned} x_{n+1} &= [a_{n+1}, x_{n+2}] = 1 + \frac{1}{x_{n+2}}, \\ q_{n+1} &= q_n a_{n+1} + q_{n-1} = q_n + q_{n-1}. \end{aligned}$$

Mit diesem Wissen gehen wir zurück zu den Ungleichungen (8.3), die wir zunächst nach x_{n+1} und x_{n+2} auflösen:

$$\begin{aligned} \frac{2q_n - q_{n-1}}{q_n} &> x_{n+1} = 1 + \frac{1}{x_{n+2}}, \\ \frac{2q_{n+1} - q_n}{q_{n+1}} &> x_{n+2}. \end{aligned} \tag{8.4}$$

Aus (8.4) bekommen wir eine untere Schranke für x_{n+2}, so daß

$$2 - \frac{q_n}{q_n + q_{n-1}} = \frac{2q_{n+1} - q_n}{q_{n+1}} > x_{n+2} > \frac{1}{\frac{2q_n - q_{n-1}}{q_n} - 1} = \frac{q_n}{q_n - q_{n-1}}$$

Mit $1 \geq t := q_{n-1}/q_n > 0$ gibt das den Widerspruch

$$2 > \frac{q_n}{q_n + q_{n-1}} + \frac{q_n}{q_n - q_{n-1}} = \frac{1}{1+t} + \frac{1}{1-t} = \frac{2}{1-t^2} > 2.$$

(3) Sei $\frac{p}{q} = [a_0, \ldots, a_n]$ eine Kettenbruchentwicklung der rationalen Zahl p/q. Wir haben die folgende Freiheit:

$$\frac{p}{q} = [a_0, \ldots, a_n] = \begin{cases} [a_0, \ldots, a_n - 1, 1] & \text{wenn } a_n \geq 2 \\ [a_0, \ldots, a_{n-1} + 1] & \text{wenn } a_n = 1. \end{cases}$$

Damit dürfen wir uns die Parität von n wünschen. Wir richten es so ein, daß es ein $0 \leq \delta < 1$ gibt mit

$$x = \frac{p}{q} + (-1)^n \frac{\delta}{2q^2}.$$

Wir schreiben nun mit der Notation von Satz 8.6 für ein geeignetes $\xi \in \mathbb{R}$

$$x = [a_0, \ldots, a_n, \xi] = \frac{p_n \xi + p_{n-1}}{q_n \xi + q_{n-1}}$$

Das ξ bestimmt sich eindeutig aus

$$\xi = \frac{x q_{n-1} - p_{n-1}}{-x q_n + p_n},$$

so daß einzig der Fall $p = p_n = x q_n = x q$ gesondert zu betrachten ist. In diesem Fall ist bereits $x = \frac{p}{q}$, und es ist nichts zu tun. Aus

$$(-1)^n \frac{\delta}{2q^2} = x - \frac{p}{q} = \frac{p_n \xi + p_{n-1}}{q_n \xi + q_{n-1}} - \frac{p_n}{q_n} = \frac{-\det(M_n)}{q_n(q_n \xi + q_{n-1})} = \frac{(-1)^n}{q(q\xi + q_{n-1})}$$

folgt

$$\delta(q\xi + q_{n-1}) = 2q$$

und weiter

$$\xi = \frac{2}{\delta} - \frac{q_{n-1}}{q} > 2 - 1 = 1,$$

weil $\delta < 1$ und $q_{n-1} < q_n = q$. Wir setzen nun

$$\xi_i = [a_{i+1}, \ldots, a_n, \xi],$$

so daß $\xi_i > a_{i+1} \geq 1$ und

$$x = [a_0, \ldots, a_n, \xi] = [a_0, \ldots, a_i, \xi_i].$$

Hieraus liest man ab, daß $a_{i+1} = \lfloor \xi_i \rfloor$ und somit nach Satz 8.12 dies den Anfang der Kettenbruchentwicklung von x beschreibt. $\square$

Teil 2. Arithmetik in Restklassenringen

9. Der Chinesische Restsatz

9.1. **Algebraisch: Teilbarkeit und Kongruenzrechnung.** Wir haben bisher Kongruenzrechnungen als Rechnung in ganzen Zahlen oder allenfalls noch mit Brüchen, die zum Modulus teilerfremde Nenner haben, betrachtet. Wir gehen nun einen Schritt konsequenter vor.

Sei $m \geq 1$ eine natürliche Zahl. Wir wissen bereits aus Bemerkung 4.2, daß Kongruenz modulo m eine Äquivalenzrelation auf $\mathbb{Z}$ definiert. Wir schreiben nun

$$[a] = \{b \in \mathbb{Z} \ ; \ b \equiv a \pmod{m}\}$$

für die Äquivalenzklasse modulo m, die die ganze Zahl a enthält. Diese besteht aus allen Zahlen der Form $a + mx$ mit $x \in \mathbb{Z}$, kurz

$$[a] = a + m\mathbb{Z}.$$

Definition 9.1. Der Ring der **Restklassen modulo** m ist der Ring

$$\mathbb{Z}/m\mathbb{Z}$$

bestehend aus den Äquivalenzklassen der Kongruenzrelation modulo m. Addition und Multiplikation sind vertreterweise definiert (und wohldefiniert!):

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

Bemerkung 9.2. (1) Die Menge $\mathbb{Z}/m\mathbb{Z}$ ist in der Tat ein kommutativer Ring mit Eins [1]. Das haben wir, wenn wir die Kongruenzrechnung verinnerlicht haben, eigentlich schon immer gewußt: wir rechnen mit Vertretern, aber entscheidend ist nur die Kongruenzklasse modulo m. Die Wahl des Vertreters spielt für das Ergebnis als Kongruenzklasse keine Rolle, siehe Proposition 4.3.

(2) Der Übergang $a \mapsto [a]$ zur Restklasse modulo m ist ein Ringhomomorphismus $\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$. Die Homomorphieeigenschaft besagt gerade

$$[a + b] = [a] + [b],$$
$$[ab] = [a] \cdot [b],$$

also daß man in $\mathbb{Z}/m\mathbb{Z}$ vertreterweise rechnen darf und muß. **Rechnen in $\mathbb{Z}/m\mathbb{Z}$ ist nichts anderes als das Rechnen mit Kongruenzen modulo m.**

Proposition 9.3. *Sei m eine natürliche Zahl.*

(1) Der Ring $\mathbb{Z}/m\mathbb{Z}$ hat genau m Elemente.

(2) Für jedes $a \in \mathbb{Z}$ sind die Elemente

$$[a], [a + 1], \ldots, [a + m - 1]$$

die m verschiedenen Elemente von $\mathbb{Z}/m\mathbb{Z}$.

(3) Für $a \in \mathbb{Z}$ ist $m \mid a \iff [a] = 0 \in \mathbb{Z}/m\mathbb{Z}$.

Beweis. Trivial. □

9.2. **Algebraisch: Faktorisierung in Primpotenzen — der Chinesische Restsatz.** Seien $m \mid n$ natürliche Zahlen und $a, b \in \mathbb{Z}$. Wenn $n \mid a - b$, dann gilt auch $m \mid a - b$. Jede Kongruenzklasse modulo n ist also in einer einzigen Kongruenzklasse modulo m enthalten. Bezeichnen wir vorübergehend zur Abgrenzung diese Klassen mit $[a]_n$ bzw. $[a]_m$, dann heißt das $[a]_n \subseteq [a]_m$ als Teilmenge von $\mathbb{Z}$. Wenn wir die Äquivalenzklassen aber als Elemente der jeweiligen Restklassenringe auffassen, dann bedeutet diese Beobachtung, daß

$$\mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$$
$$[a]_n \mapsto [a]_m$$

wohldefiniert ist. Außerdem ist es ein Ringhomomorphismus.

Definition 9.4. Seien A und B Ringe (mit Eins). Dann ist auf der Produktmenge

$$A \times B = \{(a, b) \ ; \ a \in A, b \in B\}$$

durch komponentenweise Addition und Multiplikation, also

$$(a_1, b_1) + (a_2, b_2) := (a_1 + a_2, b_1 + b_2)$$
$$(a_1, b_1) \cdot (a_2, b_2) := (a_1 \cdot a_2, b_1 \cdot b_2),$$

eine Ringstruktur (mit Eins $(1, 1)$) definiert. Die nötigen Rechengesetze erbt $A \times B$ automatisch von seinen Faktoren A und B. Der Ring $A \times B$ wird **Produkt** von A und B genannt.

Satz 9.5 (Chinesischer Restsatz)**.** *Seien n und m teilerfremde natürliche Zahlen. Dann ist*

$$\mathbb{Z}/nm\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$$
$$[a]_{nm} \mapsto ([a]_m, [a]_n)$$

ein Isomorphismus von Ringen.

Beweis. Die angegebene Abbildung ist wohldefiniert und offensichtlich mit Addition und Multiplikation verträglich. Es bleibt einzusehen, daß die Abbildung bijektiv ist. Beide Seiten haben nm Elemente. Daher reicht es nur eine der beiden Eigenschaften „injektiv“ oder „surjektiv“ nachzuweisen. Wir machen beides, um zu illustrieren, wie das nette Abzählargument Arbeit spart!

Angenommen $[a]_{nm}$ und $[b]_{nm}$ gehen auf das gleiche Paar. Dann ist $[a]_n = [b]_n$, also $n \mid b - a$ und genauso $m \mid b - a$. Damit ist $b - a$ ein gemeinsames Vielfaches von n und m. Wegen $(n, m) = 1$ ist $\mathrm{kgV}(n, m) = nm$, also $nm \mid b - a$, siehe Lemma 3.22, oder eben $[a]_{nm} = [b]_{nm}$. Die Abbildung ist also injektiv. (Es hätte gereicht, aus $[a]_{nm}$ geht auf 0 auf $[a]_{nm} = 0$ zu schließen.)

Wir zeigen nun Surjektivität. Sei $[a]_m$ und $[b]_n$ gegeben. Dann suchen wir ein $x \in \mathbb{Z}$, so daß $x \equiv a \pmod{m}$ und gleichzeitig $x \equiv b \pmod{n}$ gilt. Weil n, m teilerfremd sind, gibt es nach dem Lemma von Bézout 3.10 ganze Zahlen u, v mit

$$1 = un + vm.$$

Dann tut $x = aun + bvm$ das Gewünschte:

$$[x]_n = [aun + bvm]_n = [bvm]_n = [b(1 - un)]_n = [b]_n,$$
$$[x]_m = [aun + bvm]_m = [aun]_m = [a(1 - vm)]_m = [a]_m. \qquad \square$$

Korollar 9.6. *Sei $m = \prod_{i=1}^{r} m_i$ ein Produkt paarweise teilerfremder Faktoren. Dann ist*

$$\mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/m_1\mathbb{Z} \times \ldots \times \mathbb{Z}/m_r\mathbb{Z}$$
$$[a]_m \mapsto ([a]_{m_i})_{1 \le i \le r}$$

ein Isomorphismus von Ringen.

Insbesondere, sei $n = \prod_{i=1}^{r} p_i^{e_i}$ die Primfaktorzerlegung von $n \in \mathbb{N}$ mit paarweise verschiedenen Primzahlen p_i. Dann gilt

$$\mathbb{Z}/n\mathbb{Z} \simeq \mathbb{Z}/p_1^{e_1} \times \ldots \times \mathbb{Z}/p_r^{e_r}\mathbb{Z}.$$

Beweis. Per Induktion nach der Anzahl der Faktoren r aus Satz 9.5. Der Induktionsschritt benutzt die zwei Faktoren m_1 und $m' := m_2 \cdot \ldots \cdot m_r$. Diese sind teilerfremd, weil ein gemeinsamer Primteiler $p \mid (m_1, m')$ als Teiler des Produkts m' ein Teiler von einem der Faktoren m_i, $2 \le i \le r$ wäre und damit dann $(m_1, m_i) \neq 1$, Widerspruch. $\square$

Beispiel 9.7. Aus Satz 9.5 folgt sofort

$$\mathbb{Z}/6\mathbb{Z} \times \mathbb{Z}/35\mathbb{Z} \simeq \mathbb{Z}/10\mathbb{Z} \times \mathbb{Z}/21\mathbb{Z} \simeq \mathbb{Z}/14\mathbb{Z} \times \mathbb{Z}/15\mathbb{Z},$$

weil alle drei Gruppen nach dem Chinesischen Restsatz isomorph zu $\mathbb{Z}/210\mathbb{Z}$ sind.

Korollar 9.8. *Sei $m = \prod_{i=1}^r m_i$ ein Produkt paarweise teilerfremder Faktoren, und seien $a_i \in \mathbb{Z}$ für $i = 1, \ldots, r$. Dann haben die Kongruenzen*

$$X \equiv a_i \pmod{m_i}$$

für alle $i = 1, \ldots, r$ eine eindeutige simultane Lösung $[\xi] \in \mathbb{Z}/m\mathbb{Z}$. Das heißt, es gibt $\xi \in \mathbb{Z}$, so daß

$$[\xi] = \xi + m\mathbb{Z} = \{x \in \mathbb{Z} \ ; \ x \equiv a_i \pmod{m_i} \text{ für alle } 1 \leq i \leq r\}$$

und dieses ξ ist eindeutig bis auf Kongruenz modulo m (weil eben nur die Restklasse $[\xi]$ eindeutig ist).

Beweis. Das folgt sofort aus Korollar 9.6, denn die Lösungsmenge entspricht der Menge aller $x \in \mathbb{Z}$, so daß $[x]_m$ unter dem Isomorphismus des Chinesischen Restsatzes auf das Tupel der $[a_i]_{m_i}$ für alle $1 \leq i \leq r$ abgebildet wird. Nach Korollar 9.6 gibt es so eine Restklasse, und sie ist auch eindeutig. □

Bemerkung 9.9. Wenn man den Beweis von Korollar 9.8 zu seinem argumentativen Kern zurückverfolgt, bekommt man das folgende Verfahren zur Bestimmung der Lösung. In der Notation des Korollars setzen wir für $i = 1, \ldots, r$

$$m_i' = \prod_{j \neq i} m_j,$$

so daß $m = m_i \cdot m_i'$ und $(m_i, m_i') = 1$. Wir wenden den euklidischen Algorithmus an, um $u_i, v_i \in \mathbb{Z}$ zu finden mit

$$u_i m_i + v_i m_i' = 1$$

so wie das Lemma von Bezout, Lemma 3.10, dies verspricht. Dann gilt

$$e_i := v_i m_i' \equiv \begin{cases} 0 & \pmod{m_j} \text{ für alle } j \neq i, \\ 1 - u_i m_i \equiv 1 & \pmod{m_i}. \end{cases}$$

Aus den Idempotenten e_i kombinieren wir linear eine Lösung ξ_0 der simultanen Kongruenzen $X \equiv a_i \pmod{m_i}$ als

$$\xi_0 = \sum_{i=1}^r a_i e_i.$$

Die Menge aller Lösungen aus $\mathbb{Z}$ ist dann $[\xi_0] = \xi_0 + m\mathbb{Z}$.

Übungsaufgaben zu §9

Übungsaufgabe 9.1. Seien n, m natürliche Zahlen. Sei $d = (n, m)$ und $D = [a, b]$. Dann gilt

$$\mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z} \simeq \mathbb{Z}/d\mathbb{Z} \times \mathbb{Z}/D\mathbb{Z}.$$

10. Einheiten in Restklassenringen

10.1. **Der Chinesische Restsatz für Einheiten.** Wenn $a \equiv b \pmod{m}$, dann gibt es $k \in \mathbb{Z}$ mit $b = a + km$ und daher

$$(a, m) = (b, m).$$

Daher ist das Folgende wohldefiniert.

Definition 10.1. Eine **teilerfremde Restklasse** modulo m ist ein $[a] \in \mathbb{Z}/m\mathbb{Z}$ mit $(a, m) = 1$, also eine Restklasse in der ein (äquivalent jeder) Vertreter zu m teilerfremd ist.

Bemerkung 10.2. Die Lösungstheorie der Kongruenzgleichung $aX \equiv 1 \pmod{m}$ besagt, daß es genau dann eine Lösung gibt, wenn $(a, m) = 1$, also wenn $[a]$ eine teilerfremde Restklasse ist. Es gibt dann $x \in \mathbb{Z}$ mit $[a] \cdot [x] = 1$.

Wir erinnern: Elemente eines Rings R mit einem multiplikativen Inversen heißen **Einheiten** und bilden eine Gruppe

$$R^\times = \{a \in R \; ; \; \exists x \in R : ax = 1\}$$

bezüglich Multiplikation.

Proposition 10.3. *Sei $m \in \mathbb{N}$. Die Menge der zu m teilerfremden Reste bildet eine abelsche Gruppe*

$$(\mathbb{Z}/m\mathbb{Z})^\times$$

bezüglich der Multiplikation. Diese Gruppe hat die Ordnung $\varphi(m)$.

Beweis. Die Einheiten eines jeden Rings bilden eine Gruppe. Stellt man die Einheiten in $\mathbb{Z}/m\mathbb{Z}$ durch Division mit Rest eindeutig mit einem Vertreter r in $0 \leq r \leq m-1$ dar, dann ergibt sich

$$(\mathbb{Z}/m\mathbb{Z})^\times = \{[r] \in \mathbb{Z}/m\mathbb{Z} \; ; \; 0 \leq r \leq m-1, \; (r, m) = 1\}$$

und dies ist genau die Menge, deren Mächtigkeit die Eulersche φ-Funktion definiert. □

Satz 10.4 (Chinesischer Restsatz für Einheiten)**.** *Seien n und m teilerfremde natürliche Zahlen. Dann ist*

$$\begin{aligned} (\mathbb{Z}/nm\mathbb{Z})^\times &\to (\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times \\ [a]_{nm} &\mapsto ([a]_m, [a]_n) \end{aligned}$$

ein Isomorphismus von abelschen Gruppen.

Beweis. Das ist nichts anderes als die Einschränkung des Isomorphismus des Chinesischen Restsatzes, Satz 9.5, auf die Einheitengruppe der Ringe

$$(\mathbb{Z}/nm\mathbb{Z})^\times \simeq (\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z})^\times$$

verbunden mit der Beobachtung, daß für alle Ringe A, B gilt $(A \times B)^\times = A^\times \times B^\times$.

Etwas direkter: eine Restklasse $[a]$ ist teilerfremd modulo mn, wenn sie keinen gemeinsamen Teiler mit nm also weder mit m noch mit n hat. Dies zeigt, daß die (Produkte der) Einheitengruppen durch den besagten Isomorphismus bijektiv aufeinander abgebildet werden. □

Korollar 10.5. *Sei $m = \prod_{i=1}^r m_i$ ein Produkt paarweise teilerfremder Faktoren. Dann ist*

$$\begin{aligned} (\mathbb{Z}/m\mathbb{Z})^\times &\to (\mathbb{Z}/m_1\mathbb{Z})^\times \times \ldots \times (\mathbb{Z}/m_r\mathbb{Z})^\times \\ [a]_m &\mapsto ([a]_{m_i})_{1 \leq i \leq r} \end{aligned}$$

ein Isomorphismus abelscher Gruppen.

Insbesondere, sei $n = \prod_{i=1}^r p_i^{m_i}$ die Primfaktorzerlegung von $n \in \mathbb{N}$ mit paarweise verschiedenen Primzahlen p_i. Dann gilt

$$(\mathbb{Z}/n\mathbb{Z})^\times \simeq (\mathbb{Z}/p_1^{m_1}\mathbb{Z})^\times \times \ldots \times (\mathbb{Z}/p_r^{m_r}\mathbb{Z})^\times$$

Beweis. Klar. □

Korollar 10.6. *Die Eulersche φ-Funktion ist multiplikativ.*

Beweis. Das wissen wir bereits mit einem kombinatorischen Beweis seit Korollar 7.23, aber es folgt auch für teilerfremde n und m auf algebraische Weise aus

$$\varphi(nm) = \#(\mathbb{Z}/nm\mathbb{Z})^\times = \#\left(\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}\right)^\times = \#(\mathbb{Z}/m\mathbb{Z})^\times \cdot \#(\mathbb{Z}/n\mathbb{Z})^\times = \varphi(m)\varphi(n). \quad \square$$

Bemerkung 10.7. (1) Mit der alternativen Definition $\varphi(m) = \#(\mathbb{Z}/m\mathbb{Z})^\times$ hat man eine algebraische und konzeptionelle Definition der Eulerschen φ-Funktion. Die Multiplikativität, die wir in Korollar 7.23 hintenrum und auf undurchsichtige Art und Weise erhalten haben, findet nun als numerische Konsequenz des Chinesischen Restsatzes für Einheiten eine befriedigend natürliche Erklärung.

(2) Wir reproduzieren nun durch ein arithmetisches Argument die Formel aus Beispiel 7.24 für die Eulersche φ-Funktion. Sei p eine Primzahl. Dann gilt für alle $m \geq 1$

$$\varphi(p^m) = (p-1)p^{m-1},$$

denn genau die p^{m-1} durch p teilbaren Reste $1 \leq r \leq p^m$ sind nicht zu p^m teilerfremd.

Sei $n = \prod_{i=1}^r p_i^{m_i}$ die Primfaktorzerlegung von $n \in \mathbb{N}$. Es gilt dann aufgrund der Multiplikativität von φ die Formel

$$\varphi(n) = \prod_{i=1}^r (p_i - 1)p_i^{m_i - 1}.$$

Satz 10.8 (Euler — kleiner Fermat)**.** *Sei $m \in \mathbb{N}$ und $a \in \mathbb{Z}$ teilerfremd zu m. Dann gilt*

$$a^{\varphi(m)} \equiv 1 \pmod{m}.$$

Beweis. Multiplikation mit $[a]$ ist eine Bijektion auf der Menge $(\mathbb{Z}/m\mathbb{Z})^\times$ der teilerfremden Reste. Daher gilt

$$a^{\varphi(m)} \cdot \prod_{x \in (\mathbb{Z}/m\mathbb{Z})^\times} x \equiv \prod_{x \in (\mathbb{Z}/m\mathbb{Z})^\times} ax \equiv \prod_{x \in (\mathbb{Z}/m\mathbb{Z})^\times} x \pmod{m}.$$

Kürzen des Faktors $\prod_{x \in (\mathbb{Z}/m\mathbb{Z})^\times} x$ liefert die Behauptung. □

Bemerkung 10.9. Da $\varphi(p) = p-1$ gilt für Primzahlen p, ist Satz 5.30 ein Spezialfall von Satz 10.8.

Bemerkung 10.10. Der kleine Satz des Fermat stellt ein notwendiges Primzahlkriterium dar. Sei ein $n \in \mathbb{N}$ gegeben und $2 \leq a < n$. Dann ist n höchstens dann Primzahl, wenn

$$a^{n-1} \equiv 1 \pmod{n}$$

gilt. Wenn n keine Primzahl ist, dann gilt entweder $d = (a, n) \neq 1$ und man hat mit d sogar einen Teiler von n gefunden. Oder $(a, n) = 1$ und a^{n-1} nimmt heuristisch gesehen jeden der möglichen $\varphi(n)$ teilerfremden Werte mit Gleichverteilung zufällig an. Die Kongruenz $a^{n-1} \equiv 1 \pmod{n}$ hat dann eine Wahrscheinlichkeit von $1/\varphi(n)$.

Nach t Versuchen mit in $\mathbb{Z}$ multiplikativ unabhänigigen a (also zum Beispiel den ersten t Primzahlen), bei denen stets der Rest $a^{n-1} \equiv 1 \pmod{n}$ herauskommt, ist die Zahl n nach dieser Heuristik nur noch mit Wahrscheinlichkeit $1/\varphi(n)^t$ keine Primzahl. Leider kann man auf diese Weise nicht beweisen, daß eine Zahl Primzahl ist.

Definition 10.11. Man nennt eine natürliche Zahl $n \geq 2$ eine **Fermat–Pseudoprimzahl zur Basis** $a \in \mathbb{N}$, wenn

(i) n keine Primzahl ist,
(ii) a teilerfremd zu n ist,
(iii) und $a^{n-1} \equiv 1 \pmod{n}$.

Beispiel 10.12. Beispiele für Fermat–Pseudoprimzahlen zur Basis 2 sind (bis 10.000):

$$341, 561, 645, 1105, 1387, 1729, 1905, 2047, 2465, 2701, 2821,$$
$$3277, 4033, 4369, 4371, 4681, 5461, 6601, 7957, 8321, 8481, 8911, \ldots$$

Bemerkung 10.13. Eine **Carmichael-Zahl** ist eine natürliche Zahl $n \geq 2$, die bezüglich aller zu n teilerfremden Zahlen a eine Fermat–Pseudoprimzahl zur Basis a ist. Die Carmichael-Zahlen sind also genau die natürlichen Zahlen, für die der Primzahltest durch den kleinen Fermat versagt

(es sei denn, man zieht den extrem unwahrscheinlichen Fall, daß die gewählte Basis a ein Teiler von n ist).

Beispiele für Carmichael-Zahlen sind (bis 100.000)

$$561, 1105, 1729, 2465, 2821, 6601, 8911, 10585, 15841, 29341,$$
$$41041, 46657, 52633, 62745, 63973, 75361, \ldots$$

Theorem 10.14 (Alford, Granville, Pomerance 1994)**.** *Es gibt unendlich viele Carmichael-Zahlen.*

Genauer gibt es ein x_0, so daß für alle $x \geq x_0$ die Zählfunktion der Carmichaelzahlen abgeschätzt werden kann zu

$$\#\{n \ ; \ n \leq x \text{ und } n \text{ ist Carmichaelzahl}\} \geq x^{2/7}.$$

10.2. **Primitivwurzeln.** Wir sind an der Struktur von $(\mathbb{Z}/n\mathbb{Z})^\times$ als abelsche Gruppe interessiert. Der Chinesische Restsatz reduziert diese Frage auf den Fall einer Primpotenz $n = p^m$. Man bestimmt durch Ausprobieren:

n	$\varphi(n)$				Vertreter	erzeugt von:
2	1	$(\mathbb{Z}/2\mathbb{Z})^\times$	$\simeq$	0	$\{1\}$	
3	2	$(\mathbb{Z}/3\mathbb{Z})^\times$	$\simeq$	$\mathbb{Z}/2\mathbb{Z}$	$\{1,2\}$	$[2]$
4	2	$(\mathbb{Z}/4\mathbb{Z})^\times$	$\simeq$	$\mathbb{Z}/2\mathbb{Z}$	$\{1,3\}$	$[3]$
5	4	$(\mathbb{Z}/5\mathbb{Z})^\times$	$\simeq$	$\mathbb{Z}/4\mathbb{Z}$	$\{1,2,3,4\}$	$[2]$
6	2	$(\mathbb{Z}/6\mathbb{Z})^\times$	$\simeq$	$\mathbb{Z}/2\mathbb{Z}$	$\{1,5\}$	$[5]$
7	6	$(\mathbb{Z}/7\mathbb{Z})^\times$	$\simeq$	$\mathbb{Z}/6\mathbb{Z}$	$\{1,2,3,4,5,6\}$	$[3]$
8	4	$(\mathbb{Z}/8\mathbb{Z})^\times$	$\simeq$	$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$	$\{1,3,5,7\}$	$[-1], [5]$
9	6	$(\mathbb{Z}/9\mathbb{Z})^\times$	$\simeq$	$\mathbb{Z}/6\mathbb{Z}$	$\{1,2,4,5,7,8\}$	$[2]$

Notation 10.15. Sei p eine Primzahl. Der Ring der Restklassen $\mathbb{Z}/p\mathbb{Z}$ hat

$$(\mathbb{Z}/p\mathbb{Z})^\times = \mathbb{Z}/p\mathbb{Z} \setminus \{0\},$$

denn für jedes a mit $p \nmid a$ hat $[a] \in \mathbb{Z}/p\mathbb{Z}$ ein Inverses. Damit ist dieser Restklassenring sogar ein Körper, und wird als solcher mit

$$\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$$

bezeichnet. Dies ist (bis auf eindeutige Isomorphie) der einzige Körper mit p Elementen.

Definition 10.16. Sei $m \geq 1$ und $[a] \in (\mathbb{Z}/m\mathbb{Z})^\times$. Die Ordnung von $[a]$ ist die kleinste natürliche Zahl $\operatorname{ord}(a) > 0$ mit

$$[a]^{\operatorname{ord}(a)} = 1 \in (\mathbb{Z}/m\mathbb{Z})^\times.$$

Wir schreiben $\operatorname{ord}(a)$ statt $\operatorname{ord}([a])$, um Klammern zu sparen, bemerken aber auch, daß die Notation besser von m abhängen sollte, denn $\operatorname{ord}(a)$ tut dies auch!

Bemerkung 10.17. Die Ordnung von $[a]$ in $(\mathbb{Z}/m\mathbb{Z})^\times$ ist natürlich nichts anderes ein Spezialfall des Konzepts der Ordnung des Elements in einer Gruppe.

Lemma 10.18. *Sei $m \geq 1$ und $[a] \in (\mathbb{Z}/m\mathbb{Z})^\times$.*

(1) Für $n \in \mathbb{Z}$ gilt $[a]^n = 1 \iff \operatorname{ord}(a) \mid n$.

(2) Es ist $\operatorname{ord}(a)$ ein Teiler von $\varphi(m)$.

Beweis. (1) Wir schreiben nach Division mit Rest $n = q \operatorname{ord}(a) + r$ mit $0 \leq r < \operatorname{ord}(a)$. Dann gilt

$$1 \equiv a^n = a^{q \operatorname{ord}(a) + r} = (a^{\operatorname{ord}(a)})^q \cdot a^r \equiv a^r \pmod{m}.$$

Wenn $r > 0$ gilt, dann ist $\operatorname{ord}(a)$ nicht minimal. Also gilt $r = 0$ und damit $\operatorname{ord}(a) \mid n$.

Aussage (2) folgt sofort aus $a^{\varphi(m)} \equiv 1 \pmod{m}$ nach Satz 10.8 und Aussage (1) . □

Definition 10.19. Eine **Primitivwurzel modulo** der Primzahl p ist ein $w \in \mathbb{Z}$, so daß $[w] \in (\mathbb{Z}/p\mathbb{Z})^\times$ die Ordnung $p - 1$ hat.

Satz 10.20. *Sei K ein Körper und $P(X) \in K[X]$ ein Polynom $P(X) \neq 0$.*

(1) $P(X)$ hat höchstens $\deg(P)$ viele Nullstellen in K.

(2) Wenn $P(X) = Q(X)R(X)$ in $K[X]$ und $P(X)$ hat $\deg(P)$ verschiedene Nullstellen in K, dann hat auch $Q(X)$ genau $\deg(Q)$ verschiedene Nullstellen in K.

Beweis. (1) Wir zeigen dies per Induktion nach $\deg(P)$. Der Induktionsanfang behandelt konstante Polynome $\neq 0$, die damit keine Nullstelle haben. Zu $a \in K$ führt Polynomdivision von $P(X)$ durch $(X - a)$ zu Polynomen $q(X)$ und r mit

$$P(X) = q(X)(X - a) + r$$

und mit $\deg(r) < \deg(X - a) = 1$, also $r \in K$. Ist $a \in K$ Nullstelle von $P(X)$, dann folgt

$$r = P(a) - q(a)(a - a) = 0.$$

Also gilt dann $P(X) = (X - a)q(X)$.

Weil $\deg(q) = \deg(P) - 1$ ist, hat $q(X)$ per Induktion höchstens $\deg(P) - 1$ viele Nullstellen. Nullstellen von $P(X)$ sind Nullstellen von $q(X)$ und a, und das sind höchstens $(\deg(P)-1)+1 = \deg(P)$-viele.

(2) Aus $P(a) = Q(a)R(a)$ folgt, daß die Nullstellen von $P(X)$ genau die Nullstellen von $Q(X)$ und die Nullstellen von $R(X)$ sind. Wenn $P(X)$ nach Voraussetzung $\deg(P)$ viele Nullstellen hat und $R(X)$ nach (1) höchstens $\deg(R)$ viele, dann hat $Q(X)$ mindestens

$$\deg(P) - \deg(R) = \deg(Q)$$

viele Nullstellen in K. Weil nach (1) dies auch die obere Schranke für die Anzahl der Nullstellen von $Q(X)$ ist, folgt (2). □

Satz 10.21 (Gauß)**.** *Für alle $d \mid p - 1$ gibt es genau $\varphi(d)$ Elemente in $\mathbb{F}_p^\times$ der Ordnung d. Insbesondere gibt es Primtivwurzeln modulo p, und zwar genau $\varphi(p - 1)$ viele.*

Beweis. Wir setzen für jedes $d \mid p - 1$

$$A(d) = \{a \in \mathbb{F}_p^\times \ ; \ \operatorname{ord}(a) = d\}$$

und $a(d) = \#A(d)$.

Für jedes $m \mid p-1$ ist $X^m - 1$ nach (5.1) ein Teiler von $X^{p-1} - 1$. Nach dem kleinen Satz des Fermat, Satz 5.30, hat $X^{p-1} - 1$ die $p - 1$ verschiedenen Nullstellen $[1], \ldots, [p - 1]$ in $\mathbb{F}_p$. Nach Satz 10.20 hat damit $X^m - 1$ genau m verschiedene Nullstellen in $\mathbb{F}_p$. Nach Lemma 10.18 sind die Nullstellen von $X^m - 1$ genau die $A(d)$ mit $d \mid m$. Daher gilt (die Vereinigung der $A(d)$ ist disjunkt)

$$m = \#\{a \in \mathbb{F}_p \ ; \ \text{Nullstelle von } X^m - 1\} = \#\bigcup_{d|m} A(d) = \sum_{d|m} a(d).$$

Nach der Möbiusschen Umkehrformel (die gilt auch eingeschränkt auf Werte $m \mid p-1$) gilt dann für alle $m \mid p - 1$

$$a(m) = \sum_{d|m} d \cdot \mu(\frac{m}{d}) = \varphi(m)$$

nach Korollar 7.23.

Für alle Primzahlen p ist $p - 1 \geq 1$ und daher $\varphi(p - 1) \geq 1$. Dies zeigt die Existenz von Primitivwurzeln modulo p. □

Bemerkung 10.22. (1) Satz 10.21 besagt zwar die Existenz von Primitivwurzeln, aber kein bekannter Beweis gibt Hinweise dazu, wie eine Primitivwurzel zu finden ist.

(2) Emil Artin hat 1927 vermutet, daß zu jedem $a \in \mathbb{N}$, das keine Quadratzahl ist, unendlich viele Primzahlen p existieren, so daß a eine Primitivwurzel modulo p ist. Das ist nur unter der Annahme der Verallgemeinerten Riemannschen Vermutung bewiesen. Ohne Zusatzannahmen ist die Vermutung für keine natürliche Zahl bekannt.

Abbildung 3 plottet für Primzahlen bis 10.000 die Funktion

$$k(p) = \min\{a \ ; \ a \in \mathbb{N} \text{ Primitivwurzel modulo } p\}.$$

Das benötigte Maximum ist 31 (bei $p = 5881$) und $a \leq 10$ reicht oft.

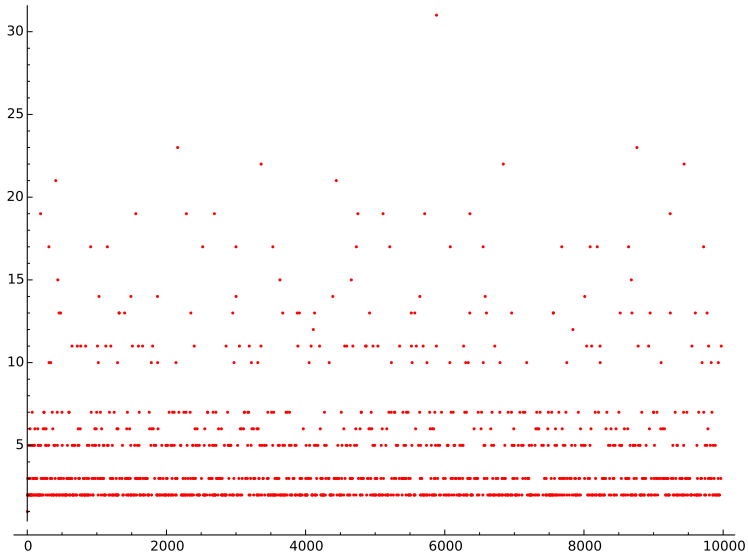


Abbildung 3. Die kleinste Primitivwurzel modulo p.

Theorem 10.23. *Sei p eine Primzahl.*

(1) Die multiplikative Gruppe $\mathbb{F}_p^\times$ ist zyklisch.

(2) Sei w Primitivwurzel modulo p. Dann definiert $\exp_w(n) = [w]^n$ einen Isomorphismus von Gruppen

$$\exp_w : \mathbb{Z}/(p-1)\mathbb{Z} \to \mathbb{F}_p^\times.$$

Beweis. (1) folgt aus (2) zusammen mit der Existenz von Primitivwurzeln aus Satz 10.21.

(2) Die Abbildung ist wohldefiniert nach dem kleinen Fermat, Satz 5.30, denn aus $n \equiv m \pmod{p-1}$ folgt $n = m + k(p-1)$ und

$$[w]^n = [w]^{m+k(p-1)} = [w]^m \cdot ([w^{p-1}])^k = [w]^m.$$

Ein Gruppenhomomorphismus ist die Abbildung aufgrund der Potenzgesetze. Es bleibt zu zeigen, daß für Primitivwurzeln w der so konstruierte Gruppenhomomorphismus ein Isomorphismus ist. Weil beide Gruppen gleich groß sind, nämlich

$$\#(\mathbb{Z}/(p-1)\mathbb{Z}) = p - 1 = \#\mathbb{F}_p^\times,$$

reicht es aus, injektiv oder surjektiv zu zeigen.

Die Abbildung ist injektiv genau dann, wenn $\operatorname{ord}([w]) = p - 1$: aus $[w]^r = [w]^s$ mit $0 \leq r < s \leq p-1$ folgt sonst $[w]^{s-r} = 1$ mit $0 < s - r < p - 1$, ein Widerspruch zu $\operatorname{ord}([w]) = p - 1$. □

Bemerkung 10.24. (1) Die Repräsentanten der Erzeuger der zyklischen Gruppe $\mathbb{F}_p^\times$ sind genau die Primitivwurzeln modulo p.

(2) Sei a eine Primitivwurzel modulo p. Die Notation $\exp_a(n) = a^n$ für den Isomorphismus $\mathbb{Z}/(p-1)\mathbb{Z} \to \mathbb{F}_p^\times$ erinnert uns daran, daß ein Homomorphismus von einer additiven Gruppe $\mathbb{Z}/(p-1)\mathbb{Z}$ in eine multiplikative Gruppe $\mathbb{F}_p^\times$ vorliegt.

10.3. **Struktur der Einheitengruppe modulo Primpotenzen.** Nachdem die multiplikative Struktur von $\mathbb{F}_p^\times$ geklärt ist, wenden wir uns $(\mathbb{Z}/p^n\mathbb{Z})^\times$ zu.

Proposition 10.25. *Sei p eine Primzahl und seien $a, b \in \mathbb{Z}$ teilerfremd zu p mit $p \mid a-b$. Wenn $p = 2$ fordern wir sogar $p^2 \mid a-b$. Dann gilt:*

$$v_p(a^p - b^p) = v_p(a-b) + 1.$$

Beweis. Sei $v_p(a-b) = n$, es gibt also $x \in \mathbb{Z}$ teilerfremd zu p mit $b = a + x \cdot p^n$. Nach Voraussetzung ist $n \geq 1$, und wenn $p = 2$ sogar $n \geq 2$. Wir müssen zeigen, daß p^{n+1} die genaue p-Potenz ist, die in $a^p - b^p$ aufgeht. Es gilt

$$a^p - b^p = (a-b)(a^{p-1} + a^{p-2}b + \ldots + ab^{p-2} + b^{p-1}).$$

Zu zeigen ist, daß der zweite Faktor $\sum_{i=0}^{p-1} a^{p-1-i}b^i$ durch p, aber nicht durch p^2 teilbar ist. Dazu rechnen wir mit dem Binomischen Lehrsatz

$$\begin{aligned}
\sum_{i=0}^{p-1} a^{p-1-i}b^i &= \sum_{i=0}^{p-1} a^{p-1-i}(a+xp^n)^i = \sum_{i=0}^{p-1} a^{p-1-i} \sum_{j=0}^{i} \binom{i}{j} a^{i-j}(xp^n)^j \\
&\equiv \sum_{i=0}^{p-1} a^{p-1-i}\big(a^i + i \cdot a^{i-1}xp^n\big) \pmod{p^2} \\
&\equiv pa^{p-1} + xa^{p-2}p^n \cdot \sum_{i=0}^{p-1} i \equiv pa^{p-1} + xa^{p-2}p^n \frac{p(p-1)}{2} \pmod{p^2}.
\end{aligned}$$

Der zweite Summand hat für ungerade p einen extra Faktor p aus $\frac{p(p-1)}{2}$ und ist daher $\equiv 0 \pmod{p^2}$. Für $p = 2$ reicht nach Voraussetzung bereits p^n, so daß der zweite Summand immer $\equiv 0 \pmod{p^2}$ ist. Der erste Summand pa^{p-1} hingegen ist durch p, aber nicht durch p^2 teilbar. □

Korollar 10.26. *Sei p eine Primzahl und $n \in \mathbb{N}$. Seien $a, b \in \mathbb{Z}$ mit $p \mid a-b$. Dann gilt*

$$p^n \mid a-b \quad \Longrightarrow \quad p^{n+1} \mid a^p - b^p.$$

Beweis. Das folgt aus dem Beweis von Proposition 10.25. Die genaueren Voraussetzungen werden für diese Teilaussage nicht gebraucht. □

Korollar 10.27. *Sei p eine Primzahl und $n \in \mathbb{N}$ mit $p^n \geq 3$. Dann gilt in $(\mathbb{Z}/p^n\mathbb{Z})^\times$*

$$\begin{aligned}
\operatorname{ord}(1+p) &= p^{n-1} \quad \textit{wenn } p \geq 3, \\
\operatorname{ord}(1+p^2) &= p^{n-2} \quad \textit{wenn } p = 2.
\end{aligned}$$

Beweis. Sei zunächst $p \neq 2$ und $a = 1 + p$. Dann zeigt Proposition 10.25 mit $a = a$ und $b = 1$, daß

$$p^n \mid a^{p^{n-1}} - 1.$$

Daher folgt $\operatorname{ord}(a) \mid p^{n-1}$ aus Lemma 10.18 (1). Es reicht nun zu zeigen, daß $a^{p^{n-2}} \not\equiv 1 \pmod{p^n}$ ist. Aber das folgt auch aus Proposition 10.25.

Für $p = 2$ und $a = 1 + p^2 = 5$ argumentiert man genauso. □

Satz 10.28. *Sei p eine Primzahl und $n \in \mathbb{N}$. Zu $a \in \mathbb{Z}$ teilerfremd zu p definieren wir die Folge $(a_k)_{k\in\mathbb{N}}$ rekursiv durch $a_0 = a$ und $a_{k+1} = (a_k)^p$ für alle $n \geq 0$. Dann konvergiert die Folge der Kongruenzklassen $[a_k]$ in $(\mathbb{Z}/p^n\mathbb{Z})^\times$.*

Bemerkung 10.29. Unter *Konvergenz* in der endlichen Menge $(\mathbb{Z}/p^n\mathbb{Z})^\times$ verstehen wir hier, daß die Folge schlußendlich konstant wird.

Beweis. Nach dem kleinen Fermat, Satz 5.30, gilt $a_{k+1} \equiv a_k \pmod{p}$ für alle k. Insbesondere gilt $a_1 \equiv a_0 \pmod{p}$. Per Induktion zeigen wir $p^k \mid a_k - a_{k-1}$ für alle $k \geq 1$. In der Tat ist für $k = 1$ gerade $p \mid a_1 - a_0$ und dann mit Korollar 10.26

$$p^{k+1} \mid a_k^p - a_{k-1}^p = a_{k+1} - a_k$$

für alle k. Ab $k = n - 1$ wird die Folge der Reste in $(\mathbb{Z}/p^n\mathbb{Z})^\times$ konstant. □

Satz 10.30. *Sei p eine Primzahl und $n \geq 1$. Es gibt einen Gruppenhomomorphismus*

$$\omega : \mathbb{F}_p^\times \to (\mathbb{Z}/p^n\mathbb{Z})^\times,$$

so daß für alle $a \in \mathbb{F}_p^\times$ gilt:

(i) $\omega(a)^p = \omega(a)$,
(ii) $a \equiv \omega(a) \pmod{p}$.

Beweis. Zu $a \in \mathbb{F}_p^\times$ wählen wir einen Vertreter $x \in \mathbb{Z}$. Nach Satz 10.28 konvergiert die Folge der $[x^{p^k}] \in (\mathbb{Z}/p^n\mathbb{Z})^\times$. Wir setzen daher

$$\omega(a) := \lim_{k\to\infty} [x^{p^k}].$$

Wenn y ein anderer Vertreter für a ist, dann gilt $p \mid x - y$ und nach Korollar 10.26 wie im Beweis von Satz 10.28

$$\lim_{k\to\infty} [x^{p^k}] = \lim_{k\to\infty} [y^{p^k}].$$

Die Definition von $\omega(a)$ ist daher unabhängig von den Wahlen. Zu $a, b \in \mathbb{F}_p^\times$ wählen wir Vertreter $x, y \in \mathbb{Z}$ und für $ab \in \mathbb{F}_p^\times$ wählen wir den Vertreter xy. Dann gilt für alle $k \in \mathbb{N}$

$$[(xy)^{p^k}] = [x^{p^k}] \cdot [y^{p^k}],$$

so daß

$$[ab] = \lim_{k\to\infty} [(xy)^{p^k}] = \lim_{k\to\infty} [x^{p^k}] \cdot \lim_{k\to\infty} [y^{p^k}] = [a] \cdot [b].$$

Damit haben wir einen Gruppenhomomorphismus konstruiert. Eigenschaft (i) folgt, weil die Folge $[x^{p^k}]$ modulo p betrachtet konstant ist. Da in $\mathbb{F}_p^\times$ überdies $a^p = a$ gilt, folgt dieselbe Relation aufgrund der Homomorphie von $\omega(-)$ auch für $\omega(a)$ als Element von $(\mathbb{Z}/p^n\mathbb{Z})^\times$. □

Definition 10.31. Sei p eine Primzahl und $n \geq 1$. Zu $a \in \mathbb{F}_p^\times$ nennt man das in Satz 10.30 konstruierte Element $\omega(a)$ den **Teichmüllervertreter**.

Theorem 10.32. *Sei $p \geq 3$ ein Primzahl und $n \in \mathbb{N}$.*

(1) *Die Kongruenzklasse von $1 + p$ erzeugt in $(\mathbb{Z}/p^n\mathbb{Z})^\times$ eine zyklische Untergruppe der Ordnung p^{n-1}.*
(2) *$(\mathbb{Z}/p^n\mathbb{Z})^\times$ ist zyklisch von der Ordnung $(p-1)p^{n-1}$.*
(3) *Sei $\zeta \in \mathbb{Z}$ eine Primitivwurzel modulo p. Dann ist $\omega(\zeta)\cdot(1+p)$ ein Erzeuger von $(\mathbb{Z}/p^n\mathbb{Z})^\times$, wobei $\omega(\zeta)$ der Teichmüllervertreter von $\zeta \pmod{p}$ ist.*

Beweis. (1) Das folgt sofort aus Korollar 10.27. Aussage (2) folgt aus Aussage (3).

(3) Wir müssen zeigen, daß $\operatorname{ord}(\omega(\zeta) \cdot (1 + p)) = (p - 1)p^{n-1}$. Wir definieren einen Gruppenhomomorphismus

$$\begin{aligned} \psi : \mathbb{Z}/(p-1)\mathbb{Z} \times \mathbb{Z}/p^{n-1}\mathbb{Z} &\to (\mathbb{Z}/p^n\mathbb{Z})^\times \\ ([x], [y]) &\mapsto \omega(\zeta)^x (1+p)^y. \end{aligned}$$

Wenn $\psi([x], [y]) = 1$, dann ist

$$\omega(\zeta)^x = (1 + p)^{-y} \in (\mathbb{Z}/p^n\mathbb{Z})^\times$$

ein Element von Ordnung ein Teiler des ggT von $\operatorname{ord}(\omega(\zeta)) = p-1$ und $\operatorname{ord}([1+p]_{p^n}) = p^{n-1}$, also gleich $[1]$. Aus

$$\omega(\zeta)^x = [1] = (1+p)^{-y}$$

folgt $p-1 \mid x$ und $p^{n-1} \mid y$. Somit hat ψ trivialen Kern und ist injektiv. Weil beide Seiten die gleiche Ordnung haben, ist ψ sogar ein Isomorphismus. Jetzt können wir bequem die Ordnung von $\omega(\zeta)\cdot(1+p)$ als Ordnung von

$$([1],[1]) \in \mathbb{Z}/(p-1)\mathbb{Z} \times \mathbb{Z}/p^{n-1}\mathbb{Z}$$

bestimmen. Das ist das kgV der Ordnungen in den beiden Faktoren, also von $p-1$ und p^{n-1}. Das ist $(p-1)p^{n-1}$ und war zu zeigen. □

Theorem 10.33. *Sei $n \geq 2$ eine natürliche Zahl.*

(1) $5 = 1+2^2$ *erzeugt in* $(\mathbb{Z}/2^n\mathbb{Z})^\times$ *eine zyklische Untergruppe der Ordnung* 2^{n-2}.

(2) *Es gilt*

$$(\mathbb{Z}/2^n\mathbb{Z})^\times \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{n-2}\mathbb{Z},$$

wobei $[-1]$ *den ersten Faktor und* $[5]$ *den zweiten Faktor erzeugt.*

Beweis. (1) Das folgt sofort aus Korollar 10.27. Für Aussage (2) betrachten wir den Gruppenhomomorphismus

$$\begin{aligned} \psi : \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{n-2}\mathbb{Z} &\to (\mathbb{Z}/2^n\mathbb{Z})^\times \\ ([x],[y]) &\mapsto [-1]^x[5]^y. \end{aligned}$$

Dann ist ψ injektiv, weil

$$[1] = [-1]^x[5]^y \equiv [-1]^x \pmod 4$$

bedeutet, daß x gerade ist. Daraufhin hat man sogar

$$[1] = [5]^y,$$

woraus nach Korollar 10.27 bereits $2^{n-2} \mid y$ folgt.

Weil beide Seiten die gleiche Ordnung haben, ist ψ sogar ein Isomorphismus. □

Übungsaufgaben zu §10

Übungsaufgabe 10.1. Zeigen Sie, daß 561 eine Carmichael-Zahl ist.

Übungsaufgabe 10.2. Zeigen Sie das folgende Kriterium für Carmichael-Zahlen. Eine natürliche Zahl n ist Carmichael-Zahl genau dann, wenn jeder Primteiler $p \mid n$ nur einmal in n vorkommt (n ist quadratfrei) und $p-1$ ein Teiler von $n-1$ ist.

Übungsaufgabe 10.3. Sei $k(p)$ die kleinste positive Primitivwurzel modulo p. Bestimmen Sie

$$\max_{p \leq 10^6} k(p)$$

und

$$\min\{p \ ; \ k(p) \geq 100\}$$

mit Hilfe eines Computeralgebrasystems.

11. Arithmetik in der Kryptographie

11.1. **RSA.** Das RSA-Verfahren zur Verschlüsselung wurde in den 1970er Jahren von Rivest, Shamir und Adleman entwickelt. Das RSA-Verfahren gehört zur **Public Key Kryptographie**.

Von Public Key Kryptographie spricht man bei Protokollen zur Nachrichtenübermittlung, bei denen die benötigten Schlüssel zur Verschlüsselung öffentlich ist, aber keine in vernünftiger Zeit brauchbaren Hinweise die geheimen zur Entschlüsselung nötigen Schlüssel verrät. Jeder Teilnehmer, der gerne Post bekommen möchte, hängt seinen persönlichen Verschlüsselungsmechanismus an eine öffentliche Pinnwand. Dort muß man nachsehen, wenn man verschlüsselte Post verschicken möchte. Hat man die Post erst einmal verschlüsselt, kann man als Absender auch nicht mehr Korrektur lesen. Das kann nur noch der Empfänger.

Insbesondere werden bei Kommunikation mit Public Key Protokollen zwischen zwei Parteien für jede Richtung ein eigener Schlüssel benutzt. Es ist jeweils nur dem Empfänger bekannt, wie die Nachrichten zu entschlüsseln sind.

Definition 11.1 (RSA-Verschlüsselung)**.** Wir beschreiben nun, wie Alice eine RSA-codierte Nachricht an Bob übermittelt.

- Der Empfänger Bob macht für alle, die ihm eine Nachricht schicken wollen, einen öffentlichen Schlüssel verfügbar: das sind bei RSA
  (i) ein Modulus n, genannt der **RSA-Modul**, und
  (ii) ein **Verschlüsselungsexponent** e (von *encrypt*, engl. für *verschlüsseln*).
  Dabei gilt, und das ist nicht öffentlich,
  $$n = pq$$
  ist Produkt zweier Primzahlen $p \neq q$, und e ist teilerfremd zu $\varphi(n) = (p-1)(q-1)$.
- Alice codiert ihren Text auf einfache Weise als Ziffernfolge, z.B. durch eine öffentliche umkehrbare Funktion von Buchstaben zu Zahlen $1, \ldots, 26$ (bei Sonderzeichen entsprechend mehr). Dann unterteilt sie die Ziffernfolge in Blöcke, so daß der maximal mögliche Wert in einem Block $< n$ ist. Diese Ziffernfolge mit Blockstruktur nennen wir den **Klartext**.
  Sei B ein Block des Klartextes. Dann übermittelt Alice den Wert C, bestimmt als
  $$C \equiv B^e \pmod{n}$$
  und zwar den eindeutigen Repräsentanten $0 \leq C < n$. Die Ziffernfolge der C-Blöcke nennen wir den **Geheimtext**.
- Der Geheimtext kommt bei Bob an, der in der Zwischenzeit nicht untätig war. Mit dem erweiterten Euklidischen Algorithmus, das ist die Version, die es erlaubt, lineare Kongruenz-Gleichungen zu lösen, berechnet Bob den **Entschlüsselungsexponenten** d (von *decrypt*, engl. für *entschlüsseln*), der durch
  $$de \equiv 1 \pmod{\varphi(n)}$$
  bestimmt ist. Bob kennt ja nicht nur n, sondern auch p, q, die in der Faktorisierung $n = pq$ auftreten. Also kennt Bob
  $$\varphi(n) = (p-1)(q-1)$$
  und kann d leicht bestimmen. Auf die Geheimtextblöcke C wendet Bob nun
  $$B' \equiv C^d \pmod{n}$$
  an und bestimmt dabei wieder für B' den Repräsentanten im Bereich $0 \leq B' < n$.
- Es folgt fast unmittelbar aus dem Satz von Euler–Fermat, Satz 10.8, daß
  $$B' \equiv C^d \equiv (B^d)^e \equiv B^{de} \equiv B \pmod{n}$$
  ist. Damit gilt aufgrund der Wahl der minimal nicht-negativen Repräsentanten:
  $$B' = B,$$

und Bob kann seine Nachricht lesen.

Leider kann man ja nicht wissen, ob $(B, n) = 1$ gilt, also die Voraussetzungen von Satz 10.8 gelten. Weil n quadratfrei ist, brauchen wir uns darum aber nach Lemma 11.2 nicht sorgen.

Lemma 11.2. *Sei n quadratfrei. Dann gilt für alle $a \in \mathbb{Z}$ und alle $m \equiv 1 \pmod{\varphi(n)}$:*

$$a^m \equiv a \pmod{n}.$$

Beweis. Wenn $(a, n) = 1$ ist alles klar nach Satz 10.8. Der Gehalt des Lemmas betrifft die Fälle mit $(a, n) \neq 1$. Nach dem Chinesischen Restsatz reicht es aus, wenn für alle $p \mid n$ gilt

$$a^m \equiv a \pmod{p}.$$

Wenn $p \mid a$, dann steht da $0 \equiv 0$. Wenn $p \nmid a$, schreiben wir $m = 1 + k\varphi(n) = 1 + (p-1)k\varphi(n/p)$ und

$$a^m = a^{1+(p-1)k\varphi(n/p)} \equiv a \cdot (a^{p-1})^{k\varphi(n/p)} \equiv a \pmod{p}. \qquad \square$$

Bemerkung 11.3 (Praktikabilität). Das Protokoll RSA ist auf einfachen Computern implementierbar. Man braucht weder viel Speicherplatz noch große Rechenpower. Verschlüsseln und Entschlüsseln beruhen auf der Berechnung von Potenzen $\pmod{n}$, und das geht wie folgt sehr schnell. Wir schreiben den Exponenten e in Binärschreibweise

$$e = a_0 + a_1 \cdot 2 + \ldots a_r \cdot 2^r$$

mit $a_i \in \{0, 1\}$. Dann berechnen wir durch iteriertes Quadrieren $B_0 = B$ und $B_{i+1} \equiv B_i^2 \pmod{n}$ die Werte

$$B_i \equiv B^{2^i} \pmod{n}.$$

Wenn $I = \{i \ ; \ a_i = 1\}$, dann ergibt sich

$$B^e \equiv \prod_{i \in I} B_i \pmod{n},$$

worin auch nach jeder partiellen Multiplikation modulo n reduziert wird.

Die ganze Zeit bleiben die zu speichernden Zahlen im Bereich $< n$, und es werden maximal $1 + 2\log_2 n$-viele Multiplikationen solcher Zahlen mit anschließender Reduktion modulo n benötigt. Das geht insgesamt polynomial in $\log(n)$ und das ist billig.

Bemerkung 11.4 (Sicherheit). Unter Sicherheit des RSA-Protokolls versteht man, daß niemand in vernünftiger Zeit ohne den Entschlüsselungsexponenten aus dem Geheimtext auf den Klartext schließen kann. Dabei muß man berücksichtigen, daß jeder Angreifer Paare aus Klartext und passendem Geheimtext zum Üben selbst herstellen kann, denn der Verschlüsselungsexponent ist ja öffentlich. Das darf also nicht helfen.

Wenn man glaubt, daß eine RSA-verschlüsselte Botschaft nicht von einem Angreifer entschlüsselt werden kann, geht man die Wette ein, daß es der Zahl n nicht anzusehen ist, wie sie in $p \cdot q$ zu faktorisieren ist. Genauer wettet man auch noch darauf, daß die Kenntnis des Exponenten e, der ja teilerfremd zu $\varphi(n)$ sein muß und daher etwas über p und q weiß, beim Faktorisieren von n nicht hilft.

Die aktuell (2017) empfohlene Schlüssellänge beträgt 2048 Bit, also $\log_2(n) \approx 2048$. in der Praxis wird RSA in der reinen hier besprochenen Form laut Wikipedia nicht angewandt.

Das folgende Lemma ist mathematisch unsauber, denn wir sagen nicht, was *praktisch äquivalent* bedeutet. Man müßte alle Rechenoperationen bewerten und so einem Verfahren, das aus Anfangsdaten A die Enddaten B berechnet, einen Begriff der Kosten zuordnen. Dann einigt man sich noch, wieviel etwas kosten darf, um als praktisch berechenbar zu gelten. Dann erst hat das Lemma einen Sinn.

Lemma 11.5. *Sei n das Produkt zweier Primzahlen p und q. Die Faktorisierung $n = pq$ zu kennen ist praktisch äquivalent zur Kenntnis von $\varphi(n)$.*

Beweis. Wenn man $n = pq$ kennt, dann ist $\varphi(n) = (p-1)(q-1)$, und Multiplizieren selbst großer Zahlen gilt als schnell. Sei umgekehrt n und $\varphi(n)$ bekannt. Dann kennt man

$$p + q = n + 1 - \varphi(n).$$

Aus $p + q$ und pq kann man mittels

$$p, q = \frac{p+q}{2} \pm \sqrt{\left(\frac{p+q}{2}\right)^2 - pq}$$

die Werte von p und q berechnen. Das geht auch schnell. □

11.2. **Diffie-Hellman Schlüsselaustausch und der diskrete Logarithmus.** Bei der verschlüsselten Kommunikation über einen unsicheren Kanal bei Anwendung eines symmetrischen Kryptosystems, in dem also die beiden Partner Alice und Bob mit demselben Verschlüsselungsverfahren arbeiten, ist man darauf angewiesen, ein gemeinsames Geheimnis der Gesprächspartner zu generieren. Da dieser Geheimnisaustausch über den einzig verfügbaren unsicheren Kanal stattfinden muß, ist das besondere Problem dabei, der lauschenden Eve durch die übermittelten Informationen keinen Anhaltspunkt über das Geheimnis zu verraten.

$$\begin{tikzcd} & \boxed{\text{Eve}} \arrow[d, "\text{lauscht}"] & \\ \boxed{\text{Alice}} \arrow[rr, leftrightarrow] & & \boxed{\text{Bob}} \end{tikzcd}$$

Der Diffie–Hellmann Schlüsselaustausch basierend auf dem diskreten Logarithmus bietet die folgende Lösung an.

- **Alice** wählt eine große Primzahl p und eine Primitivwurzel w für $(\mathbb{Z}/p\mathbb{Z})^\times$. Diese Daten
$$p, w$$
werden veröffentlicht, für alle sichtbar an die Pinnwand gehängt.
- Jetzt überlegt sich **Alice** ein eigenes Geheimnis $\alpha \in \mathbb{Z}/(p-1)\mathbb{Z}$, berechnet
$$a = \exp_w(\alpha) = w^\alpha \in (\mathbb{Z}/p\mathbb{Z})^\times$$
und übermittelt dies an Bob.
- **Bob** überlegt sich seinerseits ein eigenes Geheimnis $\beta \in \mathbb{Z}/(p-1)\mathbb{Z}$, berechnet
$$b = \exp_w(\beta) = w^\beta \in (\mathbb{Z}/p\mathbb{Z})^\times$$
und übermittelt dies an Alice.
- Das gemeinsame Geheimnis berechnen Alice und Bob in $(\mathbb{Z}/p\mathbb{Z})^\times$ als
$$b^\alpha = (w^\beta)^\alpha = w^{\alpha\beta} = (w^\alpha)^\beta = a^\beta,$$
und zwar kann Alice b^α berechnen, denn Alice kennt b und α, und Bob kennt a und β und berechnet damit a^β.
- **Eve** hat nun a, b sowie w und p zur Verfügung. Die Wette auf die Sicherheit des gemeinsamen Geheimnisses von Alice und Bob besteht in dem Vertrauen, daß man aus a und w nicht in vernünftiger Zeit α berechnen kann. Dieses Problem nennt man das Problem des **diskreten Logarithmus**, denn die Notation
$$a = \exp_w(\alpha)$$
mit dem Isomorphismus $\exp_w : \mathbb{Z}/(p-1)\mathbb{Z} \xrightarrow{\sim} (\mathbb{Z}/p\mathbb{Z})^\times$ schlägt für den dazu inversen Isomorphismus die Notation
$$\log_w : (\mathbb{Z}/p\mathbb{Z})^\times \xrightarrow{\sim} \mathbb{Z}/(p-1)\mathbb{Z}.$$
Die Lauschaufgabe, die Eve zu bewältigen hat, verlangt die Berechnung
$$\alpha = \log_w(a)$$

des diskreten Logarithmus zur Basis w.

Bemerkung 11.6. RSA und Diffie-Hellmann mit dem diskreten Logarithmus in $(\mathbb{Z}/p\mathbb{Z})^\times$ benutzen im Wesentlichen die **multiplikative algebraische Gruppe** $\mathbb{G}_{\mathrm{m}}$. Für jeden Ring R liefert diese eine Gruppe

$$\mathbb{G}_{\mathrm{m}}(R) = R^\times$$

und das „algebraisch" bezieht sich darauf, daß die Formeln für Multiplikation und das Inverse durch algebraische Formeln gegeben werden. Die Multiplikation kann damit schnell berechnet werden, sofern wir in R schnell rechnen können. Der diskrete Logarithmus hingegen ist im algemeinen schwer zu berechnen.

Wenn man die algebraische Gruppe wechselt, etwa zu einer elliptischen Kurve, dann bekommt man Varianten der Protokolle RSA und Diffie-Hellmann.

11.3. **Primzahltests.** Um RSA-Parameter zu generieren oder um einen Diffie-Hellman Schlüsselaustausch zu vollziehen, muß man große Primzahlen erzeugen können. Einen Test, der entscheidet, ob eine Zahl n Primzahl ist, nennt man einen **Primzahltest**.

Die aktuell schnellsten Primzahltests sind probabilistisch. Diese Tests können in vertretbarer Laufzeit nur mit verschwindend geringer Fehlerwahrscheinlichkeit entscheiden, ob n Primzahl ist. Es gibt zwar mit dem AKS-Primzahltest[25] mittlerweile einen Test, der deterministisch (n ist dann beweisbar Primzahl) in einer Zeit, die polynomial in $\log(n)$ ist, arbeitet, aber dieser ist immer noch langsamer als die probabilistischen Tests.

Lemma 11.7. *Seien $n, m \in \mathbb{N}$. Dann gilt*

$$\#\{x \in \mathbb{Z}/n\mathbb{Z}\ ;\ mx = 0\} = \mathrm{ggT}(n, m).$$

Beweis. Sei $d = (n, m)$ und $n = d\nu$ und $m = d\mu$. Dann ist $(\nu, \mu) = 1$ und es gilt für $a \in \mathbb{Z}$, daß

$$n \mid mx \iff \nu \mid \mu x \iff \nu \mid x.$$

Daher haben wir

$$\{x \in \mathbb{Z}/n\mathbb{Z}\ ;\ mx = 0\} = \{[0], [\nu], [2\nu], \ldots, [(d-1)\nu]\}$$

und das sind d Elemente. □

Lemma 11.8. *Sei A eine abelsche Gruppe, und sei $x \in A$ ein Element der Ordnung 2^t für ein $t \geq 1$. Dann ist* $\mathrm{ord}(x^2) = 2^{t-1}$.

Beweis. Sei $y = x^2$. Es gilt $y^n = 1$ genau dann, wenn $x^{2n} = 1$. Das ist äquivalent zu $2^t \mid 2n$ und damit zu $2^{t-1} \mid n$. Somit folgt die Behauptung. □

Nun können wir die zahlentheoretische Aussage, die dem Primzahltest nach Miller und Rabin zugrunde liegt, beweisen.

Satz 11.9 (Miller–Rabin–Primzahltest)**.** *Sei $n \geq 3$ eine ungerade natürliche Zahl und $n - 1 = 2^r m$ mit m ungerade. Dann ist n eine Primzahl genau dann, wenn für alle $0 < a < n$ gilt:*

$$a^m \equiv 1 \pmod{n} \quad \text{oder es gibt } 0 \leq s \leq r-1 \text{ mit} \quad a^{2^s m} \equiv -1 \pmod{n}.$$

Genauer: wenn n keine Primzahl ist und $n \neq 9$, dann erfüllen höchstens $1/4$ aller $0 < a < n$ diese Bedingungen.

Beweis. Schritt 1: Wenn n Primzahl ist, dann ist $(\mathbb{Z}/n\mathbb{Z})^\times$ nach Theorem 10.23 zyklisch von Ordnung $n - 1$. Demnach gilt für $b = a^m$

$$b^{2^r} = a^{n-1} \equiv 1 \pmod{n}$$

und damit $\mathrm{ord}(b) \mid 2^r$. Es gibt also ein $t \in \mathbb{N}_0$ mit $\mathrm{ord}(b) = 2^t$. Wenn $\mathrm{ord}(b) = 1$, dann ist $a^m \equiv 1 \pmod{n}$. Wenn $\mathrm{ord}(b) > 1$, dann gibt es im Wesentlichen nach Lemma 11.8 ein $0 \leq s \leq r - 1$

[25] Manindra Agrawal, Neeraj Kayal und Nitin Saxena, 2002.

mit $\operatorname{ord}(b^{2^s}) = 2$. In einer zyklischen Gruppe gerader Ordnung gibt es genau ein Element der Ordnung 2, und hier ist das $-1 \in (\mathbb{Z}/n\mathbb{Z})^\times$. Daher gilt dann

$$a^{2^s m} = b^{2^s} \equiv -1 \pmod{n}.$$

Sei nun n keine Primzahl. Wenn $p \mid (a, n)$ für eine Primzahl p, dann ist stets $a^{2^s m} \equiv 0 \not\equiv \pm 1 \pmod{p}$. Dies beweist schon das Primzahlkriterium.

Schritt 2: Wir wollen aber auch die genauere Aussage beweisen. Dazu betrachten wir die Menge U der unbrauchbaren a (unbrauchbar, um damit „n ist nicht Primzahl“ zu beweisen). Sei für $0 \le s \le r-1$

$$U_s = \begin{cases} \{a \in (\mathbb{Z}/n\mathbb{Z})^\times \ ; \ a^{2^s m} = [-1]\} & 0 \le s \le r-1, \\ \{a \in (\mathbb{Z}/n\mathbb{Z})^\times \ ; \ a^m = [1]\} & s = -1. \end{cases}$$

Die Mengen U_s sind disjunkt, weil aus $s < t$ und $a \in U_s \cap U_t$ der Widerspruch

$$[1] = [(-1)^{2^{t-s}}] = [-1]^{2^{t-s}} = (a^{2^s m})^{2^{t-s}} = a^{2^t m} = [-1]$$

folgt. Es ist $U = \bigcup_{s=-1}^{r-1} U_s$ die disjunkte Vereinigung der U_s.

Sei $\varphi_s : (\mathbb{Z}/n\mathbb{Z})^\times \to (\mathbb{Z}/n\mathbb{Z})^\times$ die Multiplikation mit $2^s m$ (im Sinne der multiplikativen Gruppenverknüpfung), also auf Vertretern das Potenzieren

$$\varphi_s([x]) = [x^{2^s m}].$$

Wir setzen für $0 \le s \le r-1$

$$H_s = \{x \in (\mathbb{Z}/n\mathbb{Z})^\times \ ; \ x^{2^s m} = [1]\} = \ker(\varphi_s)$$

und $H_{-1} := H_0 = U_{-1}$. Die Menge U_s ist eine Nebenklasse nach H_s oder leer. Wir bestimmen nun, wann welcher Fall eintritt.

Schritt 3: Sei $n = \prod_{i=1}^t p_i^{e_i}$, und sei $p_i - 1 = 2^{r_i} m_i$ mit ungeradem m_i. Dann gibt es nach dem Chinesischen Restsatz, speziell Korollar 10.5, und Theorem 10.32 einen Isomorphismus

$$\begin{aligned} (\mathbb{Z}/n\mathbb{Z})^\times &\simeq \prod_{i=1}^t (\mathbb{Z}/p_i^{e_i}\mathbb{Z})^\times \simeq \prod_{i=1}^t \mathbb{Z}/(p_i-1)p_i^{e_i-1}\mathbb{Z} \\ &\simeq \prod_{i=1}^t (\mathbb{Z}/2^{r_i}\mathbb{Z} \times \mathbb{Z}/m_i p_i^{e_i-1}\mathbb{Z}). \end{aligned} \tag{11.1}$$

Unter diesem Isomorphismus geht die Klasse $[-1]$ in jedem Faktor auf die einzige nichttriviale Klasse der Ordnung 2, das ist

$$([2^{r_i-1}], 0) \in \mathbb{Z}/2^{r_i}\mathbb{Z} \times \mathbb{Z}/m_i p_i^{e_i-1}\mathbb{Z}.$$

Sei $\rho = \min_{1 \le i \le t}\{r_i\}$. Es folgt für $0 \le s \le r-1$

$$\#U_s = \begin{cases} \#H_s & s \le \rho - 1 \\ 0 & s \ge \rho. \end{cases}$$

Schritt 4: Der Homomorphismus φ_s ist unter dem Isomorphismus (11.1) komponentenweise die Multiplikation mit $2^s m$. Aus Lemma 11.7 schließen wir

$$\#H_s = \prod_{i=1}^t \operatorname{ggT}(2^s m, (p_i-1)p_i^{r_i-1}) = \prod_{i=1}^t \operatorname{ggT}(2^s m, p_i - 1),$$

weil $p_i \nmid m \mid n-1$ für all $i = 1, \dots, t$, denn $p_i \mid n$. Wir setzen $R = \sum_{i=1}^t r_i$ und rechnen für $s \leq \rho - 1$

$$\begin{aligned}\frac{\#H_s}{\#(\mathbb{Z}/n\mathbb{Z})^\times} &= \prod_{i=1}^t \frac{\mathrm{ggT}(2^s m, p_i - 1)}{(p_i-1)p_i^{e_i-1}} = \prod_{i=1}^t \frac{\mathrm{ggT}(2^s m, 2^{r_i} m_i)}{2^{r_i} m_i p_i^{e_i-1}} = \prod_{i=1}^t \frac{\mathrm{ggT}(m, m_i)}{2^{r_i - s} m_i p_i^{e_i-1}} \\ &= \frac{1}{2^{R-st}} \cdot \prod_{i=1}^t \frac{\mathrm{ggT}(m, m_i)}{m_i p_i^{e_i-1}} \leq \frac{1}{2^{R-st}}. \end{aligned} \tag{11.2}$$

Damit können wir den Anteil der unbrauchbaren Restklassen abschätzen (indem wir die Summe der reziproken 2-er Potenzen auffüllen) zu

$$\frac{\#U}{\#(\mathbb{Z}/n\mathbb{Z})^\times} = \sum_{s=-1}^{\rho-1} \frac{\#U_s}{\#(\mathbb{Z}/n\mathbb{Z})^\times} = \sum_{s=-1}^{\rho-1} \frac{\#H_s}{\#(\mathbb{Z}/n\mathbb{Z})^\times} = \left(\frac{1}{2^R} + \sum_{s=0}^{\rho-1} \frac{1}{2^{R-ts}}\right) \cdot \prod_{i=1}^t \frac{\mathrm{ggT}(m, m_i)}{m_i p_i^{e_i-1}} \tag{11.3}$$

$$\leq \frac{1}{2^R} + \sum_{s=0}^{\rho-1} \frac{1}{2^{R-ts}} = \frac{1}{2^R} \cdot \left(1 + \sum_{s=0}^{\rho-1} 2^{ts}\right) \leq 2^{t(\rho-1)+1-R} = \frac{1}{2^{(R-t\rho)+(t-1)}}.$$

Wenn $(R - t\rho) + (t-1) \geq 2$, dann folgt aus (11.3) die Behauptung

$$\frac{\#U}{\#(\mathbb{Z}/n\mathbb{Z})^\times} \leq \frac{1}{2^{(R-t\rho)+(t-1)}} \leq \frac{1}{4}.$$

Schritt 5: Es gilt nun, die Aussage $\frac{\#U}{\#(\mathbb{Z}/n\mathbb{Z})^\times} \leq \frac{1}{4}$ im Falle von $(R - t\rho) + (t-1) < 2$ zu beweisen. Da $R \geq t\rho$ und $t \geq 1$, bleiben die Fälle

(a) $t = 1$ also $\rho = r_1$ und $R = \rho = t\rho$.
(b) $t = 2$ und $R = t\rho$.

Schritt 6: Sei $t = 1$, also $n = p_1^{e_1}$ eine ungerade Primpotenz mit $e_1 \geq 2$. Dann ist $p_1 - 1$ ein Teiler von $p^{e_1} - 1$, also $m_1 \mid m$ und damit $m_1 = \mathrm{ggT}(m_1, m)$. Daher folgt aus (11.3) unter Benutzung von $\rho = r_1 = R$:

$$\frac{\#U}{\#(\mathbb{Z}/n\mathbb{Z})^\times} = \left(\frac{1}{2^\rho} + \sum_{s=0}^{\rho-1} \frac{1}{2^{\rho-s}}\right) \frac{1}{p_1^{e_1-1}} = \frac{1}{p_1^{e_1-1}}.$$

Die einzige Ausnahme ist hier somit $n = 3^2 = 9$. (In diesem Fall ist der Anteil auch nur $1/3$).

Schritt 7: Sei nun $t = 2$ und $R = t\rho$. Dann ist $n = p_1^{e_1} p_2^{e_2}$ und $r_1 = r_2 = \rho \geq 1$. Der erste Faktor in (11.3) lautet nun

$$\frac{1}{2^{2\rho}} + \sum_{s=0}^{\rho-1} \frac{1}{2^{2\rho-2s}} = \frac{1}{2^{2\rho}} \cdot \left(1 + \sum_{s=0}^{\rho-1} 2^{2s}\right) = \frac{1}{2^{2\rho}} \cdot \left(1 + \frac{2^{2\rho}-1}{3}\right) = \frac{1 + 2^{1-2\rho}}{3} \leq \frac{1 + 2^{1-2}}{3} = \frac{1}{2}.$$

Der andere Faktor in (11.3) hat wegen $\mathrm{ggT}(m, m_i) \mid m_i$ die Form $1/M$ mit ungeradem

$$M = \prod_{i=1}^2 \frac{m_i}{\mathrm{ggT}(m, m_i)} \cdot p_i^{e_i-1} \in \mathbb{Z}.$$

Wenn $M > 1$, dann zeigt dies $\frac{\#U}{\#(\mathbb{Z}/n\mathbb{Z})^\times} \leq \frac{1}{4}$ und wir sind fertig. Wir nehmen daher $e_1 = e_2 = 1$ und $m_i \mid m$ für $i = 1, 2$ an. Dann gilt

$$1 \equiv 2^r m + 1 = n = p_1 p_2 = (2^\rho m_1 + 1)(2^\rho m_2 + 1) \equiv 2^\rho m_2 + 1 \pmod{m_1}$$

und weil m_1 ungerade ist:

$$m_1 \mid m_2.$$

Das gilt symmetrisch genauso: $m_2 \mid m_1$. Daher ist $m_1 = m_2$ und liefert den Widerspruch

$$p_1 = 2^\rho m_1 + 1 = 2^\rho m_2 + 1 = p_2. \qquad \square$$

Bemerkung 11.10. Die Notation sei wie in Satz 11.9. Testet man eine natürliche Zahl n bezüglich einem $0 < a < n$

$$a^m \equiv 1 \pmod{n}$$

oder es gibt $0 \le s \le r-1$ mit

$$a^{2^s m} \equiv -1 \pmod{n},$$

dann erhält man

(a) Negativ — Das geht nicht: die Zahl n ist bewiesenermaßen nicht Primzahl.
(b) Positiv — Das geht: es bleibt eine Fehlerwahrscheinlichkeit von $\le 1/4$, wenn wir behaupten, daß n eine Primzahl ist.

Wenn man nun annimmt, daß man k-Iterationen des Tests mit **unabhängigen** a (was auch immer das bedeutet) macht und alle gehen positiv aus, dann bleibt eine Restfehlerwahrscheinlichkeit von $\le 1/4^k$, daß n entgegen des Ausgangs des Tests doch keine Primzahl ist.

Satz 11.11 (Agrawal, Kayal und Saxena)**.** *Sei $n \in \mathbb{N}$. Sei r eine Primzahl mit $r \nmid n$ und $\ell \mid r-1$ ein Primfaktor, so daß*

(i) $\ell > A := \lceil 2\sqrt{r}\log_2(n)\rceil$,
(ii) $n^{\frac{r-1}{\ell}} \not\equiv 1 \pmod{r}$,
(iii) $n > A$.

Dann ist n eine Primzahl genau dann, wenn

(a) Für alle ganzen Zahlen $0 \le a < A$ gilt

$$(X+a)^n \equiv X^n + a \pmod{(n, X^r - 1)}.$$

(b) n hat keine Primfaktoren $\le A$.
(c) n ist keine Potenz: für alle $2 \le k \le \log_2(n)$ ist $\sqrt[k]{n} \notin \mathbb{Z}$.

Beweis. Schritt 0: Wenn n eine Primzahl ist und (i)-(iii) gelten, dann folgt (a)–(c) sofort. Wir zeigen jetzt die viel interessantere Umkehrung.

Schritt 1: Sei p ein Primteiler von n. Wegen (ii) können wir p so wählen, daß

$$p^{\frac{r-1}{\ell}} \not\equiv 1 \pmod{r}. \tag{11.4}$$

Sei $d = \operatorname{ord}(p)$ als Element in $\mathbb{F}_r^\times$. Es gilt $d \mid r-1$ nach dem kleinen Satz des Fermat. Wenn $\ell \nmid d$, dann ist bereits $d \mid \frac{r-1}{\ell}$ im Widerspruch zu (11.4). Daher folgt $\ell \mid d$ und somit

$$d \ge \ell > A.$$

Schritt 2: Da $p \nmid r$, ist das Polynom $X^r - 1$ separabel und $X^r - 1$ hat als Nullstellen die Gruppe μ_r der r-ten Einheitswurzeln in $\overline{\mathbb{F}}_p$. Die irreduziblen Faktoren von $X^r - 1$ entsprechen den Galois-Orbits der Galoiswirkung auf μ_r. Die Galoisgruppe wird vom Frobenius $x \mapsto x^p$ erzeugt, also entspricht der Untergruppe

$$G = \langle p \rangle \subseteq \operatorname{Aut}(\mu_r) = \mathbb{F}_r^\times.$$

Es gilt $\#G = d$. Sei $\zeta \in \mu_r$ eine primitive r-te Einheitswurzel und $\exp_\zeta : \mathbb{Z}/r\mathbb{Z} \simeq \mu_r$. Dann sind die Galoisorbits der Form $\{e \cdot p^s \ ; \ 1 \le s \le d\}$ für $e \not\equiv 0 \pmod{r}$, oder der 1-elementige Orbit $\{0\}$. Also hat $X^r - 1$ einen Linearfaktor mit Wurzel 1, also $X - 1$, und ansonsten $m = \frac{r-1}{d}$-viele Faktoren vom Grad d. Es gilt

$$R := \mathbb{F}_p[X]/(X^r - 1) \simeq \mathbb{F}_p \times \prod_{i=1}^{m} \mathbb{F}_{p^d}.$$

Schritt 3: Wir betrachten zu $N \in \mathbb{N}$ nun die folgenden Familien λ_N und σ_N von Gruppenendomorphismen $R^\times \to R^\times$, die für $f = f(X) \in R^\times$ definiert sind durch:

$$\lambda_N(f) = f^N,$$
$$\sigma_N(f(X)) = f(X^N).$$

Die Abbildung σ_N ist die Einschränkung eines Ringendomorphismus $f(X) \mapsto f(X^N)$, und dieser ist wohldefiniert, weil

$$X^r - 1 \mid (X^N)^r - 1.$$

Dann gilt für alle $N, M \in \mathbb{N}$, wie man sofort nachrechnet,

$$\lambda_{NM} = \lambda_N \circ \lambda_M \quad \text{und} \quad \sigma_{NM} = \sigma_N \circ \sigma_M.$$

Schritt 4: Für alle $0 \le a < A$ ist nach Schritt 2 der Linearfaktor $X + a$ in $\mathbb{F}[X]$ teilerfremd zu $X^r - 1$. Andernfalls wäre $-a$ eine Nullstelle von $X^r - 1$ in $\mathbb{F}_p$, und in Schritt 2 haben wir gesehen, daß es nur die Nullstelle 1 gibt. Wenn $1 \equiv -a \pmod{p}$, dann ist $p \mid a + 1$ und somit $p \le a + 1 \le A$ im Widerspruch zur Voraussetzung. Es folgt

$$X + a \in R^\times$$

für alle $0 \le a < A$.

Schritt 5: Wir setzen weiter

$$U = \langle X + a \ ; \ 0 \le a < A \rangle \subseteq R^\times$$

und

$$\begin{aligned} K &= \{k \in \mathbb{N} \ ; \ \lambda_k(u) = \sigma_k(u) \text{ für alle } u \in U\} \\ &= \{k \in \mathbb{N} \ ; \ (X + a)^k \equiv X^k + a \pmod{(n, X^r - 1)} \text{ für alle } 0 \le a < A\}. \end{aligned}$$

Offensichtlich ist K multiplikativ abgeschlossen und nach Voraussetzung ist

$$n \in K.$$

Es ist $\lambda_p = \sigma_p$ gerade der Frobenius-Automorphismus von R eingeschränkt auf $R^\times$. Daher gilt auch

$$p \in K.$$

Schritt 6: Wir fixieren einen Faktor $\pi : R \twoheadrightarrow \mathbb{F}_{p^d} = \mathbb{F}_p[X]/h(X)$, der zum irreduziblen Faktor $h(X) \mid X^r - 1$ vom Grad d gehört. Zu einer Teilmenge $B \subseteq [0, A)$ setzen wir

$$P_B(X) := \prod_{b \in B} (X - [b]) \in \mathbb{F}_p[X].$$

Angenommen es gibt disjunkte Teilmengen $B \coprod C \subseteq [0, A)$, so daß

$$P_B(X) \equiv P_C(X) \pmod{(p, h(X))}.$$

Wegen Division mit Rest durch $h(X)$ im Polynomring $\mathbb{F}_p[X]$ und

$$\deg(h) = d > A \ge \max\{\deg(P_B), \deg(P_C)\}$$

gilt bereits $P_B(X) = P_C(X)$ in $\mathbb{F}_p[X]$. Damit haben die Polynome P_B und P_C die gleichen Nullstellen in $\mathbb{F}_p$, das sind gerade die Bilder von B bzw. C in $\mathbb{F}_p$. Weil p als Primteiler von n nach Voraussetzung $> A$ ist, sind die Linearfaktoren $X - [a]$ zu $0 \le a < A$ in $\mathbb{F}_p[X]$ paarweise verschieden. Daher gilt $B = C$.

Es folgt, daß das Bild $\pi(U) \subseteq \mathbb{F}_{p^d}^\times$ eine Gruppe mit mindestens

$$\#\pi(U) \ge 2^A = \#\{\pi(P_B(X)) \ ; \ B \subseteq [0, A)\} > 2^{2\sqrt{r}\log_2(n)} = n^{2\sqrt{r}}$$

Elementen ist.

Schritt 7: Wir behaupten nun: Wenn $N, M \in K$ und $N \equiv M \pmod{r}$, dann gilt sogar $N \equiv M \pmod{\#\pi(U)}$.

Da $X^r \equiv 1$ im Ring R, ist $\sigma_N = \sigma_M$. Dann gilt für alle $f(X) \in U$

$$f(X)^N = f(X^N) = f(X^M) = f(X)^M \in R^\times.$$

Daher gilt für alle $x \in \pi(U)$

$$x^{N-M} = 1 \in \mathbb{F}_{p^d}^\times.$$

Da die Gruppe $\pi(U)$ als Untergruppe der zyklischen Gruppe $\mathbb{F}_{p^d}^\times$ zyklisch ist, folgt $\#\pi(U) \mid N - M$.

Schritt 8: Unter den Elementen $n^\alpha p^\beta \in K$ mit $0 \le \alpha, \beta < \sqrt{r}$, das sind $(1+\lfloor\sqrt{r}\rfloor)^2 > (\sqrt{r})^2 = r$ viele, gibt es nach dem Schubfachprinzip zwei verschiedene $(\alpha, \beta) \neq (\gamma, \delta)$ mit

$$n^\alpha p^\beta \equiv n^\gamma p^\delta \pmod{r}.$$

Nach Schritt 7 gilt dann sogar schon

$$n^\alpha p^\beta \equiv n^\gamma p^\delta \pmod{\#\pi(U)}.$$

Da aber beide Produkte nach Voraussetzung und Schritt 6 abgeschätzt werden können zu

$$1 \le n^\alpha p^\beta, n^\gamma p^\delta < n^{\sqrt{r}} p^{\sqrt{r}} \le n^{2\sqrt{r}} \le \#\pi(U)$$

muß sogar in $\mathbb{Z}$

$$n^\alpha p^\beta = n^\gamma p^\delta$$

gelten. Nun folgt aus der eindeutigen Primfaktorzerlegung, daß n eine Potenz von p sein muß. Nach Voraussetzung kann dies nur $n = p$ sein, und n ist Primzahl. □

Bemerkung 11.12. Man muß noch die Frage beantworten, wie man zu einem n ein passendes r und ℓ wie im AKS-Algorithmus bekommt.

Bemerkung 11.13. Die Bedeutung des AKS-Algorithmus liegt in seiner in der Stellenlänge von n polynomialen Laufzeit, was für einen deterministischen Primzahltest vorher nicht bekannt war. Genauer hat der AKS-Test die Laufzeit $\mathrm{O}((\log n)^{12+\varepsilon})$. Dies wurde von Lenstra und Pomerance auf $\mathrm{O}((\log n)^{6+\varepsilon})$ verbessert. Zwar sind probabilistische Tests bisher noch schneller, können aber nur mit einer praktisch verschwindenden Restwahrscheinlichkeit ausschließen, daß die Eingabe n nicht doch nur eine Pseudo-Primzahl für den jeweiligen Test ist. Das ist theoretisch/qualitativ ein riesiger Unterschied, der für die Praxis ignoriert werden kann und wird.

12. Quadratische Reste

12.1. **Das quadratische Reziprozitätsgesetz.** Das quadratische Reziprozitätsgesetz steht am Beginn der modernen Zahlentheorie. Vermutet von Euler, wurde das quadratische Reziprozitätsgesetz zuerst von Gauß am 8. April 1796 per Induktion bewiesen. In der Folge entwickelte Gauß alleine mindestens 8 publizierte Beweise. Auf der Website von Franz Lemmermeyer findet man mindestens 246 publizierte Beweise (in ≥ 30 Beweisfamilien).

Definition 12.1. Sei p eine Primzahl.

(1) Eine ganze Zahl $a \in \mathbb{Z}$ (oder ihre Restklasse $[a] \in \mathbb{F}_p$) heißt **quadratischer Rest** modulo p, falls $p \nmid a$ und es ein $b \in \mathbb{Z}$ gibt mit

$$[a] = [b]^2 \in \mathbb{F}_p.$$

Wenn $p \nmid a$ und a kein quadratischer Rest ist, dann heißt die Zahl a (oder ihre Restklasse $[a] \in \mathbb{F}_p$) **quadratischer Nichtrest** modulo p.

(2) Sei nun $p \neq 2$ eine ungerade Primzahl. Wir definieren das **Legendre-Symbol**

$$\left(\frac{-}{p}\right) : \mathbb{F}_p^\times \to \{\pm 1\}$$

durch

$$[a] \mapsto \left(\frac{[a]}{p}\right) := \begin{cases} 1 & \text{falls } [a] \text{ quadratischer Rest modulo } p, \\ -1 & \text{falls } [a] \text{ quadratischer Nichtrest modulo } p. \end{cases}$$

Wir treffen die Konvention, daß für eine ganze Zahl a mit Restklasse $[a] \in \mathbb{F}_p$

$$\left(\frac{a}{p}\right) := \begin{cases} \left(\frac{[a]}{p}\right) & \text{falls } p \nmid a, \\ 0 & \text{falls } p \mid a \end{cases} \in \{-1, 0, 1\}.$$

Bemerkung 12.2. (1) Alternativ ist $a \in \mathbb{Z}$ mit $p \nmid a$ ein quadratischer Rest, wenn die Kongruenzgleichung

$$X^2 \equiv a \pmod{p}$$

eine Lösung besitzt. Das Legendre-Symbol gibt demnach über die Lösbarkeit einfacher quadratischer Kongruenzgleichungen Auskunft.

(2) Das Legendre-Symbol $\left(\frac{a}{p}\right)$ hängt nur von der Restklasse von a modulo p ab. Daher ist es nicht gefährlich, nur ein Symbol zu verwenden egal ob $a \in \mathbb{Z}$ oder $a \in \mathbb{Z}/p\mathbb{Z}$ gemeint ist.

Theorem 12.3. *Es gilt das* ***quadratische Reziprozitätsgesetz****:*

(1) Seien p, $q > 2$ verschiedene ungerade Primzahlen. Dann gilt

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

(2) Erster Ergänzungssatz. Sei $p > 2$ eine ungerade Primzahl. Dann gilt

$$\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}} = \begin{cases} 1 & \textit{falls } p \equiv 1 \pmod{4}, \\ -1 & \textit{falls } p \equiv 3 \pmod{4}. \end{cases}$$

(3) Zweiter Ergänzungssatz. Sei $p > 2$ eine ungerade Primzahl. Dann gilt

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

Bemerkung 12.4. Seien p, $q > 2$ verschiedene ungerade Primzahlen. Man merke sich das quadratische Reziprozitätsgesetz in der folgenden Form:

(1) Es gilt $\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right)$ außer wenn $p \equiv q \equiv 3 \pmod{4}$. In diesem Fall gilt $\left(\frac{p}{q}\right) = -\left(\frac{q}{p}\right)$.
(2) Es ist -1 ein quadratischer Rest modulo p genau für $p \equiv 1 \pmod{4}$.
(3) Es ist 2 ein quadratischer Rest modulo p genau für $p \equiv \pm 1 \pmod{8}$.

Bemerkung 12.5. Das Besondere am Reziprozitätsgesetz besteht in den folgenden Eigenschaften:

- es verknüpft zwei *reziproke* Fragen: modulo p mit modulo q.
- es ist *globaler* Natur, d.h. es geht über Kongruenzrechnungen hinaus.
- es besitzt in der Klassenkörpertheorie eine moderne Verallgemeinerung (Poitou–Tate–Dualität).
- es ist algorithmisch extrem effizient (via Jacobi–Symbol): es ist schneller entscheidbar, ob $X^2 \equiv a \pmod{p}$ lösbar ist, als die Lösung selbst zu bestimmen.

Bemerkung 12.6. (1) Für eine ungerade Primzahl $p = 2n + 1$ sind

$$\frac{p-1}{2} = n \quad \text{und} \quad \frac{p^2-1}{8} = \frac{n(n+1)}{2}$$

ganze Zahlen. Die Exponenten im Theorem 12.3 sind also ganze Zahlen.

(2) Der erste Ergänzungssatz besagt, daß -1 ein quadratischer Rest modulo $p > 2$ genau für Primzahlen $p \equiv 1(4)$ ist. Das haben wir bereits in den Sätzen 5.35 und 5.37 bewiesen.

(3) Die beiden Ergänzungssätze stellen eine Reziprozität $p \leftrightarrow 8$ dar, die p modulo 8 mit dem Verhalten von 2 und -1 modulo p verknüpfen. Wegen der speziellen Rolle der 2 in der Arithmetik ist diese Beziehung etwas komplizierter. In der Tat ist

$$(\mathbb{Z}/8\mathbb{Z})^* = \{1, 3, 5, -1\} \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$$

und hat 3 Untergruppen von Index 2, nämlich die von 3, von 5 und die von -1 erzeugte Untergruppe. Sei p eine ungerade Primzahl. Dann gilt für $[p] \in (\mathbb{Z}/8\mathbb{Z})^\times$:

$$[p] \in \begin{cases} \{[1],[5]\} \\ \{[1],[-1]\} \\ \{[1],[3]\} \end{cases} \iff \begin{cases} \left(\frac{-1}{p}\right) = 1 \\ \left(\frac{2}{p}\right) = 1 \\ \left(\frac{-2}{p}\right) = 1 \end{cases}$$

Bevor wir einen Beweis des quadratischen Reziprozitätsgesetzes führen, betonen wir durch ein paar Beispiele seine algorithmische Effizienz.

Beispiel 12.7. Die Zahl $p = 1889$ ist eine Primzahl. Kann man die Gleichung

$$x^2 \equiv 67 \pmod{1889}$$

mit $x \in \mathbb{Z}$ lösen? Die Zahl 67 ist auch eine Primzahl. Nach dem Reziprozitätsgesetz gilt

$$\left(\frac{67}{1889}\right) = \left(\frac{1889}{67}\right) = \left(\frac{13}{67}\right) = \left(\frac{67}{13}\right) = \left(\frac{2}{13}\right) = (-1)^{\frac{13^2-1}{8}} = (-1)^{\frac{5^2-1}{8}} = -1.$$

Die Gleichung läßt sich also mit $x \in \mathbb{Z}$ nicht lösen.

Beispiel 12.8. Die Zahl $p = 1999$ ist eine Primzahl. Kann man die Gleichung

$$x^2 \equiv 83 \pmod{1999}$$

mit $x \in \mathbb{Z}$ lösen? Die Zahl 83 ist auch eine Primzahl. Nach dem Reziprozitätsgesetz gilt:

$$\left(\frac{83}{1999}\right) = -\left(\frac{1999}{83}\right) = -\left(\frac{7}{83}\right) = \left(\frac{83}{7}\right) = \left(\frac{-1}{7}\right) = (-1)^{\frac{7-1}{2}} = -1.$$

Die Gleichung läßt sich also mit $x \in \mathbb{Z}$ nicht lösen.

Bemerkung 12.9. In den obigen Beispielen hatte man Glück, und die auftretenden Reste waren wieder Primzahlen. Wenn dies nicht so ist, braucht man die Multiplikativität des Legendre-Symbols, die wir in Korollar 12.12 beweisen werden.

Proposition 12.10 (Euler)**.** *Sei p eine ungerade Primzahl und a eine ganze Zahl, dann gilt für alle $a \in \mathbb{Z}$*

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

Beweis. Wenn eine Seite 0 ist, so offensichtlich auch die andere. Daher können wir $p \nmid a$ annehmen. Nach dem kleinen Satz des Fermat gilt

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

Daher gilt

$$a^{\frac{p-1}{2}} \equiv \pm 1 \pmod{p},$$

und wir müssen nur zeigen:

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

Sei zunächst $\left(\frac{a}{p}\right) = 1$. Dann gibt es $b \in \mathbb{Z}$ mit $a \equiv b^2 \pmod{p}$ und dann gilt

$$a^{\frac{p-1}{2}} \equiv b^{p-1} \equiv 1 \pmod{p}.$$

Dies zeigt die eine Richtung. Für die andere Richtung nehmen wir an, daß $a^{\frac{p-1}{2}} \equiv 1$ modulo p. Sei w eine Primitivwurzel modulo p, also ein Erzeuger der zyklischen Gruppe $\mathbb{F}_p^\times$, der nach Satz 10.21 und Theorem 10.23 existiert. Dann ist $a \equiv w^r$ für ein r und

$$1 \equiv a^{\frac{p-1}{2}} \equiv w^{r\frac{p-1}{2}} \pmod{p}$$

zeigt, daß

$$p - 1 = \operatorname{ord}(w) \mid r\frac{p-1}{2}.$$

Daraus folgt, $r = 2s$ ist gerade. Offensichtlich ist dann $a \equiv (w^s)^2 \pmod{p}$ und $\left(\frac{a}{p}\right) = 1$. □

Bemerkung 12.11. Das Argument in Kurzform: In der zyklischen Gruppe $\mathbb{F}_p^\times = (\mathbb{Z}/p\mathbb{Z})^*$ der Ordung $p-1$ sind die 2-fachen (multiplikativ: Quadrate) genau die Elemente der Ordnung $\frac{p-1}{2}$.

Als direktes Korollar gibt es den ersten Ergänzungssatz.

Beweis des ersten Ergänzungssatzes. Da ist nichts mehr zu beweisen. Proposition 12.10 sagt das genau aus für $a = -1$. □

Korollar 12.12. *Sei $p > 2$ eine ungerade Primzahl. Dann ist das Legendre-Symbol multiplikativ:*

$$\left(\frac{-}{p}\right) : \mathbb{F}_p^\times \to \{\pm 1\}$$

ist ein surjektiver Gruppenhomomorphismus. Genauer für $a, b \in \mathbb{Z}$ gilt

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

Beweis. Die Multiplikativität der Formel $a^{\frac{p-1}{2}}$ aus Proposition 12.10 ist evident. Es bleibt die Surjektivität zu zeigen, mit anderen Worten: für jede ungerade Primzahl p gibt es quadratische Nichtreste. Wegen $x^2 = (-x)^2$ gilt

$$(\mathbb{F}_p)^2 = \{[x^2] \ ; \ x \in \mathbb{Z} \text{ und } 0 \leq x \leq \frac{p-1}{2}\},$$

und das sind höchstens $\frac{p+1}{2}$ viele Quadrate. Es bleiben mindestens $p - \frac{p+1}{2} = \frac{p-1}{2} \geq 1$ viele Nichtquadrate. □

Korollar 12.13. *Sei $p > 2$ eine ungerade Primzahl. Dann sind genau die Hälfte der Reste in $\mathbb{F}_p^\times$ Quadrate.*

Beweis. Die Menge der Quadrate $(\mathbb{F}_p^\times)^2$ ist der Kern des Homomorphismus, der durch das Legendre-Symbol gegeben ist. Das Bild des Legendre-Symbols hat Ordnung 2 nach Korollar 12.12, und das ist nach dem Homomorphiesatz der Index des Kerns. Nach dem Satz von Lagrange gilt dann

$$\frac{\#(\mathbb{F}_p^\times)^2}{\#\mathbb{F}_p^\times} = \frac{1}{(\mathbb{F}_p^\times : (\mathbb{F}_p^\times)^2)} = \frac{1}{2}.$$ □

Beispiel 12.14. Kann man die Gleichung

$$x^2 \equiv 2012 \pmod{30167}$$

mit $x \in \mathbb{Z}$ lösen? Nach dem Chinesischen Restsatz und der Primfaktorzerlegung $30167 = 97 \cdot 311$ reicht es, Lösungen zu den unabhängigen Gleichungen

$$x^2 \equiv 2012 \equiv 72 = 2 \cdot 6^2 \pmod{97}$$

und

$$x^2 \equiv 2012 \equiv 146 = 2 \cdot 73 \pmod{311}$$

zu studieren. Es gilt nun nach dem zweiten Ergänzungssatz

$$\left(\frac{2012}{97}\right) = \left(\frac{2}{97}\right) = (-1)^{\frac{97^2-1}{8}} = 1,$$

da $97 \equiv 1$ modulo 8, und die erste Gleichung ist lösbar. Weiter gilt

$$\left(\frac{2}{311}\right) = (-1)^{\frac{311^2-1}{8}} = 1,$$

da $311 \equiv -1$ modulo 8, und so

$$\left(\frac{2012}{311}\right) = \left(\frac{146}{311}\right) = \left(\frac{2}{311}\right) \cdot \left(\frac{73}{311}\right) = 1 \cdot \left(\frac{311}{73}\right) = \left(\frac{19}{73}\right) = \left(\frac{73}{19}\right) = \left(\frac{16}{19}\right) = 1.$$

Die Gleichung ist also lösbar! (Da 30167 nur von 2 Primzahlen geteilt wird, gibt es 4 verschiedene Lösungen modulo 30167.) In der Tat ist

$$7553^2 - 2012 = 1891 \cdot 30167.$$

Das rechnet einem die Computeralgebra heutzutage natürlich blitzschnell. Aber wenn man wissen will, wie der Computer das macht, dann braucht man das Reziprozitätsgesetz.

12.2. **Ein erster Beweis mittels Gauß-Lemma.** Sei $p > 2$ eine ungerade Primzahl. Die ganzen Zahlen

$$-\frac{p-1}{2}, -\frac{p-3}{2}, \ldots, -1, 0, 1, \ldots, \frac{p-3}{2}, \frac{p-1}{2}$$

sind ein Vertretersystem für $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. Dabei ist

$$(\mathbb{Z}/p\mathbb{Z})^* = \{[-\frac{p-1}{2}], \ldots, [-2], [-1], [1], [2], \ldots, [\frac{p-1}{2}]\}.$$

Sei $a \in \mathbb{Z}$. Wir wählen $\varepsilon_i \in \{-1, 0, 1\}$ und $b_i \in \mathbb{Z}$ für $1 \leq i \leq \frac{p-1}{2}$ durch

$$a \cdot i \equiv \varepsilon_i b_i \pmod{p}$$

und

$$1 \leq b_i \leq \frac{p-1}{2}.$$

Falls $p \mid a$, so gilt $\varepsilon_i = 0$, und b_i ist beliebig. Falls hingegen $p \nmid a$, dann sind die Zahlen $\varepsilon_i \in \{-1, 1\}$ und b_i eindeutig.

Lemma 12.15 (Gauß)**.** *Sei $(a, p) = 1$. Mit den obigen Bezeichnungen gilt*

$$\left(\frac{a}{p}\right) = \prod_{i=1}^{\frac{p-1}{2}} \varepsilon_i.$$

Beweis. Wir zeigen zuerst, daß die Abbildung

$$\begin{aligned} \{1, \ldots, \frac{p-1}{2}\} &\to \{1, \ldots, \frac{p-1}{2}\} \\ i &\mapsto b_i \end{aligned}$$

bijektiv ist. Beide Seiten haben $(p-1)/2$ Elemente. Für bijektiv reicht demnach injektiv. Angenommen $1 \leq i < j \leq \frac{p-1}{2}$ und $b = b_i = b_j$. Dann ist

$$(ai)^2 \equiv b_i^2 = b_j^2 \equiv (aj)^2 \pmod{p},$$

also $p \mid a^2(i-j)(i+j)$. Das geht nur mit $i = j$.

Wir berechnen nun das Legendre-Symbol mit Proposition 12.10 als

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} = \prod_{i=1}^{\frac{p-1}{2}} \frac{a\cdot i}{i} \equiv \prod_{i=1}^{\frac{p-1}{2}} \varepsilon_i \cdot \frac{b_i}{i} \equiv \prod_{i=1}^{\frac{p-1}{2}} \varepsilon_i \pmod{p}$$

und aus der Kongruenz wird eine Gleichheit, da beide Seiten in $\{-1,0,1\}$ enthalten sind. □

Beweis von Theorem 12.3. Sei $p \nmid a$. Wir setzen für $1 \le i \le \frac{p-1}{2}$ mit den obigen Notationen

$$a\cdot i = \varepsilon_i \cdot b_i + p\cdot c_i$$

mit eindeutigem $c_i \in \mathbb{Z}$. Es gilt dann wegen $0 < 2b_i/p < 1$:

$$\lfloor \frac{2ai}{p} \rfloor = \lfloor \frac{2\varepsilon_i b_i}{p} \rfloor + 2c_i = \begin{cases} 2c_i & \text{falls } \varepsilon_i = 1, \\ 2c_i - 1 & \text{falls } \varepsilon_i = -1. \end{cases}$$

Somit folgt unmittelbar

$$\varepsilon_i = (-1)^{\lfloor \frac{2ai}{p} \rfloor}$$

und wegen Lemma 12.15

$$\left(\frac{a}{p}\right) = (-1)^{\sum_{i=1}^{\frac{p-1}{2}} \lfloor \frac{2ai}{p} \rfloor}.$$

Jetzt müssen wir die 2 loswerden. Sei dazu a ungerade. Dann

$$\begin{aligned} \left(\frac{a}{p}\right) = \left(\frac{a+p}{p}\right) &= \left(\frac{2}{p}\right)\left(\frac{\frac{a+p}{2}}{p}\right) \\ &= \left(\frac{2}{p}\right)\cdot(-1)^{\sum_{i=1}^{\frac{p-1}{2}} \lfloor \frac{ai+pi}{p} \rfloor} \\ &= \left(\frac{2}{p}\right)\cdot(-1)^{\sum_{i=1}^{\frac{p-1}{2}} i + \lfloor \frac{ai}{p} \rfloor} \\ &= \left(\frac{2}{p}\right)\cdot(-1)^{\frac{p^2-1}{8}}\cdot(-1)^{\sum_{i=1}^{\frac{p-1}{2}} \lfloor \frac{ai}{p} \rfloor}. \end{aligned} \tag{12.1}$$

Für $a = 1$ folgt

$$1 = \left(\frac{1}{p}\right) = \left(\frac{2}{p}\right)\cdot(-1)^{\frac{p^2-1}{8}}\cdot(-1)^{\sum_{i=1}^{\frac{p-1}{2}} \lfloor \frac{i}{p} \rfloor} = \left(\frac{2}{p}\right)\cdot(-1)^{\frac{p^2-1}{8}}$$

und damit der zweite Ergänzungssatz. Dieser läßt dann (für a ungerade) in (12.1) weiter schließen:

$$\left(\frac{a}{p}\right) = (-1)^{\sum_{i=1}^{\frac{p-1}{2}} \lfloor \frac{ai}{p} \rfloor}.$$

Sei $q > 2$ eine andere ungerade Primzahl, dann gilt somit

$$\left(\frac{p}{q}\right)\cdot\left(\frac{q}{p}\right) = (-1)^{\sum_{i=1}^{\frac{p-1}{2}} \lfloor \frac{qi}{p} \rfloor + \sum_{j=1}^{\frac{q-1}{2}} \lfloor \frac{pj}{q} \rfloor}$$

und es bleibt zu zeigen (es wird sogar Gleichheit gelten), daß

$$\sum_{i=1}^{\frac{p-1}{2}} \lfloor \frac{qi}{p} \rfloor + \sum_{j=1}^{\frac{q-1}{2}} \lfloor \frac{pj}{q} \rfloor \equiv \frac{p-1}{2}\cdot\frac{q-1}{2} \pmod{2}.$$

Wir betrachten die Box von Paaren (i,j)

$$S = \{i \in \mathbb{Z}\ ;\ 1 \le i \le \frac{p-1}{2}\} \times \{j \in \mathbb{Z}\ ;\ 1 \le j \le \frac{q-1}{2}\}$$

und partitionieren nach

$$S_1 = \{(i,j) \in S \ ; \ qi > pj\}$$

$$S_2 = \{(i,j) \in S \ ; \ qi < pj\}$$

Dann ist S die disjunkte Summe $S = S_1 \sqcup S_2$, denn $qi = pj$ widerspricht der eindeutigen Primfaktorzerlegung.

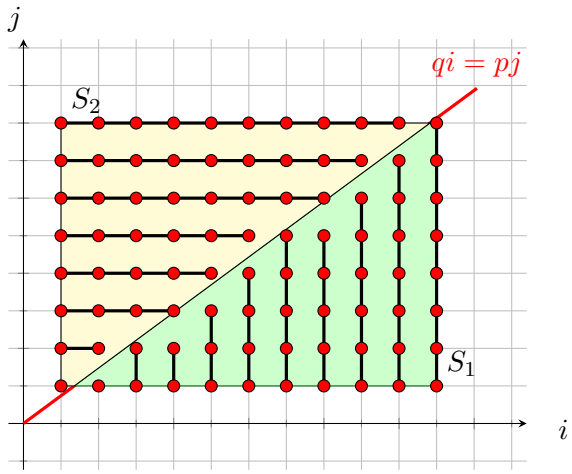


Abbildung 4. Zerlegung von S in S_1 und S_2 am Beispiel $p = 23$ und $q = 17$.

Offenbar ist

$$\#S = \frac{p-1}{2} \cdot \frac{q-1}{2}.$$

Bei festem i ist $(i,j) \in S_1$ genau dann, wenn $1 \leq j \leq \min\{\frac{q-1}{2}, \lfloor \frac{qi}{p} \rfloor\}$. Weil $i \leq \frac{p-1}{2}$, gilt

$$\lfloor \frac{qi}{p} \rfloor \leq \lfloor \frac{q(p-1)}{2p} \rfloor = \lfloor \frac{q}{2} - \frac{q}{2p} \rfloor \leq \frac{q-1}{2}.$$

Somit zählen wir S_1 als Summe über die Anzahl der zu jedem i passenden j:

$$\#S_1 = \sum_{i=1}^{\frac{p-1}{2}} \lfloor \frac{qi}{p} \rfloor.$$

Per Symmetrie gilt $\lfloor \frac{pj}{q} \rfloor = \min\{\frac{p-1}{2}, \lfloor \frac{pj}{q} \rfloor\}$ bei $j \leq \frac{q-1}{2}$, und wir zählen S_2 umgekehrt als Summe über die Anzahl der zu jedem j passenden i:

$$\#S_2 = \sum_{j=1}^{\frac{q-1}{2}} \lfloor \frac{pj}{q} \rfloor.$$

Diese Rechnung zeigt

$$\left(\frac{p}{q}\right) \cdot \left(\frac{q}{p}\right) = (-1)^{\sum_{i=1}^{\frac{p-1}{2}} \lfloor \frac{qi}{p} \rfloor + \sum_{j=1}^{\frac{q-1}{2}} \lfloor \frac{pj}{q} \rfloor} = (-1)^{\#S_1 + \#S_2} = (-1)^{\#S} = (-1)^{\frac{p-1}{2} \cdot \frac{q-1}{2}}$$

und komplettiert so den Beweis des Reziprozitätsgesetzes. $\square$

12.3. **Ein zweiter Beweis mittels Gauß-Summen.** Der zweite Beweis des quadratischen Reziprozitätsgesetzes benutzt ein wenig Körpertheorie.

Sei p eine **ungerade** Primzahl. Sei

$$\zeta = e^{2\pi/p} \in \mathbb{C}$$

die ausgezeichnete primitive komplexe p-te Einheitswurzel (jede andere tut es übrigens auch). Wir rechnen im p-ten Kreisteilungskörper

$$\mathbb{Q}(\zeta) \subseteq \mathbb{C}.$$

Genauer rechnen wir im Teilring

$$\mathbb{Z}[\zeta] \subseteq \mathbb{Q}(\zeta)$$

der ganzzahligen Linearkombinationen der ζ^i für $0 \leq i \leq p-1$.

Da $\zeta^p = 1$ gilt, ist für alle $a \in \mathbb{F}_p$ der Ausdruck

$$\zeta^a = e^{2\pi\frac{a}{p}} \in \mathbb{Q}(\zeta)$$

wohldefiniert. Außerdem gilt

$$0 = \frac{\zeta^p - 1}{\zeta - 1} = \sum_{i=0}^{p-1} \zeta^i = 1 + \sum_{a \in \mathbb{F}_p^\times} \zeta^a.$$

Lemma 12.16. *Sei $n \in \mathbb{N}$. Dann ist*

$$\sum_{a \in \mathbb{F}_p^\times} \zeta^{na} = \begin{cases} -1 & p \nmid n \\ p-1 & p \mid n. \end{cases}$$

Beweis. Wenn $p \mid n$, dann sind alle Summanden gleich 1, und davon gibt es $p-1$-viele. Wenn $p \nmid n$, dann ist Mutliplikation mit $[n]$ bijektiv auf $\mathbb{F}_p^\times$. Damit ist

$$\sum_{a \in \mathbb{F}_p^\times} \zeta^{na} = \sum_{a \in \mathbb{F}_p^\times} \zeta^a = -1. \qquad \square$$

Lemma 12.17. *Es gilt $\sum_{x \in \mathbb{F}_p^\times} \left(\frac{x}{p}\right) = 0$.*

Beweis. Nach Korollar 12.13 ist die Hälfte der Werte 1 und die andere -1. $\square$

Wir brauchen die folgende Anleihe aus der Algebra.

Lemma 12.18. *Der Ring $\mathbb{Z}[\zeta]$ hat als abelsche Gruppe eine Basis gegeben durch $1, \zeta, \zeta^2, \ldots, \zeta^{p-2}$. Insbesondere gilt*

(1) Sei q eine Primzahl. Ein Element

$$z = a_0 + a_1\zeta + \ldots + a_{p-2}\zeta^{p-2} \in \mathbb{Z}[\zeta]$$

(mit $a_i \in \mathbb{Z}$ für alle i) ist genau dann in $\mathbb{Z}[\zeta]$ durch q teilbar, wenn $q \mid a_i$ für alle i.

(2) Für $n, m \in \mathbb{Z}$ mit $n \equiv m \pmod{q\mathbb{Z}[\zeta]}$ gilt schon $n \equiv m \pmod{q}$.

(3) $(z+w)^q \equiv z^q + w^q \pmod{q\mathbb{Z}[\zeta]}$ für alle $z, w \in \mathbb{Z}[\zeta]$.

Beweis. Der Ringhomomorphismus

$$\mathbb{Z}[X] \to \mathbb{Z}[\zeta]$$

definiert durch $X \mapsto \zeta$ ist per Definition von $\mathbb{Z}[\zeta]$ surjektiv. Sei $P(X) \in \mathbb{Z}[X]$ im Kern, d.h. $P(\zeta) = 0$. Aus der Algebra wissen wir, daß das p-te Kreisteilungspolynom

$$\Phi_p(X) = X^{p-1} + \ldots + X + 1$$

irreduzibel ist und die Nullstelle ζ hat. Damit ist $P(X)$ ein Vielfaches von $\Phi_p(X)$ wenigstens im Hauptidealring $\mathbb{Q}[X]$. Sei $Q(X) \in \mathbb{Q}[X]$ mit $P(X) = \Phi_p(X)Q(X)$. Nach dem Gauß-Lemma

über die Multiplikativität des Inhalts[26] folgt, daß bereits $Q(X) \in \mathbb{Z}[X]$. Damit ist nach dem Homomorphiesatz

$$\mathbb{Z}[X]/(\Phi_p(X)) \simeq \mathbb{Z}[\zeta].$$

Damit hat $\mathbb{Z}[\zeta]$ als abelsche Gruppe die Basis $1, \zeta, \zeta^2, \ldots, \zeta^{p-2}$.

Aussage (1) folgt sofort aus der Eindeutigkeit der Darstellung von z bzw. z/q in der gegebenen Basis. Aussage (2) ist ein Spezialfall von (1).

Für (3) erinnern wir daran, daß für alle $1 \leq i \leq q$ gilt $q \mid \binom{q}{i}$. Daher folgt aus dem Binomischen Lehrsatz, daß es ein Polynom $F(A, B) \in \mathbb{Z}[A, B]$ mit ganzen Koeffizienten gibt, so daß

$$(A + B)^q = A^q + B^q + q \cdot F(A, B).$$

Spezialisieren wir $X \mapsto z$ und $B \mapsto w$, so folgt (3). □

Bemerkung 12.19. Der Ring $\mathbb{Z}[\zeta]$ hat die algebraische Struktur

$$\mathbb{Z}[X]/(X^{p-1} + \ldots + X + 1) \simeq \mathbb{Z}[\zeta],$$

wobei der Isomorphismus durch $X \mapsto \zeta$ festgelegt ist. Die Menge

$$\{\zeta^a \ ; \ a \in \mathbb{F}_p^\times\}$$

ist eine Normalbasis für die Galoiserweiterung $\mathbb{Q}(\zeta)/\mathbb{Q}$ mit Galoisgruppe $\mathbb{F}_p^\times$. Durch einen Ansatz aus der Darstellungstheorie endlicher Gruppen, der Mittelung gegen einen Charakter der Gruppe, wird die folgende Definition motiviert.

Definition 12.20. Wir definieren die **Gauß-Summe** als

$$\tau = \sum_{a \in \mathbb{F}_p^\times} \left(\frac{a}{p}\right) \zeta^a.$$

Außerdem vereinbaren wir noch die Notation

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

Satz 12.21. *Sei q eine ungerade Primzahl. Es gilt*

(1) $\tau^2 = p^*$.

(2) $\tau^q \equiv \left(\frac{q}{p}\right) \tau \pmod{q\mathbb{Z}[\zeta]}$.

Beweis. (1) Wir rechnen

$$\tau^2 = \Big(\sum_{a \in \mathbb{F}_p^\times} \left(\frac{a}{p}\right) \zeta^a\Big) \cdot \Big(\sum_{b \in \mathbb{F}_p^\times} \left(\frac{b}{p}\right) \zeta^b\Big) = \sum_{a,b \in \mathbb{F}_p^\times} \left(\frac{ab}{p}\right) \zeta^{a+b}.$$

Wir substituieren $b = xa$ und rechnen weiter

$$\begin{aligned}
&= \sum_{x \in \mathbb{F}_p^\times} \sum_{a \in \mathbb{F}_p^\times} \left(\frac{xa^2}{p}\right) \zeta^{a(1+x)} = \sum_{x \in \mathbb{F}_p^\times} \left(\frac{x}{p}\right) \sum_{a \in \mathbb{F}_p^\times} (\zeta^{(1+x)})^a \\
&= \left(\frac{-1}{p}\right) \cdot (p-1) + \sum_{x \in \mathbb{F}_p^\times, x \neq -1} \left(\frac{x}{p}\right) \cdot (-1) && \text{(Lemma 12.16)} \\
&= \left(\frac{-1}{p}\right) \cdot p - \sum_{x \in \mathbb{F}_p^\times} \left(\frac{x}{p}\right) = p^*. && \text{(Lemma 12.17)}
\end{aligned}$$

[26] Erinnerung: der Inhalt eines Polynoms in $\mathbb{Z}[X]$ ist der ggT der Koeffizienten.

(2) Wir rechnen nun

$$\begin{aligned}
\tau^q &= \left(\sum_{a\in\mathbb{F}_p^\times}\left(\frac{a}{p}\right)\zeta^a\right)^q \equiv \sum_{a\in\mathbb{F}_p^\times}\left(\frac{a}{p}\right)^q\zeta^{qa} \pmod{q\mathbb{Z}[\zeta]} && \text{(Lemma 12.18)}\\
&= \sum_{a\in\mathbb{F}_p^\times}\left(\frac{a}{p}\right)\zeta^{qa} = \left(\frac{q}{p}\right)\cdot\sum_{a\in\mathbb{F}_p^\times}\left(\frac{qa}{p}\right)\zeta^{qa}\\
&= \left(\frac{q}{p}\right)\cdot\sum_{a\in\mathbb{F}_p^\times}\left(\frac{a}{p}\right)\zeta^{a} = \left(\frac{q}{p}\right)\tau.
\end{aligned}$$

□

Beweis von Theorem 12.3. Seien p und q verschiedene ungerade Primzahlen. Dann rechnen wir in $\mathbb{Z}[\zeta]$ modulo $q\mathbb{Z}[\zeta]$

$$\begin{aligned}
(-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}\cdot\left(\frac{p}{q}\right)\cdot p^* &\equiv \left(\left(\frac{-1}{p}\right)\cdot p\right)^{\frac{q-1}{2}}\cdot p^* \pmod{q} && \text{(Korollar 12.12)}\\
&= (p^*)^{\frac{q+1}{2}} = \tau^{q+1} && \text{(Satz 12.21(1))}\\
&\equiv \left(\frac{q}{p}\right)\tau^2 \pmod{q\mathbb{Z}[\zeta]} && \text{(Satz 12.21(2))}\\
&= \left(\frac{q}{p}\right)p^*.
\end{aligned}$$

Aus Lemma 12.18 (2) schließen wir

$$(-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}\cdot\left(\frac{p}{q}\right)\cdot p^* \equiv \left(\frac{q}{p}\right)p^* \pmod{q}.$$

Weil p^* teilerfremd zu q ist, kann man kürzen zu

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

Weil beide Seiten nun in $\{-1,0,1\}$ sind, gilt sogar Gleichheit als ganze Zahlen:

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

Das war zu zeigen.

Beweis des zweiten Ergänzungssätzes: Wir rechnen im Ring $\mathbb{Z}[i] = \mathbb{Z}[X]/(X^2+1)$

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

also modulo p in $\mathbb{F}_p[X]/(X^2+1)$

$$\left(\frac{2}{p}\right) \equiv 2^{\frac{p-1}{2}} = (-i)^{\frac{p-1}{2}}\cdot(1+i)^{p-1} \equiv (-i)^{\frac{p-1}{2}}(1+i^p)/(1+i) \pmod{p}.$$

Das hängt nur noch von der Restklasse $\bar{p}\in\mathbb{Z}/8\mathbb{Z}$ ab. Die Primzahlen

$$p = 3,5,7,17$$

decken alle ungeraden Restklassen modulo 8 ab. Weiter rechnet man leicht, daß

$$\begin{aligned}
2 &\equiv 3^2 \pmod{7}\\
2 &\equiv 6^2 \pmod{17}
\end{aligned}$$

und 2 ist weder Quadrat modulo (3) noch modulo (5). Also stimmen $\left(\frac{2}{p}\right)$ und $(-1)^{\frac{p^2-1}{8}}$ für die vier ausgesuchten Primzahlen überein. Weil beide Seiten nur von $p \pmod 8$ abhängen, beweist dies den zweiten Ergänzungssatz. □

12.4. **Das Jacobi-Symbol.** Das algorithmische Hauptproblem in der Berechnung des Legendre-Symbols besteht im Faktorisierungsschritt. Bevor man $\left(\frac{a}{p}\right)$ durch Anwendung des Reziprozitätsgesetzes „umdrehen“ kann, muß man a in Primfaktoren zerlegen. Das ist algorithmisch schwierig. Das Jacobi-Symbol setzt das Legendre-Symbol fort, aber hat das Faktorisierungsproblem nicht.

Definition 12.22. Das **Jacobi-Symbol** $\left(\frac{-}{n}\right)$ für eine **ungerade** natürliche Zahl n mit Primfaktorzerlegung $n = \prod_{i=1}^r p_i^{e_i}$ ist definiert als

$$\left(\frac{a}{n}\right) = \prod_{i=1}^r \left(\frac{a}{p_i}\right)^{e_i} \in \{1, 0, -1\}$$

für alle $a \in \mathbb{Z}$. Für $n = 1$ betrachten wir das auftretende Produkt als das leere Produkt und geben den Wert 1, außer im Fall $a = 0$:

$$\left(\frac{a}{1}\right) = \begin{cases} 1 & a \neq 0 \\ 0 & a = 0. \end{cases}$$

Bemerkung 12.23. (1) Wenn a quadratischer Rest ist modulo n, dann ist a auch quadratischer Rest modulo aller Primfaktoren $p_i \mid n$. Dann ist offensichtlich $\left(\frac{a}{n}\right) = 1$.

(2) Es ist aber $\left(\frac{2}{15}\right) = \left(\frac{2}{3}\right) \cdot \left(\frac{2}{5}\right) = (-1) \cdot (-1) = 1$ obwohl 2 kein quadratischer Rest modulo 15 ist.

(3) Das Jacobi-Symbol $\left(\frac{a}{n}\right)$ hängt offensichtlich nur von $[a]_n \in \mathbb{Z}/n\mathbb{Z}$ ab.

(4) Ist $n = p$ eine Primzahl, dann stimmt das Jacobi-Symbol $\left(\frac{a}{p}\right)$ mit dem entsprechenden Legendre-Symbol überein.

Proposition 12.24. *Das Jacobi-Symbol ist multiplikativ: für alle $a, b \in \mathbb{Z}$ und alle $n \in \mathbb{N}$ gilt*

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

Insbesondere definiert das Jacobi-Symbol einen Gruppenhomomorphismus

$$(\mathbb{Z}/n\mathbb{Z})^\times \to \{\pm 1\}.$$

Beweis. Das folgt, weil $\left(\frac{-}{n}\right)$ das Produkt vollständig multiplikativer Funktionen $\left(\frac{-}{p_i}\right)$ ist. □

Theorem 12.25 (Reziprozität für das Jacobi-Symbol)**.** *Seien n und m teilerfremde ungerade natürliche Zahlen. Dann gilt:*

(1) $\left(\frac{n}{m}\right) \cdot \left(\frac{m}{n}\right) = (-1)^{\frac{n-1}{2} \cdot \frac{m-1}{2}}$,

(2) $\left(\frac{-1}{n}\right) = (-1)^{\frac{n-1}{2}}$,

(3) $\left(\frac{2}{n}\right) = (-1)^{\frac{n^2-1}{8}}$.

Beweis. (2) Das beweisen wir per Induktion über die Anzahl der Primfaktoren in n. Es reicht, für a, b ungerade zu zeigen, daß

$$\frac{a-1}{2} + \frac{b-1}{2} \equiv \frac{ab-1}{2} \pmod 2.$$

Das ist äquivalent zu $4 \mid (ab-1) - (a-1) - (b-1) = (a-1)(b-1)$ und das ist ok, weil a, b ungerade sind.

(3) Das beweisen wir wieder per Induktion über die Anzahl der Primfaktoren in n. Es reicht, für a, b ungerade zu zeigen, daß

$$\frac{a^2-1}{8} + \frac{b^2-1}{8} \equiv \frac{(ab)^2-1}{8} \pmod 2.$$

Das ist äquivalent zu $16 \mid ((ab)^2-1) - (a^2-1) - (b^2-1) = (a^2-1)(b^2-1)$ und das ist ok, weil a, b ungerade sind.

(1) Das Jacobi-Symbol $\left(\frac{m}{n}\right)$ ist per Definition multiplikativ in n und nach Proposition 12.24 auch in m. Damit ist auch $\left(\frac{n}{m}\right) \cdot \left(\frac{m}{n}\right)$ multiplikativ in n und m. Nach der Rechnung für (1) ist auch

$$(-1)^{\frac{n-1}{2} \cdot \frac{m-1}{2}}$$

multiplikativ in ungeraden n und m. Damit reicht es, die Aussage für Primzahlen zu beweisen. Wenn $n = p \neq q = m$ Primzahlen sind, dann folgt dies aus dem quadratischen Reziprozitätsgesetz, Theorem 12.3. □

Bemerkung 12.26. Die Berechnung des Jacobi-Symbols $\left(\frac{n}{m}\right)$ funktioniert nun wie folgt. Man spaltet aus n die Faktoren 2 und -1 ab. Diese werden nach den Ergänzungssätzen berechnet. Danach bleibt $\left(\frac{n}{m}\right)$ mit ungeraden n und m übrig. Man geht mit dem entsprechenden Vorzeichen zum Reziproken $\left(\frac{m}{n}\right)$ über und reduziert m modulo n per Division mit Rest analog zum Vorgehen beim Euklidischen Algorithmus. Das Verfahren besitzt eine Laufzeit, die ähnlich effizient wie die des Euklidischen Algorithmus ist!

Beispiel 12.27. Wir rechnen zwei Beispiele:

(1)

$$\left(\frac{101}{167}\right) = \left(\frac{167}{101}\right) = \left(\frac{66}{101}\right) = \left(\frac{2}{101}\right)\left(\frac{33}{101}\right) = -\left(\frac{33}{101}\right) = -\left(\frac{101}{33}\right) = -\left(\frac{2}{33}\right) = -1.$$

(2)

$$\begin{aligned}\left(\frac{137}{349}\right) &= \left(\frac{349}{137}\right) = \left(\frac{75}{137}\right) = \left(\frac{137}{75}\right) = \left(\frac{62}{75}\right) = \left(\frac{2}{75}\right)\left(\frac{31}{75}\right) = -\left(\frac{31}{75}\right) \\ &= \left(\frac{75}{31}\right) = \left(\frac{13}{31}\right) = \left(\frac{31}{13}\right) = \left(\frac{5}{13}\right) = \left(\frac{13}{5}\right) = \left(\frac{3}{5}\right) = \left(\frac{5}{3}\right) = \left(\frac{2}{3}\right) = -1.\end{aligned}$$

12.5. **Das Kronecker-Symbol.** Auch das Jacobi-Symbol läßt sich noch erweitern.

Definition 12.28. Das **Kronecker-Symbol** $\left(\frac{a}{n}\right)$ für $a, n \in \mathbb{Z}$, $n \neq 0$, mit Primfaktorzerlegung $n = \varepsilon \cdot \prod_{i=1}^r p_i^{e_i}$, hier ist $\varepsilon \in \{\pm 1\}$, ist definiert durch

$$\left(\frac{a}{n}\right) = \left(\frac{a}{\varepsilon}\right) \cdot \prod_{i=1}^r \left(\frac{a}{p_i}\right)^{e_i},$$

wobei für die ungeraden p_i mit $\left(\frac{-}{p_i}\right)$ das Legendre-Symbol gemeint ist und die weiteren Faktoren durch

$$\left(\frac{a}{1}\right) = \begin{cases} 1 & a \neq 0, \\ 0 & a = 0, \end{cases} \qquad \left(\frac{a}{-1}\right) = \begin{cases} -1 & a < 0, \\ 0 & a = 0, \\ 1 & a > 0, \end{cases} \qquad \left(\frac{a}{2}\right) = \begin{cases} -1 & a \equiv 3, 5 \pmod 8, \\ 0 & 2 \mid a, \\ 1 & a \equiv 1, 7 \pmod 8 > 0 \end{cases}$$

erklärt sind.

Bemerkung 12.29. Wir notieren offensichtliche Eigenschaften des Kronecker-Symbols.

(1) Wenn n eine ungerade natürliche Zahl ist, dann stimmt das Kronecker-Symbol $\left(\frac{-}{n}\right)$ mit dem Jacobi-Symbol überein.

(2) Das Kronecker-Symbol $\left(\frac{a}{n}\right)$ ist genau dann 0, wenn $(a,n) \neq 1$, also a und n nicht teilerfremd sind, oder im Fall $\left(\frac{0}{\pm 1}\right)$.

Proposition 12.30. *Das Kronecker-Symbol ist vollständig multiplikativ in Zähler und Nenner.*

(1) Für alle $a, b \in \mathbb{Z}$ und $n \in \mathbb{Z}$, $n \neq 0$ gilt:

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

(2) Für alle $a, b \in \mathbb{Z}$, $a \neq 0 \neq b$, und $n \in \mathbb{Z}$ gilt:

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

Beweis. (1) Das Kroneckersymbol ist Produkt vollständig multiplikativer Funktionen der Form $\left(\frac{\ }{-1}\right)$, und $\left(\frac{\ }{2}\right)$ und $\left(\frac{\ }{p}\right)$ mit ungeradem p.

Aussage (2) folgt unmittelbar aus der Definition, weil die Primfaktorzerlegung von ab das Produkt der Primfaktorzerlegungen von a und von b ist. □

Proposition 12.31. *Das Kronecker-Symbol ist modular im folgenden Sinne.*

(1) Für alle $n \neq 0$ und alle $a, b \in \mathbb{Z}$, sofern bei $n < 0$ die Zahlen a und b das gleiche Vorzeichen haben:

$$\left(\frac{a}{n}\right) = \left(\frac{b}{n}\right), \quad \textit{falls} \quad \begin{cases} a \equiv b \pmod{4|n|} & \textit{wenn } n \equiv 2 \pmod 4, \\ a \equiv b \pmod{|n|} & \textit{sonst}\ . \end{cases}$$

(2) Für alle $n, m > 0$ und alle $a \in \mathbb{Z}$, $a \neq 0$, mit $a \not\equiv 3 \pmod 4$:

$$\left(\frac{a}{n}\right) = \left(\frac{a}{m}\right), \quad \textit{falls} \quad \begin{cases} n \equiv m \pmod{4|a|} & \textit{wenn } a \equiv 2 \pmod 4, \\ n \equiv m \pmod{|a|} & \textit{sonst}\ . \end{cases}$$

Beweis. (1) folgt aus den Modularitätseigenschaften des Legendre-Symbols sowie $\left(\frac{a}{2}\right) = \left(\frac{b}{2}\right)$, wenn $a \equiv b \pmod 8$. Der extra Faktor 4 sorgt für $8 \mid n$, also $a \equiv b \pmod 8$. Wenn $v_2(n) = 2$, dann tritt der Faktor $\left(\frac{\ }{2}\right)$ quadratisch auf und hebt sich weg.

(2) Wenn $(a,n) \neq 1$, dann ist ein Primteiler von (a,n) auch Primteiler von m, somit beide Kronecker-Symbole gleich 0. Andernfalls ist (2) eine unmittelbare Folge von (1) und Reziprozität aus Theorem 12.32 (sowie einer lästigen Vorzeichenüberlegung). □

Theorem 12.32 (Reziprozität für das Kronecker-Symbol)**.** *Seien n und m teilerfremde ganze Zahlen ungleich 0. Schreiben wir $n = 2^r \cdot n'$ und $m = 2^s \cdot m'$ mit ungeraden n' und m', dann gilt:*

$$\left(\frac{n}{m}\right) \cdot \left(\frac{m}{n}\right) = \varepsilon \cdot (-1)^{\frac{n'-1}{2} \cdot \frac{m'-1}{2}}$$

mit $\varepsilon = -1$, falls $n < 0$ und $m < 0$, und $\varepsilon = 1$ sonst.

Beweis. Beide Seiten sind vollständig multiplikativ in n und m. Daher reichen die Fälle, wenn n und m Primzahlen oder ± 1 sind. Dies folgt aus dem Quadratischen Reziprozitätsgesetz Theoem 12.3 für das Legendre-Symbol, insbesondere den Ergänzungssätzen und der Definition des Kronecker-Symbols für $\left(\frac{\ }{-1}\right)$ und $\left(\frac{\ }{2}\right)$. □

Bemerkung 12.33. Zu $D \equiv 0, 1 \pmod 4$ ist $\chi(n) = \left(\frac{D}{n}\right)$ auf $n > 0$ multiplikativ, und genauer ein Dirichlet-Charakter zum Führer D. Das bedeutet, daß χ eigentlich ein Gruppenhomomorphismus

$$\chi : (\mathbb{Z}/D\mathbb{Z})^\times \to \{\pm 1\}$$

ist, der auf $\mathbb{N}$ durch 0 für nicht zu D teilerfremde Zahlen fortgesetzt wurde. Wenn D kein Quadrat in $\mathbb{Z}$ ist, dann gehört dazu nach Galoistheorie ein quadratischer Zwischenkörper $K/\mathbb{Q}$

der zyklotomischen Erweiterung $\mathbb{Q}(\zeta)/\mathbb{Q}$, wobei ζ eine primitive D-te Einheitswurzel ist. Der Zwischenkörper ist genau $K = \mathbb{Q}(\sqrt{D})$.

Dirichlet-Charaktere spielen in Kapitel §13 eine wichtige Rolle. Quadratische Zahlkörper lernen wir in Kapitel 16 kennen.

Übungsaufgaben zu §12

Übungsaufgabe 12.1. Zeigen Sie, daß die folgenden Abbildungen Gruppenhomomorphismen sind.

(1) $\varepsilon : (\mathbb{Z}/4\mathbb{Z})^\times \to \mathbb{Z}/2\mathbb{Z}$ definiert für $n \in \mathbb{Z}$ teilerfremd zu 4 durch

$$\varepsilon([n]) = \frac{n-1}{2} \pmod 2.$$

(2) $\omega : (\mathbb{Z}/8\mathbb{Z})^\times \to \mathbb{Z}/2\mathbb{Z}$ definiert für $n \in \mathbb{Z}$ teilerfremd zu 8 durch

$$\omega([n]) = \frac{n^2-1}{8} \pmod 2.$$

13. Primzahlen in arithmetischen Folgen

Wir wissen seit langem, daß es in $\mathbb{N}$ unendlich viele Primzahlen gibt. Wie geht die gleiche Frage für bestimmte unendliche Teilmengen von $\mathbb{N}$ aus?

Definition 13.1. Eine **arithmetische Folge** ist eine Folge $(a_n)_{n\in\mathbb{N}}$ der Form

$$a_n = d \cdot n + c.$$

Hier soll für die Zahlentheorie natürlich $c, d \in \mathbb{Z}$ sein.

Offensichtlich bilden die arithmetischen Folgen zu festem $d \geq 1$ und $1 \leq c \leq d$ eine Partition von $\mathbb{N}$, nämlich die Partition nach den Kongruenzklassenmodulo d, bzw. nach $[c] \in \mathbb{Z}/d\mathbb{Z}$. Da wir wissen, daß es in $\mathbb{N}$ unendlich viele Primzahlen gibt, liegt es nahe zu fragen, ob dies für Primzahlen in arithmetischen Folgen auch gilt.

Offensichtlich muß mindestens eine der Kongruenzklassen $[c]$ bei festem $d \geq 1$ unendlich viele Primzahlen enthalten. Außerdem gilt für n, daß

$$(c, d) \mid c \cdot n + d,$$

so daß nur bei teilerfremden c, d unendlich viele Primzahlen $\equiv c \pmod d$ sein können. Die Antwort auf diese Verteilungsfragen ist bemerkenswert:

Theorem 13.2 (Satz von Dirichlet über Primzahlen in arithmetischen Folgen)**.** *Seien m, a teilerfremd und $m \geq 1$. Dann gibt es unendlich viele Primzahlen p mit*

$$p \equiv a \pmod m.$$

13.1. Algebraische Beweise in Spezialfällen.

Wir haben zwei Beweise dafür gegeben, daß es unendlich viele Primzahlen gibt. Wir diskutieren nun Varianten des Euklidischen Beweises, die Teilresultate von Theorem 13.2 hergeben.

Bemerkung 13.3. Varianten des Eulerschen Beweises liefern mehr, nämlich Gleichverteilung! Für jedes feste $d \geq 2$ hat die Menge der Primzahlen in der arithmetischen Folge $(dn + c)_{n\in\mathbb{N}}$ eine Dirichlet–Dichte. Diese ist 0, wenn $(d, c) \neq 1$, und sie ist für alle anderen Fälle mit $(d, c) = 1$ gleich! Mit diesen analytischen Methoden befassen wir uns im Anschluß.

Satz 13.4. *Es gibt unendlich viele Primzahlen $p \equiv -1 \pmod 4$.*

Beweis. Angenommen $p_1, \ldots, p_s$ wäre eine endliche Liste aller Primzahlen $\equiv -1 \pmod 4$. Dann betrachten wir

$$N = 4 \prod_{i=1}^{s} p_i - 1.$$

Dann gilt $N \equiv -1 \pmod{p_i}$, also $p_i \nmid N$ für alle $i = 1, \ldots, s$ und auch $2 \nmid N$. Daher sind alle Primfaktoren von N kongruent zu 1 modulo 4. Dann muß aber N als Produkt seiner Primfaktoren auch $\equiv 1 \pmod 4$ sein. Widerspruch. □

Satz 13.5. *Es gibt unendlich viele Primzahlen* $p \equiv 1 \pmod 4$.

Beweis. Angenommen $p_1, \ldots, p_s$ wäre eine endliche Liste aller Primzahlen $\equiv 1 \pmod 4$. Dann betrachten wir

$$N = (2\prod_{i=1}^{s} p_i)^2 + 1.$$

Dann gilt $N \equiv 1 \pmod{p_i}$, also $p_i \nmid N$ für alle $i = 1, \ldots, s$ und auch $2 \nmid N$. Daher sind alle Primfaktoren von N kongruent zu -1 modulo 4.

Da $N > 1$ ist, gibt es einen Primteiler $\ell \mid N$, und $\ell \equiv -1 \pmod 4$. Dann gilt

$$(2\prod_{i=1}^{s} p_i)^2 \equiv N - 1 \equiv -1 \pmod{\ell}$$

im Widerspruch zu Satz 5.35. □

Das letzte Argument läßt sich verallgemeinern.

Satz 13.6. *Sei* ℓ *eine Primzahl. Dann gibt es unendlich viele Primzahlen* $p \equiv 1 \pmod{\ell}$.

Beweis. Angenommen $p_1, \ldots, p_s$ wäre eine endliche Liste aller Primzahlen $\equiv 1 \pmod{\ell}$. Dann betrachten wir

$$P = \ell \prod_{i=1}^{s} p_i$$

und daraus

$$N = P^{\ell-1} + P^{\ell-2} + \ldots + P + 1.$$

Sei q ein Primteiler von N. Da $(N, P) = 1$, ist $q \nmid P$ und damit q kein Primteiler von der Liste, also $q \not\equiv 1 \pmod{\ell}$, d.h.

$$\ell \nmid q - 1.$$

Weil ℓ Primzahl ist, folgt $(\ell, q-1) = 1$. Wir wählen nach Bézout eine $\mathbb{Z}$-Linearkombination

$$\ell x + (q-1)y = 1.$$

Außerdem beobachten wir, daß q ein Teiler von $N(P-1) = P^\ell - 1$ ist. Dann folgt mit Satz 5.30

$$P \equiv P^{\ell x + (q-1)y} \equiv (P^\ell)^x \cdot (P^{(q-1)})^y \equiv 1^x \cdot 1^y \equiv 1 \pmod{q}.$$

Dann aber ist

$$0 \equiv N \equiv P^{\ell-1} + P^{\ell-2} + \ldots + P + 1 \equiv 1^{\ell-1} + \ldots + 1 \equiv \ell \pmod{q}.$$

Aus $q \mid \ell$ folgt $q = \ell$, aber das geht auch nicht, da $\ell \mid P$ und $q \nmid P$. Widerspruch. □

Bemerkung 13.7. (1) Anstelle des Polynoms X^2+1 in Satz 13.5, das zu den 4. Einheitswurzeln gehört, nehmen wir in Satz 13.6 das Polynom $X^{\ell-1}+\ldots+X+1$, das zu den ℓ-ten Einheitswurzeln gehört. Dies ist ein deutliches Indiz des Zusammenhangs zwischen Primzahlen in arithmetischen Folgen und der Galoistheorie der zyklotomischen Körper.

(2) Im Beweis von Satz 13.6 haben wir implizit (x oder y ist negativ!) von der Erweiterung des Kongruenzbegriffs auf Brüche mit zum Modulus teilerfremdem Nenner Gebrauch gemacht, siehe Bemerkung 4.8.

Bemerkung 13.8. Wir schließen diesen Abschnitt mit zwei Bemerkungen zu jüngeren Entwicklungen in der additiven Zahlentheorie.

(1) Der Satz von Dirichlet gibt Auskunft über Primzahlen in arithmetischen Folgen. Man kann die Frage verdrehen und auf eine sehr spannende Frage stoßen, die erst kürzlich gelöst worden ist. Gibt es arithmetische Folgen in der Menge der Primzahlen? Ein Beispiel der Länge 5 ist:
$$5, 17, 29, 41, 53.$$
Natürlich kann eine arithmetische Folge $a_n = dn + c$, $d \neq 0$ aus Primzahlen nicht unendlich lang sein. Dann würde nämlich das lineare Polynom $P(X) = dX + c$ für alle $n \in \mathbb{N}$ nur Primzahlen als Werte annehmen im Widerspruch zu Satz 5.21.
Green und Tao haben 2004 gezeigt, daß es für jedes $k \in \mathbb{N}$ ein $d \geq 1$ und eine arithmetische Folge
$$p_0 = c,\ p_1 = d + c,\ \ldots,\ p_k = dk + c$$
von Primzahlen gibt.

(2) Die Goldbach–Vermutung (1742) besagt, daß jede gerade natürliche Zahl > 2 die Summe zweier Primzahlen ist. Diese so einfache Frage ist bis heute (April 2017) nicht entschieden. Allerdings hat Helfgott 2013 einen Beweis für die ternäre Goldbach–Vermutung gefunden: jede ungerade Zahl > 5 ist Summe dreier Primzahlen.

13.2. **Charaktere und Dualität für endliche abelsche Gruppen.** Für den Beweis von Theorem 13.2 benötigen wir algebraische und analytische Werkzeuge. Interessanterweise haben die algebraischen dabei etwas mit der diskreten Fouriertransformation zu tun, also eine Art diskrete Analysis.

Notation 13.9. Wir notieren abelsche Gruppen hier multiplikativ, da wir bereits die Gruppe $(\mathbb{Z}/m\mathbb{Z})^\times$ für Dirichlet-Charaktere als Anwendung im Blick haben.

Definition 13.10. Sei G eine endliche abelsche Gruppe. Ein **(komplexwertiger) Charakter** auf G ist ein Gruppenhomomorphismus
$$\chi : G \to \mathbb{C}^\times,$$
das heißt für alle $x, y \in G$ gilt $\chi(xy) = \chi(x)\chi(y)$.

Beispiel 13.11. Sei n ungerade. Das Jacobi-Symbol definiert einen Charakter
$$\left(\frac{-}{n}\right) : (\mathbb{Z}/n\mathbb{Z})^\times \to \{\pm 1\} \subseteq \mathbb{C}^\times, \qquad a \mapsto \left(\frac{a}{n}\right).$$

Bemerkung 13.12. Sind χ und ψ Charaktere auf G, dann ist das Produkt $\chi\psi$ definiert als der Charakter
$$(\chi\psi)(x) := \chi(x)\psi(x) \qquad \text{für alle } x \in G.$$
Das neutrale Element bezüglich der Multiplikation von Charakteren ist der triviale Charakter
$$\mathbf{1} : G \to \mathbb{C}^\times$$
mit $\mathbf{1}(x) = 1$ für alle $x \in G$. Der zum Charakter χ inverse Charakter χ^{-1} ist damit
$$\chi^{-1}(x) = \frac{1}{\chi(x)}.$$
Folglich bilden die Menge der Charaktere der endlichen abelschen Gruppe G
$$\hat{G} := \{\chi\ ;\ \chi : G \to \mathbb{C}^\times \text{ Charakter}\}$$
mit der angegebenen Multiplikation selbst eine Gruppe, die **Charaktergruppe** von G.

Bemerkung 13.13. Wir beschränken uns auf endliche abelsche Gruppen G, obwohl die Theorie auch für beliebige kompakte Gruppen funktioniert.

(1) Jeder Charakter $\chi : G \to \mathbb{C}$ auf einer endlichen Gruppe G nimmt seine Werte in den Einheitswurzeln an. Eine komplexe Zahl $z \in \mathbb{C}$ ist eine (n-te)-Einheitswurzel, wenn es ein $n \in \mathbb{N}$ gibt mit $z^n = 1$. Aus der Polarkoordinatendarstellung liest man ab, daß die n-ten Einheitswurzeln eine endliche Untergruppe von $\mathbb{C}^\times$ bilden:

$$\mu_n := \mu_n(\mathbb{C}) := \{\zeta = e^{2\pi i \frac{a}{n}} \ ; \ a = 0, 1, \ldots, n-1\}.$$

Mit $n = \#G$ gilt für jedes $x \in G$ nach dem Satz von Lagrange $x^n = 1$, also

$$\chi(x)^n = \chi(x^n) = \chi(1) = 1.$$

(2) Der zu einem Charakter $\chi : G \to \mathbb{C}^\times$ konjugierte Charakter $\overline{\chi}$ ist definiert durch

$$\overline{\chi}(x) = \overline{\chi(x)} \qquad \text{für alle } x \in G.$$

Das Inverse χ^{-1} stimmt mit dem komplex konjugierten Charakter überein, weil für alle möglichen Werte $\zeta \in \mu_n$ gilt: $\overline{\zeta} = \zeta^{-1}$.

(3) Später wird die folgende Identität im Polynomring $\mathbb{C}[T]$ nützlich sein:

$$\prod_{\zeta \in \mu_n} (1 - \zeta T) = 1 - T^n.$$

Beide Seiten sind Polynome in T vom selben Grad mit der gleichen Nullstellenmenge μ_n. Daher unterscheiden sie sich nur um eine multiplikative Konstante. Diese ist 1, wie man durch Auswertung in $T = 0$ feststellt.

Beispiel 13.14. Jeder Charakter χ auf $\mathbb{Z}/n\mathbb{Z}$ nimmt Werte in μ_n an. Die Auswertung in einem $x \in \mathbb{Z}/n\mathbb{Z}$ liefert dabei einen Gruppenhomomorphismus

$$\widehat{\mathbb{Z}/n\mathbb{Z}} \to \mu_n, \quad \chi \mapsto \chi(x)$$

allein aufgrund der Definitionen der beteiligten Gruppenstruktur. Wir werten nun in $x = [1]$ aus. Da $[1]$ ein Erzeuger ist, wird jeder Charakter χ auf $\mathbb{Z}/n\mathbb{Z}$ durch $\chi([1])$ eindeutig festgelegt, denn $\chi([a]) = \chi([1])^a$. Dies bedeutet, daß die Auswertung in $[1]$ injektiv ist.

Wir zeigen nun durch eine Konstruktion, daß Auswertung in $[1]$ sogar surjektiv und damit ein Gruppeisomorphismus ist. Zu $\zeta \in \mu_n$ ist die Abbildung

$$\exp_\zeta : \mathbb{Z}/n\mathbb{Z} \to \mathbb{C}^\times, \quad \exp_\zeta(a) = \zeta^a$$

ein Charakter mit $\exp_\zeta([1]) = \zeta$.

Jetzt betrachten wir Charaktere auf μ_n. Zu jedem $k \in \mathbb{Z}$ gibt es

$$\psi_k : \mu_n \to \mathbb{C}^\times, \quad \psi_k(z) = z^k.$$

Der Charakter ψ_k hängt von k nur modulo n ab, denn $\zeta^n = 1$ für alle $\zeta \in \mu_n$. Daher definiert $k \mapsto \psi_k$ einen Gruppenhomomorphismus

$$\mathbb{Z}/n\mathbb{Z} \to \widehat{\mu_n}.$$

Dieser ist injektiv, denn $\psi_k = \mathbf{1}$ impliziert $1 = \psi_k(e^{2\pi i/n}) = e^{2\pi i k/n}$ und damit $n \mid k$. Aus dem Folgenden schließen wir, daß es sich sogar um einen Isomorphismus handelt: jeder Charakter auf μ_n ist von der Form ψ_k für ein eindeutiges $k \in \mathbb{Z}/n\mathbb{Z}$.

Proposition 13.15 (Fortsetzbarkeit)**.** *Sei G eine endliche abelsche Gruppe. Jeder Charakter auf einer Untergruppe H ist die Einschränkung eines Charakters auf G.*

Beweis. Wir beweisen das per Induktion nach dem Index $(G : H)$. Es reicht daher für ein $x \in G$ den Charakter $\chi \in \hat{H}$ auf die Untergruppe $\langle H, x\rangle$ fortzusetzen. Sei $m > 0$ minimal mit $x^m \in H$. Ein solches m gibt es, weil $x^{\#G} = 1 \in H$. Sei $a = x^m$ und $z = \chi(a)$. In $\mathbb{C}$ kann man beliebige Wurzeln ziehen (Polarkoordinaten!). Daher gibt es ein $\zeta \in \mathbb{C}$ mit $\zeta^m = z$. Wir setzen nun χ fort durch

$$\tilde{\chi}(hx^k) = \chi(h)\zeta^k, \quad h \in H, k \in \mathbb{Z}.$$

Das ist offensichtlich eine Fortsetzung und ein Gruppenhomomorphismus $\tilde{\chi} : \langle H, x\rangle \to \mathbb{C}^\times$, sofern wir uns nur überzeugen können, daß $\tilde{\chi}$ überhaupt wohldefiniert ist.

Wenn $hx^k = h'x^l$ für $h, h' \in H$ und $k, l \in \mathbb{Z}$, dann ist

$$x^{k-l} = h'h^{-1} \in H,$$

und die Minimalität von m zeigt $m \mid k-l$. Sei $k-l = mt$, dann gilt $h' = h(h'h^{-1}) = hx^{mt} = ha^t$ und

$$\chi(h)\zeta^k = \chi(h)\zeta^{mt+l} = \chi(h)z^t\zeta^l = \chi(ha^t)\zeta^l = \chi(h')\zeta^l.$$

Damit ist $\tilde{\chi}$ wohldefiniert. □

Satz 13.16. *Sei G eine endliche abelsche Gruppe. Dann gilt*

$$\#G = \#\hat{G}.$$

Beweis. Sei H eine Untergruppe von G. Einschränkung auf H definiert einen Gruppenhomomorphismus

$$\mathrm{res} : \hat{G} \to \hat{H}, \qquad \mathrm{res}(\chi) = \chi|_H.$$

Proposition 13.15 besagt nichts anderes, als daß res surjektiv ist. Sei $\pi : G \to G/H$ die Quotientenabbildung. Dann haben wir einen induzierten Homomorphismus

$$\pi^* : \widehat{G/H} \to \ker\big(\mathrm{res} : \hat{G} \to \hat{H}\big), \qquad \chi \mapsto \chi \circ \pi.$$

Nach der Eigenschaft des Quotienten induziert jeder Charakter $\chi \in \ker(\mathrm{res})$, also ein Charakter mit $\chi(H) = 1$ eindeutig einen Charakter

$$G/H \to \mathbb{C}^\times,$$

dessen BIld unter π^* er ist. Damit ist π^* surjektiv, und injektiv ist π^* sowieso, weil π surjektiv ist. Man kann daher $\ker(\mathrm{res})$ mit $\widehat{G/H}$ identifizieren. Es folgt

$$\#\hat{G} = \#\,\mathrm{im}(\mathrm{res}) \cdot \#\ker(\mathrm{res}) = \#\hat{H} \cdot \#\widehat{G/H},$$

und wegen des Satzes von Lagrange $\#G = \#H \cdot \#(G/H)$ reicht es daher den Satz für H und G/H zu beweisen.

Der Induktionsschirtt bei Induktion nach $\#G$ führt mit obigem Argument solange zum Ziel, wie es in G eine echte Untergruppe gibt. Es reicht daher, den Satz für zyklische Gruppen zu beweisen, denn nicht zyklische Gruppen G haben für jedes $x \in G$, $x \neq 1$ die echte Untergruppe $H = \langle x\rangle$.

Im zyklischen Fall ist G isomorph zu $\mathbb{Z}/n\mathbb{Z}$ für ein n. Diesen Fall haben wir bereits in Beispiel 13.14 behandelt. □

Wie üblich gibt es eine natürliche Abbildung in das doppelte Dual, hier der Homomorphismus

$$\iota : G \to \hat{\hat{G}}, \quad \iota(g) = \big(\varphi \mapsto \varphi(g)\big).$$

Korollar 13.17. *Sei G eine endliche abelsche Gruppe. Dann ist $\iota : G \to \hat{\hat{G}}$ ein Isomorphismus.*

Beweis. Der Homomorphismus ist injektiv, denn $\iota(x) = 1$ bedeutet, daß $\varphi(x) = 1$ für alle $\varphi \in \hat{G}$. Dann würden alle Charaktere über $G/\langle x\rangle$ faktorisieren, also

$$\#G = \#\hat{G} = \#\hat{G/\langle x\rangle} = \#(G/\langle x\rangle) = \#G/\,\mathrm{ord}(x).$$

Daraus folgt $\mathrm{ord}(x) = 1$ und $x = 1$.

Es reicht nun, die Ordnungen zu vergleichen: $\#G = \#\hat{G} = \#\hat{\hat{G}}$ nach Satz 13.16. □

Integrale über Charaktere sind besonders einfach. Wenn es sich nicht um den trivialen Charakter handelt, dann oszilliert der Charakter und mittelt sich zu 0 im Integral. Das besagt die folgende Proposition.

Proposition 13.18. *Sei G eine endliche abelsche Gruppe. Für jeden Charakter $\chi \in \hat{G}$ gilt:*

$$\sum_{x\in G} \chi(x) = \begin{cases} \#G & \text{falls } \chi = \mathbf{1}, \\ 0 & \text{sonst.} \end{cases}$$

Beweis. Die Formel ist trivial für den trivialen Charkter $\chi = \mathbf{1}$. Wir nehmen daher an, daß es ein $a \in G$ gibt mit $\chi(a) \neq 1$. Dann ist

$$\chi(a) \cdot \sum_{x\in G} \chi(x) = \sum_{x\in G} \chi(a)\chi(x) = \sum_{x\in G} \chi(ax) = \sum_{x\in G} \chi(x),$$

und das geht wegen $\chi(a) - 1 \neq 0$ nur falls $\sum_{x\in G} \chi(x) = 0$. □

Sei G eine endliche abelsche Gruppe. Die Menge der $\mathbb{C}$-wertigen Abbildungen auf G

$$R(G, \mathbb{C}) := \{\varphi \ ; \ \varphi : G \to \mathbb{C}\}$$

ist durch punktweise Addition und Skalarmultiplikation ein $\mathbb{C}$-Vektorraum. Wir statten $R(G, \mathbb{C})$ mit der sesqui-linearen Bilinearform

$$\langle \varphi, \psi \rangle := \frac{1}{\#G} \sum_{x\in G} \varphi(x)\overline{\psi(x)}$$

aus. In der (offensichtlichen) Basis aus den charakteristischen Funktionen zu $g \in G$:

$$\mathbf{1}_g(x) = \begin{cases} 1 & \text{falls } g = x, \\ 0 & \text{sonst,} \end{cases}$$

wird unter dem durch die Basis induzierten Isomorphismus $R(G, \mathbb{C}) \simeq \mathbb{C}^{\#G}$ aus φ das Tupel der Werte $(\varphi(x))_{x\in G}$

$$\varphi = \sum_{x\in G} \varphi(x)\mathbf{1}_x.$$

Aus der Bilinearform $\langle \varphi, \psi \rangle$ wird die um den Faktor $1/\#G$ skalierte hermitesche Standardform. Insbesondere is $\langle -, - \rangle$ ein hermitesches Skalarprodukt auf $R(G, \mathbb{C})$.

Proposition 13.19. *Sei G eine endliche abelsche Gruppe. Die Menge der Charaktere $\hat{G}$ ist eine Orthogonalbasis von $R(G, \mathbb{C})$.*

Beweis. Wegen $\#\hat{G} = \#G = \dim_{\mathbb{C}} R(G, \mathbb{C})$ aus Satz 13.16 reicht es zu zeigen, daß die Charaktere zueinander Orthogonal und normiert sind. Das leistet Proposition 13.20. □

Proposition 13.20 (Orthogonalitätsrelation)**.** *Seien ϕ und ψ Charaktere auf der endlichen abelschen Gruppe G. Dann gilt*

$$\langle \phi, \psi \rangle = \begin{cases} 1 & \text{falls } \phi = \psi, \\ 0 & \text{sonst.} \end{cases}$$

Beweis. Seien ϕ und ψ Charaktere. Aus Proposition 13.18 angewandt auf $\chi = \varphi\overline{\psi}$ folgt

$$\langle \phi, \psi \rangle = \frac{1}{\#G} \sum_{x\in G} \varphi(x)\overline{\psi(x)} = \frac{1}{\#G} \sum_{x\in G} \chi(x) = \begin{cases} 1 & \text{falls } \phi = \psi, \\ 0 & \text{sonst.} \end{cases}$$ □

Ein beliebiges $f \in R(G, \mathbb{C})$ läßt sich somit eindeutig als Linearkombination $f = \sum_\chi a_\chi \chi$ der Charaktere schreiben. An die Koeffizienten $a_\chi \in \mathbb{C}$ kommt man auf die übliche Weise heran:

$$\langle f, \chi_0 \rangle = \langle \sum_\chi a_\chi \chi, \chi_0 \rangle = \sum_\chi a_\chi \langle \chi, \chi_0 \rangle = a_{\chi_0}.$$

Speziell für die Charakteristischen Funktionen $\mathbf{1}_g$ und einen Charakter χ gilt

$$\langle \mathbf{1}_g, \chi\rangle = \frac{1}{\#G}\sum_{x\in G}\mathbf{1}_g(x)\overline{\chi}(x) = \frac{1}{\#G}\sum_{x\in G}\mathbf{1}_g(x)\chi^{-1}(x) = \frac{\chi^{-1}(g)}{\#G},$$

also

$$\mathbf{1}_g(x) = \sum_{\chi}\frac{\chi^{-1}(g)}{\#G}\chi(x). \tag{13.1}$$

Wir spezialisieren (13.1) für den Fall **modularer** oder **Dirichlet-Charaktere**, also Charaktere für die multiplikative Gruppe $(\mathbb{Z}/m\mathbb{Z})^\times$. Zunächst definieren wir den folgenden Notationsmißbrauch für einen Dirichlet-Charakter $\chi : (\mathbb{Z}/m\mathbb{Z})^\times \to \mathbb{C}^\times$, die Fortsetzung durch 0 auf $\mathbb{Z}$ mittels

$$\chi(a) = \begin{cases} \chi([a]) & \text{falls } (a,m)=1, \\ 0 & \text{sonst.} \end{cases}$$

Satz 13.21. *Sei $a \in \mathbb{Z}$ teilerfremd zu m. Die Charakteristische Funktion $\mathbf{1}_{[a]}$ der Kongruenzklasse $[a] \subseteq \mathbb{Z}$ hat die folgende Darstellung als Linearkombination*

$$\mathbf{1}_{[a]}(n) = \frac{1}{\varphi(m)}\cdot\sum_{\chi}\chi^{-1}([a])\cdot\chi(n) \qquad \textit{für alle } n\in\mathbb{Z},$$

wobei die Summe über alle Charaktere auf $(\mathbb{Z}/m\mathbb{Z})^\times$ läuft.

Beweis. Das ist ein Spezialfall von (13.1). □

Sei $a \in \mathbb{Z}$ teilerfremd zu m. Wir wollen Primzahlen in $[a] = a + m\mathbb{Z}$ zählen, also die Funktion

$$\pi([a];x) := \#\{p \le x \ ;\ p \text{ ist Primzahl und } p \equiv a \pmod m\} = \sum_{p\le x}\mathbf{1}_{[a]}(p)$$

kontrollieren. Hier wird sich die Zerlegung aus Satz 13.21 als nützlich erweisen. Mit

$$\pi(\chi;x) := \sum_{p\le x}\chi(p)$$

ergibt sich

$$\pi([a];x) = \sum_{p\le x}\mathbf{1}_{[a]}(p) = \sum_{p\le x}\frac{1}{\varphi(m)}\cdot\sum_{\chi}\chi^{-1}([a])\cdot\chi(p) = \frac{1}{\varphi(m)}\cdot\sum_{\chi}\chi^{-1}([a])\cdot\pi(\chi;x).$$

13.3. **Dirichlet-Reihen.** Wir wenden uns nun den analytischen Werkzeugen zu. Allerdings verweisen wir für die Beweise oft auf [Ser73] Kapite VI.

Wir erinnern zunächst an die Interpretation von Potenzen mit komplexem Exponenten. Für $x \in \mathbb{R}_>0$ und $s = \sigma + it \in \mathbb{C}$ (mit Realteil σ und Imaginärteil t gilt

$$x^s := e^{s\log(x)} = \sum_{k\ge 0}\frac{1}{k!}(s\log(x))^k.$$

Der Absolutbetrag der Potenz wird zum Abschätzen oft gebraucht und berechnet sich als

$$|x^s| = |e^{(\sigma+it)\log(x)}| = |e^{\sigma\log(x)}|\cdot|e^{it\log(x)}| = e^{\sigma\log(x)} = x^\sigma.$$

Definition 13.22. Die **(formale) Dirichlet-Reihe** oder **L-Reihe** zu einer arithmetischen Funktion $f : \mathbb{N} \to \mathbb{C}$ ist die Reihe

$$L(f,s) := \sum_{n\ge 1}\frac{f(n)}{n^s}.$$

Beispiel 13.23. Die Dirichlet-Reihe zu $\mathbf{1}(n) = 1$ für alle n ist die Riemannsche Zeta-Funktion

$$\zeta(s) = \sum_{n\geq 1} \frac{1}{n^s}.$$

Die Dirichlet-Reihe entwickelt ihre Stärke erst, wenn wir sie als holomorphe Funktion in $s \in \mathbb{C}$ auf einem Teil der komplexen Ebene betrachten. Den Realteil von $s = \sigma + it$ bezeichnen wir mit $\Re(s) = \sigma$, den Imaginärteil mit $\Im(s) = t$.

Potenzreihen haben einen Konvergenzkreis(-radius), und Dirichlet-Reihen haben einen Konvergenzhalbraum und eine Konvergenzabszisse. Das zitieren wir im Wesentlichen durch die folgenden Propositionen.

Proposition 13.24. *Wenn $L(f, s_0)$ konvergiert, dann konvergiert $L(f, s)$ für alle s mit $\Re(s) > \Re(s_0)$ zu einer auf dem Halbraum $\{s \ ; \ \Re(s) > \Re(s_0)\}$ holomorphen Funktion.*

Beweis. [Ser73] VI §2 Corollary 1. □

Proposition 13.25. *Sei $f : \mathbb{N} \to \mathbb{C}$ eine arithmetische Funktion mit reellen Werten $f(n) \geq 0$ für alle $n \in \mathbb{N}$. Sei $\rho \in \mathbb{R}$, so daß $L(f, s)$ auf dem Halbraum $\{s \ ; \ \Re(s) > \rho\}$ konvergiert aber nicht auf einem größeren Halbraum. Dann hat $L(f, s)$ eine Singularität in $s = \rho$.*

Beweis. [Ser73] VI §2 Proposition 7. □

Beispiel 13.26. Für relle $s = 1 + \varepsilon$ mit $\varepsilon > 0$ erhält man die Konvergenz der Reihe zu $\zeta(1 + \varepsilon)$ durch Vergleich mit dem Integral. Es konvergiert $\sum_n n^{-s}$ genau dann, wenn das uneigentliche Integral $\int_1^\infty x^{-s} dx$ einen endlichen Wert hat. Wegen

$$\int_1^\infty x^{-s} dx = \lim_{N\to\infty} \int_1^N x^{-s} ds = \lim_{N\to\infty} \left[\frac{1}{1-s} x^{1-s}\right]_1^N = \frac{1}{s-1}$$

gilt das bei $s = 1 + \varepsilon$ mit reellem $\varepsilon > 0$.

Bei $s = 1$ wird die formale Dirichlet-Reihe der Zeta-Funktion zur harmonischen Reihe und divergiert. Damit hat $\zeta(s)$ bei $s = 1$ eine Singularität, wir sehen gleich, daß es ein Pol erster Ordnung ist.

Proposition 13.27. *Sei $f : \mathbb{N} \to \mathbb{C}$ eine arithmetische Funktion.*

(1) Ist f beschränkt, dann konvergiert $L(f, s)$ sicher für $\Re(s) > 1$.
(2) Sind die Partialsummen $\sum_{k=n}^m f(k)$ für alle n, m beschränkt, dann konvergiert $L(f, s)$ sogar für $\Re(s) > 0$.

Beweis. [Ser73] VI §2 Proposition 8+9. □

13.4. **Eulerprodukte.** Sei f eine mulitplikative Funktion. Wir definieren den L-Faktor zur Primzahl p als

$$L_p(f, s) = \lambda_p(f, p^{-s}) = 1 + f(p)p^{-s} + f(p^2)p^{-2s} + \ldots = \sum_{k\geq 0} f(p^k)p^{-ks}.$$

Wenn f sogar vollständig multiplikativ ist, dann gilt $f(p^k) = f(p)^k$ und die geometrische Reihe zeigt

$$L_p(f, s) = \frac{1}{1 - f(p)p^{-s}}.$$

Proposition 13.28 (Eulerprodukt)**.** *Sei f eine beschränkte mulitplikative Funktion.*

Es konvergieren $L(f, s)$ sowie $L_p(f, s)$ für alle Primzahlen p für $\Re(s) > 1$ und dort stimmt $L(f, s)$ mit dem dort konvergenten Eulerprodukt überein:

$$L(f, s) = \prod_p L_p(f, s).$$

Beweis. [Ser73] VI §3 Lemma 4+5. □

Korollar 13.29. *Ist f vollständig multiplikativ, dann gilt genauer für $\Re(s) > 1$*

$$L(f,s) = \prod_p \frac{1}{1 - f(p)p^{-s}}.$$

Beweis. Das folgt aus der Form der L-Faktoren $L_p(f,s)$. □

Satz 13.30. *Die Zetafunktion $\zeta(s)$ ist holomorph für $\Re(s) > 1$ mit Eulerprodukt*

$$\zeta(s) = \prod_p \frac{1}{1 - p^{-1}}.$$

Für $\Re(s) > 0$ besitzt die Zetafunktion eine meromorphe Fortsetzung

$$\zeta(s) = \frac{1}{s-1} + \Delta(s)$$

mit $\Delta(s)$ holomorph in $\Re(s) > 0$.

Beweis. Der erste Teil folgt sofort aus Proposition 13.28, weil $\zeta(s)$ zur vollständig multiplikativen Funktion **1** gehört.

Zur meromorphen Fortsetzung approximieren wir die Summe durch das Integral $\int_1^\infty t^{-s}dt$ und rechnen zunächst für $\Re(s) > 1$:

$$\begin{aligned}\zeta(s) &= \sum_{n\geq 1} \frac{1}{n^s} = \int_1^\infty t^{-s}dt + \sum_{n\geq 1}\Big(\frac{1}{n^s} - \int_n^{n+1} t^{-s}dt\Big)\\ &= \frac{1}{1-s}t^{1-s}\Big|_{t=1}^{t=\infty} + \sum_{n\geq 1}\int_n^{n+1}(\frac{1}{n^s} - t^{-s})dt = \frac{1}{s-1} + \sum_{n\geq 1}\Delta_n(s)\end{aligned}$$

mit $\Delta_n(s) = \int_n^{n+1}(\frac{1}{n^s} - t^{-s})dt$. Die Funktionen $\Delta_n(s)$ sind holomorph für $\Re(s) > 0$. Damit reicht es zu zeigen, daß auf jeder kompakten Teilmenge $K \subseteq \{s\ ;\ \Re(s) > 0\}$ die Konvergenz $\Delta(s) = \sum_n \Delta_n(s)$ uniform ist.

Wir wählen $M, \varepsilon > 0$ mit $|s| \leq M$ und $\Re(s) \geq \varepsilon$ für alle $s \in K$. Dann folgt die uniforme Konvergenz aus der Abschätzung (die relevante Ableitung ist $\frac{d}{dt} - t^{-s} = st^{-s-1}$):

$$\begin{aligned}|\Delta_n(s)| &\leq \sup_{t\in[n,n+1]} |\frac{1}{n^s} - \frac{1}{t^s}| = \sup_{t\in[n,n+1]} |\int_n^t s\tau^{-s-1}d\tau|\\ &\leq \sup_{t\in[n,n+1]} \big(|n-t| \cdot \sup_{\tau\in[n,t]} |s\tau^{-s-1}|\big) = \frac{|s|}{n^{\Re(s)+1}} \leq M \cdot \frac{1}{n^{1+\varepsilon}}.\end{aligned}$$

Damit majorisieren wir die Reihe $\Delta(s) = \sum_n \Delta_n(s)$ durch die konvergente Reihe

$$M \cdot \sum_n \frac{1}{n^{1+\varepsilon}} = M \cdot \zeta(1+\varepsilon).$$ □

Jetzt fixieren wir einen modulus m und betrachten Dirichlet-Charaktere modulo m. Diese definieren via Fortsetzung durch 0 multiplikative Funktionen auf $\mathbb{N}$ und damit L-Reihen

$$L(\chi, s) = \sum_{n\geq 1} \frac{\chi(n)}{n^s}.$$

Proposition 13.31. *Sei χ ein nichttrivialer Dirichlet-Charakter modulo m. Dann hat $L(\chi,s)$ für $\Re(s) > 1$ das Eulerprodukt*

$$L(\chi, s) = \prod_p \frac{1}{1 - \chi(p)p^{-s}}$$

und definiert eine holomorphe Funktion auf $\Re(s) > 0$.

Beweis. Das Eulerprodukt folgt aus Proposition 13.28 und die holomorphe Fortsetzung nach $\Re(s) > 0$ folgt aus Proposition 13.27, weil die Partialsummen aufgrund Proposition 13.18 nur endlich viele Werte annehmen. □

Beispiel 13.32. Sei $\chi_4 : (\mathbb{Z}/4\mathbb{Z})^\times \to \mathbb{C}^\times$ der einzige nichttriviale Dirichlet-Charakter modulo 4. Das bedeutet

$$\chi_4(n) = \begin{cases} 1 & \text{falls } n \equiv 1 \pmod 4, \\ 0 & \text{falls } 2 \mid n, \\ -1 & \text{falls } n \equiv -1 \pmod 4. \end{cases}$$

Den Wert $L(\chi_4, 1)$ bestimmen wir mittels der Potenzreihe des Arcustangens. Zur Erinnerung:

$$\frac{d}{dx}\arctan(x) = \frac{1}{\left(\frac{d}{dx}\tan\right)(\arctan(x))} = \frac{1}{1+x^2} = \sum_{k\geq 0}(-x^2)^k.$$

Integration und $\arctan(0) = 0$ liefert die Potenzreihe des Arcustangens:

$$\arctan = \sum_{k\geq 0}(-1)^k \frac{x^{2k+1}}{2k+1}.$$

Daraus schließen wir auf

$$L(\chi_4, 1) = \sum_{k\geq 1}\frac{\chi_4(k)}{k} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} \pm \ldots = \arctan(1) = \frac{\pi}{4}.$$

Die Fortsetzung durch 0 des trivialen Dirichlet-Charakters modulo m soll nicht mit der Funktion $\mathbf{1}$ verwechselt werden, die zur Riemannschen Zetafunktion führt. Daher bezeichnen wir diese Funktion mit $\mathbf{1}_m$.

Proposition 13.33. *Für die Fortsetzung $\mathbf{1}_m$ des trivialen Charakters gilt*

$$L(\mathbf{1}_m, s) = \zeta(s) \cdot \prod_{p\mid m}(1 - p^{-s}).$$

Insbesondere ist $L(\mathbf{1}_m, s)$ meromorph auf $\Re(s) > 0$ mit einzigem Pol bei $s = 1$ der Ordnung 1.

Beweis. Das Eulerprodukt folgt aus Proposition 13.28. Nur für $p \mid n$ weicht $\mathbf{1}_m(p) = 0$ von 1 ab, so daß genau diese Eulerfaktoren fehlen. Die Aussage zur Fortsetzung folgt aus Satz 13.30, weil die endlich vielen Faktoren $1 - p^{-s}$ holomoph sind, und überdies keine Nullstelle bei $s = 1$ haben und so den einfachen Pol bei $s = 1$ nicht zerstören. □

Wir multiplizieren nun die L-Reihen zu Dirichlet-Charakteren auf

$$\zeta_m(s) := \prod_\chi L(\chi, s).$$

Weil jeder der Faktoren $L(\chi, s)$ ein Eulerprodukt hat, besitzt auch $\zeta_m(s)$ ein Eulerprodukt. Dieses bestimmen wir als nächstes. Dazu setzen wir für jede Primzahl $p \nmid m$

$$H_p := \langle p \rangle \subseteq (\mathbb{Z}/m\mathbb{Z})^\times$$

und $f_p = \#H_p$ sowie $r_p = \varphi(m)/f_p$.

Lemma 13.34. *Sei p eine Primzahl mit $p \nmid m$. Dann gilt*

$$\prod_\chi (1 - \chi(p)T) = (1 - T^{f_p})^{r_p}$$

mit dem Produkt über alle Charaktere χ auf $(\mathbb{Z}/m\mathbb{Z})^\times$.

Beweis. In die Formel gehen nur Werte bei p ein, also die Einschränkung $\chi|_{H_p}$. Um zu bestimmen, wie oft jeder Charakter φ auf H_p als Einschränkung vorkommt, betrachten wir

$$\mathrm{res} : \widehat{(\mathbb{Z}/m\mathbb{Z})^\times} \to \hat{H}_p$$

wie im Beweis von Satz 13.16. Die Menge $\{\chi\ ;\ \mathrm{res}(\chi) = \varphi\}$ ist eine Nebenklasse unter dem Kern von res. Es folgt, daß jede Einschränkung genau r_p-mal angenommen wird, denn es gilt

$$r_p = \#((\mathbb{Z}/m\mathbb{Z})^\times)/\#H_p = \#\big(\widehat{(\mathbb{Z}/m\mathbb{Z})^\times}/H_p\big) = \#\ker(\mathrm{res}).$$

Die Gruppe H_p ist zyklisch von Ordnung f_p mit Erzeuger p. Aus Beispiel 13.14 folgt, daß die Charaktere von H_p auf p als Werte jede der f_p-ten Einheitswurzeln genau einmal annehmen. Zusammengenommen liefert das

$$\prod_{\chi\in\widehat{(\mathbb{Z}/m\mathbb{Z})^\times}} (1-\chi(p)T) = \Big(\prod_{\chi\in\hat{H}_p}(1-\chi(p)T)\Big)^{r_p} = \Big(\prod_{\zeta\in\mu_{f_p}}(1-\zeta T)\Big)^{r_p} = (1-T^{f_p})^{r_p},$$

wobei wir für die letzte Gleichung auf Bemerkung 13.13 verweisen. □

Satz 13.35. *(1) Die Funktion $\zeta_m(s)$ hat für $\Re(s) > 1$ das Eulerprodukt*

$$\zeta_m(s) = \prod_{p\nmid m}\Big(\frac{1}{1-p^{-sf_p}}\Big)^{r_p}$$

und ist damit eine Dirichlet-Reihe mit nichtnegativen ganzzahligen Koeffizienten.

(2) Die Funktion $\zeta_m(s)$ ist meromorph in $\Re(s) > 0$ mit einem einfachen Pol bei $s = 1$.

(3) Für alle $\chi \neq 1$ haben wir nicht verschwindende L-Werte

$$L(\chi, 1) \neq 0.$$

Beweis. (1) Das Eulerprodukt folgt sofort aus Lemma 13.34. Die Dirichlet-Reihe ist das Produkt für alle $p \nmid m$ von Potenzen der Reihen

$$1 + p^{-sf_p} + p^{-2sf_p} + \dots$$

und haben daher nichtnegative ganzzahlige Koeffizienten.

(2) Wegen Proposition 13.33 und Proposition 13.31 ist $\zeta_m(s)$ meromorph in $\Re(s) > 0$ mit höchstens einem Pol erster Ordnung in $s = 1$. Genauer ist (2) äquivalent zu (3), denn

$$\mathrm{ord}_{s=1}(\zeta_m(s)) = \sum_\chi \mathrm{ord}_{s=1}(L(\chi, s)) = 1 + \sum_{\chi\neq\mathbf{1}} \mathrm{ord}_{s=1}(L(\chi, s)).$$

Angenommen (2) ist falsch. Dann ist $\zeta_m(s)$ holomorph bei $s = 1$, und wegen (1) und Proposition 13.25 hat dann $\zeta_m(s)$ eine in $\Re(s) > 0$ konvergente Dirichlet-Reihe. Aber für alle $p \nmid m$ ist $f_p \leq \varphi(m)$, daher dominiert der L-Faktor zur Primzahl p die Reihe

$$1 + p^{-s\varphi(m)} + p^{-2s\varphi(m)} + \dots.$$

Insgesamt dominiert damit $\zeta_m(s)$ die L-Reihe $L(\mathbf{1}_m, \varphi(m)s) = \sum_{(n,m)=1} \frac{1}{n^{\varphi(m)s}}$. Diese hat einen Pol bei $s = 1/\varphi(m)$ nach Proposition 13.33, und dies steht im Widerspruch zur Konsequenz, daß $\zeta_m(s)$ für alle $\Re(s) > 0$ konvergent ist. Dies zeigt (2) und damit auch (3). □

13.5. **Dirichlet-Dichte.** Ein analytisches Maß für den asymptotischen Anteil (die relative Häufigkeit) einer Teilmenge der Primzahlen ist die Dirichlet-Dichte, sofern der sie definierende Limes existiert.

Notation 13.36. Funktionen $f(s)$ und $g(s)$, die für reelle $1 < s < 1 + \varepsilon$ definiert sind, nennen wir asymptotisch gleich für $s \searrow 1$, wenn

$$f(s) \sim g(s) :\Longleftrightarrow \lim_{s\searrow 1} f(s)/g(s) = 1.$$

Proposition 13.37. *Es gilt für $s \searrow 1$*

$$\sum_p p^{-s} \sim \log \zeta(s) \sim \log \frac{1}{s-1}.$$

Beweis. Weil $\zeta(s) - \frac{1}{s-1}$ nach Satz 13.30 für s nahe 1 beschränkt ist und $\zeta(s)$ selbst unbeschränkt, folgt die zweite asymptotische Gleichheit.

Mittels des Eulerprodukts für $\zeta(s)$ und der Potenzreihe $\log(\frac{1}{1-t}) = \sum_{k\geq 1} \frac{t^k}{k}$ berechnen wir für reelle $s > 1$:

$$\log \zeta(s) - \sum_p p^{-s} = \log\Big(\prod_p \frac{1}{1-p^{-s}}\Big) - \sum_p p^{-s} = \sum_p \Big(\log(\frac{1}{1-p^{-s}}) - p^{-s}\Big) = \sum_{p,k\geq 2} \frac{1}{kp^{sk}}$$

Hieraus folgt

$$\begin{aligned} 0 \leq \log \zeta(s) - \sum_p p^{-s} &= \sum_{p,k\geq 2} \frac{1}{kp^{sk}} \leq \sum_{p,k\geq 2} \frac{1}{p^{sk}} = \sum_p \frac{1}{p^{2s}} \cdot \frac{1}{1-p^{-s}} = \sum_p \frac{1}{p^s(p^s-1)} \\ &\leq \sum_p \frac{1}{p(p-1)} \leq \sum_{n\geq 2} \frac{1}{n(n-1)} = \sum_{n\geq 2} \frac{1}{n-1} - \frac{1}{n} = 1. \end{aligned}$$

und damit die erste asymptotische Gleichheit. □

Definition 13.38. Die **Dirichlet–Dichte** einer Teilmenge A aller Primzahlen ist, sofern der Limes existiert, die reelle Zahl $\delta(A) \in [0,1]$ definiert durch

$$\delta(A) = \lim_{s\searrow 1} \frac{\sum_{p\in A} \frac{1}{p^s}}{\sum_p \frac{1}{p^s}}.$$

Bemerkung 13.39. Aus Proposition 13.37 folgt, daß sich die Dirichlet–Dichte einer Menge A von Primzahlen nicht ändert, wenn man A um endlich viele Primzahlen verändert. Endliche Mengen haben Dirichlet–Dichte 0.

Für komplexe Zahlen ist der Logarithmus nur bis auf ganzzahlige Vielfache von $2\pi i$ eindeutig definiert. Wir legen uns daher mittels der Potenzreihe $\log(\frac{1}{1-t}) = \sum_{k\geq 1} \frac{t^k}{k}$ auf einen Hauptteil fest.

Proposition 13.40. *Sei χ ein Dirichlet-Charakter modulo m. Für $s \searrow 1$ bleibt die Differenz*

$$\sum_p \frac{\chi(p)}{p^s} - \log L(\chi, s)$$

beschränkt und

$$\log L(\chi, s) \sim \begin{cases} \log \frac{1}{s-1} & \textit{falls } \chi = \mathbf{1}, \\ \log L(\chi, 1) & \textit{sonst.} \end{cases}$$

Beweis. Für den trivialen Charakter $\chi = \mathbf{1}_m$ unterscheidet sich $L(\mathbf{1}_m, s)$ von $\zeta(s)$ nur um einen in $s = 1$ holomorphen Faktor ohne Nullstelle, nämlich Faktoren $1 - p^{-s}$ für $p \mid m$. Die Logarithmen unterscheiden sich daher um einen beschränkten Summanden und haben somit die gleiche Asymptotik. Die Behauptung folgt aus Proposition 13.37.

Für nichttriviale Charaktere ist $L(\chi, 1) \neq 0$ nach Satz 13.35 und die zweite Asymptotik ist trivial. Die erste folgt wie in Proposition 13.37.

$$\begin{aligned} |\log L(\chi, s) - \sum_p \frac{\chi(p)}{p^s}| &= |\log\Big(\prod_p \frac{1}{1-\chi(p)p^{-s}}\Big) - \frac{\chi(p)}{p^s}| = |\sum_p \Big(\log(\frac{1}{1-\chi(p)p^{-s}}) - \frac{\chi(p)}{p^s}\Big)| \\ &= |\sum_{p,k\geq 2} \frac{\chi(p)^k}{kp^{sk}}| \leq \sum_{p,k\geq 2} \frac{1}{kp^{sk}} \leq 1. \end{aligned}$$

□

Bemerkung 13.41. In Proposition 13.37 ist es entscheidend, daß $L(\chi, 1) \neq 0$ für $\chi \neq \mathbf{1}$, damit der Logarithmus beschränkt bleibt und die Asymptotik nicht beeinflußt.

Wir beweisen nun Theorem 13.2 in der folgenden Form. Weil endliche Mengen von Primzahlen Dirichlet-Dichte 0 haben, folgt auch das ursprüngliche Theorem 13.2.

Theorem 13.42 (Satz von Dirichlet)**.** *Seien m, a teilerfremd und $m \geq 1$. Die Menge der Primzahlen p mit*

$$p \equiv a \pmod{m}.$$

hat die Dirichlet–Dichte $1/\varphi(m)$.

Beweis. Wir benutzen Satz 13.21, um die Bedingung $p \equiv a \pmod{m}$ in eine Summe von Dirichlet-Charakteren zu übersetzen. Damit folgt

$$\sum_{\substack{p \equiv a \\ (\mathrm{mod}\ m)}} \frac{1}{p^s} = \sum_p \frac{\mathbf{1}_{[a]}(p)}{p^s} = \frac{1}{\varphi(m)} \sum_\chi \Big(\chi^{-1}(a) \sum_p \frac{\chi(p)}{p^s} \Big) \sim \frac{1}{\varphi(m)} \log \frac{1}{s-1},$$

denn nach Proposition 13.40 sind die Beiträge für nicht-triviale Charaktere beschränkt und verlieren gegen den Beitrag des trivialen Charakter bei $s \searrow 1$. Die behauptete Dirichlet-Dichte von $1/\varphi(m)$ folgt sofort. □

Bemerkung 13.43. Für die Gleichverteilung der Primzahlen auf die zu m teilerfremden Restklassen modulo m sind letztendlich die folgenden Fakten verantwortlich:

- Die L-Werte $L(\chi, 1)$ für nichttriviale Dirichlet-Charaktere χ sind endlich und von 0 verschieden. Damit tragen nur die trivialen Dirichlet-Charaktere zur Asymptotik bei.
- Der Koeffizient in der Formel von Satz 13.21 des trivialen Charakters hängt nicht von der ausgewählten Restklasse modulo m ab.

14. Diophantische Gleichungen in endlichen Körpern

14.1. **Endliche Körper.** Um über Diophantische Gleichungen über endlichen Körpern sprechen zu können, muß man zunächst endliche Körper verstehen.

Proposition 14.1. *Sei $\mathbb{F}$ ein endlicher Körper. Dann gibt es eine eindeutige Primzahl p und einen Unterkörper $\mathbb{F}_0 \subseteq \mathbb{F}$ mit p Elementen, der eindeutig mit $\mathbb{F}_p$ identifizierbar ist. Die Anzahl $q = \#\mathbb{F}$ der Elemente von $\mathbb{F}$ ist eine p-Potenz.*

Beweis. Es gibt einen einzigen Ringhomomorphismus

$$f : \mathbb{Z} \to \mathbb{F},$$

denn es muß $f(-n) = -f(n)$ gelten und $f(1) = 1$ legt $f(n)$ für $n \geq 0$ durch

$$f(n) = n \cdot 1 = 1 + \ldots + 1 \qquad (\text{mit } n \text{ Summanden})$$

eindeutig fest. Dies ist in der Tat, wie man sofort nachrechnet, ein Ringhomomorphismus.

Da $\mathbb{F}$ endlich ist, hat f einen nichttrivialen Kern. Der Kern ist ein Ideal von $\mathbb{Z}$. Weil $\mathbb{Z}$ ein Hauptidealring ist, gibt es ein eindeutiges $p > 0$ mit

$$f(n) = 0 \iff p \mid n.$$

Angenommen p wäre keine Primzahl, dann gibt es $p = ab$ mit $a, b > 1$ und $0 = f(p) = f(a)f(b)$. In einem Körper ist ein Produkt nur dann 0, wenn einer der Faktoren 0 ist. Daher gilt oBdA $f(a) = 0$. Dann gilt $p \mid a$ und $a \mid p$, daher $p = a$ und $b = 1$, ein Widerspruch. Daher ist p Primzahl und $\mathbb{F}$ enthält das Bild, also nach dem Homomorphiesatz

$$\mathrm{im}(f) \simeq \mathbb{Z}/\ker(f) = \mathbb{F}_p.$$

Dies ist der behauptete Unterkörper $\mathbb{F}_0$ mit p Elementen. Dieser ist eindeutig, denn als Unterkörper enthält $\mathbb{F}_0$ die 1 und damit $\operatorname{im}(f)$. Es gilt also $\mathbb{F}_p \simeq \operatorname{im}(f) \subseteq \mathbb{F}_0$, und aus Anzahlgründen gilt Gleichheit.

Die Einschränkung der Körpermultiplikation von $\mathbb{F}$

$$\mathbb{F}_p \times \mathbb{F} \to \mathbb{F}$$

macht aus $\mathbb{F}$ einen $\mathbb{F}_p$-Vektorraum. Da $\mathbb{F}$ endlich ist, ist $r = \dim_{\mathbb{F}_p}(\mathbb{F})$ endlich. Als $\mathbb{F}_p$-Vektorraum ist damit $\mathbb{F} \simeq (\mathbb{F}_p)^r$. Demnach hat $\mathbb{F}$ genau $q = p^r$ Elemente. □

Definition 14.2. Die **Charakteristik** eines Körpers k ist das minimale $n > 0$ mit $n \cdot 1 = 0 \in k$, oder per Konvention $= 0$, wenn kein endliches solches n existiert.

Bemerkung 14.3. Der Beweis von Proposition 14.1 zeigt, daß die Charakteristik eines Körpers stets 0 oder eine Primzahl ist. Für endliche Körper $\mathbb{F}$ ist dies die Primzahl p, so daß $\mathbb{F}_p \subseteq \mathbb{F}$.

Proposition 14.4. *Sei k ein Körper der Charakteristik p. Dann ist $F : k \to k$ definiert durch*

$$F(x) = x^p$$

für alle $x \in k$ ein Körperendomorphismus. Für endliche k ist F ein Automorphismus.

Beweis. Die Abbildung F ist offensichtlich verträglich mit Multiplikation. Bezüglich Addition bemühen wir den Binomischen Lehrsatz

$$(a+b)^p = \sum_{i=0}^{p} \binom{p}{i} a^i b^{p-i}$$

und, weil $p \mid \binom{p}{i}$ für alle $1 \le i \le p-1$, erhalten $F(a+b) = F(a) + F(b)$.

Jeder Körperhomomorphismus ist injektiv. Für endliche Körper ist F daher surjektiv. □

Proposition 14.5. *Sei $\mathbb{F}$ ein endlicher Körper mit $q = p^r$ Elementen. Dann ist für alle $a \in \mathbb{F}$*

$$a^q = a$$

und

$$X^q - X = \prod_{a \in \mathbb{F}} (X - a).$$

Beweis. Für $a = 0$ gilt $0^q = 0$, und für $a \neq 0$ ist Multiplikation mit a eine Bijektion auf der multiplikativen Gruppe $\mathbb{F}^\times = \mathbb{F} \setminus \{0\}$. Daher gilt

$$a^q \cdot \prod_{x \in \mathbb{F}^\times} x = a \cdot \prod_{x \in \mathbb{F}^\times} ax = a \cdot \prod_{x \in \mathbb{F}^\times} x$$

und Kürzen mit $\prod_{x \in \mathbb{F}^\times} x \neq 0$ zeigt $a^q = a$.

Damit hat das Polynom $X^q - X$ alle q Elemente aus $\mathbb{F}_q$ als Nullstellen. Dies sind dann nach Satz 10.20 alle Nullstellen von $X^q - X$. Genauer liefert der Beweis von Satz 10.20, daß man nach und nach für jede Nullstelle a den Linearfaktor $X - a$ abspalten darf. Dies zeigt die Faktorisierung von $X^q - X$. □

Theorem 14.6. *Für jede Primzahlpotenz $q = p^r$ gibt es bis auf Isomorphie genau einen Körper $\mathbb{F}_q$ mit q Elementen.*

Beweis. Wir verwenden den folgenden Fakt aus der Körpertheorie. Es gibt eine algebraisch abgeschlossene Erweiterung $\mathbb{F}_p \subseteq \bar{\mathbb{F}}_p$ in die jede endliche Körpererweiterung $\mathbb{F}_p \subseteq \mathbb{F}$ eingebettet werden kann: $\mathbb{F}$ ist isomorph zu einem Unterkörper von $\bar{\mathbb{F}}_p$.

Dann ist aber alles klar, denn das Bild von $\mathbb{F}$ in $\bar{\mathbb{F}}_p$ besteht nach Proposition 14.5 aus allen $a \in \bar{\mathbb{F}}_p$ mit $a^q = a$. Dies zeigt die Eindeutigkeit.

Die Existenz folgt, da der Fixkörper des Automorphismus $F^r : \bar{\mathbb{F}}_p \to \bar{\mathbb{F}}_p$ aus den Lösungen der Polynomgleichung $X^q = X$ besteht, von denen es q-viele gibt, weil $X^q - X$ ein Polynom ohne

mehrfache Nullstelle ist. Das sieht man weil die formale Ableitung $\frac{d}{dX}(X^q - X) = qX^{q-1} - 1$ zu $X^q - X$ teilerfremd ist. □

Theorem 14.7. *Die Gruppe $\mathbb{F}_q^\times$ ist zyklisch von Ordnung $q-1$.*

Beweis. Mit Proposition 14.5 ist der Beweis formal der gleiche wie der Beweis nach Satz 10.21 und Theorem 10.23, daß $\mathbb{F}_p^\times$ zyklisch ist. □

14.2. **Chevalley–Warning.** Sei p eine Primzahl und sei $\mathbb{F}_q$ ein endlicher Körper mit $q = p^r$ Elementen. Für ein Polynom $f \in \mathbb{F}_q[X_1, \dots, X_n]$ definieren wir

$$S(f) = \sum_{(x_1,\dots,x_n)\in\mathbb{F}_q^n} f(x_1, \dots, x_n) \qquad \in \mathbb{F}_q,$$

also die Summe über alle Funktionswerte auf dem affinen Raum $\mathbb{A}^n(\mathbb{F}_q) = (\mathbb{F}_q)^n$.

Bemerkung 14.8. (1) Der Wert $S(f)$ ist gewissermaßen das Integral über die Polynomfunktion

$$f : \mathbb{A}^n(\mathbb{F}_q) \to \mathbb{F}_q$$

bezüglich des diskreten Maßes.

(2) Das Integral S ist $\mathbb{F}_q$-linear: für $a, b \in \mathbb{F}_q$ und Polynome $f, g \in \mathbb{F}_q[X_1, \dots, X_n]$ gilt

$$S(af + bg) = aS(f) + bS(g).$$

(3) Es gilt das Fubini-Theorem: Seien $f(X_1, \dots, X_r)$ und $g(X_{r+1}, \dots, X_n)$ Polynome in disjunkten Variablen. Dann gilt

$$S(f(X_1, \dots, X_r)g(X_{r+1}, \dots, X_n)) = S(f) \cdot S(g).$$

Das ist trivial!

(4) Für $n > 0$ hingegen hat X^n Träger in $\mathbb{F}_q^\times \subseteq \mathbb{A}^1(\mathbb{F}_q)$ und im Besonderen ist dort ein Charakter

$$(-)^n : \mathbb{F}_q^\times \to \mathbb{F}_q^\times.$$

Dieser Charakter ist integrabel, wie jede Funktion, weil ja alles endliche Summen sind. Dies steht im Kontrast zum Fall reeller oder komplexer Polynome, wo X^n auf $\mathbb{R}^\times$ oder $\mathbb{C}^\times$ auch Charaktere sind, aber keine endliche $\mathscr{L}^p$-Norm haben.

Integrale über Charaktere und Produkte mit Charakteren spielen in der Darstellungstheorie eine wichtige Rolle. Die Fourier-Transformation ist ein Beispiel.

Die Funktion X^0 soll die konstante Funktion 1 bedeuten.

Lemma 14.9. *Für alle ganzen Zahlen $n \geq 0$ gilt*

$$S(X^n) = \begin{cases} 0 & & q-1 \nmid n, \\ q-1 & \textit{wenn} & q-1 \mid n,\ n > 0, \\ q & & n = 0. \end{cases}$$

Beweis. Für $n = 0$ summieren wir 1 über die Anzahl von $\mathbb{F}_q$ und erhalten q. Wenn $q-1 \mid n$ und $n > 0$, dann ist X^n die charakteristische Funktion von $\mathbb{F}_q^\times$, also 1 genau dann, wenn $x \neq 0$ und 0 sonst. Die Summe über diese Werte ist dann $q-1$.

Sei daher $n > 0$ und $q-1 \nmid n$. Wir wählen einen Erzeuger ϑ der multiplikativen Gruppe $\mathbb{F}_q^\times$. Nach Voraussetzung ist nun $\vartheta^n \neq 1$ und daher

$$S(X^n) = \sum_{x\in\mathbb{F}_q^\times} x^n = \sum_{i=0}^{q-2} \vartheta^{ni} = \frac{\vartheta^{n(q-1)} - 1}{\vartheta^n - 1} = 0,$$

weil $\vartheta^{n(q-1)} = (\vartheta^{q-1})^n = 1^n = 1$. □

Wir führen die folgende Notation für die simultane Nullstellenmenge von Polynomen f_α ein:

$$V(f_1,\dots,f_s) := \{(x_1,\dots,x_n) \in (\mathbb{F}_q)^n \ ; \ f_i(x_1,\dots,x_n) \text{ für alle } \alpha = 1,\dots,s\}.$$

Wir schreiben ein Monom in Multiindexschreibweise für den Multiindex $I = (i_1,\dots,i_n)$ als

$$X^I := X_1^{i_1} \cdot \ldots \cdot X_n^{i_n}$$

mit Grad $|I| = \sum_{\alpha=1}^n i_\alpha$. Ein Polynom $f = \sum_I a_I X^I \in \mathbb{F}_q[X_1,\dots,X_n]$ hat dann den Grad

$$\deg(f) = \max\{|I| \ ; \ a_I \neq 0\}.$$

Theorem 14.10 (Chevalley–Warning)**.** *Seien $f_\alpha \in \mathbb{F}_q[X_1,\dots,X_n]$, $f_\alpha \neq 0$ für $1 \leq \alpha \leq s$ Polynome mit*

$$\sum_{\alpha=1}^s \deg(f_\alpha) < n.$$

Dann ist

$$\#V(f_1,\dots,f_s) \equiv 0 \pmod p.$$

Beweis. Die Funktion $1 - X^{q-1}$ ist die charakteristische Funktion von $0 \in \mathbb{A}^1(\mathbb{F}_q)$. Bei 0 ist $1-0^{q-1} = 1$ und für $a \neq 0$ ist $1-a^{q-1} = 1-1 = 0$. Daraus ergibt sich sofort die charakteristische Funktion von $V := V(f_1,\dots,f_s) \subseteq \mathbb{A}^n(\mathbb{F}_q)$ als

$$\mathbf{1}_V = \prod_{\alpha=1}^s (1 - f_\alpha^{q-1}(X_1,\dots,X_n)),$$

denn jeder Faktor nimmt nur die Werte 0 und 1 an, und genauer

$$\begin{aligned} \mathbf{1}_V(x_1,\dots,x_n) = 1 &\iff 1 - f_\alpha(x_1,\dots,x_n)^{q-1} = 1 \text{ für alle } \alpha \\ &\iff f_\alpha(x_1,\dots,x_n) = 0 \text{ für alle } \alpha \iff (x_1,\dots,x_n) \in V. \end{aligned}$$

Die Anzahl der simultanen Nullstellen berechnet sich daher als

$$\#V = \sum_{(x_1,\dots,x_n)\in\mathbb{A}^n(\mathbb{F}_q)} \mathbf{1}_V(x_1,\dots,x_n) = S(\mathbf{1}_V),$$

und es bleibt zu zeigen, daß

$$S(\mathbf{1}_V) = 0 \in \mathbb{F}_q.$$

Wenn eines der f_α konstant ist, dann nach Voraussetzung mit einem Wert $\neq 0$, so daß $V = \emptyset$ ist. In diesem Fall ist nichts weiter zu zeigen.

Wir nehmen daher nun an, daß kein Polynom f_α konstant ist. Der Grad von $\mathbf{1}_V$ ist nach Voraussetzung

$$\deg(\mathbf{1}_V) = \sum_\alpha \deg(1 - f_\alpha(x_1,\dots,x_n)^{q-1}) = (q-1)\sum_\alpha \deg(f_\alpha) < (q-1)n.$$

Daher ist $\mathbf{1}_V$ eine Linearkombination von Monomen $X^I = X_1^{i_1} \cdot \ldots \cdot X_n^{i_n}$ mit $\sum_\beta i_\beta < (q-1)n$ und $i_\beta \geq 0$ für alle β. Es reicht wegen der Linearität von $S(-)$ nun $S(X^I) = 0$. Mittels Fubini folgt dies aus

$$S(X^I) = \prod_{\beta=1}^n S(X_\beta^{i_\beta}) = 0,$$

weil insbesondere mindestens ein $i_\beta < q-1$ sein muß, somit der entsprechende Faktor $S(X_\beta^{i_\beta}) = 0$ ist nach Lemma 14.9. $\square$

Korollar 14.11. *Seien $f_\alpha \in \mathbb{F}_q[X_1, \dots, X_n]$ für $1 \le \alpha \le s$ Polynome mit*

$$\sum_{\alpha=1}^{s} \deg(f_\alpha) < n$$

und ohne konstante Terme. Dann haben die $f_1, \dots, f_s$ eine gemeinsame nichttriviale Nullstelle, d.h. verschieden von $(0, \dots, 0)$.

Beweis. Es ist $(0, \dots, 0)$ eine triviale gemeinsame Nullstelle, weil nach Voraussetzung die konstanten Terme 0 sind. Also ist

$$\#V(f_1, \dots, f_s) \ge 1.$$

Da nach Theorem 14.10 aber $p \mid \#V(f_1, \dots, f_s)$, folgt sofort $\#V(f_1, \dots, f_s) \ge p$. Es muß also weitere Nullstellen geben! □

Wir erinnern: eine quadratische Form Q in n Variablen ist ein Polynom $Q \in \mathbb{F}_q[X_1, \dots, X_n]$, das Linearkombination von Monomen vom Grad 2 ist.

Korollar 14.12. *Jede quadratische Form Q über $\mathbb{F}_q$ in $n \ge 3$ Variablen hat nichttriviale Nullstellen.*

Beweis. Das folgt aus $\deg(Q) = 2 < 3 \le n$ und dem Korollar 14.11. □

14.3. **Eine Anwendung.** Die Reichardt[27]-Lind[28]-Kurve wird durch die Gleichung

$$X^4 - 17 = 2Y^2$$

beschrieben. Wir diskutieren die Lösbarkeit der Gleichung mittels quadratischem Reziprozitätsgesetz und der Technik aus dem Beweis von Chevalley-Warning.

Theorem 14.13. *Die Gleichung der Reichardt-Lind-Kurve hat keine rationalen Lösungen: es gibt keine $x, y \in \mathbb{Q}$ mit*

$$x^4 - 17 = 2y^2.$$

Jedoch gibt es

(a) reelle Lösungen: $x, y \in \mathbb{R}$ mit $x^4 - 17 = 2y^2$,
(b) für jedes $m \in \mathbb{N}$ Lösungen modulo m: Elemente $x, y \in \mathbb{Z}/m\mathbb{Z}$ mit $x^4 - 17 = 2y^2$ in $\mathbb{Z}/m\mathbb{Z}$.

Bemerkung 14.14. (1) Die Reichardt-Lind-Kurve zeigt, daß es nicht ausreicht, eine Diophantische Gleichung modulo beliebigem $m \in \mathbb{N}$ zu betrachten.

(2) Sei R ein Ring. Wir setzen

$$C(R) = \{(x, y) \in R^2 \ ; \ x^4 - 17 = 2y^2\}.$$

Dann behauptet Theorem 14.13, daß $C(\mathbb{Q}) = \emptyset$, obwohl $C(\mathbb{R}) \neq \emptyset$ und für alle $m \in \mathbb{Z}$ auch $C(\mathbb{Z}/m\mathbb{Z}) \neq \emptyset$.

(3) Zu einem Ringhomomorphismus $f : A \to B$ gehört eine Abbildung

$$\begin{aligned} C(f) : C(A) &\to C(B) \\ (x, y) &\mapsto (f(x), f(y)). \end{aligned}$$

Lemma 14.15. *Für alle Primzahlen $p \neq 2, 17$ und alle $n \ge 1$ ist*

$$C(\mathbb{Z}/p^{n+1}\mathbb{Z}) \to C(\mathbb{Z}/p^n\mathbb{Z})$$

induziert von $[a]_{p^{n+1}} \mapsto [a]_{p^n}$ eine surjektive Abbildung.

[27] Hans Reichardt (1908–1991), deutscher Mathematiker, Assistent von Carl Ludwig Siegel an der Goethe-Universität Frankfurt (1934).

[28] Carl-Erik Lind, Doktorarbeit 1940, Universität Uppsala.

Beweis. Sei $(a,b) \in C(\mathbb{Z}/p^n\mathbb{Z})$ eine Lösung modulo p^n. Seien $x_0, y_0 \in \mathbb{Z}$ Repräsentanten von a und b. Dann gibt es $c \in \mathbb{Z}$ mit

$$x_0^4 - 17 - 2y_0^2 = c \cdot p^n.$$

Wir betrachten die Taylorentwicklung der Gleichung

$$f(X,Y) = X^4 - 17 - 2Y^2$$

nahe (x_0, y_0) mit dem Ansatz

$$\begin{aligned} x &= x_0 + \xi p^n \\ y &= y_0 + \eta p^n. \end{aligned}$$

Die Bedingung für $(x,y) \in C(\mathbb{Z}/p^{n+1}\mathbb{Z})$ ist nun

$$\begin{aligned} 0 \equiv f(x,y) &= f(x_0,y_0) + \frac{\partial f}{\partial X}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial Y}(x_0,y_0)(y-y_0) \pmod{p^{n+1}} \\ &\equiv c \cdot p^n + 4x_0^3 \cdot \xi p^n - 4y_0 \cdot \eta p^n \pmod{p^{n+1}}, \end{aligned} \tag{14.1}$$

denn $p^n \mid x - x_0$ und $y - y_0$ und daher sind die Terme mindestens quadratischer Ordnung in der Talorentwicklung mindestens durch p^{2n} teilbar. (Man muß sich überlegen, daß die Nenner der Taylorentwicklung durch die Vorfaktoren, die beim Ableiten entstehen, kompensiert werden und daher die Koeffizienten der Taylorentwicklung ganzzahlig sind. Das ist hier klar, gilt aber allgemein.)

Die Gleichung (14.1) ist nun äquivalent zu

$$c \equiv 4y_0 \cdot \eta - 4x_0^3 \cdot \xi \pmod{p}.$$

Dies hat in jedem Fall, egal wie c aussieht, eine Lösung $\xi, \eta \in \mathbb{Z}$, falls $p \nmid 4x_0^3$ oder $p \nmid 4y_0$. Die gerade Primzahl $p = 2$ braucht wieder eine Sonderbehandlung. Sei also $p > 2$. Dann haben wir nur noch Probleme, wenn $x_0 \equiv y_0 \equiv 0 \pmod{p}$. Dann ist $(0,0) \in C(\mathbb{F}_p)$ und das bedeutet $p = 17$. Für $p \neq 2, 17$ haben wir also gezeigt, was wir wollen. □

Beweis von Theorem 14.13. Schritt 1: Offensichtlich gibt es reelle Lösungen, zum Beispiel $x = 3$ und $y = \sqrt{32}$.

Schritt 2: Aus dem Chinesischen Restsatz folgt unmittelbar für m mit Primfaktorzerlegung $m = \prod_{i=1}^r p_i^{e_i}$ eine Bijektion

$$C(\mathbb{Z}/m\mathbb{Z}) \xrightarrow{\sim} C(\mathbb{Z}/p_1^{e_1}) \times \ldots \times C(\mathbb{Z}/p_r^{e_r}),$$

denn für $x, y \in \mathbb{Z}$ gilt unter dem Chinesischen Restsatz $\mathbb{Z}/m\mathbb{Z} = \prod_i \mathbb{Z}/p_i^{e_i}\mathbb{Z}$

$$f([x]_m, [y]_m) \mapsto \left(f([x]_{p_i^{e_i}}, [y]_{p_i^{e_i}})\right)_{i=1,\ldots,r}.$$

und somit entspricht eine Lösung modulo m (linke Seite 0) einem Tupel von Lösungen modulo $p_i^{e_i}$ für alle $i = 1, \ldots, r$ (rechte Seite 0).

Wir können uns also auf einen Modulus $m = p^n$ beschränken, der eine Primzahlpotenz ist.

Schritt 3: Wenn $p \neq 2, 17$, dann ist nach Lemma 14.15 die Abbildung

$$C(\mathbb{Z}/p^n\mathbb{Z}) \twoheadrightarrow C(\mathbb{F}_p)$$

surjektiv. Es reicht daher, eine Lösung modulo p zu finden.

Schritt 4: Für jedes $p \neq 2, 17$ gibt es Lösungen modulo p: die Menge $C(\mathbb{F}_p)$ ist nicht leer. Diesen Fall unterteilen wir weiter:

(a) Als erstes probieren wir $x = 3$. Dann gilt $y^2 \equiv (81 - 17)/2 \equiv 32 \equiv 4^2 \cdot 2 \pmod{p}$. Das hat genau dann eine Lösung, wenn $\left(\frac{2}{p}\right) = 1$, also für $p \equiv \pm 1 \pmod{8}$.

(b) Als zweites probieren wir $x = 1$. Dann gilt $y^2 \equiv (1 - 17)/2 \equiv -2 \cdot 2^2 \pmod{p}$. Das hat genau dann eine Lösung, wenn $\left(\frac{-2}{p}\right) = 1$, also für $p \equiv 1$ oder $3 \pmod{8}$.

(c) Für den verbleibenden Fall $p \equiv 5 \pmod 8$ verweisen wir auf Beipiel 14.16.

Schritt 5: Für $p = 2$ setzen wir $y = 0$ und zeigen, daß 17 eine 4-te Potenz modulo 2^n für alle n ist. Dies folgt aus Theorem 10.33, denn 4-te Potenzen sind alle $\equiv 1 \pmod{16}$.

Schritt 6: Für $p = 17$ setzen wir $x = 1$ und zeigen, daß -8 ein Quadrat ist modulo 17^n für alle n. Das reicht für $n = 1$ und da gilt

$$\left(\frac{-8}{17}\right) = \left(\frac{-1}{17}\right) \cdot \left(\frac{2}{17}\right) = 1 \cdot 1 = 1.$$

Schritt 7: Wir zeigen nun durch einen Widerspruchsbeweis, daß es keine rationalen Lösungen gibt. Sei $x, y \in \mathbb{Q}$ mit $x^4 - 17 = 2y^2$ und als gekürzte Brüche $x = \frac{a}{b}, y = \frac{c}{d}$. Dann gilt

$$a^4 d^2 - 17 b^4 d^2 = 2c^2 b^4.$$

Wir haben $d^2 \mid d^2(a^4 - 17b^4) = 2c^2b^4$ und da $(d, c) = 1$, folgt $d^2 \mid 2b^4$. Da d^2 und b^4 Quadratzahlen sind, hat die 2-er Potenz in d^2 gerade Exponenten und ist daher schon in b^4 enthalten. Es gilt also sogar $d^2 \mid b^4$. Weiter haben wir $b^4 \mid b^4(2c^2 + 17d^2) = a^4d^2$. Da $(b, a) = 1$, folgt bereits $b^4 \mid d^2$. Insgesamt gilt daher $d^2 = b^4$ und somit

$$a^4 - 17b^4 = 2c^2.$$

Es gilt $(a, c) = 1$, denn ein Primteiler $p \mid (a, c)$ erfüllt

$$p^2 \mid a^4 - 2c^2 = 17b^4.$$

Als Teiler von a ist $p \nmid b$, und somit folgt $p^2 \mid 17$, Widerspruch.

Schritt 8: Nun arbeiten wir modulo 17. Es gilt $a, c \not\equiv 0 \pmod{17}$, denn $a^4 \equiv 2c^2 \pmod{17}$ und wegen $(a, c) = 1$ können nicht beide durch 17 teilbar sein.

Sei nun $p \mid c$ ein ungerader Primteiler. Dann ist $a^4 \equiv 17b^4 \pmod p$ und $a, b \not\equiv 0 \pmod p$. Also ist $17 \equiv (a^2/b^2)^2 \pmod p$ ein quadratischer Rest. Aus dem quadratischen Reziprozitätsgesetz folgt

$$1 = \left(\frac{17}{p}\right) = \left(\frac{p}{17}\right).$$

Es gilt nach den Ergänzungssätzen auch

$$\left(\frac{-1}{17}\right) = \left(\frac{2}{17}\right) = 1.$$

Daher ist jeder Faktor, auch das Vorzeichen ε, in der Primfaktorzerlegung von $c = \varepsilon \cdot \prod p_i$ ein quadratischer Rest modulo 17. Es folgt

$$\left(\frac{c}{17}\right) = \left(\frac{\varepsilon}{17}\right) \cdot \prod \left(\frac{p_i}{17}\right).$$

Daher gibt es $\gamma \in \mathbb{Z}$ mit $17 \nmid \gamma$ und $c \equiv \gamma^2 \pmod{17}$. Daraus folgt

$$a^4 \equiv 2\gamma^4 \pmod{17}$$

und $2 \equiv (a/\gamma)^4 \pmod{17}$ ist sogar eine 4-te Potenz modulo 17. Dann gilt aber nach dem kleinen Fermat

$$1 \equiv (a/\gamma)^{16} \equiv 2^4 \equiv -1 \pmod{17},$$

und das ist der gesuchte Widerspruch. □

Beispiel 14.16. Sei p eine Primzahl $p \equiv 5 \pmod 8$. Insbesondere gilt

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

Wir zeigen nun, daß die Gleichung

$$X^4 - 17 = 2Y^2$$

modulo p stets eine Lösung hat. Dies ist eine Gleichung in 2 Variablen vom Grad 4. Chevalley-Warning sagt direkt hier gar nichts.

Schritt 1: Wir probieren $x = 0$ aus. Dann muß man $17 \equiv -2y^2 \pmod{p}$ lösen. Das geht, wenn

$$\left(\frac{17}{p}\right) = \left(\frac{-2y^2}{p}\right) = \left(\frac{-2}{p}\right) = -1$$

ist. Wir sind also fertig, wenn 17 kein quadratischer Rest modulo p ist (und $p \equiv 5 \pmod{8}$), und nehmen deshalb nun das Gegenteil an: $\left(\frac{17}{p}\right) = 1$.

Schritt 2: Wie im Beweis von Theorem 14.10 berechnen wir die Anzahl der Lösungen als

$$S(1 - (X^4 - 17 - 2Y^2)^{p-1}).$$

Das zu „integrierende" Polynom hat Grad $\leq 4(p-1)$. Nur Terme der Form $X^{r(p-1)}Y^{s(p-1)}$ mit $r, s \geq 1$ tragen bei. Wenn man ausmultipliziert, dann hat man also mindestens $\frac{p-1}{2}$-mal den Summanden $-2Y^2$ und mindestens $\frac{p-1}{4}$-mal den Faktor X^4 zu wählen. Damit bleiben nur noch $\frac{p-1}{4}$-viele Faktoren und nur die Monome

(a) $X^{p-1}Y^{p-1}$ mit Koeffizient

$$-\binom{p-1}{\frac{p-1}{2}} \cdot \binom{\frac{p-1}{2}}{\frac{p-1}{4}} \cdot (-17)^{\frac{p-1}{4}} \cdot (-2)^{\frac{p-1}{2}}$$

(b) $X^{2(p-1)}Y^{p-1}$ mit Koeffizient

$$-\binom{p-1}{\frac{p-1}{2}} \cdot (-2)^{\frac{p-1}{2}}$$

treten auf.

Schritt 3: Bei $r > 0$ gilt

$$S(X^{r(p-1)}Y^{(p-1)}) \equiv S(X^{r(p-1)}) \cdot S(Y^{(p-1)}) \equiv (-1)^2 \equiv 1 \pmod{p}.$$

Daraus ergibt sich, daß

$$S(1 - (X^4 - 17 - 2Y^2)^{p-1}) \equiv -\binom{p-1}{\frac{p-1}{2}} \cdot (-2)^{\frac{p-1}{2}} \cdot \left(1 + \binom{\frac{p-1}{2}}{\frac{p-1}{4}} \cdot (-17)^{\frac{p-1}{4}}\right) \pmod{p}.$$

Die Binomialkoeffizienten sind nicht durch p teilbar. Da wir nur an der Existenz von Lösungen interessiert sind, müssen wir nun zeigen, daß

$$\binom{\frac{p-1}{2}}{\frac{p-1}{4}} \cdot (-17)^{\frac{p-1}{4}} \not\equiv -1 \pmod{p}.$$

Sei $p = 8n + 5$. Dann ist $\frac{p-1}{4} = 2n + 1$ ungerade. Zu zeigen ist also

$$(\frac{p-1}{2})! \cdot 17^{\frac{p-1}{4}} \not\equiv ((2n+1)!)^2 \pmod{p}.$$

Da nach Voraussetzung 17 ein quadratischer Rest modulo p ist, reicht es aus, wenn $(\frac{p-1}{2})!$ kein quadratischer Rest modulo p ist.

Schritt 4: Es ist $\frac{p-1}{2} = 4n + 2$ gerade, also

$$\left((\frac{p-1}{2})!\right)^2 = \prod_{i=-\frac{p-1}{2}, i\neq 0}^{\frac{p-1}{2}} i \equiv (p-1)! \equiv -1 \pmod{p}$$

nach dem Satz von Wilson. Damit ist

$$\left(\frac{(\frac{p-1}{2})!}{p}\right) \equiv \left((\frac{p-1}{2})!\right)^{\frac{p-1}{2}} \equiv (-1)^{\frac{p-1}{4}} \equiv (-1)^{2n+1} \equiv -1 \pmod{p}.$$

Damit ist $(\frac{p-1}{2})!$ kein quadratischer Rest modulo p, und der Nachweis der Lösbarkeit der Reichardt-Lind-Gleichung modulo p ist erbracht.

Teil 3. Arithmetik in quadratischen Zahlkörpern

15. Gausssche ganze Zahlen

Das Beispiel der ganzen Gaußschen Zahlen haben wir bereits implizit in Satz 5.41 kennengelernt, als es um natürliche Zahlen ging, die eine Darstellbarkeit als Summe zweier Quadrate besitzen. Bei dieser Frage hilft die Zerlegung

$$x^2 + y^2 = (x + iy)(x - iy),$$

die wir nun konzeptioneller untersuchen möchten.

15.1. **Arithmetik in $\mathbb{Z}[i]$.** Wir zeigen zunächst, daß der Ring der Gaußschen ganzen Zahlen ein euklidischer Hauptidealring ist. Dazu bestimmen wir die Einheiten. Das Hilfsmittel hierzu ist die Norm.

Definition 15.1. Eine **Gaußsche Zahl** ist eine komplexe Zahl der Form $z = x+yi$ mit $x, y \in \mathbb{Q}$. Eine **ganze Gaußsche Zahl** ist eine Gaußsche Zahl $z = x + yi$ mit $x, y \in \mathbb{Z}$.

Über die algebraischen Eigenschaften der (ganzen) Gaußschen Zahlen gibt die folgende Proposition Auskunft.

Proposition 15.2. *Die Menge der Gaußschen Zahlen ist ein Teikörper von $\mathbb{C}$:*

$$\mathbb{Q}(i) := \{x + yi \in \mathbb{C} \ ; \ x, y \in \mathbb{Q}\}.$$

Als $\mathbb{Q}$-Vektorraum hat $\mathbb{Q}(i)$ die Basis $1, i$: für jedes $z \in \mathbb{Q}(i)$ sind die Koordinaten $x, y \in \mathbb{Q}$ mit $z = x + yi$ eindeutig.

Die Menge der ganzen Gaußschen Zahlen ist ein Unterring in $\mathbb{Q}[i]$:

$$\mathbb{Z}[i] := \{a + bi \ ; \ a, b \in \mathbb{Z}\}.$$

Die zugrundeliegende abelsche Gruppe ist frei mit Basis $1, i$: für jedes $z \in \mathbb{Z}[i]$ gibt es eindeutige $a, b \in \mathbb{Z}$ mit $z = a + bi$.

Beweis. Das ist alles klar bis eventuell auf die Tatsache, daß man in $\mathbb{Q}(i)$ jedes $z = x + iy \neq 0$ invertieren kann. Das Inverse existiert in $\mathbb{C}$ und liegt bereits in $\mathbb{Q}(i)$ wegen

$$\frac{1}{z} = \frac{1}{x + yi} = \frac{x - yi}{(x + yi)(x - yi)} = \frac{x - yi}{x^2 + y^2} = \frac{x}{x^2 + y^2} - \frac{y}{x^2 + y^2} i. \qquad \square$$

Korollar 15.3. *$\mathbb{Z}[i]$ ist ein Integritätsring.*

Beweis. Unterringe von Körpern sind nullteilerfrei und somit Integritätsringe. $\square$

Proposition 15.4. *Die komplexe Konjugation $\mathbb{C} \to \mathbb{C}$, kurz **Konjugation** genannt,*

$$z = x + yi \mapsto \bar{z} := x - yi$$

induziert einen Körperautomorphismus $\mathbb{Q}(i) \to \mathbb{Q}(i)$ und einen Ringautomorphismus

$$\mathbb{Z}[i] \to \mathbb{Z}[i].$$

Beweis. Das ist klar: die komplexe Konjugation ist als $\mathbb{R}$-linearer Automorphismus von $\mathbb{C}$ bekannt und bildet offenbar die Teilmengen $\mathbb{Z}[i] \subseteq \mathbb{Q}(i) \subseteq \mathbb{C}$ wieder bijektiv in sich ab. $\square$

Bemerkung 15.5. Es gilt für $z = x + yi \in \mathbb{Q}(i)$ (bzw. $z = a + bi \in \mathbb{Z}[i]$)

$$z = \bar{z} \iff z = x \in \mathbb{Q} \quad (\text{bzw. } z = a \in \mathbb{Z}).$$

Definition 15.6. Aus der Konjugation entsteht die **Norm** $N : \mathbb{Q}(i) \to \mathbb{Q}$ definiert durch

$$N(z) = z\bar{z}.$$

In der Tat ist $N(z) \in \mathbb{Q}$ wegen

$$\overline{N(z)} = \overline{z\bar{z}} = \bar{z}\bar{\bar{z}} = \bar{z}z = z\bar{z} = N(z).$$

Mit $z = x + yi$ sieht man dies auch wegen

$$N(x + yi) = (x + yi)(x - yi) = x^2 + y^2.$$

Bemerkung 15.7. Hier sind einige Eigenschaften der Norm und $z, w \in \mathbb{Q}(i)$.

(1) Die Norm ist multiplikativ:

$$N(zw) = zw\overline{zw} = zw\bar{z}\bar{w} = z\bar{z} \cdot w\bar{w} = N(z)N(w).$$

(2) Es gilt $N(z) = 0$ genau dann, wenn $z = 0$ gilt: eine Summe reeller Quadrate ist nur 0, wenn jeder Summand 0 ist. Dabei definiert die Norm einen Gruppenhomomorphismus

$$N : \mathbb{Q}(i)^\times \to \mathbb{Q}^\times,$$

der auch als Norm bezeichnet wird.

(3) Die Norm respektiert die ganzzahlige Struktur: für $z = a + bi \in \mathbb{Z}[i]$ ist

$$N(a + bi) = a^2 + b^2 \in \mathbb{Z}.$$

(4) Die Norm nimmt auf $\mathbb{Q}(i)$ und damit erst recht auf $\mathbb{Z}[i]$ nur nichtnegative Werte an.

Proposition 15.8. *Für $z \in \mathbb{Z}[i]$ gilt:*

$$z \text{ ist Einheit in } \mathbb{Z}[i] \iff N(z) \text{ ist Einheit in } \mathbb{Z}.$$

Beweis. Sei z eine Einheit mit Inversem w. Aus $zw = 1$ in $\mathbb{Z}[i]$ folgt $N(z)N(w) = 1$ in $\mathbb{Z}$, und $N(z)$ ist Einheit. Wenn umgekehrt $N(z)$ eine Einheit in $\mathbb{Z}$ ist, dann gibt es $u \in \mathbb{Z}$ mit $N(z)u = 1$. Aber dann folgt

$$z(\bar{z}u) = (z\bar{z})u = N(z)u = 1$$

und z ist Einheit. □

Proposition 15.9. *Die Gruppe der Einheiten $\mathbb{Z}[i]^\times$ besteht aus $\{\pm 1, \pm i\}$ und ist zyklisch von Ordnung 4.*

Beweis. Die Einheiten von $\mathbb{Z}$ sind ± 1. Nach Proposition 15.8 ist $a + bi \in \mathbb{Z}[i]$ Einheit, wenn

$$a^2 + b^2 = \pm 1.$$

Da $a^2 + b^2 \geq 0$ ist, kommt -1 nicht in Frage. Die Gleichung $a^2 + b^2 = 1$ hat nur die ganzzahligen Lösungen

$$(a, b) \in \{(1, 0), (-1, 0), (0, 1), (0, -1)\}$$

entsprechend $a + bi \in \{1, -1, i, -i\}$. Damit wird $\mathbb{Z}[i]^\times$ von i erzeugt und ist zyklisch. □

Bemerkung 15.10. (1) Wie in jedem Integritätsring definieren wir für $z, w \in \mathbb{Z}[i]$ die Teilbarkeit

$$z \mid w \iff \text{es gibt } z' \in \mathbb{Z}[i] \text{ mit } w = zz'.$$

Wenn $z = n \in \mathbb{Z}$ und $w = a + bi \in \mathbb{Z}[i]$, dann ist

$$n \mid a + bi \iff n \mid a \text{ und } n \mid b,$$

wie man sofort sieht: ein $z' = a' + b'i$ mit $zz' = w$ führt zu $a + bi = na' + nb'i$.

(2) Ohne weiteres Wissen zur Arithmetik in $\mathbb{Z}[i]$ können wir mit der Teilbarkeitsrelation nicht viel anfangen. Die guten Eigenschaften der Teilbarkeitsrelation in $\mathbb{Z}$ sind Konsequenzen der Existenz von ggT und kgV und folgen damit letztlich, weil $\mathbb{Z}$ ein Hauptidealring ist.

(3) Wir erinnern an den Begriff der assoziierten Elemente. Das sind hier Elemente $z, w \in \mathbb{Z}[i]$, die sich multiplikativ nur um eine Einheit unterscheiden, also $w \in \{z, -z, iz, -iz\}$.

Satz 15.11. *Der Ring $\mathbb{Z}[i]$ ist euklidisch bezüglich der Norm*

$$N : \mathbb{Z}[i] \to \mathbb{N}_0$$

als euklidischer Normfunktion.

Beweis. Wir müssen zu beliebigem $z \in \mathbb{Z}[i]$ und $0 \neq d \in \mathbb{Z}[i]$ ein q und $r \in \mathbb{Z}[i]$ finden mit

$$z = dq + r$$

und $r = 0$ oder $N(r) < N(d)$. Dazu betrachten wir

$$\frac{z}{d} = x + yi \in \mathbb{Q}(i)$$

und wählen $q = a + bi \in \mathbb{Z}[i]$ derart, daß

$$|x - a| \text{ und } |y - b| \leq \frac{1}{2},$$

indem wir die Koordinaten x, y auf ganze Zahlen runden. Dann gilt für $r = z - dq$

$$\begin{aligned} N(r) &= N(d) \cdot N(\frac{z}{d} - q) = N(d) \cdot N((x-a) + (y-b)i) \\ &= N(d) \cdot ((x-a)^2 + (y-b)^2) \leq N(d)(1/4 + 1/4) < N(d). \end{aligned}$$

□

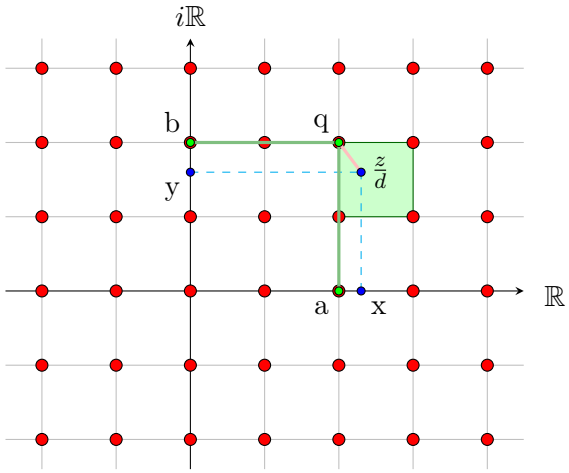


Abbildung 5. Division mit Rest für ganze Gaußsche Zahlen. Pythagoras: das Abstandsquadrat von $\frac{z}{d}$ nach q ist $N(\frac{z}{d} - q) = ((x-a)^2 + (y-b)^2)$.

Korollar 15.12. *Der Ring $\mathbb{Z}[i]$ ist ein Hauptidealring.*

Beweis. Euklidische Ringe sind Hauptidealringe. □

Korollar 15.13. *Der Ring $\mathbb{Z}[i]$ hat eindeutige Primfaktorzerlegung, d.h., $\mathbb{Z}[i]$ ist faktoriell. Es gibt den ggT und das kgV.*

Beweis. Hauptidealringe sind faktoriell und haben damit ggT und kgV. □

Bemerkung 15.14. Die euklidische Division mit Rest im Ring $\mathbb{Z}[i]$ ist nicht nur eine Existenzaussage, sondern sogar algorithmisch bestimmbar. Damit gibt es auch für den Ring $\mathbb{Z}[i]$ einen euklidischen Algorithmus für den ggT.

15.2. **Die Primzahlen der Gaußschen Zahlen.** Nachdem wir nun wissen, daß die Gaußschen ganzen Zahlen als großer Bruder der ganzen Zahlen auch eigene Primzahlen besitzen, wollen wir nun zunächst eine Übersicht über diese erhalten und sodann verstehen, wie Primzahlen p in $\mathbb{Z}$ in Primzahlen in $\mathbb{Z}[i]$ zerfallen.

Proposition 15.15. *Sei $\pi \in \mathbb{Z}[i]$ ein Primelement. Dann gibt genau eine Primzahl $p \in \mathbb{Z}$ mit*

$$\pi \mid p.$$

Für dieses p gilt genau einer der beiden Fälle:

(a) $N(\pi) = p^2$ *und* π *ist zu* p *assoziiert:* $\pi \in \{\pm p, \pm ip\}$.
(b) $N(\pi) = p$ *und* π *und* $\bar{\pi}$ *sind Primelemente, und es gilt* $\pi \cdot \bar{\pi} = p$.

Beweis. Es ist $\pi \mid N(\pi)$. Als Primelement teilt π dann einen der Primfaktoren der Primfaktorzerlegung von $N(\pi)$ als Element in $\mathbb{Z}$. Dies zeigt die Existenz einer Primzahl p mit $\pi \mid p$.

Angenommen $\ell \neq p$ sei eine weitere Primzahl mit $\pi \mid \ell$. Nach dem Lemma von Bézout, Lemma 3.10, gibt es dann $u, v \in \mathbb{Z}$ mit $pu + \ell v = 1$. Dann teilt aber $\pi \mid pu + \ell v = 1$ und π ist Einheit, Widerspruch.

Sei $p = \pi z$ mit $z \in \mathbb{Z}[i]$. Dann ist

$$p^2 = N(p) = N(\pi) \cdot N(z),$$

also $N(\pi)$ ein Teiler in $\mathbb{Z}$ von p^2. Da kommen nur p und p^2 in Frage, weil die Norm stets nichtnegativ ist und aus $N(\pi) = 1$ folgen würde, daß π eine Einheit ist.

- Wenn $N(\pi) = p^2$, dann ist $N(z) = 1$ und z ist eine Einheit. Damit sind π und p assoziiert in diesem Fall.
- Wenn $N(\pi) = p$, dann ist $p = \pi \cdot \bar{\pi}$. Primelemente werden unter Ringautomorphismen wieder auf Primelemente abgebildet, da Primelementsein ausschließlich durch ringtheoretische Eigenschaften festgelegt ist. Damit ist auch $\bar{\pi}$ ein Primelement von $\mathbb{Z}$. □

Wir wollen nun wissen, wie man der Primzahl $p \in \mathbb{Z}$ ansehen kann, wann welcher Fall eintritt. Der Fall (a) in Proposition 15.15 tritt genau dann auf, wenn p auch Primelement in $\mathbb{Z}[i]$ ist. Der Fall (b) unterteilt sich weiter, je nachdem ob π und $\bar{\pi}$ im Prinzip verschiedene oder das gleiche Primelement sind. Im Prinzip bedeutet hier 'bis auf Assoziiertheit', weil das für Teilbarkeitsfragen die entscheidende Klassifizierung ist.

Theorem 15.16 (Primzerlegungsgesetz)**.** *Sei* p *eine Primzahl in* $\mathbb{Z}$. *Dann ist* p *in* $\mathbb{Z}[i]$

$$\begin{array}{lcl} \textit{Produkt zweier assoziierter Primelemente} & \iff & p = 2, \\ \textit{Produkt zweier nicht assoziierter Primelemente} & \iff & p \equiv 1 \pmod 4, \\ \textit{Primelement} & \iff & p \equiv 3 \pmod 4. \end{array}$$

Beweis. Es gilt $N(1+i) = 2$, also $2 = (1+i)(1-i)$ und $(1-i) = -i(1+i)$. Außerdem ist $1+i$ ein Primelement, da irreduzibel: ein $1+i = zw$ bedeutet $2 = N(1+i) = N(z)N(w)$, und z oder w hat Norm 1 und ist Einheit. Damit ist $p = 2$ richtig einsortiert und wir dürfen nun $p \geq 3$ annehmen.

Die Primzahl p ist in $\mathbb{Z}[i]$ keine Einheit. Daher gibt es einen Primfaktor $\pi \mid p$. Für dieses $\pi = a + bi$ ist p die eindeutige Primzahl aus Proposition 15.15. Es gibt also zwei Fälle zu betrachten:

Wenn $p = \pi\bar{\pi} = a^2 + b^2$, dann ist $p \equiv 1 \pmod 4$, weil Quadrate $\equiv 0, 1 \pmod 4$ sind, und wir $p = 2$ schon ausgeschlossen haben. Ist umgekehrt $p \equiv 1 \pmod 4$, dann gibt es nach dem quadratischen Reziprozitätsgesetz, genauer nach Satz 5.37, ein $x \in \mathbb{Z}$ mit $x^2 \equiv -1 \pmod p$. Dann gilt in $\mathbb{Z}[i]$

$$p \mid x^2 + 1 = (x+i)(x-i).$$

Weil p aber in $\mathbb{Z}[i]$ kein Teiler von $x \pm i$ ist, kann p nicht Primelement von $\mathbb{Z}[i]$ sein. Damit scheidet Fall (a) in Proposition 15.15 aus und p zerfällt als Produkt zweier Primelemente.

Es bleibt zu zeigen, daß für $p \equiv 1 \pmod 4$ die beiden Faktoren in $p = \pi\bar{\pi}$ nicht zueinander assoziiert sind. Angenommen $\bar{\pi} = u\pi$ mit u Einheit. Dann ist $u \in \{\pm 1, \pm i\}$ und

$$p = \pi\bar{\pi} = u\pi^2 = u(a^2 - b^2 + 2abi).$$

- Wenn $u = \pm 1$, dann folgt $2ab = 0$ und daher $p = \pm a^2$ oder $\pm b^2$. Das ist keine Primzahl in $\mathbb{Z}$, Widerspruch.
- Wenn $u = \pm i$, dann ist $p = \pm 2ab$ und p gerade. Aber $p = 2$ ist bereits behandelt. □

Bemerkung 15.17. Es folgt aus Proposition 15.15, daß die Primfaktoren von Primzahlen $p \in \mathbb{Z}$, über die das Zerlegunsgesetz, Theorem 15.16, Auskunft gibt, alle Primelemente von $\mathbb{Z}[i]$ beschreiben (natürlich bis auf assoziierte Primelemente).

15.3. **Anwendungen.** Wir beschreiben nun ein paar Anwendungen der Arithmetik in $\mathbb{Z}[i]$. Als erstes erhalten wir einen neuen Beweis des Fermat'schen Zweiquadratesatzes.

Satz 15.18 (Zwei–Quadrate–Satz)**.** *Eine natürliche Zahl n ist Summe zweier Quadrate $n = a^2 + b^2$ mit $a, b \in \mathbb{Z}$ genau dann, wenn jeder Primfaktor $p \mid n$ mit $p \equiv 3 \pmod 4$ mit geradem Exponenten in der Primfaktorzerlegung von n auftritt.*

Beweis. Die Zahl n läßt eine solche Zerlegung zu genau dann, wenn $n = N(a + ib)$ eine Norm eines Elements $a + ib \in \mathbb{Z}[i]$ ist. Um zu bestimmen, welche Zahlen Normen sind, zerlegen wir $a + ib$ als Produkt von Primfaktoren. Es ist $n \in N(\mathbb{Z}[i])$, wenn n Produkt von Normen von Primelementen $N(\pi)$ ist. Aber solche Normen sind nach Proposition 15.15 und Theorem 15.16

$$\begin{aligned} N(1+i) &= (1+i)(1-i) = 2, \\ N(\pi) &= p, \quad \text{wenn } p \equiv 1 \pmod 4, \\ N(\pi) &= p^2, \quad \text{wenn } p \equiv 3 \pmod 4. \end{aligned}$$

□

Es sei an die charakteristische Funktion $\mathbf{1}_{[a]}(n)$ zu einer Restklasse $[a] \in \mathbb{Z}/m\mathbb{Z}$ aus Kapitel §13 erinnert. Wir definieren eine modifizierte Teilersumme zu $[a]$ als

$$d_{a,m}(n) := \sum_{\substack{d \mid n \\ d \equiv a \pmod m}} 1 = \sum_{d \mid n} \mathbf{1}_{[a]}(d)$$

und können dadurch einen Ausdruck für die Darstellungszahlen, das sind die Anzahl von Darstellungen von n als Wert einer quadratischen Form bei ganzzahligen Werten, angeben.

Satz 15.19 (Jacobi 1828)**.** *Sei n eine natürliche Zahl. Die Anzahl der Darstellungen $n = a^2 + b^2$ mit $a, b \in \mathbb{Z}$ ist*

$$4(d_{1,4}(n) - d_{3,4}(n)).$$

Beweis. Der Beweis beruht auf der eindeutigen Primfaktorzerlegung in $\mathbb{Z}[i]$. Wir definieren eine Funktion $r : \mathbb{N} \to \mathbb{C}$ durch

$$r(n) = \frac{1}{4} \cdot \#\{a + ib \in \mathbb{Z}[i] \ ; \ N(a + bi) = n\}$$

Die Funktion r zählt im Prinzip die Anzahl der Lösungen $n = a^2 + b^2$ mit ganzen Zahlen $a, b \in \mathbb{Z}$. Der Faktor $4 = \#\mathbb{Z}[i]^\times$ trägt dazu bei, die 4 assoziierten Lösungen, bei der $a + bi$ mit einer Einheit $u \in \mathbb{Z}[i]^\times$ multipliziert wird, nur einfach zu zählen. Wir müssen also zeigen, daß

$$r(n) = d_{1,4}(n) - d_{3,4}(n).$$

Das Jacobi-Symbol $\left(\frac{-1}{n}\right)$ definiert eine Funktion $\left(\frac{-1}{\ }\right) : \mathbb{N} \to \mathbb{C}$ durch

$$\left(\frac{-1}{n}\right) = \begin{cases} 1 & n \equiv 1 \pmod 4 \\ 0 & 2 \mid n \\ -1 & n \equiv -1 \pmod 4. \end{cases}$$

Dies ist nichts anderes als $\chi_4 : \mathbb{N} \to \mathbb{C}$, die Fortsetzung durch 0 des nichttrivialen Dirichlet-Charakters modulo 4 gegeben durch $\chi_4 : (\mathbb{Z}/4\mathbb{Z})^\times = \pm 1 \to \mathbb{C}^\times$. Daher ist die Funkton $\chi_4(n) = \left(\frac{-1}{n}\right)$ ist vollständig multiplikativ und

$$d_{1,4}(n) - d_{3,4}(n) = \sum_{d \mid n} \chi_4(d) = (\chi_4 * \mathbf{1})(n).$$

Damit ist $d_{1,4}(n) - d_{3,4}(n)$ multiplikativ nach Proposition 7.5.

Wir zeigen nun, daß auch die Funktion $r(n)$ multiplikativ ist. Seien n und m teilerfremde natürliche Zahlen. Die Abbildung von Teilmengen von $\mathbb{Z}[i]$

$$\begin{aligned}\{a+bi\ ;\ N(a+bi)=n\}\times\{c+di\ ;N(c+di)=m\} &\to \{x+yi\ ;N(x+yi)=nm\}\\ (z,w) &\mapsto x+yi=zw\end{aligned}$$

ist aufgrund der eindeutigen Primfaktorzerlegung $4:1$, denn die Zerlegung in Faktoren mit Norm n respektive m ist eindeutig bis auf eine Einheit als Faktor. Damit ergeben genau die 4 Paare $(uz, u^{-1}w)$ mit $u\in\mathbb{Z}[i]^\times$ das gleiche Produkt. Dies zeigt, daß

$$4\cdot r(nm)=\frac{4r(n)\cdot 4r(m)}{4},$$

und somit ist $r(n)$ multiplikativ.

Um den Satz zu beweisen, dürfen wir nun annehmen, daß $n=p^m$ eine Primzahlpotenz ist.

- Wenn $p=2$ ist, dann ist $a+bi=u\cdot(1+i)^m$ mit $u\in\mathbb{Z}[i]^\times$. Die Darstellung ist also eindeutig bis auf die Wahl von u und
$$r(2^m)=1=d_{1,4}(2^m)-d_{3,4}(2^m)$$
- Wenn $p\equiv 3 \pmod 4$, dann ist $a+bi=u\cdot p^{m/2}$ und das geht genau für $2\mid m$ und dann auch nur eindeutig bis auf die Wahl von u. Es folgt
$$r(p^m)=\left\{\begin{array}{ll}1 & 2\mid m\\ 0 & 2\nmid m\end{array}\right\}=d_{1,4}(p^m)-d_{3,4}(p^m).$$
- Wenn $p\equiv 1 \pmod 4$, dann gibt es zwei verschiedene nichtassoziierte Primelemente von Norm p, nämlich π und $\bar\pi$ mit $N(\pi)=\pi\cdot\bar\pi=p$. Dann folgt
$$a+bi=u\cdot\pi^\alpha\cdot\bar\pi^{m-\alpha}$$
und davon gibt es $m+1$-viele Möglichkeiten bis auf Multiplikation mit $u\in\mathbb{Z}[i]^\times$. Es folgt
$$r(p^m)=m+1=d_{1,4}(p^m)-d_{3,4}(p^m).$$ □

Beispiel 15.20. Als nächste Anwendung diskutieren wir die Diophantische Gleichung

$$Y^2=X^3-1,$$

also die ganzzahligen Lösungen dieser Gleichung. Sei $(x,y)\in\mathbb{Z}^2$ eine Lösung. Mit Hilfe der ganzen Gaußschen Zahlen $\mathbb{Z}[i]$ finden wir die Faktorisierung

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

Sei π ein Primelement von $\mathbb{Z}[i]$, das den ggT der Faktoren $y+i$ und $y-i$ teilt. Dann teilt

$$\pi\mid i\big((y-i)-(y+i)\big)=2.$$

Damit ist bis auf Assoziierte $\pi=1+i$ das einzig mögliche Primelement, das den ggT der Faktoren teilt. Aber aus

$$2=N(\pi)\mid N(y+i)=y^2+1=x^3$$

folgt x gerade, und weiter

$$y^2+1\equiv 0 \pmod 4.$$

Das ist nicht möglich, weil 0 und 1 die einzigen quadratischen Reste modulo 4 sind.

Die Faktoren $y+i$ und $y-i$ sind daher teilerfremd in $\mathbb{Z}[i]$. Sei $x=\prod_{\alpha=1}^r\pi_\alpha^{m_\alpha}$ die Primfaktorzerlegung von x in $\mathbb{Z}[i]$. Dann ist

$$(y+i)(y-i)=x^3=\prod_{\alpha=1}^r\pi_\alpha^{3m_\alpha}.$$

Wegen der Eindeutigkeit der Primfaktorzerlegung in $\mathbb{Z}[i]$ müssen sich die Primfaktoren von x^3 zur Primfaktorzerlegung der Faktoren $y\pm i$ aufteilen. Da $y+i$ und $y-i$ teilerfremd sind, kommt

mit einem Faktor π_α automatisch die ganze Potenz $\pi_\alpha^{3m_\alpha}$ zum selben Faktor. Wir schließen, daß es nach entsprechendem Umnummerieren ein $1 \le s \le r$ und ein $u \in \mathbb{Z}[i]^\times$ gibt, so daß

$$y + i = u \prod_{\alpha=1}^{s} \pi_\alpha^{3m_\alpha}$$
$$y - i = u^{-1} \prod_{\alpha=s+1}^{r} \pi_\alpha^{3m_\alpha}.$$

Damit sind die Faktoren $y \pm i$ bis auf eine Einheit u dritte Potenzen in $\mathbb{Z}[i]$. Sei also $a + bi \in \mathbb{Z}[i]$ mit

$$y + i = u(a + bi)^3 = u(a^3 - 3ab^2 + (3a^2b - b^3)i).$$

- Wenn $u = \pm 1$, dann folgt per Koeffizientenvergleich
$$b(3a^2 - b^2) = \pm 1,$$
somit $b = \pm 1$ und $3a^2 = \pm 1 + b^2 = \pm 1 + 1$. Dies geht nur für $a = 0$ und $b = \pm 1$.
- Wenn $u = \pm i$, dann folgt per Koeffizientenvergleich
$$a(a^2 - 3b^2) = \pm 1,$$
somit $a = \pm 1$ und $3b^2 = a^2 - (\pm 1) = 1 - (\pm 1)$. Dies geht nur für $b = 0$ und $a = \pm 1$.

In jedem Fall ist $a + bi$ selbst eine Einheit. Folglich sind $y \pm i$ auch Einheiten, somit

$$y = 0$$

und dann $x = 1$. wir haben damit gezeigt, daß es außer der Lösung

$$(x, y) = (1, 0)$$

keine weiteren ganzzahligen Lösungen gibt.

Als dritte Anwendung parametrisieren wir pythagoräische Tripel. Dazu benötigen wir als erstes einen Spezialfall von **Hilberts Satz 90**. Varianten eines Satzes aus Hilberts berühmten Zahlbericht, der dort die Nummer 90 hat, werden mit dieser merkwürdigen, aber üblichen Terminologie bezeichnet.

Satz 15.21 (Hilbert 90 für $\mathbb{Q}(i)/\mathbb{Q}$)**.** *Ein $z \in \mathbb{Q}(i)$ hat $N(z) = 1$ genau dann, wenn es ein $w \in \mathbb{Q}[i]$ gibt mit*

$$z = w/\bar{w}.$$

Beweis. Wenn $z = w/\bar{w}$, dann ist

$$N(z) = N(w/\bar{w}) = N(w)/N(\bar{w}) = (w\bar{w})/(\bar{w}\bar{\bar{w}}) = 1.$$

Sei umgekehrt z von Norm 1. Dann setzen wir $w = 1 + z$ und rechnen

$$\frac{w}{\bar{w}} = \frac{1+z}{1+\bar{z}} = \frac{(1+z)z}{(1+\bar{z})z} = \frac{(1+z)z}{z + N(z)} = \frac{(1+z)z}{z+1} = z.$$

Bleibt einzig der Fall $z = -1$, der zu $w = 0$ führt und die obige Rechnung fehlerhaft macht. Hier hilft uns $w = i$

$$\frac{i}{\bar{\bar{i}}} = \frac{i}{-i} = -1. \qquad \square$$

Definition 15.22. Ein **pythagoräisches Tripel** ist ein Tripel (a, b, c) natürlicher Zahlen mit

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

Ein **primitives pythagoräisches Tripel** ist ein **pythagoräisches Tripel** (a, b, c) mit

(i) $2 \mid b$,

(ii) a, b, c sind paarweise teilerfremd.

Bemerkung 15.23. (1) Unsere Definition schließt die offensichtlichen ganzzahligen Lösungen mit $abc = 0$ aus.

(2) Mit (a, b, c) ist auch (b, a, c) ein pythagoräisches Tripel. Wenn beide a und b ungerade sind, folgt das unmögliche $c^2 \equiv 2 \pmod 4$. Es ist daher wenigstens einer der beiden a und b gerade. Durch Vertauschen darf man annehmen, daß b gerade ist.

(3) Wenn zwei der drei Zahlen a, b, c einen gemeinsamen Teiler d haben, dann teilt d jede davon wegen $a^2 + b^2 = c^2$. Durch Kürzen von d erhält man ein neues pythagoräisches Tripel mit dann paarweise teilerfremden a, b, c.

Wir sind daher nur an **primitiven pythagoräischen Tripeln** interessiert. Aus diesen entstehen alle pythagoräischen Tripel durch Skalieren und Vertauschen von a und b.

Satz 15.24. *Die Menge der primitiven pythagoräischen Tripel wird parametrisiert vermöge*

$$a = x^2 - y^2, \quad b = 2xy, \quad c = x^2 + y^2$$

mit teilerfremden ganzen Zahlen $x > y > 0$, die nicht beide ungerade sind.

Beweis. Zu einem primitiven pythagoräischen Tripel $(a, b, c) \in \mathbb{N}^3$ betrachten wir

$$z = \frac{a}{c} + \frac{b}{c} i \in \mathbb{Q}(i)$$

mit

$$N(z) = \frac{N(a + bi)}{N(c)} = \frac{a^2 + b^2}{c^2} = 1.$$

Nach Satz 15.21 gibt es ein $w \in \mathbb{Q}(i)$ mit $z = \frac{w}{\bar{w}}$. Wenn $\lambda \in \mathbb{Q}^\times$, dann ist

$$\frac{w}{\bar{w}} = \frac{\lambda w}{\overline{\lambda w}}.$$

Wir dürfen also w so skalieren, daß $w = x + yi$ mit teilerfremden ganzen Zahlen x, y und $x \geq 0$. Es gilt dann

$$z = \frac{x + yi}{x - yi} = \frac{x^2 - y^2}{x^2 + y^2} + \frac{2xy}{x^2 + y^2} i,$$

also mit der linken Seite in gekürzter Form

$$\begin{aligned} \frac{a}{c} &= \frac{x^2 - y^2}{x^2 + y^2} \\ \frac{b}{c} &= \frac{2xy}{x^2 + y^2}. \end{aligned} \tag{15.1}$$

Weil $(x, y) = 1$, ist $(x^2 + y^2, 2xy) \mid 2$. Der Fall $(x^2 + y^2, 2xy) = 2$ trifft genau dann ein, wenn x und y beide ungerade sind. In diesem Fall wäre b ungerade, da die 2 in $2xy/(x^2 + y^2)$ dann gekürzt werden könnte. Da aber b gerade ist, muß $x^2 + y^2$ bereits zu $2xy$ teilerfremd und nicht beide x, y ungerade sein. Dann ist aber $x^2 + y^2$ ungerade und

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

Damit ist auch die rechte Seite in (15.1) gekürzt, woraus die behauptete Parametrisierung

$$a = x^2 - y^2, \quad b = 2xy, \quad c = x^2 + y^2 \tag{15.2}$$

folgt. Die Nebenbedingungen $a, b > 0$ erzwingen $x > y > 0$, weil wir sowieso $x \geq 0$ gewählt haben. Damit haben wir gezeigt, daß (15.2) eine surjektive Abbildung

$$\left\{ (x, y) \in \mathbb{N}^2 \ ; \quad \begin{matrix} x > y > 0 \\ x, y \text{ teilerfremd} \\ \text{nicht beide ungerade} \end{matrix} \right\} \to \left\{ (a, b, c) \ ; \quad \begin{matrix} \text{primitives} \\ \text{pyth. Tripel} \end{matrix} \right\}$$

definiert. Wir müssen nun die Eindeutigkeit zeigen. Sei (x', y') mit den genannten Nebenbedingungen für (x, y) eine weitere Parametrisierung des primitiven pythagoräischen Tripels (a, b, c). Wir setzen $w = x + yi$ und $w' = x' + y'i$ und finden

$$\frac{w}{\bar{w}} = \frac{a}{c} + \frac{b}{c}i = \frac{w'}{\bar{w}'}$$

Daraus folgt

$$\frac{w'}{w} = \frac{\bar{w}'}{\bar{w}}$$

Daher ist $\lambda = \frac{w'}{w} \in \mathbb{Q}^\times$, oder

$$x' + y'i = \lambda x + \lambda y i.$$

Weil x, y teilerfremd sind, darf λ keinen nicht-trivialen Primfaktor im Nenner haben, und weil x', y' teilerfremd sind auch keinen im Zähler. Daher ist $\lambda = \pm 1$. Weil $x, y, x', y' > 0$ folgt $\lambda = 1$ und die Eindeutigkeit ist gezeigt. □

16. Quadratische Zahlkörper

Die Gaußschen Zahlen $\mathbb{Q}(i)$ beschreiben die Algebra von $\sqrt{-1}$. Die Arithmetik der *ganzen* Gaußschen Zahlen hilft bei Fragen zu Werten der quadratische Form

$$x^2 + y^2 = (x + iy)(x - iy).$$

16.1. **Quadratische Zahlkörper.** Nun ersetzen wir -1 durch ein beliebiges $d \in \mathbb{Z}$ und sehen, was wir bekommen.

Notation 16.1. Sei $d \in \mathbb{Z}$, $d \neq 0$ eine ganze Zahl. Von den zwei Quadratwurzeln von d in $\mathbb{C}$ wählen wir $\sqrt{d} \in \mathbb{C}$ wie folgt aus:

$$\sqrt{d} = \begin{cases} \sqrt{d} > 0 & \text{wenn } d > 0 \text{ und daher } \sqrt{d} \in \mathbb{R}, \\ i\sqrt{-d} & \text{wenn } d < 0 \text{ mit } \sqrt{-d} > 0. \end{cases}$$

Proposition 16.2. *Sei $d \in \mathbb{Z}$ kein Quadrat einer ganzen Zahl. Der von 1 und $\sqrt{d}$ aufgespannte $\mathbb{Q}$-Untervektorraum von $\mathbb{C}$*

$$\mathbb{Q}(\sqrt{d}) = \{z = x + y\sqrt{d} \ ; \ x, y \in \mathbb{Q}\}$$

ist ein Unterkörper von $\mathbb{C}$ mit $\mathbb{Q}$-Basis $1, \sqrt{d}$.

Beweis. Die folgenden Rechnungnen zeigen, daß $\mathbb{Q}(\sqrt{d})$ ein Unterring von $\mathbb{C}$ ist:

$$(x + y\sqrt{d}) + (z + w\sqrt{d}) = (x + z) + (y + w)\sqrt{d},$$
$$(x + y\sqrt{d}y) \cdot (z + w\sqrt{d}) = (xw + dyw) + (xw + zy)\sqrt{d}.$$

Weil d kein Quadrat in $\mathbb{Z}$ ist, gilt auch $\sqrt{d} \notin \mathbb{Q}$ nach Korollar 5.19. Damit sind $1, \sqrt{d}$ linear unabhängig über $\mathbb{Q}$ und $\dim_{\mathbb{Q}}(\mathbb{Q}(\sqrt{d}) = 2$.

Es fehlt noch die Existenz eines Inversen Elements in $\mathbb{Q}(\sqrt{d})$. Das inverse Element existiert in $\mathbb{C}$, und genauer besteht unsere Aufgabe darin zu zeigen, daß das Inverse aus $\mathbb{C}$ schon in $\mathbb{Q}(\sqrt{d})$ liegt. Das folgt aus derselben Rechnung wie im Fall der Gaußschen Zahlen:

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

Diese Rechnung ist gültig (keine 0 im Nenner), weil mit $x + y\sqrt{d} \neq 0$ auch $x - y\sqrt{d} \neq 0$. Andernfalls wäre $\sqrt{d} \in \mathbb{Q}$ und das haben wir bereits ausgeschlossen. □

Beispiel 16.3. (1) Die Gaußschen Zahlen $\mathbb{Q}(i)$ erhält man als Spezialfall $d = -1$.

(2) Wenn $D = n^2 d$ für ein $n \in \mathbb{Z}$, $n \neq 0$, dann ist $\mathbb{Q}(\sqrt{D}) = \mathbb{Q}(\sqrt{d})$, denn für $x, y \in \mathbb{Q}$ gilt
$$x + y\sqrt{D} = x + ny\sqrt{d}.$$

Definition 16.4. (1) Ein **quadratischer Zahlkörper** K ist ein Unterkörper $\mathbb{Q} \subseteq K \subseteq \mathbb{C}$, der als $\mathbb{Q}$-Vektorraum die Dimension 2 hat.
(2) Ein **reell-quadratischer Zahlkörper** ist ein quadratischer Zahlkörper $K \subseteq \mathbb{C}$, der sogar $K \subseteq \mathbb{R}$ erfüllt.
(3) Ein **imaginär-quadratischer Zahlkörper** ist ein quadratischer Zahlkörper $K \subseteq \mathbb{C}$, der nicht in $\mathbb{R}$ enthalten ist.

Bemerkung 16.5. Sei $d \in \mathbb{Z}$ kein Quadrat in $\mathbb{Z}$. Der quadratische Zahlkörper $\mathbb{Q}(\sqrt{d})$ ist
(1) reell-quadratisch $\iff d > 0$,
(2) imaginär-quadratisch $\iff d < 0$.

Die beiden Fälle reell- bzw. imaginär-quadratische Zahlkörper unterscheiden sich in manchen Aspekten grundsätzlich. Wir werden also im Folgenden öfters diese Dichotomie sehen.

Proposition 16.6. *Sei $d \in \mathbb{Z}$ kein Quadrat in $\mathbb{Z}$, und sein $K = \mathbb{Q}(\sqrt{d})$. Die **Konjugation** auf K ist definiert für alle $x, y \in \mathbb{Q}$ durch*
$$z = x + y\sqrt{d} \mapsto \widetilde{z} := x - y\sqrt{d}.$$
Die Konjugation ist der einzige nichttriviale Körperautomorphismus von K. Insbesondere ist die Konjugation unabhängig von der Wahl von d mit $K = \mathbb{Q}(\sqrt{d})$.

Beweis. Man rechnet sofort nach, daß Addition und Multiplikation mit der Konjugation verträglich sind. Außerdem ist das Bild der $\mathbb{Q}$-Basis $1, \sqrt{d}$ unter Konjugation die $\mathbb{Q}$-Basis $1, -\sqrt{d}$. Weil die Konjugation $\mathbb{Q}$-linear ist, ist sie folglich bijektiv.

Jeder Körperautomorphismus fixiert $1 \in K$ und daher alle $z \in \mathbb{Q} \subseteq K$. Insbesondere ist ein Körperautomorphismus eine $\mathbb{Q}$-lineare Abbildung.

Weiter muß ein Körperautomorphismus $\sqrt{d}$ auf eine Nullstelle des Polynoms $T^2 - d$ abbilden. Das ist $\sqrt{d}$ und dann haben wir die Identität, oder $-\sqrt{d}$ und dann haben wir die Konjugation, denn der Effekt auf der $\mathbb{Q}$-Basis bestimmt den Körperautomorphismus eindeutig. □

Bemerkung 16.7. (1) Wenn K imaginär-quadratisch ist, dann ist die Konjugation die Einschränkung der komplexen Konjugation: mit $d < 0$ gilt
$$\widetilde{\sqrt{d}} = -\sqrt{d} = -i\sqrt{-d} = \overline{i\sqrt{-d}} = \overline{\sqrt{d}}.$$
Wenn K reell-quadratisch ist, dann ist die komplexe Konjugation auf K trivial.
(2) Sei K ein quadratischer Zahlkörper. Dann gilt für $z \in K$:
$$\widetilde{z} = z \iff z \in \mathbb{Q}.$$

Satz 16.8. *Quadratische Zahlkörper haben die folgende Klassifikation.*
(1) Jeder quadratische Zahlkörper ist von der Form
$$\mathbb{Q}(\sqrt{d}) \subseteq \mathbb{C}$$
für ein quadratfreies $d \in \mathbb{Z}$, $d \neq 0, 1$.
(2) Seien $d, e \in \mathbb{Z}$ keine Quadrate in $\mathbb{Z}$. Dann gilt
$$\mathbb{Q}(\sqrt{d}) = \mathbb{Q}(\sqrt{e}) \iff e/d \text{ ist ein Quadrat in } \mathbb{Q}.$$
Wenn d und e quadratfrei sind, dann ist dies äquivalent zu $e = d$.

Beweis. (1) Sei $1, \alpha$ eine Basis des quadratischen Zahlkörpers $K \subseteq \mathbb{C}$ als $\mathbb{Q}$-Vektrorraum. Dann ist $\alpha^2 \in K$ eine $\mathbb{Q}$-Linearkombination von 1 und α: es gibt $x, y \in \mathbb{Q}$ mit
$$\alpha^2 = x\alpha + y.$$

Folglich ist α Nullstelle eines quadratischen Polynoms mit Koeffizienten aus $\mathbb{Q}$, nämlich

$$T^2 - xT - y \in \mathbb{Q}[T].$$

Durch quadratisches Ergänzen und Skalieren mittels einer lineare Substitution $\beta = q(\alpha - \frac{x}{2})$ erhält man eine $\mathbb{Q}$-Basis $1, \beta$ von K mit

$$\beta^2 = q^2(y + \frac{x^2}{4}) =: d,$$

wobei $d \in \mathbb{Z}$ quadratfrei ist bei richtiger Wahl von $q \in \mathbb{Q}$. Damit ist $K = \mathbb{Q}(\sqrt{d})$ wie behauptet. Die Fäll $d = 0$ und 1 treten nicht auf, weil sonst $1, \beta$ keine $\mathbb{Q}$-Basis von K sein kann.

(2) Wenn e/d ein Quadrat ist, folgt die Gleichheit $\mathbb{Q}(\sqrt{d}) = \mathbb{Q}(\sqrt{e})$ aus Beispiel 16.3.

Wir nehmen nun $\mathbb{Q}(\sqrt{d}) = \mathbb{Q}(\sqrt{e})$ an. Die Konjugation ist eindeutig und hat als $\mathbb{Q}$-linearer Endomorphismus einen eindimensionalen Eigenraum zum Eigenwert -1. Wegen $\widetilde{\sqrt{d}} = -\sqrt{d}$ und $\widetilde{\sqrt{e}} = -\sqrt{e}$ sind beide Wurzeln im Eigenraum zum Eigenwert -1 und unterscheiden sich daher nur um ein rationales Vielfaches: $\sqrt{e} = q\sqrt{d}$ mit $q \in \mathbb{Q}$. Daraus folgt durch Quadrieren die Behauptung. □

Definition 16.9. Sei K ein quadratischer Zahlkörper.

(1) Die **Norm** $N = N_{K/\mathbb{Q}} : K \to \mathbb{Q}$ ist definiert für $z \in K$ durch

$$N(z) = z \cdot \widetilde{z}.$$

(2) Die **Spur** (trace) $\mathrm{tr} = \mathrm{tr}_{K/\mathbb{Q}} : K \to \mathbb{Q}$ ist definiert für $z \in K$ durch

$$\mathrm{tr}(z) = z + \widetilde{z}.$$

Bemerkung 16.10. Hier sind einige Eigenschaften von Norm uns Spur.

(1) Norm und Spur sind invariant unter Konjugation und daher wohldefiniert wegen Bemerkung 16.7.

(2) In Koordinaten $z = x + y\sqrt{d} \in \mathbb{Q}(\sqrt{d})$ gilt

$$N(z) = (x + y\sqrt{d}) \cdot (x - y\sqrt{d}) = x^2 - dy^2$$
$$\mathrm{tr}(z) = (x + y\sqrt{d}) + (x - y\sqrt{d}) = 2x.$$

(3) Die Norm ist multiplikativ:

$$N(zw) = (zw)\widetilde{zw} = zw\widetilde{z}\widetilde{w} = (z\widetilde{z})(w\widetilde{w}) = N(z)N(w).$$

Die Spur ist offensichtlich eine $\mathbb{Q}$-Linearform.

(4) Es gilt $N(z) = 0$ genau dann, wenn $z = 0$. Zu $z \neq 0$ gibt es z^{-1} und $N(z)N(z^{-1}) = N(1) = 1$, also ist $N(z) \neq 0$. Damit induziert die Norm einen Gruppenhomomorphismus

$$N : K^\times \to \mathbb{Q}^\times.$$

Proposition 16.11. *Sei K ein quadratischer Zahlkörper. Sei $z \in K$ ein beliebiges Element. Dann ist z Nullstelle des rationalen quadratischen Polynoms*

$$T^2 - \mathrm{tr}(z)T + N(z) \in \mathbb{Q}[T].$$

Beweis. Das folgt aus der offensichtlichen Rechnung

$$z^2 - (z + \widetilde{z})z + z\widetilde{z} = 0.$$ □

16.2. **Der Ganzzahlring.** Um Arithmetik (Zahlentheorie) in einem quadratischen Zahlkörper betreiben zu können, braucht es eine Verallgemeinerung der ganzen Zahlen.

Notation 16.12. Seien R ein Ring und $\alpha \in R$ ein Element.

(1) Die Auswertung von α im ganzzahligen Polynom
$$P(X) = a_0X^d + a_1X^{d-1} + \ldots + a_{d-1}X + a_d \in \mathbb{Z}[X],$$
also mit Koeffizienten $a_i \in \mathbb{Z}$ für alle $0 \leq i \leq d$, ist das Ringelement
$$P(\alpha) = a_0\alpha^d + a_1\alpha^{d-1} + \ldots + a_{d-1}\alpha + a_d \in R.$$

(2) Der von $\alpha \in R$ erzeugte Unterring wird mit $\mathbb{Z}[\alpha] \subseteq R$ bezeichnet. Es gilt
$$\mathbb{Z}[\alpha] = \{P(\alpha)\ ;\ P \in \mathbb{Z}[X]\}.$$
Dies ist ein Unterring, denn für Polynome $P, Q \in \mathbb{Z}[X]$ gilt
$$(P+Q)(\alpha) = P(\alpha) + Q(\alpha), \quad \text{und} \quad (P \cdot Q)(\alpha) = P(\alpha)Q(\alpha).$$
Genauer ist $\mathbb{Z}[\alpha]$ der kleinste Unterring von R, der α enthält.

Proposition 16.13. *Sei $d \in \mathbb{Z}$ kein Quadrat in $\mathbb{Z}$. Es gilt*
$$\mathbb{Z}[\sqrt{d}] = \{a + b\sqrt{d}\ ;\ a, b \in \mathbb{Z}\}.$$
Die zugrundeliegende Gruppe ist frei mit Basis $1, \sqrt{d}$: für jedes $z \in \mathbb{Z}[\sqrt{d}]$ gibt es eindeutige a, b mit $z = a + b\sqrt{d}$.

Beweis. Die Elemente $a+b\sqrt{d}$ mit $a, b \in \mathbb{Z}$ gehören zu $\mathbb{Z}[\sqrt{d}]$. Wir müssen zeigen, daß die Menge dieser Elemente abgeschlossen ist unter Addition und Multiplikation. Das ist klar. □

Beispiel 16.14. Sei $d \in \mathbb{Z}$ kein Quadrat in $\mathbb{Z}$, und sei $n \in \mathbb{Z}$, $n \neq 0$. Dann ist
$$\mathbb{Z}[\sqrt{dn^2}] \subseteq \mathbb{Z}[\sqrt{d}],$$
genauer besteht $\mathbb{Z}[\sqrt{dn^2}]$ aus den $a + b\sqrt{d}$, $a, b \in \mathbb{Z}$ mit $n \mid b$.

Definition 16.15. Ein Element $\alpha \in R$ eines Rings R ist **ganz algebraisch über** $\mathbb{Z}$, wenn es ein Polynom $P(X) \in \mathbb{Z}[X]$ mit ganzen Koeffizienten, Leitkoeffizient 1 und $P(\alpha) = 0$ gibt, d.h. es gibt $a_i \in \mathbb{Z}$ mit
$$\alpha^d + a_1\alpha^{d-1} + \ldots + a_{d-1}\alpha + a_d = 0.$$
Ein solches Polynom $P(X)$ wird auch **Ganzheitsrelation** für α genannt.

Beispiel 16.16. Die wichtigsten Beispiele ganzer algebraischer Zahlen sind die folgenden.

(1) Die Elemente von $\mathbb{Z}$ sind ganz algebraisch über $\mathbb{Z}$, denn $a \in \mathbb{Z}$ ist Nullstelle von $X - a$.

(2) Sei $d \in \mathbb{Z}$ kein Quadrat in $\mathbb{Z}$. Die Elemente $\alpha = a + b\sqrt{d}$ mit $a, b \in \mathbb{Z}$ sind ganz über $\mathbb{Z}$, denn
$$\alpha^2 - 2a\alpha + a^2 - db^2 = (\alpha - a)^2 - db^2 = 0.$$

(3) Sei $d \equiv 1 \pmod 4$. Dann ist $\omega = \frac{1+\sqrt{d}}{2}$ ganz über $\mathbb{Z}$, denn $\frac{1-d}{4} \in \mathbb{Z}$ und
$$\omega^2 - \omega + \frac{1-d}{4} = 0.$$

(4) Der goldene Schnitt $\varphi = \frac{1+\sqrt{5}}{2}$ ist Nullstelle von $X^2 - X - 1 = 0$ und damit ganz über $\mathbb{Z}$.

Satz 16.17. *Sei R ein Unterring eines Körpers K und $\alpha \in R$. Dann sind äquivalent:*

(a) Das Element α ist ganz algebraisch über $\mathbb{Z}$.

(b) Der Unterring $\mathbb{Z}[\alpha] \subseteq R$ ist als abelsche Gruppe endlich erzeugt.

(c) Es gibt eine endlich erzeugte abelsche Untergruppe $M \subseteq R$, $M \neq 0$, mit
$$\alpha \cdot M \subseteq M.$$

Beweis. (a) $\Longrightarrow$ (b): Sei $P(X) = X^d + a_1X^{d-1} + \ldots + a_{d-1}X + a_d$ eine Ganzheitsrelation für α. Sei $d = \deg(P)$ der Grad von P. Dann gilt

$$\mathbb{Z}[\alpha] = \langle 1, \ldots, \alpha^{d-1}\rangle_{\mathbb{Z}}.$$

Es reicht offenbar $\alpha^n \in \langle 1, \ldots, \alpha^{d-1}\rangle_{\mathbb{Z}}$ für alle n zu zeigen. Dies zeigen wir per Induktion. Für $n < d$ ist nichts zu zeigen. Für $n \geq d$ gilt durch Auswertung von $X^{n-d}P(X)$ in α:

$$\alpha^n = -a_1\alpha^{n-1} - \ldots - a_{d-1}\alpha^{n-d+1} - a_0\alpha^{n-d}$$

und das ist per Induktionsannahme bereits in $\langle 1, \ldots, \alpha^{d-1}\rangle_{\mathbb{Z}}$.

(b) $\Longrightarrow$ (c): Das ist trivial: $M = \mathbb{Z}[\alpha]$ funktioniert.

(c) $\Longrightarrow$ (a): Seien $v_1, \ldots, v_n \in M$ Erzeuger als abelsche Gruppe. Nach Voraussetzung gibt es $a_{ij} \in \mathbb{Z}$, $1 \leq i, j \leq n$ mit

$$\alpha \cdot v_i = \sum_{j=1}^{n} a_{ij}v_j \qquad \text{für alle } 1 \leq i \leq n.$$

Sei $E \in \mathrm{M}_n(K)$ die Einheitsmatrix. Wir betrachten die Matrix $A = \alpha \cdot E - (a_{ij})$ und den Vektor $v = (v_1, \ldots, v_n)^t$. Dann gilt

$$Av = 0.$$

Mit der adjunkten Matrix $A^{\#}$, mit welcher die Relation $A^{\#}A = \det(A) \cdot E$ gilt, folgt

$$0 = A^{\#}Av = \det(A) \cdot v \in K^n.$$

Weil M nichttrivial ist, gibt es $v_i \neq 0$ und damit folgt $\det(A) = 0$. Jetzt berechnen wir $\det(A)$ mit der Leibnitz-Formel und sortieren nach Potenzen von α, die auftreten. Wir erhalten ein Polynom $P(X) \in \mathbb{Z}[X]$ mit Leitkoeffizient 1, Grad n und $P(\alpha) = \det(A) = 0$: nur die Hauptdiagonale mit dem Beitrag $(\alpha - a_{11}) \cdot \ldots \cdot (\alpha - a_{nn})$ liefert einen Beitrag vom höchsten Grad n, und dieser hat offenbar Leitkoeffizient 1. Dieses $P(X)$ ist die gesuchte Ganzheitsrelation. □

Korollar 16.18. *Sei R ein Unterring eines Körpers, der als abelsche Gruppe endlich erzeugt ist. Dann sind alle Elemente von R ganz algebraisch über $\mathbb{Z}$.*

Beweis. Satz 16.17 (c) mit $M = R$. □

Korollar 16.19. *Sei R ein Unterring eines Körpers. Summen und Produkte über $\mathbb{Z}$ ganz algebraischer Elemente aus R sind wieder ganz algebraisch über $\mathbb{Z}$.*

Beweis. Seien $\alpha, \beta \in R$ ganz algebraisch über $\mathbb{Z}$, und zwar α mit einer Ganzheitsrelation vom Grad d und β mit einer vom Grad e. Dann ist wird der von α und β in R erzeugte Unterring $\mathbb{Z}[\alpha, \beta] \subseteq R$ als abelsche Gruppe von den Monomen

$$\alpha^i\beta^j$$

mit $0 \leq i < d$ und $0 \leq j < e$ aufgespannt. Nach Korollar 16.18 sind alle Elemente von $\mathbb{Z}[\alpha, \beta]$ ganz algebraisch über $\mathbb{Z}$. Dazu gehören insbesondere $\alpha + \beta$ und $\alpha\beta$. □

Korollar 16.20. *Sei R ein Unterring eines Körpers. Die Menge*

$$\mathfrak{o} = \{\alpha \in R \ ; \ \alpha \textit{ ganz algebraisch über } \mathbb{Z}\}$$

ist ein Unterring $\mathfrak{o} \subseteq R$.

Beweis. Das folgt sofort aus Korollar 16.19. □

Definition 16.21. Der Ring der **ganzen algebraischen Zahlen** im quadratischen Zahlkörper K ist der Ring

$$\mathfrak{o}_K = \{\alpha \in K \ ; \ \alpha \text{ ganz algebraisch über } \mathbb{Z}\}.$$

Proposition 16.22. *Sei K ein quadratischer Zahlkörper. Es gilt $\mathfrak{o}_K \cap \mathbb{Q} = \mathbb{Z}$.*

Beweis. Sei $z = x/y \in \mathbb{Q} \cap \mathfrak{o}_K$ mit teilerfremden ganzen Zahlen x, y eine ganze algebraische rationale Zahl. Sei $P(X) = X^d + a_1X^{d-1} + \ldots + a_{d-1}X + a_d$ eine Ganzheitsrelation für z. Dann gilt

$$0 = y^d P(\frac{x}{y}) = x^d + y(a_1x^{d-1} + a_2x^{d-2}y + \ldots + a_{d-1}xy^{d-2} + a_dy^{d-1}).$$

Daher teilt $y \mid x^d$ und das ist für teilerfremde x, y nur mit $y = \pm 1$ möglich. Somit liegt $z = x/y$ bereits in $\mathbb{Z}$. □

Proposition 16.23. *Sei K ein quadratischer Zahlkörper. Dann ist $z \in \mathfrak{o}_K$ genau dann, wenn $N(z)$ und $\operatorname{tr}(z)$ ganze Zahlen sind.*

Beweis. Sei $z \in \mathfrak{o}_k$ ein ganz algebraisches Element, und sei $P(X) = X^d + a_1X^{d-1} + \ldots + a_{d-1}X + a_d$ eine Ganzheitsrelation für z. Weil die Koeffizienten in $\mathbb{Z}$ liegen, gilt

$$P(\widetilde{z}) = \widetilde{P(z)} = \widetilde{0} = 0$$

und $\bar{z}$ ist auch ganz algebraisch über $\mathbb{Z}$. Damit folgt aus Korollar 16.19, daß $N(z) = z\widetilde{z}$ und $\operatorname{tr}(z) = z + \widetilde{z}$ ganz algebraisch sind und damit aus $\mathbb{Z}$ nach Proposition 16.22.

Sei nun umgekehrt ein $z \in K$ gegeben, dessen Norm und Spur aus $\mathbb{Z}$ sind. Dann folgt aus Proposition 16.11, daß $T^2 - \operatorname{tr}(z)T + N(z) = 0$ eine Ganzheitsrelation für z ist. □

Im folgenden Satz tritt der Fall $d \equiv 0 \pmod 4$ tritt nicht auf, weil d als quadratfrei angenommen wird.

Satz 16.24. *Sei $K = \mathbb{Q}(\sqrt{d})$ ein quadratischer Zahlkörper mit $d \in \mathbb{Z}$ quadratfrei. Der Ring der ganzen Zahlen in K besteht aus*

$$\mathfrak{o}_k = \begin{cases} \mathbb{Z}[\sqrt{d}] &= \{a + b\sqrt{d}\ ;\ a, b \in \mathbb{Z}\} & \text{falls } d \equiv 2, 3 \pmod 4, \\ \mathbb{Z}[\frac{1+\sqrt{d}}{2}] &= \{x + y\sqrt{d}\ ;\ 2x, 2y, x + y \in \mathbb{Z}\} & \text{falls } d \equiv 1 \pmod 4. \end{cases}$$

Beweis. Wir wissen $\sqrt{d} \in \mathfrak{o}_K$ bzw. $\frac{1+\sqrt{d}}{2} \in \mathfrak{o}_K$ bereits aus Beispiel 16.16 und daher mit Korollar 16.20 auch, daß stets $\mathbb{Z}[\sqrt{d}]$ bzw. $\mathbb{Z}[\frac{1+\sqrt{d}}{2}]$ in $\mathfrak{o}_K$ enthalten sind. Es bleibt, die umgekehrte Inklusion der äußersten Mengen nachzuweisen.

Sei dazu $z = x + y\sqrt{d} \in \mathfrak{o}_K$. Aus $\operatorname{tr}(z) = 2x$ und $N(z) = x^2 - dy^2$ folgt mit Proposition 16.23, daß $2x \in \mathbb{Z}$ und $x^2 - dy^2 \in \mathbb{Z}$. Damit folgt $d(2y)^2 = (\operatorname{tr}(z))^2 - 4N(z) \in \mathbb{Z}$. Weil d quadratfrei ist, darf $2y$ keinen Nenner haben, also $2y \in \mathbb{Z}$.

Angenommen $x = a/2$ und $y = b/2$ mit a oder b ungerade. Dann folgt aus

$$a^2 - db^2 = 4N(z) \equiv 0 \pmod 4,$$

daß entweder beide a, b gerade oder beide ungerade sein müssen. Sonst: wenn a ungerade und b gerade, dann ist $1 \equiv a^2 - db^2 \equiv 0 \pmod 4$ ein Widerspruch. Wenn a gerade und b ungerade, dann ist $-d \equiv a^2 - db^2 \equiv 0 \pmod 4$ ein Widerspruch. Der Fall a und b ungerade bedeutet

$$0 \equiv a^2 - db^2 \equiv 1 - d \pmod 4$$

und kann also nur bei $d \equiv 1 \pmod 4$ auftreten. Das zeigt bereits alles. □

16.3. **Ordnungen.** Der Name Ordnung ist meiner Meinung nach unglücklich gewählt aber Standard. Es geht nicht um eine Anordnung in irgendeinem Sinne.

Definition 16.25. Eine Ordnung im quadratischen Zahlkörper K ist ein Unterring $\mathfrak{o} \subseteq K$, der

(i) als abelsche Gruppe endlich erzeugt ist,
(ii) und den Körper K als $\mathbb{Q}$-Vektorraum erzeugt.

Proposition 16.26. *Sei K ein quadratischer Zahlkörper.*

(1) Der Ring der ganzen algebraischen Zahlen $\mathfrak{o}_K$ ist eine Ordnung von K.
(2) Jede Ordnung von K ist in $\mathfrak{o}_K$ enhalten.

Beweis. (1) Sei $K = \mathbb{Q}(\sqrt{d})$ mit $d \in \mathbb{Z}$ quadratfrei. Damit folgt die Aussage aus Satz 16.24, denn mit

$$\omega = \begin{cases} \sqrt{d} & \text{falls } d \equiv 2,3 \pmod 4, \\ \frac{1+\sqrt{d}}{2} & \text{falls } d \equiv 1 \pmod 4. \end{cases}$$

erzeugen 1 und ω den Ring $\mathfrak{o}_K$ als abelsche Gruppe. Eigenschaft (ii) ist trivial.

(2) Wenn $\mathfrak{o} \subseteq K$ eine Ordnung ist, dann erfüllen alle $\alpha \in \mathfrak{o}$ das Kriterium aus Satz 16.17 bezüglich $M = \mathfrak{o}$. Daher sind alle Elemente aus $\mathfrak{o}$ ganz algebraisch, folglich $\mathfrak{o} \subseteq \mathfrak{o}_K$. □

Bemerkung 16.27. Man nennt $\mathfrak{o}_K$ wegen Proposition 16.26 auch die **Maximalordnung** von K.

Proposition 16.28. *Sei $d \in \mathbb{Z}$ quadratfrei und $\mathfrak{o}$ eine Ordnung im quadratischen Zahlkörper $K = \mathbb{Q}(\sqrt{d})$. Sei ω wie im Beweis von Proposition 16.26. Dann gilt mit dem Index als abelsche Gruppen $c = (\mathfrak{o}_K : \mathfrak{o})$, daß*

$$\mathfrak{o} = \mathbb{Z} + c \cdot \mathbb{Z}\omega.$$

Beweis. Die Darstellung als $u + v\omega$ mit $u, v \in \mathbb{Z}$ ist für jedes Element von $\mathfrak{o}_K$ eindeutig. Die Menge

$$I = \{v \in \mathbb{Z}\ ;\ \text{ es gibt } u \in \mathbb{Z} \text{ mit } u + v\omega \in \mathfrak{o}\}$$

ist ein Ideal von $\mathbb{Z}$.

Da $\mathfrak{o} \neq \mathbb{Z}$ gilt, folgt $I \neq (0)$. Sei $c > 0$ mit $I = \mathbb{Z}c$. Sei $t = b + c\omega \in \mathfrak{o}$ ein Zeuge für $c \in I$. Wegen $1 \in \mathfrak{o}$ können wir $t = c\omega$, also $b = 0$ wählen. Dann ist offensichtlich

$$\mathbb{Z} + c \cdot \mathbb{Z}\omega \subseteq \mathfrak{o},$$

und wir haben die umgekehrte Inklusion zu zeigen. Sei also $z = u + v\omega \in \mathfrak{o}$ beliebig. Dann gibt es per Konstruktion ein $n \in \mathbb{Z}$ mit $v = nc$. Dann ist

$$z = u + nc\omega \in \mathbb{Z} + c \cdot \mathbb{Z}\omega.$$

Die Abbildung $\mathfrak{o}_K \to \mathbb{Z}/c\mathbb{Z}$ definiert durch

$$u + v\omega \mapsto v \pmod c$$

ist ein surjektiver Gruppenhomomorphismus mit Kern $\mathfrak{o}$. Dies zeigt $c = (\mathfrak{o}_K : \mathfrak{o})$. □

Definition 16.29. Das eindeutige c aus Proposition 16.28 nennt man **Führer (conductor)** der Ordnung $\mathfrak{o}$.

16.4. **Spurform und Diskriminante.** Wir gehen nun der Frage nach, wie groß die Maximalordnung $\mathfrak{o}_K$ eines quadratischen Zahlkörpers in einem gewissen Sinne ist.

Definition 16.30. Sei K ein quadratischer Zahlkörper. Die Spurform auf K ist definiert für $x, y \in K$ durch

$$x, y \mapsto \mathrm{tr}_{K/\mathbb{Q}}(xy).$$

Bemerkung 16.31. Sei K ein quadratischer Zahlkörper.

(1) Die Spurform ist eine symmetrische Bilinearform auf K aufgefaßt als $\mathbb{Q}$-Vektorraum. In der Tat ist die Multiplikation $(x, y) \mapsto xy$ bilinear und $\mathrm{tr}_{K/\mathbb{Q}}$ eine $\mathbb{Q}$-Linearform.

(2) Sei $\mathfrak{o} \subseteq K$ eine Ordnung. Die Spurform nimmt nach Proposition 16.23 auf $\mathfrak{o}$ nur Werte in $\mathbb{Z}$ an.

(3) Seien α, β eine $\mathbb{Z}$-Basis von $\mathfrak{o}$, also $\mathfrak{o} = \mathbb{Z}\alpha + \mathbb{Z}\beta$. Dann hat die Gram'sche Matrix A der Spurform bezüglich α, β eine Determinante $\det(A) \in \mathbb{Z}$, die von der Wahl der Basis unabhängig ist. Wir nennen diese Gram'sche Determinante die **Diskriminante** der Ordnung

$$\Delta_{\mathfrak{o}} := \det \begin{pmatrix} \mathrm{tr}(\alpha^2) & \mathrm{tr}(\alpha\beta) \\ \mathrm{tr}(\alpha\beta) & \mathrm{tr}(\beta^2) \end{pmatrix}.$$

In der Tat hat die Basiswechselmatrix S zu einer weiteren Basis Einträge in $\mathbb{Z}$, genauso wie die inverse Basiswechselmatrix S^{-1}. Damit sind $\det(S)$ und $\det(S^{-1})$ ganzzahlig und wegen

$$1 = \det(E) = \det(SS^{-1}) = \det(S) \cdot \det(S^{-1})$$

notwendigerweise $\det(S) = \pm 1$. Der Basiswechsel zur neuen Gram'schen Matrix liefert $S^t AS$ mit Determinante

$$\det(S^t AS) = \det(S^t)\det(A)\det(S) = \det(S)^2 \det(A) = \det(A).$$

Definition 16.32. Die **Diskriminante** eines quadratischen Zahlkörpers K ist die Diskriminante

$$\Delta_K = \Delta_{\mathfrak{o}_K}$$

der zugehörigen Maximalordnung.

Proposition 16.33. *Sei $d \in \mathbb{Z}$ quadratfrei. Dann ist die Diskriminante von $K = \mathbb{Q}(\sqrt{d})$ gleich*

$$\Delta_K = \begin{cases} 4d & \textit{falls } d \not\equiv 1 \pmod 4 \\ d & \textit{falls } d \equiv 1 \pmod 4. \end{cases}$$

Insbesondere ist die Diskriminante stets $\equiv 0, 1 \pmod 4$.

Beweis. Sei zunächst $d \not\equiv 1 \pmod 4$. Dann ist $1, \sqrt{d}$ eine $\mathbb{Z}$-Basis von $\mathfrak{o}_K$ und die diesbezügliche Gram'sche Matrix lautet

$$\begin{pmatrix} 2 & 0 \\ 0 & 2d \end{pmatrix}$$

mit Determinante $\Delta = 4d$.

Falls $d \equiv 1 \pmod 4$, dann ist eine $\mathbb{Z}$-Basis von $\mathfrak{o}_K$ gegeben durch 1 und $\omega = \frac{1+\sqrt{d}}{2}$. Damit ist $\operatorname{tr}(\omega) = 1$ und $\operatorname{tr}(\omega^2) = \operatorname{tr}(\frac{1+d+2\sqrt{d}}{4}) = \frac{1+d}{2}$. Die Gram'sche Matrix lautet in diesem Fall

$$\begin{pmatrix} 2 & 1 \\ 1 & \frac{1+d}{2} \end{pmatrix}$$

mit Determinante $\Delta = d$. □

Ein reell-quadratischer Zahlkörper K hat zwei Einbettungen $K \hookrightarrow \mathbb{R}$, nämlich neben der Identität auch noch $z \mapsto \widetilde{z}$. Am besten benutzt man beide!

Definition 16.34. Der **Minkowski-Raum** eines quadratischen Zahlkörpers K ist der folgende $\mathbb{R}$-Vektorraum:

$$K_{\mathbb{R}} = \begin{cases} \mathbb{R} \times \mathbb{R} & \text{falls } K \text{ reell-quadratisch ist,} \\ \mathbb{C} & \text{falls } K \text{ imaginär-quadratisch ist.} \end{cases}$$

Dazu betrachten wir die Einbettung $j : K \hookrightarrow K_{\mathbb{R}}$ und ihre Einschränkung $j : \mathfrak{o}_K \hookrightarrow K_{\mathbb{R}}$ definiert durch die Identität für imaginär-quadratische K und durch

$$j(z) = (z, \widetilde{z})$$

für reell-quadratische K.

Bemerkung 16.35. Der Minkowskiraum von K ist kanonisch isomorph zu $K \otimes_{\mathbb{Q}} \mathbb{R}$ und trägt damit sogar die Struktur einer $\mathbb{R}$-Algebra.

Lemma 16.36. *Sei K ein quadratischer Zahlkörper.*

(1) Das Bild einer $\mathbb{Z}$-Basis von $\mathfrak{o}_K$ ist eine $\mathbb{R}$-Basis von $K_{\mathbb{R}}$.

(2) Sei α, β eine $\mathbb{Z}$-Basis von $\mathfrak{o}_K$. Dann ist

$$K_{\mathbb{R}} = \bigcup_{\gamma \in \mathfrak{o}_K} j(\gamma) + \{z = t\alpha + s\beta \ ; \ 0 \le t, s < 1\}$$

eine disjunkte Überdeckung.

Beweis. (1) Da $\mathbb{Z}$-Bases sich um Matrixmultiplikation mit einer Matrix aus $\mathrm{GL}_2(\mathbb{Z})$ unterscheiden, ist es egal, für welche Basis wir die Aussage beweisen.

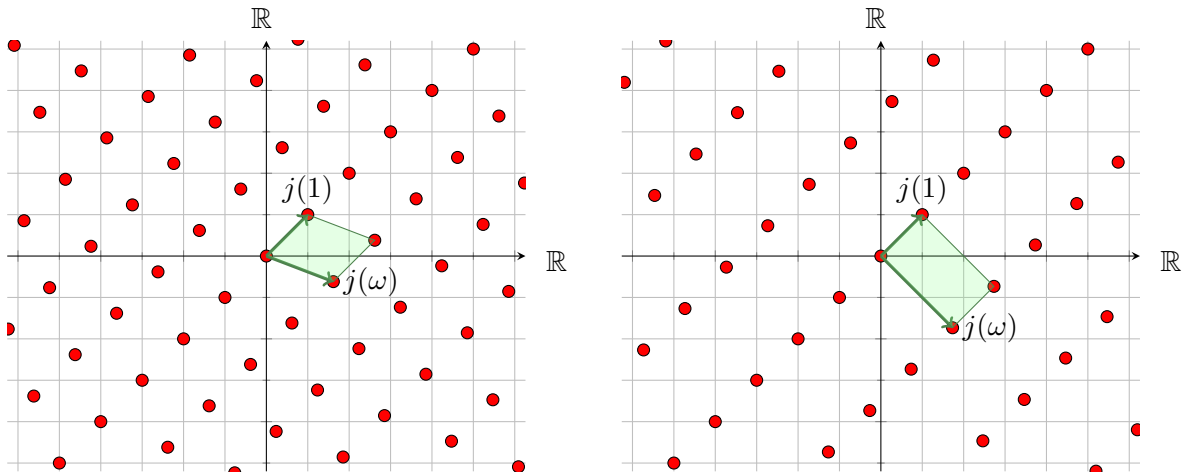
Mit dem üblichen ω geht die Basis $1, \omega$ im imaginär-quadratischen Fall auf $1, \omega$. Das ist eine $\mathbb{R}$-Basis, weil $\omega \notin \mathbb{R}$.

Wenn K reell-quadratisch ist, dann geht die Basis $1, \omega$ auf

$$j(1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad j(\omega) = \begin{pmatrix} \omega \\ \widetilde{\omega} \end{pmatrix}$$

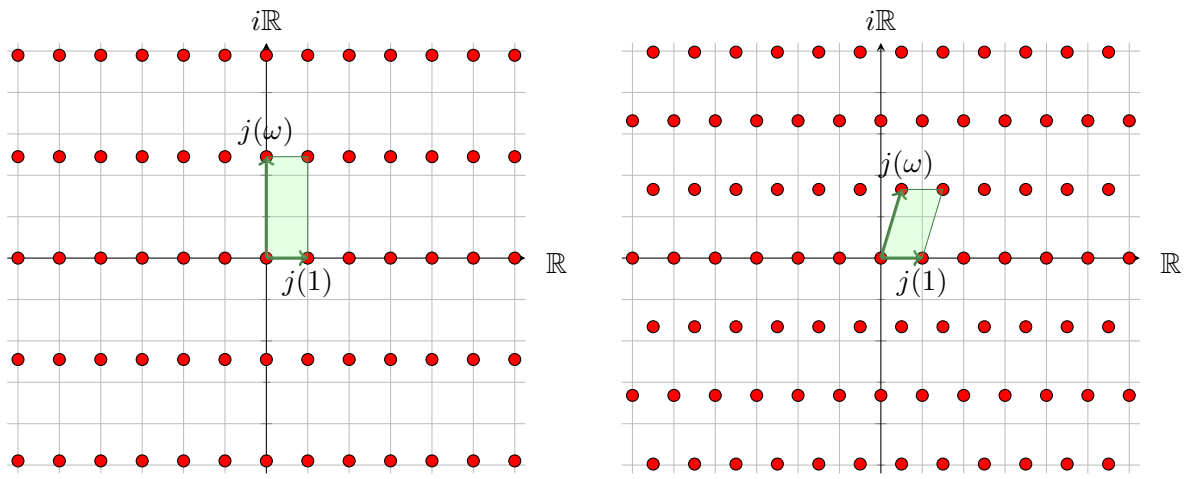
und das ist wieder eine $\mathbb{R}$-Basis, weil $\omega \neq \widetilde{\omega}$.

(2) folgt aus (1): in Koordinaten bezüglich der gegebenen Basis sieht $j(\mathfrak{o}_K)$ aus wie $\mathbb{Z}^2 \subseteq \mathbb{R}^2$ und wir sprechen von der Parkettierung der Ebene durch die Translate des Einheitsquadrats. $\square$



(a) $\mathbb{Z}[\frac{1+\sqrt{5}}{2}]$ (b) $\mathbb{Z}[\sqrt{3}]$

Abbildung 6. Der Fall reell-quadratischer Zahlkörper: ganze Zahlen als Gitter im Minkowski-Raum $\mathbb{R} \times \mathbb{R}$.



(a) $\mathbb{Z}[\sqrt{-6}]$ (b) $\mathbb{Z}[\frac{1+\sqrt{-11}}{2}]$

Abbildung 7. Der imaginär-quadratische Fall: ganze Zahlen als Gitter im Minkowski-Raum $\mathbb{C}$.

Der Minkowskiraum hat ein natürliches Flächenmaß. Das Co-Volumen des Gitters $j(\mathfrak{o}_K) \subseteq K_{\mathbb{R}}$ ist das Volumen (hier: Fläche) des Quotienten $K_{\mathbb{R}}/j(\mathfrak{o}_K)$.

Proposition 16.37. *Sei K ein quadratischer Zahlkörper. Sei $s = 1$, wenn K imaginärquadratisch und sonst $s = 0$. Die Maximalordnung hat als Gitter im Minkowskiraum das Co-Volumen*

$$\operatorname{vol}(K_{\mathbb{R}}/j(\mathfrak{o}_K)) = 2^{-s_K} \cdot \sqrt{|\Delta_K|}.$$

Beweis. Nach Lemma 16.36 ist das gesuchte Volumen gleich dem Volumen der zu einer $\mathbb{Z}$-Basis α, β von $\mathfrak{o}_K$ gehörenden **Grundmasche** (oder **Fundamentalbereich**)

$$\{x\alpha + y\beta \ ; \ 0 \leq x, y < 1\}.$$

Daraus folgt mit der 2×2-Matrix mit Spalten $j(1)$ und $j(\omega)$

$$\operatorname{vol}(K_{\mathbb{R}}/j(\mathfrak{o}_K)) = |\det(j(\alpha), j(\beta))| = \begin{cases} |\det \begin{pmatrix} \alpha & \beta \\ \widetilde{\alpha} & \widetilde{\beta} \end{pmatrix}| & K \text{ reell}, \\ |\det \begin{pmatrix} \Re(\alpha) & \Re(\beta) \\ \Im(\alpha) & \Im(\beta) \end{pmatrix}| & K \text{ imaginär}. \end{cases}$$

Im imaginär-quadratischen Fall berechnen wir die Determinante genauer zu

$$\begin{aligned} \det \begin{pmatrix} \Re(\alpha) & \Re(\beta) \\ \Im(\alpha) & \Im(\beta) \end{pmatrix} &= \det \begin{pmatrix} (\alpha + \widetilde{\alpha})/2 & (\beta + \widetilde{\beta})/2 \\ (\alpha - \widetilde{\alpha})/2 & (\beta - \widetilde{\beta})/2 \end{pmatrix} \\ &= \frac{1}{4}\big((\alpha + \widetilde{\alpha})(\beta - \widetilde{\beta}) - (\alpha - \widetilde{\alpha})(\beta + \widetilde{\beta})\big) = \frac{1}{2}(\widetilde{\alpha}\beta - \alpha\widetilde{\beta}) \\ &= -2^{-s} \cdot \det \begin{pmatrix} \alpha & \beta \\ \widetilde{\alpha} & \widetilde{\beta} \end{pmatrix}. \end{aligned}$$

Die Aussage folgt nun in beiden Fällen aus

$$\Big(\det \begin{pmatrix} \alpha & \beta \\ \widetilde{\alpha} & \widetilde{\beta} \end{pmatrix}\Big)^2 = \det \begin{pmatrix} \alpha & \widetilde{\alpha} \\ \beta & \widetilde{\beta} \end{pmatrix} \cdot \det \begin{pmatrix} \alpha & \beta \\ \widetilde{\alpha} & \widetilde{\beta} \end{pmatrix} = \det \begin{pmatrix} \alpha^2 + \widetilde{\alpha^2} & \alpha\beta + \widetilde{\alpha\beta} \\ \beta\alpha + \widetilde{\beta\alpha} & \beta^2 + \widetilde{\beta^2} \end{pmatrix} = \Delta_K,$$

denn das ist die Determinante der Spurform, berechnet zur Basis α, β. □

Bemerkung 16.38. Der Beweis von Proposition 16.37 zeigt nicht, daß Δ_K stets positiv ist, obwohl die Rechnung zeigt, daß Δ_K ein Quadrat ist. Der Ausdruck

$$\delta = \det \begin{pmatrix} \alpha & \beta \\ \widetilde{\alpha} & \widetilde{\beta} \end{pmatrix}$$

ist in K und nicht rational! Für reell-quadratische K ist $\delta \in \mathbb{R}$, aber für imaginär-quadratische K ist $\delta \in \mathbb{C} \setminus \mathbb{R}$. Nichtsdestotrotz zeigt die Rechnung, daß stets

$$\Delta_K \in K^2$$

die Diskrimante ein Quadrat im betrachteten Zahlkörper wird. Das folgt allerdings auch aus der expliziten Berechnung der Diskriminante von $\mathbb{Q}(\sqrt{d})$.

Übungsaufgabe 16.1. Sei $\Delta \equiv 0$ oder $1 \pmod 4$ und kein Quadrat in $\mathbb{Z}$. Wir setzen $\alpha = \frac{\Delta + \sqrt{\Delta}}{2}$.

(a) Finden Sie eine Ganzheitsrelation für α.

(b) Sei $\mathfrak{o} = \mathbb{Z}[\alpha]$ die von α erzeugte Ordnung in $\mathbb{Q}(\sqrt{\Delta})$. Berechnen Sie die Diskriminante von $\mathfrak{o}$ bezüglich der Spurform.

Übungsaufgabe 16.2. Sei K ein quadratischer Zahlkörper und $\mathfrak{o} \subseteq K$ eine Ordnung.

(a) Zeigen Sie, daß es ein eindeutiges $\Delta \in \mathbb{Z}$ gibt, so daß Δ kein Quadrat in $\mathbb{Z}$ ist, $\Delta \equiv 0, 1 \pmod 4$ und

$$\mathfrak{o} = \mathbb{Z}[\frac{\Delta + \sqrt{\Delta}}{2}].$$

(b) Bestimmen Sie den conductor von $\mathfrak{o}$ anhand von Δ.

17. Ganzzahlige binäre quadratische Formen

17.1. **Binäre quadratische Formen.** Nachdem der Zwei-Quadrate-Satz uns beantwortet, welche Primzahlen die Summe zweier Quadrate sind, wollen wir als Variante die Frage aufwerfen, welche Primzahlen der *Form* $x^2 + 2y^2$ mit $x, y \in \mathbb{Z}$ sind.

Definition 17.1. Eine **binäre quadratische Form** (oder **quadratische Form in zwei Variablen**) mit Koeffizienten aus $\mathbb{Z}$ ist ein homogenes Polynom vom Grad 2 der Form

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

mit $a, b, c \in \mathbb{Z}$. Die **zugehörige symmetrische Bilinearform** auf $\mathbb{Z}^2$

$$\begin{aligned} B_f(\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} u \\ v \end{pmatrix}) &:= f(x+u, y+v) - f(x,y) - f(u,v) \\ &= 2axu + b(xv + yu) + 2cyv \end{aligned}$$

hat bezüglich der Standardbasis die **Gram'sche Matrix**

$$\mathrm{G}(f) := \begin{pmatrix} 2a & b \\ b & 2c \end{pmatrix}.$$

Die abelsche Gruppe (addition von Polynomen) aller binären Quadratischen Formen bezeichnen wir mit

$$\mathrm{QF}(\mathbb{Z}) = \{aX^2 + bXY + cY^2 \ ; \ a, b, c \in \mathbb{Z}\} \simeq \mathbb{Z}^3.$$

Remark 17.2 (Polarisationsformel). Wir bekommen die quadratische Form aus der Bilinearform als

$$f(X, Y) = \frac{1}{2} B_f(\begin{pmatrix} X \\ Y \end{pmatrix}, \begin{pmatrix} X \\ Y \end{pmatrix}) = \frac{1}{2} \begin{pmatrix} X \\ Y \end{pmatrix}^t \mathrm{G}(f) \begin{pmatrix} X \\ Y \end{pmatrix}$$

zurück. Die definierende Gleichung für B_f wird damit zur Polarisationsformel für B_f.

Definition 17.3. Sei f eine ganzzahlige binäre quadratische Form.

(1) Die Form f ist **positiv definit**, wenn für alle $x, y \in \mathbb{Z}$ gilt $f(x, y) \geq 0$ mit Gleichheit nur für $(x, y) = (0, 0)$.
(2) Die Form f ist **negativ definit**, wenn für alle $x, y \in \mathbb{Z}$ gilt $f(x, y) \leq 0$ mit Gleichheit nur für $(x, y) = (0, 0)$.
(3) Die Form f ist **indefinit**, wenn f positive und negative Werte annimmt.
(4) Die Form f ist **nicht-ausgeartet**, wenn f den Wert 0 nur bei $(x, y) = (0, 0)$ annimmt.

Remark 17.4. Die folgende Aussage ist eine leichte Übungsaufgabe. Eine ganzzahlige binäre quadratische Form f ist positiv definit (bzw. negativ definit, indefinit oder nicht-ausgeartet) genau dann, wenn die zugehörige Gram'sche Matrix $\mathrm{G}(f)$ als reelle Matrix positiv definit (bzw. negativ definit, indefinit oder nicht-ausgeartet) ist.

In dieser Übungsaufgabe kommt man von $\mathbb{Z}$ nach $\mathbb{Q}$ durch Skalieren, weil f homogen ist, und zu $\mathbb{R}$, weil $\mathbb{Q}$ in $\mathbb{R}$ dicht ist.

Definition 17.5. Die **Diskriminante** einer ganzzahlige binäre quadratische Form $f(X, Y) = aX^2 + bXY + cY^2$ ist

$$D = D(f) = b^2 - 4ac.$$

Man beachte, daß $D = -\det(\mathrm{G}(f))$, und damit die Diskriminante von f das Negative der Diskrimante der Bilinearform B_f ist.

Proposition 17.6. *Die Diskriminante einer ganzzahligen binären quadratischen Form ist stets* $\equiv 0, 1 \pmod 4$.

Beweis. Die Diskriminante hat die Form $D = b^2 - 4ac$ und ist daher modulo 4 ein Quadrat, daher $\equiv 0, 1 \pmod 4$. □

Proposition 17.7. *Für eine ganzzahlige binäre quadratische Form $f(X,Y) = aX^2+bXY+cY^2$ mit Diskriminante D gilt das folgende.*

(1) $D < 0$ *und* $a > 0 \iff f$ *ist positiv definit.*
(2) $D < 0$ *und* $a < 0 \iff f$ *ist negativ definit.*
(3) $D > 0 \iff f$ *ist indefinit.*

Beweis. Das ist einfach nur das Hauptminorenkriterium in ungewohnter Form. □

17.2. **Äquivalenzklassen quadratischer Formen.**

Definition 17.8. Wir definieren eine Rechtsoperation von $\mathrm{GL}_2(\mathbb{Z})$ auf $\mathrm{QF}(\mathbb{Z})$. Für $S = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ und $f(X,Y) = aX^2 + bXY + cY^2$ setzen wir

$$f|_S(X,Y) := f(\alpha X + \beta Y, \gamma X + \delta Z) = AX^2 + BXY + CY^2$$

mit $A = f(\alpha,\gamma)$, $B = 2a\alpha\beta + b(\alpha\delta + \beta\gamma) + 2c\gamma\delta$ und $C = f(\beta,\delta)$.

Daß wir eine Rechtsoperation definiert haben sieht man sofort, wenn man die Transformation der Gram'schen betrachtet:

$$\mathrm{G}(f|_S) = S^t\,\mathrm{G}(f)S, \qquad \text{weil} \quad S\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \alpha X + \beta Y \\ \gamma X + \delta Z \end{pmatrix}.$$

Also gilt für alle $S, T \in \mathrm{GL}_2(\mathbb{Z})$ und $f \in \mathrm{QF}(\mathbb{Z})$:

$$f_{ST} = (f|_S)|_T.$$

Wir notieren die Äquivalenzrelation, die durch die Bahnen der $\mathrm{GL}_2(\mathbb{Z})$-Operation definiert wird durch $f \sim g$. Wir nennen quadratische Formen **eigentlich äquivalent** mit der Bezeichnung $f \overset{+}{\sim} g$, wenn der Basiswechsel aus $\mathrm{SL}_2(\mathbb{Z}) = \mathrm{GL}_2^+(\mathbb{Z})$ kommt.

Proposition 17.9. *Äquivalente quadratische Formen haben die gleiche Diskriminante.*

Beweis. Bis auf das Vorzeichen können wir die Determinante der Gram'schen Vergleichen. Jedes $S \in \mathrm{GL}_2(\mathbb{Z})$ hat $\det(S) \in \mathbb{Z}^\times$, also $\det(S)^2 = 1$. Daher gilt

$$D(f|_S) = -\det(\mathrm{G}(f|_S)) = -\det(S^t\,\mathrm{G}(f)S) = -\det(S)^2 \cdot \det(\mathrm{G}(f)) = D(f). \qquad \square$$

Als nächstes stellen wir einen Fundamentalbereich für die Operation von $\mathrm{SL}_2(\mathbb{Z})$ auf $\mathrm{QF}(\mathbb{Z})$ bereit.

Satz 17.10 (Reduktionssatz von Lagrange 1773)**.** *Jede positive definite ganzzahlige binäre quadratische Form ist* ***eigentlich*** *äquivalent zu genau einer quadratischen Form $aX^2 + bXY + cY^2$ mit*

$$c > a \geq b > -a \quad \text{und} \quad c = a \geq b \geq 0.$$

Quadratische Formen mit diesen Bedingungen an die Koeffizienten heißen ***reduzierte quadratische Formen****.*

Beweis. Wir wählen zunächst in einer gegebenen Äquivalenzklasse diejenige Form $f(X,Y) = aX^2 + bXY + cY^2$ mit minimalem $|b|$. Die Koeffizienten $a = f(1,0)$ und $c = f(0,1)$ sind positiv, weil f positiv definit ist. Der Basiswechsel mit $T = \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}$ liefert

$$f|_T(X,Y) = f(-Y,X) = cX^2 - bXY + aX^2,$$

so daß wir annehmen dürfen, daß $c \geq a$ gilt.

Nun betrachten wir den Basiswechsel mit $S_n = \begin{pmatrix} 1 & n \\ & 1 \end{pmatrix}$, der uns die quadratische Form

$$f|_{S_n}(X,Y) = a(X+nY)^2 + b(X+nY)Y + cY^2 = aX^2 + (2an+b)XY + (an^2+bn+c)Y^2$$

liefert. Die Extremaleigenschaft von $|b|$ besagt, daß $n \mapsto |2an+b|$ sein Minimum bei $n=0$ annimmt. Das bedeutet

$$a \geq |b|.$$

Es gilt nun die Randfälle zu untersuchen, die den Unterschied zwischen $c \geq a \geq |b|$ und den Ungleichungen im Satz ausmachen.

- Wenn $b = -a$ gilt, dann können wir mit S_{-1} den Koeffizienten b durch $b + 2a = a$ ersetzen.
- Wenn $a = c$ und $b < 0$ gilt, dann transformiert T den Koeffizienten b zu $-b$ (und tauscht die gleichen Koeffizienten a und c).

Damit haben wir gezeigt, daß jede eigentliche Äquivalenzklasse eine reduzierte quadratische Form enthält.

Wir zeigen nun die Eindeutigkeit des reduzierten Vertreters in jeder Äquivalenzklasse. Für eine positiv definite reduzierte quadratische Form $f(X,Y) = aX^2 + bXY + cY^2$ gilt

$$f(x,y) \geq (a - |b| + c)\min\{x^2, y^2\}.$$

Dies zeigt, daß für $(x,y) \neq (0,0), (0,\pm1), (\pm1,0), (\pm1,\pm1)$ der Wert

$$f(x,y) \geq 4(a - |b| + c) \geq 4c > \max\{a, c, a \pm b + c\}.$$

ist. Damit sind $a = f(\pm1,0)$ und $c = f(0,\pm)$ und $f(\pm1,\pm1) = a \pm b + c$ die vier kleinsten von 0 verschiedenen Werte von $f(X,Y)$. Diese haben die Anordnung

$$a + |b| + c \geq a - |b| + c \geq c \geq a.$$

Angenommen $f \sim g$ für eine weitere reduzierte Form $g(X,Y) = uX^2 + vXY + wY^2$. Dann gilt durch Vergleich der zwei kleinsten Werte $\neq 0$ bereits $a = u$ und $c = w$. Man beachte, wenn man die vier kleinsten Werte wegen Koinzidenz nicht unterscheiden kann, dann macht das nichts, die Gleichungen gelten weiter! Weil äquivalente Formen die gleiche Diskriminante haben, siehe Proposition 17.9, folgt

$$b^2 = D(f) + 4ac = D(g) + 4uw = v^2.$$

Somit folgt auch $b = \pm v$. Wir nehmen nun $b = -v$ an und suchen einen Widerspruch.

- Wenn $a = c$ gilt, dann folgt $b = v = 0$, weil beide nichtnegativ sind.
- Wenn $c > a$ gilt, dann ist wegen $a \geq b > -a$ und $u \geq v > -u$ in jedem Fall $a = u > |b| = |v|$. Damit ist $a - |b| + c > c$ und die beiden minimalen Werte treten nur bei $(\pm1,0)$ und $(0,\pm1)$ auf. Der Basiswechsel S, der $f \sim g$ zeigt, muß dies beherzigen. Es bleibt also nur
$$S = \begin{pmatrix} \pm1 & \\ & \pm1 \end{pmatrix}.$$
Von diesen sind nur $S = \pm E$ mit der Einheitsmatrix E aus $\mathrm{SL}_2(\mathbb{Z})$. Aber $\pm E$ stabilisiert jede quadratische Form. Daher folgt $f = g$. □

Remark 17.11. Wir haben nach Satz 17.10 die folgende Parametrisierung für eigentliche Äquivalenzklassen positiv definiter, ganzzahliger binärer quadratischer Formen:

$$\{(a,b,c)\ ;\ c > a \geq b > -a,\ \text{und } c = a \geq b \geq 0\} \xrightarrow{\sim} \{f \in \mathrm{QF}(\mathbb{Z})\ ;\ \text{ positv definit}\}/\overset{+}{\sim}$$

Auf dieser Menge operiert $\mathrm{GL}_2(\mathbb{Z})/\mathrm{SL}_2(\mathbb{Z}) \simeq \{\pm1\}$ mittels dem Basiswechsel durch $\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$ mit dem Effekt

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

Wenn $a = c$ gilt, müssen wir das mit dem Basiswechsel mittels $\begin{pmatrix} & 1 \\ -1 & \end{pmatrix}$ kombinieren, so daß die Operation trivial wird:

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

Wenn $c > a$ und $b = a$ gelten, dann müssen wir das mit dem Basiswechsel mittels $\begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix}$ kombinieren, so daß die Operation ebenfalls trivial wird:

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

Es folgt daß $f \sim g \iff f \overset{+}{\sim} g$. Das bedeutet auch, daß zu jeder positiv definiten quadratischen Form im Stabilisator ein Element mit Determinante -1 liegt.

Korollar 17.12. *Sei $D < 0$. Die Menge der Äquivalenzklassen ganzzahliger binärer quadratischer Formen mit Diskriminante D ist endlich.*

Genauer findet man im positiv definiten Fall in jeder Äquivalenzklasse einen reduzierten Vertreter $aX^2 + bXY + cY^2$ mit

$$\sqrt{\frac{-D}{3}} \geq a \geq |b|, \qquad \textit{und} \qquad c = \frac{b^2 - D}{4a}.$$

Beweis. Nach Proposition 17.7 bedeutet $D < 0$, daß die quadratischen Formen entweder positiv oder negativ definit sind. Durch Multiplikation mit -1 führen wir den negativ definiten Fall auf den positiv definiten Fall zurück. Nach Satz 17.10 schätzen wir für den reduzierten Vertreter

$$-D = 4ac - b^2 \geq 3a^2$$

ab. Dies zeigt die Abschätzung für a und b, und die Gleichung für c ist trivial. □

Definition 17.13. Die **Klassenzahl** $\mathrm{h}(D)$ einer Diskriminante $D < 0$ ist die Anzahl der Äquivalenzklassen ganzzahliger binärer quadratischer Formen mit Diskriminante D.

17.3. **Werte quadratischer Formen.**

Definition 17.14. Eine quadratische Form $f(X, Y)$ **stellt** $m \in \mathbb{Z}$ **dar**, wenn es $x, y \in \mathbb{Z}$ gibt mit $f(x, y) = m$. Wir sagen, daß m **primitiv von** f **dargestellt** wird, wenn $\mathrm{ggT}(x, y) = 1$ ist.

Lemma 17.15. *Seien $f, g \in \mathrm{QF}(\mathbb{Z})$. Wenn $f \sim g$ gilt, dann stellen f und g die gleichen $m \in \mathbb{Z}$ dar. Zusatz: Das gleiche gilt für die Zahlen die primitv dargestellt werden.*

Beweis. Sei $S \in \mathrm{GL}_2(\mathbb{Z})$ mit $f|_S = g$, und sei $m \in \mathbb{Z}$ eine Zahl die von f dargestellt wird. Dann gibt es $x, y \in \mathbb{Z}$ mit $f(x, y) = m$. Mit $\begin{pmatrix} u \\ v \end{pmatrix} = S \begin{pmatrix} x \\ y \end{pmatrix}$ folgt

$$g(u, v) = f|_S(u, v) = \frac{1}{2} \begin{pmatrix} u \\ v \end{pmatrix} S^t \, \mathrm{G}(f) S \begin{pmatrix} u \\ v \end{pmatrix} = \frac{1}{2} \begin{pmatrix} x \\ y \end{pmatrix} \mathrm{G}(f) \begin{pmatrix} x \\ y \end{pmatrix} = f(x, y) = m.$$

Der Umkehrugn folgt aufgrund der Symmetrie der Aussage.

Der Zusatz folgt aus der Beobachtung, daß unter einem Basiswechsel aus $\mathrm{GL}_2(\mathbb{Z})$ der ggT der Komponenten eines Vektors erhalten bleiben. □

Lemma 17.16. *Sei $f(X, Y) \in \mathrm{QF}(\mathbb{Z})$. Eine Zahl $m \in \mathbb{Z}$ wird genau dann primitiv von f dargestellt, wenn*

$$f \sim mX^2 + bXY + cY^2$$

für geeignete $b, c \in \mathbb{Z}$.

Beweis. Sei $m = f(\alpha, \beta)$ mit teilerfremden α, β. Nach dem Lemma von Bezout gibt es $\gamma, \delta \in \mathbb{Z}$, so daß

$$S = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{GL}_2(\mathbb{Z}).$$

Der zu S inverse Basiswechsel bildet $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ auf $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ab. Eine zu f ähnliche quadratische Form stellt m daher als Wert bei $(1,0)$ dar. Das ist nichts anderes als der Koeffizient von X^2.

Die Umkehrung folgt aus Lemma 17.15, weil $mX^2 + bXY + cY^2$ mittels $(1,0)$ den Wert m primitiv darstellt. □

Proposition 17.17. *Sei $D \in \mathbb{Z}$. Eine ungerade Zahl $m \in \mathbb{Z}$ wird genau dann primitiv von einem $f \in \mathrm{QF}(\mathbb{Z})$ von Diskriminante D dargestellt, wenn D ein Quadrat modulo m ist.*

Beweis. Angenommen m wird primitiv von $f(X,Y)$ dargestellt. Weil die Diskriminante sich bei Basiswechsel nicht ändert, dürfen wir nach Lemma 17.16 annehmen, daß $f(X,Y) = mX^2 + bXY + cY^2$. Dann ist

$$D = b^2 - 4mc \equiv b^2 \pmod{m}$$

ein Quadrat modulo m.

Sei nun umgekehrt D ein Quadrat modulo m. Es gibt daher $b \in \mathbb{Z}$ mit $D \equiv b^2 \pmod{m}$. Weil m ungerade ist, können wir annehmen, daß $D \equiv b \pmod{2}$, indem wir notfalls b durch $b+m$ ersetzen. Weil D eine Diskriminante ist, gilt $D \equiv 0, 1 \pmod{4}$, und als Quadrat ebenso $b^2 \equiv 0, 1 \pmod{4}$. Wir schließen, daß sogar

$$D \equiv b^2 \pmod{4m}$$

gilt. Damit gibt es $c \in \mathbb{Z}$ mit $D = b^2 - 4mc$. Es folgt, daß die quadratische Form

$$f(X,Y) = mX^2 + bXY + cY^2$$

Diskriminante D hat und darüberhinaus m als $f(1,0)$ darstellt. □

Theorem 17.18. *Eine ungerade Primzahl p ist von der Form x^2+2y^2 mit $x, y \in \mathbb{Z}$ genau dann, wenn $\left(\frac{-2}{p}\right) = 1$, d.h. wenn $p \equiv 1, 3 \pmod{8}$.*

Beweis. Die Diskriminante von $X^2 + 2Y^2$ ist $D = -8$. Angenommen p wird von einer quadratischen Form $f(X,Y)$ von Diskriminante -8 dargestellt, dann wird p automatisch primitiv dargestellt, denn für $\alpha, \beta \in \mathbb{Z}$ mit $p = f(\alpha, \beta)$ folgt

$$\mathrm{ggT}(\alpha,\beta)^2 \mid f(\alpha,\beta) = p,$$

so daß α und β teilerfremd sind.

Nach Proposition 17.17 wird p von einer quadratischen Form von Diskriminante -8 genau dann dargestellt, wenn -8 ein Quadrat modulo p ist, also wenn

$$1 = \left(\frac{-8}{p}\right) = \left(\frac{-2}{p}\right) = (-1)^{\frac{p-1}{2} + \frac{p^2-1}{8}}.$$

Dies sind genau die Fälle $p \equiv 1, 3 \pmod{8}$.

Die Bedingung des Theorems besagt also, daß p von irgendeiner quadratischen Form von Diskriminante -8 dargestellt werden kann. Es reicht daher nun zu zeigen, daß es bis auf Äquivalenz nur genau eine solche Form gibt, denn nach Lemma 17.15 stellen äquivalente Formen die gleichen Werte dar.

Berechnen wir also die Klassenzahl $h(-8)$ mittels Satz 17.10 und vor allem Korollar 17.12 (Weil p ein Wert ist und $D = -8 < 0$ gilt, handelt es sich um positiv definite quadratische Formen). Weil mit der dortigen Notation

$$1.633 > \sqrt{\frac{-D}{3}} \geq a \geq |b|,$$

kommt nur $a = 1$ und $b = 1$ oder 0 in Frage. Weil $b^2 \equiv D \equiv 0 \pmod{4}$, haben wir $b = 0$ und $c = -D/4a = 2$. Damit ist X^2+2Y^2 die einzige reduzierte quadratische Form mit Diskriminante -8. Es gilt $h(-8) = 1$ und dies beendet den Beweis. □

18. Die Pell-Gleichung und Kettenbrüche

18.1. **Die Pell-Gleichung.**

Definition 18.1. Unter der **Pell-Gleichung** verstehen wir für ein ganzzahliges $d > 0$ die Diophantische Gleichung

$$X^2 - dY^2 = 1.$$

Bei einer **„Diophantischen Gleichung“** werden nur Lösungen in $\mathbb{Z}$ gesucht.

Manchmal wird auch die verwandte Gleichung

$$X^2 - dY^2 = -1$$

oder sogar $X^2 - dY^2 = \pm 4$ als Pell-Gleichung bezeichnet.

Bemerkung 18.2. Wenn $d = e^2$ ein Quadrat ist, dann sucht die Pell-Gleichung nach zwei Quadraten a^2 und $(eb)^2$ mit Abstand 1. Das geht genau für $a = \pm 1$ und $b = 0$.

Wir nehmen daher im Folgenden an, daß $d \in \mathbb{N}$ kein Quadrat ist.

Bemerkung 18.3. Ganzzahlige Lösungen der Pell-Gleichung beschreibt man am besten über das Element

$$z = a + b\sqrt{d} \in \mathbb{Z}(\sqrt{d}),$$

mittels dessen die Pell-Gleichung die Form einer Normgleichung

$$N(z) = a^2 - db^2 = 1$$

annimmt. Für eine Lösung $z = a + b\sqrt{d}$ der Normgleichung ist

$$\widetilde{z} = a - b\sqrt{d} = N(z)/z = z^{-1}$$

das Inverse auch in $\mathbb{Z}[\sqrt{d}]$. Somit ist z eine Einheit in $\mathbb{Z}[\sqrt{d}]$. Genauer entsprechen Lösungen (a, b) der Pell-Gleichung als $z = a + b\sqrt{d}$ eindeutig den Einheiten von $\mathbb{Z}[\sqrt{d}]$ mit Norm 1.

Proposition 18.4. *Sei $d > 0$ kein Quadrat. Jede Lösung $(a, b) \in \mathbb{Z}^2$ mit $a, b > 0$ der Pell-Gleichung $X^2 - dY^2 = 1$ erfüllt:*

(1) Es sind a und b teilerfremd, und

(2) der gekürzte Bruch $\frac{a}{b}$ ist ein Näherungsbruch der Kettenbruchentwicklung von $\sqrt{d}$.

Beweis. (1) Der ggT von a, b teilt $a^2 - db^2 = 1$. Für Aussage (2) rechnen wir

$$\frac{a}{b} - \sqrt{d} = \frac{a - b\sqrt{d}}{b} = \frac{a^2 - db^2}{b(a + b\sqrt{d})} = \frac{1}{b(a + b\sqrt{d})} > 0.$$

Um mit Satz 8.25 (3) schließen zu können, müssen wir $\frac{a}{b} - \sqrt{d} < \frac{1}{2b^2}$ zeigen.

$$\frac{a}{b} - \sqrt{d} < \frac{1}{2b^2} \iff \frac{1}{b(a + b\sqrt{d})} < \frac{1}{2b^2} \iff 2b < \frac{a + b\sqrt{d}}{b} \iff 2 - 2\sqrt{d} < \frac{a}{b} - \sqrt{d}.$$

Nun ist die rechte Seite ist positiv, während die linke Seite negativ ist. □

18.2. **Periodische Kettenbrüche.**

Jetzt wissen wir, wo wir nach Lösungen der Pell-Gleichung suchen müssen und daß wir die Kettenbruchentwicklung von $\sqrt{d}$ verstehen müssen.

Definition 18.5. Ein **periodischer** Kettenbruch ist ein Kettenbruch

$$[a_0, a_1, \ldots],$$

so daß es ein $r > 0$ und n_0 gibt, so daß für alle $n \geq n_0$ gilt: $a_{n+r} = a_n$. Wir schreiben dies

$$[a_0, a_1, \ldots, \overline{a_{n_0}, \ldots, a_{n_0+r-1}}].$$

Ein **sofort periodischer** Kettenbruch ist ein periodischer Kettenbruch der Form

$$[\overline{a_0, a_1, \ldots, a_n}].$$

Beispiel 18.6. Der Kettenbruch des goldenen Schnitts ist sofort periodisch:

$$\varphi = \frac{1+\sqrt{5}}{2} = [\overline{1}].$$

Satz 18.7 (Euler, Lagrange)**.** *Sei $x \in \mathbb{R}$. Dann sind äquivalent:*

(a) *Die Zahl x hat eine periodische Kettenbruchentwicklung.*

(b) *Die Zahl x ist eine quadratische irrationale Zahl, d.h., $x \notin \mathbb{Q}$ liegt in einem reell-quadratischen Zahlkörper.*

Beweis. (a) $\Longrightarrow$ (b): Sei x eine reelle Zahl mit periodischer Kettenbruchentwicklung. Wenn $x = [a, y] = a + \frac{1}{y}$ mit $a \in \mathbb{Z}$, dann erzeugen x und y in $\mathbb{C}$ den gleichen Unterkörper. Damit ist x quadratisch irrational genau dann, wenn y dies ist. Wir dürfen also ohne Einschränkung annehmen, daß

$$x = [\overline{a_0, a_1, \ldots, a_n}]$$

sofort periodisch ist. Dann gilt mit der Standardnotation $M_n = \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix}$

$$x = [\overline{a_0, a_1, \ldots, a_n}] = [a_0, a_1, \ldots, a_n, x] = M_n(x) = \frac{p_n x + p_{n-1}}{q_n x + q_{n-1}}.$$

Daraus wird die quadratische Relation

$$x(q_n x + q_{n-1}) = p_n x + p_{n-1}.$$

Diese ist nichttrivial, weil $q_n \neq 0$.

(b) $\Longrightarrow$ (a): Sei x quadratisch irrational. Dann gibt es $a, b, c \in \mathbb{Z}$ mit

$$ax^2 + bx + c = 0,$$

also eine quadratische Form $f(X, Y) = aX^2 + bXY + cY^2$ mit

$$f(x, 1) = 0.$$

Die Diskriminante $D = D(f) = b^2 - 4ac$ muß $D > 0$ sein, damit die Nullstelle x reell aber nicht rational ist. Die Kettenbruchentwicklung mit Restgliedern sei

$$x = [a_0, a_1, \ldots, a_n, x_{n+1}] = \frac{p_n x_{n+1} + p_{n-1}}{q_n x_{n+1} + q_{n-1}}.$$

Die quadratische Form $f_n(X, Y) := f|_{M_n}(X, Y)$ hat dann $(x_{n+1}, 1)$ als Nullstelle, denn

$$\begin{aligned} f_n(x_{n+1}, 1) &= f(p_n x_{n+1} + p_{n-1}, q_n x_{n+1} + q_{n-1}) \\ &= (q_n x_{n+1} + q_{n-1})^2 f(\frac{p_n x + p_{n-1}}{q_n x + q_{n-1}}, 1) = (q_n x_{n+1} + q_{n-1})^2 f(x, 1) = 0. \end{aligned}$$

Wir zeigen, daß die quadratischen Formen f_n nur aus einem endlichen Vorrat stammen können und sich daher wiederholen müssen. Dann aber wiederholen sich auch die Nullstellen in der Form $(x_{n+1}, 1)$ und damit die Restglieder x_{n+1}. Sobald dies das erste Mal passiert, wird die Kettenbruchentwicklung periodisch.

Alle quadratischen Formen $f_n(X, Y) = a_n X^2 + b_n XY + c_n Y^2$ sind äquivalent und haben die gleiche Diskriminante D. Der Basiswechsel bedeutet konkret für die Koeffizienten insbesondere

$$\begin{aligned} a_n &= f(p_n, q_n), \\ c_n &= f(p_{n-1}, q_{n-1}) = a_{n-1}. \end{aligned}$$

Außerdem wissen wir, daß es ein $\delta_n \in \mathbb{R}$ mit $|\delta|_n < 1$ gibt mit

$$x = \frac{p_n}{q_n} - \frac{\delta_n}{q_n^2}.$$

Mit $F(X) = aX^2 + bX + c$ können wir per Taylorentwicklung um x die Abschätzung

$$\begin{aligned} |a_n| = |F(\frac{p_n}{q_n}) \cdot q_n^2| &= |F(x + \frac{\delta_n}{q_n^2})| \cdot q_n^2 \\ &= |F(x) + F'(x) \cdot \frac{\delta_n}{q_n^2} + \frac{1}{2}F''(x) \cdot (\frac{\delta_n}{q_n^2})^2| \cdot q_n^2 \\ &= |(2ax + b) \cdot \delta_n + a \cdot \frac{\delta_n^2}{q_n^2}| \leq |2ax + b| + |a|. \end{aligned}$$

erreichen. Damit ist $|a_n|$ nur in Abhängigkeit der Anfangsdaten a, b, x beschränkt. Dieselbe Abschätzung gilt demnach für $c_n = a_{n-1}$. Aus $b_n^2 = D + 4a_nc_n$ folgt eine obere Schranke für $|b_n|$ uniform in n. Insgesamt folgt damit, daß nur endlich viele a_n, b_n, c_n vorkommen. Dies zeigt wie bereits erwähnt die Behauptung. □

Lemma 18.8. *Wenn $x \in \mathbb{R}$ einen sofort periodischen Kettenbruch hat, dann ist $x > 1$.*

Beweis. Aus $x = [\overline{a_0, a_1, \ldots, a_n}]$ folgt

$$x = [\overline{a_0, a_1, \ldots, a_n}] = [a_0, a_1, \ldots, a_n, x].$$

Damit ist x das $n+1$-te Restglied der Kettenbruchentwicklung von x, also

$$T^{n+1}(x) = x.$$

Weil T nur Werte > 1 annimmt, folgt $x > 1$. □

Lemma 18.9. *Sei R ein Ring. Für jede Matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(R)$ gilt:*

$$\begin{pmatrix} & -1 \\ 1 & \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} & -1 \\ 1 & \end{pmatrix}^{-1} = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot (\begin{pmatrix} a & b \\ c & d \end{pmatrix}^t)^{-1}.$$

Beweis. Das rechnet man sofort nach. □

Lemma 18.10. *Sei $x = [\overline{a_0, a_1, \ldots, a_n}]$ ein sofort periodischer Kettenbruch. Dann ist*

$$y = [\overline{a_n, a_{n-1}, \ldots, a_0}]$$

gegeben durch $y = -1/\widetilde{x}$, wobei $\widetilde{x}$ das konjugierte Element im von x in $\mathbb{C}$ erzeugten reell-quadratischen Zahlkörper bezeichnet.

Beweis. Aus $x = [\overline{a_0, a_1, \ldots, a_n}] = [a_0, a_1, \ldots, a_n, x]$ folgt wie üblich mit

$$M = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix} \in \mathrm{GL}_2(\mathbb{Z}),$$

daß x ein Fixpunkt $x = M(x)$ der von M vermittelten Möbiustransformation ist. Analog folgt mit der transponierten Matrix

$$\begin{aligned} M^t &= \left(\begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix} \right)^t \\ &= \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix}^t \begin{pmatrix} a_{n-1} & 1 \\ 1 & 0 \end{pmatrix}^t \cdots \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}^t \\ &= \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n-1} & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

die Gleichung

$$y = M^t(y).$$

Wegen

$$\begin{pmatrix} & -1 \\ 1 & \end{pmatrix}(x) = -\frac{1}{x}$$

folgt (mit dem Produkt in der aufsteigenden Reihenfolge der Indizes)

$$\begin{aligned}
M^t(-\frac{1}{x}) &= M^t\begin{pmatrix} & -1 \\ 1 & \end{pmatrix}(x) = M^t\begin{pmatrix} & -1 \\ 1 & \end{pmatrix} M(x) \\
&= M^t\left(\prod_{i=0}^{n}\begin{pmatrix} & -1 \\ 1 & \end{pmatrix}\begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} & -1 \\ 1 & \end{pmatrix}^{-1}\right)\begin{pmatrix} & -1 \\ 1 & \end{pmatrix}(x) \\
&= M^t\left(\prod_{i=0}^{n} -\begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix}^{-1}\right)(-\frac{1}{x}) \qquad \text{(Lemma 18.9)} \\
&= \begin{pmatrix} (-1)^{n+1} & \\ & (-1)^{n+1} \end{pmatrix}(-\frac{1}{x}) = -\frac{1}{x}.
\end{aligned}$$

(Letzteres wegen $\begin{pmatrix} a & \\ & a \end{pmatrix}(z) = \frac{az}{a} = z$ für alle $a \neq 0$ und z.) Die Fixpunktgleichung

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}(z) = z$$

einer Möbiustransformation ist die quadratische Gleichung

$$z(cz + d) = az + b.$$

Somit gilt entweder $y = -\frac{1}{x}$, dies widerspricht $x > 1$ und $y > 1$ gemäß Lemma 18.8, oder eben

$$y = \widetilde{-1/x} = -1/\widetilde{x}. \qquad \square$$

Satz 18.11. *Ein $x \in \mathbb{R}$ hat eine sofort periodische Kettenbruchentwicklung genau dann, wenn x in einem reell-quadratischen Zahlköper liegt, $x \notin \mathbb{Q}$ und*

$$x > 1 \quad \textit{und} \quad -1 < \bar{x} < 0.$$

Beweis. Sei $x > 1$ eine quadratisch irrationale Zahl mit $-1 < \bar{x} < 0$. Satz 18.7 besagt, daß x eine periodische Kettenbruchentwicklung hat. Das bedeutet, daß die Restglieder

$$x_n = T^n(x)$$

in

$$x = [a_0, a_1, \ldots, a_{n-1}, x_n]$$

sich wiederholen: es gibt $r > 0$ und n_0, so daß für alle $n \geq n_0$ gilt

$$T^{n+r}(x) = T^n(x).$$

Wir müssen zeigen, daß wir hier T^n kürzen können. Denn dann gilt $x = T^r(x)$, und wegen

$$x = [a_0, a_1, \ldots, a_{r-1}, x]$$

hat x eine sofort periodische Kettenbruchentwicklung.

Aus $x > 1$ und $-1 < \bar{x} < 0$ folgt (sowieso)

$$T(x) > 1$$

und $\bar{x} - \lfloor x \rfloor < -1$, somit

$$-1 < \frac{1}{\bar{x} - \lfloor x \rfloor} < 0.$$

Damit vererben sich die Bedingungen $x > 1$ und $-1 < \bar{x} < 0$ auf $T(x)$. Per Induktion ist für alle $n \geq 0$

$$x_n > 1 \quad \text{und} \quad -1 < \overline{x_n} < 0.$$

Wenn $y, z > 1$ und $-1 < \bar{y}, \bar{z} < 0$, dann folgt aus $T(y) = T(z)$, daß

$$y - \lfloor y \rfloor = 1/T(y) = 1/T(z) = z - \lfloor z \rfloor.$$

Daraus folgt

$$\bar{y} - \bar{z} = \lfloor y \rfloor - \lfloor z \rfloor \in \mathbb{Z},$$

aber auch $-1 < \bar{y} - \bar{z} < 1$. Ergo $\bar{y} = \bar{z}$ und damit auch $z = y$. Wir können also T 'kürzen', und das beendet den Nachweis der sofortigen Periodizität.

Wir beweisen nun die umgekehrte Richtung. Sei $x = [\overline{a_0, a_1, \ldots, a_n}]$ sofort periodisch. Dann ist $x > 1$ nach Lemma 18.8. Außerdem hat nach Lemma 18.10 auch $-1/\bar{x}$ einen sofort periodischen Kettenbruch, somit wieder nach Lemma 18.8

$$-1/\bar{x} > 1.$$

Das übersetzt sofort in $-1 < \bar{x} < 0$. □

Korollar 18.12. *Sei $d > 0$ kein Quadrat. Sei $a_0 = \lfloor \sqrt{d} \rfloor$. Dann hat $x = a_0 + \sqrt{d}$ eine sofort periodische Kettenbruchentwicklung*

$$a_0 + \sqrt{d} = [\overline{2a_0, a_1, \ldots, a_n}]$$

und

$$\sqrt{d} = [a_0, \overline{a_1, \ldots, a_n, 2a_0}].$$

Beweis. Es gilt $x > 1$ und $-1 < \bar{x} = a_0 - \sqrt{d} < 0$ per Definition der Gauß-Klammer. Nach Satz 18.11 hat demnach $x = a_0 + \sqrt{d}$ einen sofort periodischen Kettenbruch. Dieser beginnt mit

$$\lfloor x \rfloor = \lfloor a_0 + \sqrt{d} \rfloor = 2a_0.$$

Die Kettenbruchentwicklung von $\sqrt{d}$ stimmt ab dem ersten Teilnenner mit der von $a_0 + \sqrt{d}$ überein, denn

$$T(\sqrt{d}) = \frac{1}{\sqrt{d} - a_0} = T(a_0 + \sqrt{d}).$$ □

Satz 18.13. *Sei $d > 0$ kein Quadrat. Die Pell-Gleichung $X^2 - dY^2 = 1$ hat eine Lösung*

$$z = a + b\sqrt{d} > 1.$$

Beweis. Wir nutzen die Kettenbruchentwicklung aus Korollar 18.12

$$\sqrt{d} = [a_0, \overline{a_1, \ldots, a_n, 2a_0}].$$

Damit gilt mit den üblichen Bezeichnungen

$$\sqrt{d} = [a_0, a_1, \ldots, a_n, a_0 + \sqrt{d}] = \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} (a_0 + \sqrt{d}) = \frac{p_n(a_0 + \sqrt{d}) + p_{n-1}}{q_n(a_0 + \sqrt{d}) + q_{n-1}}.$$

Ausmultiplizieren liefert

$$q_n d + (q_n a_0 + q_{n-1})\sqrt{d} = p_n a_0 + p_{n-1} + p_n\sqrt{d}.$$

Koeffizientenvergleich der Elemente aus $\mathbb{Z}[\sqrt{d}]$ bringt das Gleichungssytem

$$q_n d = p_n a_0 + p_{n-1}$$

$$p_n = q_n a_0 + q_{n-1}.$$

Die erste Gleichung mal q_n subtrahiert von p_n mal der zweiten Gleichung ergibt

$$p_n^2 - dq_n^2 = \det \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} = (-1)^{n+1}.$$

Die Periodenlänge der Kettenbruchentwicklung von $\sqrt{d}$ in unserer Darstellung ist $n + 1$. Wir sind also bei gerader Periodenlänge fertig. Sollte $n + 1$ ungerade sein, dann nehmen wir einfach die doppelte Periode als neue Periode. So wird die Periodenlänge $2n + 2$ gerade. □

Bemerkung 18.14. (1) Damit ist nun endlich Satz 19.11 vollständig bewiesen. Eine ganzzahlige Lösung $u = p_n + q_n\sqrt{d} > 1$ der Pell-Gleichung ist insbesondere eine Einheit $u \in \mathfrak{o}_K^\times$ verschieden von ± 1. Damit ist die diskrete Untergruppe $\log|\mathfrak{o}_K^\times|_\tau$ aus dem Beweis von Satz 19.11 nichttrivial.

(2) Die Situation stellt sich sogar ein bischen besser dar. Der Beweis von Satz 18.13 liefert ein effektives Verfahren, um eine Einheit $\varepsilon > 1$ in $\mathfrak{o}_K^\times$ anzugeben. Wir wissen einzig an dieser Stelle noch nicht, ob es sich um eine Fundamentaleinheit handelt.

19. Die Einheitengruppe quadratischer Zahlkörper

Wir bestimmen nun die Einheitengruppe $\mathfrak{o}_K^\times$. Die Antwort hängt wesentlich davon ab, ob der quadratische Zahlkörper K reell- oder imaginär-quadratisch ist. Ganz allgemein funktioniert aber die Charakterisierung der Einheiten über ihre Norm.

Proposition 19.1. *Sei K ein quadratischer Zahlkörper und $u \in \mathfrak{o}_K$ eine ganz algebraische Zahl aus K. Dann ist*

$$u \in \mathfrak{o}_K^\times \iff N(u) = \pm 1.$$

Beweis. Das folgt mit demselben Beweis wie bei den ganzen Gaußschen Zahlen in Proposition 15.8. □

19.1. Imaginär-quadratische Zahlkörper.

Satz 19.2. *Sei K ein imaginär-quadratischer Zahlkörper.*

(1) Die Einheitengruppe $\mathfrak{o}_K^\times$ ist eine endliche Gruppe.

(2) Sei $K = \mathbb{Q}(\sqrt{d})$ mit quadratfreiem $d \in \mathbb{Z}$, $d < 0$. Es gilt

$$\mathfrak{o}_K^\times = \begin{cases} \{\pm 1, \pm i\} \simeq \mathbb{Z}/4\mathbb{Z} & d = -1, \\ \{\pm 1, \frac{\pm 1 \pm \sqrt{-3}}{2}\} \simeq \mathbb{Z}/6\mathbb{Z} & d = -3, \\ \{\pm 1\} \simeq \mathbb{Z}/2\mathbb{Z} & \textit{sonst.} \end{cases}$$

Beweis. Sei $\zeta = a + b\sqrt{d} \in \mathfrak{o}_K^\times$. Dann sind in jedem Fall $2a, 2b \in \mathbb{Z}$ und

$$\pm 1 = N(\zeta) = a^2 - db^2 = a^2 + (-d)b^2$$

nach Proposition 19.1. Weil $d < 0$ nach Voraussetzung negativ ist, nimmt die Norm nur positive Werte an. Es gilt also $N(\zeta) = 1$ und damit $a, b \in \{-1, -1/2, 0, 1/2, 1\}$. Indem wir ζ durch $-\zeta$ ersetzen, dürfen wir $a \geq 0$ annehmen.

Fall 1: $a = 0$. Dann ist $\zeta = b\sqrt{d}$ und $N(\zeta) = 1$ bedeutet $-db^2 = 1$, also $d = -1$ und $\zeta = \pm i$.

Fall 2: $a = 1/2$. Dann ist b auch halbzahlig und $\zeta = (1 \pm \sqrt{d})/2$. Die Bedingung $N(\zeta) = 1$ liest sich dann als $(1 - d)/4 = 1$. Dies führt zu $d = -3$ und

$$\zeta = \frac{1 \pm \sqrt{-3}}{2}.$$

Fall 3: $a = 1$. Dann führt $N(\zeta) = 1 - db^2 = 1$ zu $b = 0$. In diesem Fall finden wir keine neue Einheit.

Damit haben wir alle Einheiten wie behauptet gefunden. Es bleibt zu zeigen, daß die Einheitengruppe stets zyklisch ist. Für $d = -1$ haben wir das bereits gesehen (i ist ein Erzeuger), und für $d \neq -1, -3$ ist das klar. Wenn $d = -3$, dann ist

$$\zeta_6 = \frac{1 + \sqrt{-3}}{2} = e^{2\pi i/6}$$

eine 6-te Einheitswurzel. Offensichtlich sind die Potenzen von ζ_6 auch Einheiten. Dies zeigt

$$\mathfrak{o}_{\mathbb{Q}(\sqrt{-3})}^\times = \langle \zeta_6 \rangle \simeq \mathbb{Z}/6\mathbb{Z}.$$ □

19.2. **Reell-quadratische Zahlkörper.** Sei nun $K = \mathbb{Q}(\sqrt{d})$ ein reell-quadratischer Zahlkörper mit $d \in \mathbb{Z}$ quadratfrei und $d > 0$.

Bemerkung 19.3. Sei (a, b) eine Lösung der Pell-Gleichung $X^2 - dY^2 = 1$. Dann ist $z = a + b\sqrt{d}$ ein Element von $\mathbb{Z}[\sqrt{d}]$ von Norm 1, somit nach Proposition 19.1 eine Einheit in $\mathfrak{o}_K$. Genauer ist z sogar eine Einheit im Unterring $\mathbb{Z}[\sqrt{d}] \subseteq \mathfrak{o}_K$, denn das Inverse ist $\widetilde{z} = a - b\sqrt{d}$. Wir haben

$$\{z = a + b\sqrt{d}\ ;\ (a, b) \text{ löst Pell: } X^2 - dY^2 = 1\} \subseteq \mathbb{Z}[\sqrt{d}]^\times \subseteq \mathfrak{o}_K^\times.$$

Genauer entsprechen die Lösungen der Pell-Gleichung den Elementen des Kerns der Normabbildung

$$N : \mathbb{Z}[\sqrt{d}]^\times \to \mathbb{Z}^\times = \{\pm 1\}.$$

Korollar 19.4. *Sei $K = \mathbb{Q}(\sqrt{d})$ ein reell-quadratischer Zahlkörper. Dann ist $\mathfrak{o}_K^\times$ unendlich.*

Beweis. Nach Satz 18.13 hat die Pell-Gleichung $X^2 - dY^2 = 1$ eine Lösung $z = a + b\sqrt{d} > 1$. Ein solches z liegt in $\mathfrak{o}_K^\times$ und kann nicht endliche Ordnung haben. □

Als nächstes bestimmen wir die Elemente endlicher Ordnung in $\mathfrak{o}_K^\times$. In einer abelschen Gruppe bilden die Elemente endlicher Ordnung eine Untergruppe, die **Torsionsuntergruppe**.

Proposition 19.5. *Sei K ein reell-quadratischer Zahlkörper. Dann besteht die Torsionsuntergruppe in $\mathfrak{o}_K^\times$ nur aus den Elementen ± 1.*

Beweis. Wenn $\zeta \in \mathfrak{o}_K^\times$ ein Torsionselement ist, also ein $n \in \mathbb{N}$ existiert mit $\zeta^n = 1$, dann liegt ζ in der komplexen Einbettung $K \subseteq \mathbb{R} \subseteq \mathbb{C}$ auf dem Einheitskreis. Dieser schneidet $\mathbb{R}$ nur in

$$\zeta = \pm 1.$$ □

Definition 19.6. Eine **diskrete** Untergruppe eines $\mathbb{R}^n$ ist eine Untergruppe $\Lambda \subseteq \mathbb{R}^n$, die im Sinne der reellen Topologie auf dem $\mathbb{R}$-Vektorraum $\mathbb{R}^n$ keinen Häufungspunkt hat.

Proposition 19.7. *Sei $\Lambda \subseteq \mathbb{R}^n$ eine Untergruppe. Dann sind äquivalent:*

(a) *Λ ist diskret.*

(b) *Es gibt ein $\varepsilon > 0$ gibt, so daß*

$$\{0\} = \{z \in \Lambda\ ;\ \|z\| < \varepsilon\}.$$

Dabei ist es egal, welche Norm $\|z\|$ auf $\mathbb{R}^n$ wir verwenden.

Beweis. Alle Normen auf $\mathbb{R}^n$ definieren die gleiche Topologie. Bedingung (b) sagt nur aus, daß $0 \in \Lambda$ in einer Umgebung von 0 das einzige Element der Gruppe Λ ist. Wenn Λ sich nicht in 0 häuft, dann muß (b) gelten.

Wir nehmen nun (b) an. Weil die Topologie auf $\mathbb{R}^n$ translationsinvariant ist und die Gruppe Λ durch Translation mit $\gamma \in \Lambda$ in sich überführt wird, kann man (b) um ein $\gamma \in \Lambda$ zu der Aussage

$$\{\gamma\} = \{z \in \Lambda\ ;\ \|z - \gamma\| < \varepsilon\}$$

transformieren. Damit kann aber keine Folge in Λ eine Cauchy-Folge sein, denn unter die Schranke ε werden sich die Punkte aus Λ nicht annähern. □

Proposition 19.8. *Eine diskrete Untergruppe $\Lambda \subseteq \mathbb{R}$ ist entweder $\Lambda = \{0\}$ oder es gibt ein $0 \neq x \in \mathbb{R}$ und*

$$\Lambda = \{nx\ ;\ n \in \mathbb{Z}\} \simeq \mathbb{Z}.$$

Beweis. Die Menge

$$\{z \in \Lambda\ ;\ z > 0\}$$

ist entweder leer, dann haben wir $\Lambda = \{0\}$, oder aber sie besitzt als nach unten durch 0 beschränkte Menge ein Infimum x. Weil Λ diskret ist und das Infimum ein Häufungspunkt von Λ, also ein Grenzwert einer Folge aus Λ ist, muß besagte Folge letztendlich konstant werden: das

Infimum ist ein Minimum (wird angenommen). Damit ist auch $x > 0$, denn jedes Element in der betrachteten Menge ist ja > 0.

Angnommen es gibt ein Element $z \in \Lambda \backslash \mathbb{Z} \cdot x$. Wir betrachten $\vartheta = z/x$ und setzen $n = \lfloor \vartheta \rfloor \in \mathbb{Z}$. Dann ist auch

$$z' = z - nx \in \Lambda \setminus \mathbb{Z} \cdot x,$$

aber nach Konstruktion

$$0 < z' < x.$$

Das ist ein Widerspruch zur Konstruktion von x. Folglich gilt schon $\Lambda = \mathbb{Z} \cdot x$ wie behauptet. □

Notation 19.9. Seien K ein reell-quadratischer Zahlkörper und $\tau : K \hookrightarrow \mathbb{R}$ eine Körpereinbettung (von denen es zwei gibt: die definierende $K \subseteq \mathbb{R} \subseteq \mathbb{C}$ und die dzu konjugierte Einbettung, bei der man die Konjugation von K vorschaltet). Dann bezeichnen wir den induzierten Absolutbetrag mit

$$|z|_\tau := |\tau(z)|.$$

Satz 19.10. *Sei K ein reell-quadratischer Zahlkörper mit reeller Einbettung $\tau : K \hookrightarrow \mathbb{R}$. Dann ist*

$$\mathfrak{o}_K^\times \to (\mathbb{R}, +), \qquad z \mapsto \log|z|_\tau$$

ein Gruppenhomomorphismus mit

(1) Kern $= \{\pm 1\}$,
(2) und das Bild ist eine diskrete Untergruppe von $\mathbb{R}$.

Beweis. Seien $z, w \in \mathfrak{o}_K^\times$. Dann gilt

$$\log|zw|_\tau = \log|\tau(z)\tau(w)| = \log\big(|\tau(z)| \cdot |\tau(w)|\big) = \log|z|_\tau + \log|w|_\tau.$$

(1) Der Kern besteht aus allen Elementen $z \in \mathfrak{o}_K^\times$ mit $\log|z|_\tau = 0$. Das bedeutet $|z|_\tau = 1$, also $\tau(z) = \pm 1$, oder eben $z = \pm 1$.

(2) Sei $\bar{\tau} : K \hookrightarrow \mathbb{R}$ die andere, konjugierte Einbettung nach $\mathbb{R}$. Dann gilt für alle $z \in K$

$$N(z) = z\bar{z} = \tau(z) \cdot \bar{\tau}(z).$$

Angenommen wir haben eine Folge von Einheiten $u_n \in \mathfrak{o}_K^\times$, so daß

$$\log|u_n|_\tau \to 0.$$

Weil u_n Einheit ist, folgt $\pm 1 = N(u_n) = \tau(u_n) \cdot \bar{\tau}(u_n)$ und damit

$$\log|u_n|_\tau = -\log|u_n|_{\bar{\tau}}.$$

Folglich konvergiert auch

$$\log|u_n|_{\bar{\tau}} \to 0.$$

Das bedeutet, daß für $n \gg 0$ die Werte u_n und $\bar{u}_n$ sich in einer ε-Umgebung von ± 1 aufhalten. Wegen

$$\mathrm{tr}(u_n) = u_n + \bar{u}_n \in \mathbb{Z}$$

kommen dann nur noch die Werte $-2, 0, 2$ für die Spur in Frage. Weil auch die Norm nur die Werte ± 1 annimmt, muß die Folge (u_n) letztlich nur aus Lösungen von endlich vielen quadratischen Gleichungen

$$T^2 - sT + N = 0$$

mit $N = \pm 1$ und $s \in \{-2, 0, 2\}$ bestehen. Das gilt dann auch für die Folge $(\log|u_n|_\tau)$. Eine solche Folge konvergiert nur, wenn sie letztlich konstant ist. Dies aber zeigt, daß das Bild

$$\Lambda = \log|\mathfrak{o}_K^\times|_\tau \subseteq \mathbb{R}$$

eine diskrete Untergruppe ist. □

Satz 19.11. *Die Einheitengruppe eines reell-quadratischen Zahlkörpers K hat die Struktur*

$$\mathfrak{o}_K^\times = \{\pm 1\} \times \eta^{\mathbb{Z}}$$

mit einer eindeutigen ***Fundamentaleinheit*** $\eta > 1$ *aus* $\mathfrak{o}_K^\times$.

Beweis. Nach Satz 19.10 ist $\Lambda = \mathfrak{o}_K^\times/\{\pm 1\}$ isomorph zu einer diskreten Untergruppe in $\mathbb{R}$. Nach Proposition 19.8 ist damit entweder $\Lambda = \{0\}$ oder $\Lambda \simeq \mathbb{Z}$. In letzterem Fall gibt es einen eindeutigen positiven Erzeuger, entsprechend $\eta \in \mathfrak{o}_K^\times$ mit

$$|\eta|_\tau > 1.$$

Nach Wahl des richtigen Vorzeichens gilt dann sogar $\eta > 1$ mit eindeutigem $\eta \in \mathfrak{o}_K^\times$. Die Struktur der Einheitengruppe wie angegeben folgt dann sofort.

Es bleibt zu zeigen, daß das Bild Λ tatsächlich nichttrivial ist. Das folgt sofort aus Korollar 19.4, weil ansonsten $\mathfrak{o}_K^\times$ endlich wäre. □

19.3. **Die Fundamentaleinheit eines reell-quadratischen Zahlkörpers.** Jetzt wissen wir also, daß es eine eindeutige Fundamentaleinheit $\eta > 1$ in $\mathfrak{o}_K^\times$ gibt, so daß

$$\mathfrak{o}_K^\times = \{\pm 1\} \times \eta^{\mathbb{Z}}.$$

Es bleibt anzugeben, wie man η bestimmt und wie man daraus wiederum die Pell-Gleichung vollständig löst.

Satz 19.12. *Seien $d > 0$ quadratfrei und $K = \mathbb{Q}(\sqrt{d})$. Dann gilt*

$$\mathbb{Z}[\sqrt{d}]^\times = \begin{cases} \mathfrak{o}_K^\times & d \not\equiv 5 \pmod 8 \\ \ker\left(\mathfrak{o}_K^\times \to (\mathfrak{o}_K/2\mathfrak{o}_K)^\times\right) & d \equiv 5 \pmod 8. \end{cases}$$

Insbesondere ist die Einheitengruppe $\mathbb{Z}[\sqrt{d}]^\times$ von endlichem Index in $\mathfrak{o}_K^\times$, und dieser Index ist entweder 1 oder 3.

Beweis. Schritt 1: Wenn $d \not\equiv 1 \pmod 4$, dann ist $\mathfrak{o}_K = \mathbb{Z}[\sqrt{d}]$ und ebenso für die Einheitengruppe

$$\mathbb{Z}[\sqrt{d}]^\times = \mathfrak{o}_K^\times.$$

Schritt 2: Wir nehmen daher nun $d \equiv 1 \pmod 4$ an und setzen $\omega = \frac{1+\sqrt{d}}{2}$. Dann gilt

$$\mathfrak{o}_K = \{a + b\omega \ ; \ a, b \in \mathbb{Z}\} \supseteq \mathbb{Z}[\sqrt{d}] = \{a + b\omega \in \mathfrak{o}_K \ ; \ b \text{ gerade}\}.$$

Die Inklusion

$$\mathbb{Z}[\sqrt{d}]^\times \subseteq \mathfrak{o}_K^\times$$

ist trivial. Für $z = a + b\omega \in \mathfrak{o}_K^\times$ ist das Inverse gegeben durch ($\bar{\omega} = 1 - \omega$)

$$z^{-1} = \frac{\bar{z}}{N(z)} = \pm(a + b\bar{\omega}) = \pm(a + b - b\omega).$$

Wenn $z \in \mathfrak{o}_K^\times \cap \mathbb{Z}[\sqrt{d}]$, also wenn b gerade ist, dann ist auch $z^{-1} \in \mathbb{Z}[\sqrt{d}]$. Dies zeigt

$$\mathbb{Z}[\sqrt{d}]^\times = \mathfrak{o}_K^\times \cap \mathbb{Z}[\sqrt{d}].$$

Schritt 3: Sei $z = a + b\omega \in \mathfrak{o}_K^\times$. Wenn b gerade ist, dann muß a ungerade sein, sonst wäre $2 \mid z$ und

$$4 = N(2) \mid N(z) = \pm 1.$$

Die Bedingungen modulo 2 an a und b lassen sich nun simultan als Kongruenz in $\mathfrak{o}_K$ nach dem Ideal $2\mathfrak{o}_K$ beschreiben:

$$2\mathfrak{o}_K = \{a + b\omega \ ; \ a, b \in 2\mathbb{Z}\} = 2\mathbb{Z} \oplus 2\omega\mathbb{Z}.$$

Der Faktorring $\mathfrak{o}_K/2\mathfrak{o}_K$ ist als abelsche Gruppe nichts anderes als

$$\mathfrak{o}_K/2\mathfrak{o}_K = \mathbb{Z}/2\mathbb{Z}[1] \oplus \mathbb{Z}/2\mathbb{Z}[\omega],$$

wenn wir mit $[z]$ die Restklasse von z modulo $2\mathfrak{o}_K$ bezeichnen. Da $\mathbb{Z}/2\mathbb{Z} = \mathbb{F}_2$ gilt, handelt es sich um einen 2-dimensionalen $\mathbb{F}_2$-Vektorraum, der zudem ein Ring ist. Die Quotientenabbildung $\mathfrak{o}_K \twoheadrightarrow \mathfrak{o}_K/2\mathfrak{o}_K$ induziert auf Einheiten einen Gruppenhomomorphismus

$$\mathfrak{o}_K^\times \to (\mathfrak{o}_K/2\mathfrak{o}_K)^\times.$$

Damit gilt

$$\begin{aligned}\mathbb{Z}[\sqrt{d}]^\times &= \mathfrak{o}_K^\times \cap \mathbb{Z}[\sqrt{d}] = \{a + b\omega \in \mathfrak{o}_K^\times \ ; \ b \text{ gerade}\} \\ &= \{z = a + b\omega \in \mathfrak{o}_K^\times \ ; \ z \equiv 1 \pmod{2\mathfrak{o}_K}\} = \ker\big(\mathfrak{o}_K^\times \to (\mathfrak{o}_K/2\mathfrak{o}_K)^\times\big).\end{aligned}$$

Schritt 4: Es bleibt zu zeigen, daß die Projektion $\mathfrak{o}_K^\times \to (\mathfrak{o}_K/2\mathfrak{o}_K)^\times$ modulo 2 höchstens für $d \equiv 5 \pmod 8$ nichttrivial sein kann. Es gilt

$$\mathfrak{o}_K \simeq \mathbb{Z}[\omega] \simeq \mathbb{Z}[X]/(X^2 - X + \frac{1-d}{4}).$$

Damit ist die Reduktion modulo 2

$$\mathfrak{o}_K/2\mathfrak{o}_K \simeq \mathbb{Z}[X]/(2, X^2 - X + \frac{1-d}{4}) \simeq \mathbb{F}_2[X]/(X^2 - X + \frac{1-d}{4}).$$

Wenn $d \equiv 1 \pmod 8$, dann ist $X^2 - X + \frac{1-d}{4} \equiv X(X-1) \pmod 2$ und nach dem Chinesischen Restsatz für $\mathbb{F}_2[X]$ folgt

$$\mathfrak{o}_K/2\mathfrak{o}_K \simeq \mathbb{F}_2[X]/(X(X-1)) \simeq \mathbb{F}_2[X]/(X) \times \mathbb{F}_2[X]/(X-1) \simeq \mathbb{F}_2 \times \mathbb{F}_2.$$

Wenn $d \equiv 5 \pmod 8$, dann ist $X^2 - X + \frac{1-d}{4} \equiv X^2 + X + 1 \pmod 2$ irreduzibel (hat keine Nullstelle in $\mathbb{F}_2$), nämlich das Minimalpolynom der dritten Einheitswurzeln über $\mathbb{F}_2$

$$\mathfrak{o}_K/2\mathfrak{o}_K \simeq \mathbb{F}_2[X]/(X^2 + X + 1) \simeq \mathbb{F}_4.$$

Damit ist

$$(\mathfrak{o}_K/2\mathfrak{o}_K)^\times = \begin{cases} (\mathbb{F}_2 \times \mathbb{F}_2)^\times = 1 & d \equiv 1 \pmod 8 \\ \mathbb{F}_4^\times \simeq \mathbb{Z}/3\mathbb{Z} & d \equiv 5 \pmod 8. \end{cases}$$

Nach dem Homomorphiesatz ist $\mathfrak{o}_K^\times/\mathbb{Z}[\sqrt{d}]^\times$ eine Untergruppe von $(\mathfrak{o}_K/2\mathfrak{o}_K)^\times$. Diese Untergruppe ist entweder trivial oder von Ordnung 3 (Satz von Lagrange), somit ist $\mathbb{Z}[\sqrt{d}]^\times$ vom Index 1 oder 3 in $\mathfrak{o}_K^\times$. □

Beispiel 19.13. Der Fall von Index 3 in Satz 19.12 tritt tatsächlich auf: $d = 5$ und $\varphi = \frac{1+\sqrt{5}}{2}$ ist das erste Beispiel.

Korollar 19.14. *Seien $d > 0$ quadratfrei und $K = \mathbb{Q}(\sqrt{d})$. Sei $\eta > 1$ die Fundamentaleinheit von $\mathfrak{o}_K$. Die ganzzahligen Lösungen der Pell–Gleichung*

$$X^2 - dY^2 = 1$$

sind von der Form $(x, y) \in \mathbb{Z}^2$ mit

$$x + y\sqrt{d} = \pm\eta^n$$

für ein $n \in \mathbb{Z}$, wobei die folgenden Bedingungen zu erfüllen sind:

(i) $2 \mid n$, *wenn* $N(\eta) = -1$,

(ii) $3 \mid n$, *wenn* $\eta \notin \mathbb{Z}[\sqrt{d}]$.

Bedingung (ii) tritt höchstens ein, wenn $d \equiv 5 \pmod 8$.

Beweis. Die ganzzahligen Lösungen der Pell-Gleichung sind genau die Einheiten von $\mathfrak{o}_K$, die in $\mathbb{Z}[\sqrt{d}]$ liegen und Norm 1 haben. Als Elemente von $\mathfrak{o}_K^\times$ haben diese die angegebene Form

$$x + y\sqrt{d} = \pm\eta^n.$$

Die Bedingung (i) sorgt genau dafür, daß

$$x^2 - dy^2 = N(x + y\sqrt{d}) = N(\pm\eta^n) = N(\eta)^n = 1.$$

Nach Satz 19.12 gibt es mit der Bedingung $\pm\eta^n \in \mathbb{Z}[\sqrt{d}]$ höchstens dann ein Problem, wenn $d \equiv 5 \pmod 8$ und wenn die Reduktion modulo 2

$$\mathfrak{o}_K^\times \to (\mathfrak{o}_K/2\mathfrak{o}_K)^\times$$

nichttrivial ist. Das geht nur, wenn η halbzahlig ist, also in $\mathfrak{o}_K/2\mathfrak{o}_K$ verschieden von 1. Dann erzeugt das Bild von η in $(\mathfrak{o}_K/2\mathfrak{o}_K)^\times \simeq \mathbb{Z}/3\mathbb{Z}$ diese Gruppe und $\pm\eta^n$ liegt im Kern genau dann, wenn $3 \mid n$. □

Korollar 19.15. *Seien $d > 0$ quadratfrei und $K = \mathbb{Q}(\sqrt{d})$. Zur Pell–Gleichung*

$$X^2 - dY^2 = 1$$

gibt es eine ***Fundamentallösung*** *$(a, b) \in \mathbb{Z}^2$ mit $a, b > 0$, so daß mit*

$$z = a + b\sqrt{d}$$

die Menge aller ganzzahligen Lösungen durch

$$\{(x, y) \in \mathbb{Z}^2 \ ; \ x^2 - dy^2 = 1\} = \{(x, y) \ ; \quad \text{es gibt eindeutig } n \in \mathbb{Z} : x + y\sqrt{d} = \pm(a + b\sqrt{d})^n\}$$

gegeben ist. Setzen wir

$$a_n + b_n\sqrt{d} = (a + b\sqrt{d})^n,$$

dann beschreibt $a_0 = 1$, $b_0 = 0$ und

$$a_{n+1} = aa_n + bdb_n$$

$$b_{n+1} = ba_n + ab_n$$

rekursiv die Menge aller Lösungen $(x, y) \in \mathbb{Z}^2$ mit $x, y \geq 0$.

Beweis. Das folgt sofort aus Korollar 19.14 mit der Fundamentaleinheit $\eta \in \mathfrak{o}_K^\times$ bei $K = \mathbb{Q}(\sqrt{d})$ und

$$z = \eta^e,$$

wobei

$$e = \begin{cases} 1 & \text{falls } N(\eta) = 1 \text{ und } \eta \in \mathbb{Z}[\sqrt{d}], \\ 2 & \text{falls } N(\eta) = -1 \text{ und } \eta \in \mathbb{Z}[\sqrt{d}], \\ 3 & \text{falls } N(\eta) = 1 \text{ und } \eta \notin \mathbb{Z}[\sqrt{d}], \\ 6 & \text{falls } N(\eta) = -1 \text{ und } \eta \notin \mathbb{Z}[\sqrt{d}]. \end{cases}$$

□

Bemerkung 19.16. Wenn d nicht mehr notwendigerweise quadratfrei, sondern nur noch kein Quadrat in $\mathbb{Z}$ ist, dann bekommt man für den quadratfreien Anteil $d = \delta \cdot m^2$ durch Korollar 19.14 einen Überblick über sämtliche ganzzahlige Lösungen von

$$X^2 - \delta Y^2 = 1,$$

woraus die ganzzahligen Lösungen (x, y) von

$$X^2 - dY^2 = 1$$

wegen $X^2 - dY^2 = X^2 - \delta(mY)^2$ durch eine zusätzliche Kongruenzbedingung $y \equiv 0 \pmod m$ gewonnen werden. Die Rekursionen erlauben leicht zu entscheiden, für welche Potenzen der Fundamentallösung zu δ die Kongruenz gilt.

Die Fundamentaleinheit bestimmt man algorithmisch wie folgt.

Satz 19.17 (Bestimmung der Fundamentaleinheit)**.** *Seien $d > 0$ quadratfrei, $K = \mathbb{Q}(\sqrt{d})$ und $\eta \in \mathfrak{o}_K^\times$ die Fundamentaleinheit. Wir setzen*

$$\omega = \begin{cases} \sqrt{d} & d \not\equiv 1 \pmod 4 \\ \frac{1+\sqrt{d}}{2} & d \equiv 1 \pmod 4 \end{cases}$$

und weiter

$$\vartheta = T(\omega) = \frac{1}{\omega - \lfloor \omega \rfloor}.$$

Dann hat ϑ eine sofort periodische Kettenbruchentwicklung

$$\vartheta = [\overline{a_0, \ldots, a_{n-1}}],$$

und mit der üblichen Notation

$$M := \begin{pmatrix} p_{n-1} & p_{n-2} \\ q_{n-1} & q_{n-2} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-1} & 1 \\ 1 & 0 \end{pmatrix}$$

gilt die folgende Formel für die Fundamentaleinheit:

$$\eta = q_{n-1}\vartheta + q_{n-2}.$$

Weiter gilt

$$N(\eta) = (-1)^n,$$

wobei n die Periodenlänge der Kettenbruchentwicklung von ϑ ist.

Beweis. Schritt 1: Es gilt $\vartheta + T(\omega) > 1$ und $-1 < \bar{\vartheta} = \frac{1}{\bar{\omega} - \lfloor \omega \rfloor} < 0$, weil

$$\bar{\omega} - \lfloor \omega \rfloor < -1.$$

Damit hat ϑ nach Satz 18.11 eine sofort periodische Kettenbruchentwicklung wie behauptet.

Schritt 2: Es gilt $\mathfrak{o}_K = \mathbb{Z}[\omega] = \langle 1, \omega \rangle_{\mathbb{Z}}$. Wir betrachten in K die Menge

$$\vartheta \cdot \mathfrak{o}_K = \vartheta \langle 1, \omega \rangle_{\mathbb{Z}} = \vartheta \langle 1, \omega - \lfloor \omega \rfloor \rangle_{\mathbb{Z}} = \langle \vartheta, 1 \rangle_{\mathbb{Z}}.$$

Dies ist eine freie abelsche Gruppe vom Rang 2, die stabil ist unter Multiplikation mit allen $z \in \mathfrak{o}_K$, ein $\mathfrak{o}_K$-Modul isomorph zu $\mathfrak{o}_K$:

$$z \cdot (\vartheta \cdot \mathfrak{o}_K) = \vartheta \cdot (z \cdot \mathfrak{o}_K) \subseteq \vartheta \cdot \mathfrak{o}_K.$$

Aus

$$\vartheta = M(\vartheta) = \frac{p_{n-1}\vartheta + p_{n-2}}{q_{n-1}\vartheta + q_{n-2}}$$

und der Definition von $\eta := q_{n-1}\vartheta + q_{n-2}$ folgt

$$\begin{aligned} \eta \cdot \vartheta &= p_{n-1}\vartheta + p_{n-2} \\ \eta \cdot 1 &= q_{n-1}\vartheta + q_{n-2}, \end{aligned}$$

so daß die transponierte Matrix M^t die Multiplikation mit η auf $\vartheta \cdot \mathfrak{o}_K$ in der $\mathbb{Z}$-Basis $\vartheta, 1$ darstellt. Nach Cayley-Hamilton folgt

$$(M^t)^2 - (p_{n-1} + q_{n-1})M^t + (-1)^n = 0$$

(der konstante Term ist die Determinante der Matrix) und durch Auswerten bei $1 \in \langle \vartheta, 1 \rangle_{\mathbb{Z}}$ folgt

$$\eta^2 - (p_{n-1} + q_{n-1})\eta + (-1)^n = 0.$$

Die Koeffizienten sind ganzzahlig, also

$$\eta \in \mathfrak{o}_K,$$

und der konstante Term ist die Norm: $(-1)^n = N(\eta) = \eta \cdot \bar{\eta}$. Damit ist sogar $\eta \in \mathfrak{o}_K^\times$. Weiter folgt

$$\eta = q_{n-1}\vartheta + q_{n-2} > q_{n-1} + q_{n-2} > 1.$$

Schritt 3: Wir müssen nun einsehen, daß jede Einheit $\varepsilon \in \mathfrak{o}_K^\times$ mit $\varepsilon > 1$ eine Potenz von η ist. Als Einheit operiert ε auf $\vartheta\mathfrak{o}_K$ invertierbar. Wir betrachten die Matrix bezüglich der Basis $\vartheta, 1$, also

$$(\varepsilon \cdot) = \begin{pmatrix} a & c \\ b & d \end{pmatrix} =: A \in \mathrm{GL}_2(\mathbb{Z})$$

mit

$$\begin{aligned}\varepsilon \cdot \vartheta &= a\vartheta + b\\ \varepsilon \cdot 1 &= c\vartheta + d.\end{aligned}$$

Dann ist

$$\vartheta = \frac{\varepsilon\vartheta}{\varepsilon} = \frac{a\vartheta + b}{c\vartheta + d} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}(\vartheta) = A^t(\vartheta).$$

Wenn wir zeigen können, daß für ein $r \in \mathbb{N}$

$$A^t = M^r,$$

dann gilt $A = (M^t)^r$ und nach Auswertung in $1 \in \langle \vartheta, 1\rangle_{\mathbb{Z}}$:

$$\varepsilon = \varepsilon \cdot 1 = A\begin{pmatrix} 0 \\ 1 \end{pmatrix} = (M^t)^r \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \eta^r \cdot 1 = \eta^r$$

und in der Tat ist η dann die Fundamentaleinheit.

Schritt 4: Aus $\varepsilon > 1$ und $\bar{\varepsilon} = N(\varepsilon)/\varepsilon$ folgt

$$|\bar{\varepsilon}| < 1.$$

Daher ist $\varepsilon - \bar{\varepsilon} > 0$, $\vartheta - \bar{\vartheta} > 0$ und $\varepsilon\vartheta - \bar{\varepsilon}\bar{\vartheta} > 0$ und somit

$$\begin{aligned} c &= \frac{\varepsilon - \bar{\varepsilon}}{\vartheta - \bar{\vartheta}} > 0,\\ a &= \frac{\varepsilon\vartheta - \bar{\varepsilon}\bar{\vartheta}}{\vartheta - \bar{\vartheta}} > 0.\end{aligned}$$

Schritt 5: Es gilt $\vartheta(c\vartheta + d) = \vartheta\varepsilon = a\vartheta + b$. Die Nullstellen des Polynoms

$$f(X) = cX^2 + (d - a)X - b$$

sind daher ϑ und $\bar{\vartheta}$. Aus $\vartheta > 1$ und $-1 < \bar{\vartheta} < 0$ und weil dies die einzigen Nullstellen sind, folgt

$$f(0) < 0, \quad f(-1) > 0, \quad f(1) < 0.$$

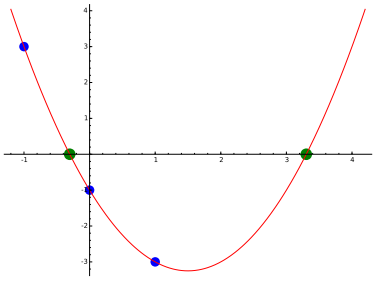


Abbildung 8. Vorzeichen von $f(X)$ aufgrund der Lage der Nullstellen $\bar{\vartheta}$ und ϑ

Das bedeutet $b = -f(0) > 0$ und $c + a - d - b = f(-1) > 0$. Damit ist auch

$$d = \frac{\det A + bc}{a} = \frac{\pm 1 + bc}{a} \geq 0.$$

Schlußendlich sehen wir

$$(a + c)(c - d) = c(c + a - d - b) - \det(A) \geq 1 \cdot 1 - (\pm 1) \geq 0,$$

also

$$c \geq d \geq 0.$$

Schritt 6: Da der ggT von a, b ein Teiler von $\det(A) = \pm 1$ ist, sind a und c teilerfremd. Wir bestimmen eine Kettenbruchentwicklung

$$\frac{a}{c} = [b_0, \ldots, b_{m-1}],$$

wobei wir durch den üblichen Trick die Parität der Länge m so wählen, daß

$$N(\varepsilon) = (-1)^m.$$

Dann ist

$$\begin{pmatrix} a & \beta \\ c & \delta \end{pmatrix} = \begin{pmatrix} b_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} b_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} b_{m-1} & 1 \\ 1 & 0 \end{pmatrix}$$

und

$$ad - bc = \det(A) = N(\varepsilon) = (-1)^m = \det(\begin{pmatrix} a & \beta \\ c & \delta \end{pmatrix}) = a\delta - \beta c.$$

Daraus folgt

$$a(d - \delta) = c(b - \beta),$$

und, weil a, c teilerfremd sind, gibt es ein $k \in \mathbb{Z}$ mit

$$\beta = b + ka$$
$$\delta = d + kc.$$

Es folgt nun aus den Abschätzungen zu den Näherungsbrüchen der Kettenbruchentwicklung, daß

$$c \geq \delta \geq 0.$$

Schritt 7: Wir zeigen nun $k = 0$ oder gleichbedeutend

$$A^t = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & \beta \\ c & \delta \end{pmatrix}. \tag{19.1}$$

Aus den bewiesenen Ungleichungen $c \geq d, \delta \geq 0$ und der Kongruenz $d \equiv \delta = d + kc \pmod{c}$ folgt dies bereits, außer in dem Fall, daß $d = c$ und $\delta = 0$ oder umgekehrt. Weil $(c, d) = 1$ und $(c, \delta) = 1$, denn beide Paare sind Zeile in einer Matrix aus $\mathrm{GL}_2(\mathbb{Z})$, folgt dann $c = 1$. Es gibt nun zwei verbleibende Fälle.

Fall $d = 0$: die Matrix $A^t = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix}$ hat Determinante ± 1 und $b > 0$, also $b = 1$. Insbesondere ist $N(\varepsilon) = -1$. Dies entspricht der Matrix aus der Kettenbruchentwicklung $a = a/c = [a]$, die wir in diesem Fall nehmen müssen. Also gilt (19.1) auch in diesem Fall.

Fall $d = 1$: die Matrix $A^t = \begin{pmatrix} a & b \\ 1 & 1 \end{pmatrix}$ hat Determinante ± 1 und $a > b$, denn

$$0 < f(-1) = c + a - d - b = a - b,$$

folglich $b = a - 1$. Insbesondere ist $N(\varepsilon) = 1$, und wir müssen die Kettenbruchentwicklung $a = a/1 = [a - 1, 1]$ nemen. Dies führt zur Matrix

$$\begin{pmatrix} a-1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} a & a-1 \\ 1 & 1 \end{pmatrix} = A^t.$$

Also gilt (19.1) auch in diesem Fall.

Schritt 8: Es gilt also $k = 0$ und $b = \beta$ und $d = \gamma$:

$$A^t = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & \beta \\ c & \delta \end{pmatrix} = \begin{pmatrix} b_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} b_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} b_{m-1} & 1 \\ 1 & 0 \end{pmatrix}$$

Mit $\vartheta = A^t(\vartheta)$ folgt, daß

$$\vartheta = [\overline{b_0, b_1, \ldots, b_{m-1}}]$$

auch eine Kettenbruchentwicklung von ϑ sein muß. Aus der Eindeutigkeit der Kettenbruchentwicklung folgt,

$$b_0, b_1, \ldots, b_{m-1} = \underbrace{a_0, a_1, \ldots, a_{n-1}}_{1}, \ldots, \underbrace{a_0, a_1, \ldots, a_{n-1}}_{r}$$

mit $r = m/n$. Damit gilt

$$\begin{aligned} A^t &= \begin{pmatrix} b_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} b_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} b_{m-1} & 1 \\ 1 & 0 \end{pmatrix} \\ &= \left(\begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-1} & 1 \\ 1 & 0 \end{pmatrix} \right)^r = M^r \end{aligned}$$

wie benötigt. □

Beispiel 19.18. Wir bestimmen die Fundamentaleinheit und die Fundamentallösung der zugehörigen Pell-Gleichung für einige quadratische Zahlkörper $K = \mathbb{Q}(\sqrt{d})$.

(1) $d = 5$: Hier ist

$$\omega = \varphi = [\overline{1}],$$

also auch $\vartheta = T(\varphi) = \varphi$ und die Fundamentaleinheit ist

$$\eta = 1 \cdot \varphi + 0 = \varphi = \frac{1 + \sqrt{5}}{2}.$$

Weiter ist

$$N(\varphi) = \varphi(1 - \varphi) = -1,$$

so daß

$$z = \varphi^6 = 9 + 4\sqrt{5}$$

zur Fundamentallösung $(5, 4)$ der Pell-Gleichung $X^2 - 5Y^2 = 1$ führt.

(2) $d = 7$: Hier ist $\omega = \sqrt{7}$, also

$$\vartheta = T(\omega) = \frac{1}{\sqrt{7} - 2} = \frac{2 + \sqrt{7}}{3} = [\overline{1, 1, 1, 4}]$$

und die Fundamentaleinheit ist

$$\eta = 9 \cdot \vartheta + 2 = 8 + 3\sqrt{7}.$$

Weiter ist

$$N(\eta) = 1,$$

so daß

$$z = \eta = 8 + 3\sqrt{7}$$

zur Fundamentallösung $(8, 3)$ der Pell-Gleichung $X^2 - 7Y^2 = 1$ führt.

(3) $d = 37$: Hier ist $\omega = \frac{1+\sqrt{37}}{2}$, also

$$\vartheta = T(\omega) = \frac{1}{\omega - 3} = \frac{5 + \sqrt{37}}{6} = [\overline{1, 1, 5}]$$

und die Fundamentaleinheit ist

$$\eta = 6 \cdot \vartheta + 1 = 6 + \sqrt{37}.$$

Weiter ist

$$N(\eta) = -1,$$

so daß

$$z = \eta^2 = 73 + 12\sqrt{37}$$

zur Fundamentallösung $(73, 12)$ der Pell-Gleichung $X^2 - 37Y^2 = 1$ führt.

(4) $d = 67$: Hier ist $\omega = \sqrt{67}$, also

$$\vartheta = T(\omega) = \frac{1}{\omega - 8} = \frac{8 + \sqrt{67}}{3} = [\overline{5, 2, 1, 1, 7, 1, 1, 2, 5, 16}]$$

und die Fundamentaleinheit ist

$$\eta = 17901 \cdot \vartheta + 1106 = 48842 + 5967\sqrt{67}.$$

Weiter ist

$$N(\eta) = 1,$$

so daß

$$z = \eta = 48842 + 5967\sqrt{67}$$

zur Fundamentallösung $(48842, 5967)$ der Pell-Gleichung $X^2 - 67Y^2 = 1$ führt.

(5) $d = 109$: Hier ist $\omega = \frac{1+\sqrt{109}}{2}$, also

$$\vartheta = T(\omega) = \frac{1}{\omega - 5} = \frac{9 + \sqrt{109}}{14} = [\overline{1, 2, 1, 1, 2, 1, 9}]$$

und die Fundamentaleinheit ist

$$\eta = 175 \cdot \vartheta + 18 = \frac{261 + 25\sqrt{109}}{2}$$

Weiter ist

$$N(\eta) = -1,$$

so daß

$$z = \eta^6 = 158070671986249 + 15140424455100\sqrt{109}$$

zur Fundamentallösung

$$(158.070.671.986.249, 15.140.424.455.100)$$

der Pell-Gleichung $X^2 - 109Y^2 = 1$ führt.

Jakob Stix, Institut für Mathematik, Goethe–Universität Frankfurt, Robert-Mayer-Strasse 6–8, 60325 Frankfurt am Main, Germany

E-mail address: stix@math.uni-frankfurt.de