OSNABRÜCKER SCHRIFTEN ZUR MATHEMATIK

Reihe V Vorlesungsskripten

Heft 146 Sommersemester 2000

Zahlentheorie

W. Bruns

Fachbereich Mathematik/Informatik
Universität Osnabrück

OSM Osnabrücker Schriften zur Mathematik

August 2000

Herausgeber	Selbstverlag der Universität Osnabrück Fachbereich Mathematik/Informatik 49069 Osnabrück
Geschäftsführer	Prof. Dr. W. Bruns
Berater:	Prof. Dr. P. Brucker (Angew. Mathematik) Prof. Dr. E. Cohors-Fresenborg (Didaktik der Mathematik) Prof. Dr. V. Sperschneider (Informatik) Prof. Dr. R. Vogt (Reine Mathematik)
Druck	Hausdruckerei der Universität Osnabrück

Weitere Reihen der OSM:

Reihe D	Mathematisch-didaktische Manuskripte
Reihe I	Manuskripte der Informatik
Reihe M	Mathematische Manuskripte
Reihe P	Preprints
Reihe U	Materialien zum Mathematikunterricht

Zahlentheorie

Winfried Bruns

Skript zur Vorlesung SS 2000
korrigierte Version SS 2007

Inhaltsverzeichnis

ABSCHNITT 0

Einige Probleme der Zahlentheorie

Die Zahlentheorie ist in ihrem Kern die Theorie der arithmetischen Beziehungen zwischen den ganzen Zahlen (die „allgemeinen Untersuchungen über die eigentlichen Beziehungen der ganzen Zahlen“, Gauß, Disquisitiones, Vorrede). Ich möchte in diesem Abschnitt einige Probleme aufzählen, die als typisch gelten können für die Fragestellungen der Zahlentheorie.

Bekannt sind jedem die Begriffe Teiler, Primzahl, größter gemeinsamer Teiler usw. (Wir werden diese Begriffe und die zugehörigen Aussagen in Abschnitt 1 exakt behandeln).

1. Da sich jede ganze Zahl eindeutig in Primfaktoren zerlegen läßt, sind die Primzahlen zentral für die Teilbarkeit der ganzen Zahlen. Ein wichtiges Problem der Zahlentheorie ist die Frage nach der „Anzahl“ der Primzahlen in allen ihren Verfeinerungen. Sei etwa

$$\pi(x) := \text{Anzahl der Primzahlen } p \leq x.$$

(a) Wie kann man das Wachstum der Funktion π beschreiben? Wie groß etwa ist $\pi(x)$? Wie „dicht“ liegen die Primzahlen?

Auf diese Fragen gibt es eine befriedigende Antwort, den Primzahlsatz von Hadamard–de la Vallée Poussin. Wir werden Vorläufer dieses Satzes in der Vorlesung kennenlernen.

(b) Es ist klar, daß eine Primzahl (abgesehen von 2 und 5) im Dezimalsystem nur die Endziffern 1, 3, 7 und 9 haben kann. Verteilen sich die Primzahlen „gleichmäßig“ auf diese Endziffern?

Die Antwort „ja“ gibt der Primzahlsatz von Dirichlet, den wir in der Vorlesung nicht behandeln werden.

(c) Gibt es unendlich viele Primzahlzwillinge, d.h. Paare (p, p') von Primzahlen mit $p' = p + 2$? Dieses Problem ist ungelöst.

(d) Das Goldbach-Problem: Läßt sich jede gerade Zahl > 4 als Summe zweier Primzahlen schreiben? Dieses Problem ist ebenfalls ungelöst, wenn es auch gewisse Teilantworten gibt.

2. Für die Längen a, b der Katheten und die Länge c der Hypotenuse eines rechtwinkligen Dreiecks gilt bekanntlich $a^2 + b^2 = c^2$. Tripel (a, b, c) *ganzer* Zahlen,

die dieser Gleichung genügen, haben bereits die Aufmerksamkeit der Babylonier (ca. 1500 v.Chr.) erregt, und diese kannten bereits eine Lösung der folgenden Aufgabe:

(a) Bestimme alle solchen Tripel! Sie wurden schließlich wurden nach den Pythagoräern benannt, einer Gruppe griechischer Philosophen und Mathematiker (ca. 580-500 v.Chr.).

(b) Ein pythagoräisches Tripel kennt sicherlich jeder: $3^2 + 4^2 = 5^2$. Dagegen dürfte niemand ganze Zahlen $a, b, c \neq 0$ mit $a^3 + b^3 = c^3$ kennen, und solche gibt es auch nicht (Euler, Legendre ca. 1780).

Dies ist ein Spezialfall des wohl berühmtesten Problems der Zahlentheorie, des *Fermatschen Problem* (Fermat (1601–1655) ist einer der Begründer der modernen Zahlentheorie): Besitzt die Gleichung $x^n + y^n = z^n$ für irgendein $n > 2$ eine nichttriviale Lösung in ganzen Zahlen? Nach mehr als 300 Jahren hat Wiles nun bewiesen, daß die Antwort „nein“ ist.

(c) ähnliche, allerdings vollständig gelöste Probleme: Welche ganze Zahlen n besitzen eine Darstellung $n =$

(i) $x^2 + y^2$

(ii) $x^2 + y^2 + z^2$ mit ganzen Zahlen x, y, z, w?

(iii) $x^2 + y^2 + z^2 + w^2$

Diese Probleme gehören zu den *Diophantischen Gleichungen*, benannt nach dem griechischen Mathematiker Diophant (ca. 250 n. Chr.).

(d) Eine weitere interessante diophantische Gleichung: Für welche ganzen x, y ist $x^2 = y^3 - 2$?

3. Viele Fragen der Zahlentheorie betreffen „spezielle“ Zahlen, so z.B. die vollkommenen Zahlen: Eine Zahl n heißt *vollkommen*, wenn sie die Summe ihrer echten Teiler ist. (Gerade an solchen und ähnlichen Zahlen entzündeten sich die Spekulationen der Zahlenmystiker). Z.B. ist 6 eine vollkommene Zahl. Ein ungelöstes Problem: Gibt es eine ungerade vollkommene Zahl? Des weiteren ungelöst: Gibt es unendlich viele gerade vollkommene Zahlen?

4. Abschließen möchte ich die Liste von typischen Problemen mit einer Kuriosität. Für die Zahlen

$$q = 3, 5, 11, 17, 41$$

gilt: $x^2 + x + q$ ist eine Primzahl für alle x, $0 \leq x \leq q - 2$. Woran liegt das? Für welche Zahlen q gilt dies sonst noch?

Zum Abschluß dieses Abschnitts wollen wir wenigstens ein Problem lösen, nachdem so viele genannt sind, nämlich das elementarste in unserer Liste, 2(a): Sicherlich gilt: $x^2 + y^2 = z^2 \iff (dx)^2 + (dy)^2 = (dz)^2$. Es genügt also

die *teilerfremden* (x, y, z) zu bestimmen. Dann aber sind x, y, z paarweise teilerfremd. Sind zwei der drei Zahlen gerade, so auch die dritte. Andererseits können nicht alle drei Zahlen ungerade sein. Folglich sind zwei ungerade, eine gerade.

Annahme: x und y sind ungerade, $x = 2u + 1$, $y = 2v + 1$. Dann ist

$$z^2 = 4(u^2 + v^2 + u + v) + 2$$

nicht durch 4 teilbar. Widerspruch. Also ist genau eine der Zahlen x oder y gerade. Bezeichne im folgenden x die gerade Zahl. Es gilt

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

und $x = 2u$, $z - y = 2v$, $z + y = 2w$ sind gerade. Aus unserer Bedingung der Teilerfremdheit folgt: v und w sind teilerfremd und nicht beide ungerade. Nun benutzen wir die Primfaktorzerlegung von x^2 und sehen, daß v und w selbst Quadrate sein müssen, $v = p^2$, $v = q^2$. Wir erhalten

$$x = 2pq,\ y = q^2 - p^2,\ z = p^2 + q^2. \qquad (*)$$

Wenn umgekehrt $v = p^2$ und $w = q^2$ teilerfremde Quadrate verschiedener Parität mit $p < q$ sind, und x, y, z mittels der Gleichung $(*)$ definiert werden, so bilden sie ein teilerfremdes pythagoräisches Tripel.

Ich möchte hier darauf verzichten, das Ergebnis zu einem Satz zusammenzufassen. (Es ist klar, daß man es in ein effektives Computerprogramm umsetzen kann). Eine einfache Folgerung ist, daß es unendlich viele teilerfremde pythagoräische Tripel gibt.

Natürlich setzt sich die Zahlentheorie nicht aus so elementaren Überlegungen wie der vorangegangenen zusammen. Vor allem werden wir gezwungen sein, den Bereich der Objekte unserer Theorie weit über die ganzen Zahlen hinaus zu erweitern und analytische und – in dieser Vorlesung vorwiegend – algebraische Hilfsmittel anzuwenden. Die ab und zu eingestreuten Jahreszahlen sollen eine gewisse zeitliche Einordnung ermöglichen.

ABSCHNITT 1

Teilbarkeit in euklidischen Ringen

Statt nur den Ring $\mathbb{Z}$ der ganzen Zahlen zu betrachten, wollen wir in diesem Abschnitt allgemeiner die Teilbarkeitstheorie in euklidischen Ringen entwickeln.

Definition. R sei ein Integritätsbereich, $a, b \in R$. Dann heißt a ein *Teiler* von b oder b ein *Vielfaches* von a, kurz $a \mid b$, wenn es ein $c \in R$ mit $b = ac$ gibt. Man nennt a und b *assoziiert*, wenn sowohl $a \mid b$ als auch $b \mid a$.

Falls $a \mid b$, so dürfen wir mit b/a das eindeutig bestimmte $c \in R$ mit $b = ac$ bezeichnen. Assoziiertheit ist eine Äquivalenzrelation; falls $a \neq 0$, sind a und b genau dann assoziiert, wenn a/b oder b/a eine Einheit in R ist. Idealtheoretisch formuliert:

$$a \mid b \iff Rb \subset Ra,$$
$$a \text{ assoziiert zu } b \iff Rb = Ra.$$

Grundlage aller Teilbarkeitsbetrachtungen im Ring der ganzen Zahlen ist die *Division mit Rest:*

Satz 1.1. *Sei $a, b \in \mathbb{Z},\ b \neq 0$. Dann existieren eindeutig bestimmte $q, r \in \mathbb{Z}$ mit folgenden Eigenschaften:*

$$a = qb + r, \qquad 0 \leq r < |b|.$$

Der Satz ist bekannt. Integritätsbereiche, in denen eine zu Satz 1.1 analoge Aussage gilt und deren Teilbarkeitstheorie sich deshalb genauso gestaltet wie die von $\mathbb{Z}$, nennt man euklidische Ringe:

Definition. Ein *euklidischer Ring* R ist ein Integritätsbereich, zu dem eine Funktion $\varphi\colon R\backslash\{0\} \to \mathbb{N}$ existiert mit folgenden Eigenschaften:

(a) Falls b das Element $a \neq 0$ teilt, ist $\varphi(b) \leq \varphi(a)$.
(b) Zu $a, b \in R,\ b \neq 0$, existieren $q, r \in R$ mit

$$a = qb + r, \qquad r = 0 \quad \text{oder} \quad \varphi(r) < \varphi(b).$$

Satz 1.1 zeigt, daß $\mathbb{Z}$ mit $\varphi(z) = |z|$ ein euklidischer Ring ist. Weitere

Beispiele. (a) Der Polynomring $K[X]$ über einem Körper K mit $\varphi = \text{grad}$ ist ein euklidischer Ring.

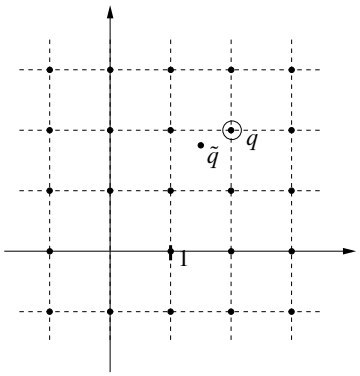

ABBILDUNG 1. Die ganzen Gaußschen Zahlen

(b) Der Ring $\mathbb{Z}[i] := \{a + bi \in \mathbb{C} : a, b \in \mathbb{Z}\}$ mit $\varphi(z) = |z|^2$ ist ebenfalls ein euklidischer Ring. Man nennt $\varphi(z)$ die *Norm* von z und schreibt dafür $N(z)$. Es gilt

$$N(wz) = N(w)N(z)$$

und daher ist Bedingung (a) offensichtlich erfüllt. Seien $a, b \in \mathbb{Z}[i]$, $b \neq 0$. Um q und r zu bestimmen, setzen wir $\tilde{q} := a/b \in \mathbb{C}$ und wählen $q \in \mathbb{Z}[i]$, so daß

$$|q - \tilde{q}| = \min\{|\tilde{q} - c| : c \in \mathbb{Z}[i]\}.$$

Dann ist $|\tilde{q} - q| \leq \sqrt{2}/2$ (siehe Abbildung 1). Für $r := a - bq$ gilt

$$|r| = |a - bq| = |b|\,|\tilde{q} - q| \leq |b| \cdot \frac{1}{2}\sqrt{2} < |b|.$$

Der Ring $\mathbb{Z}[i]$ heißt *Ring der ganzen Gaußschen Zahlen*; er wird von den ganzzahligen Punkten der komplexen Ebene gebildet.

Euklidische Ringe sind Hauptidealringe:

Satz 1.2. *Sei R ein euklidischer Ring. Dann ist jedes Ideal $\mathfrak{a}$ von R Hauptideal: es existiert ein $b \in \mathfrak{a}$ mit $\mathfrak{a} = Rb$.*

Beweis. Das Ideal $\mathfrak{a} \neq 0$ wird von jedem $b \in \mathfrak{a}$, $b \neq 0$, mit $\varphi(b) = \min\{\varphi(a) : a \in \mathfrak{a}\}$ erzeugt. Um dies einzusehen, teilen wir $a \in \mathfrak{a}$ durch b mit Rest, $a = qb + r$. Dann gehört auch $r = a - qb$ zu $\mathfrak{a}$, und nach Wahl von b kommt nur $r = 0$ in Frage. □

Man könnte nun die Teilbarkeitstheorie in euklidischen Ringen der Teilbarkeitstheorie in Hauptidealringen unterordnen; dies soll hier jedoch nicht geschehen.

Bekanntlich nennen wir *Primzahlen* diejenigen ganzen Zahlen $p > 1$, die nur durch 1 und p teilbar sind, die also multiplikativ unzerlegbar sind. Da wir uns den Namen „Primelement“ noch reservieren wollen, setzen wir fest:

Definition. R sei ein Integritätsbereich. Eine Nichteinheit $a \in R$ heißt *irreduzibel*, wenn in jeder Zerlegung $a = bc, b, c \in R$, einer der Faktoren b oder c eine Einheit ist.

Es ist leicht zu sehen, daß sich jede ganze Zahl $n > 0$ in ein Produkt von Primzahlen zerlegen läßt, und genauso leicht läßt sich dies verallgemeinern. Vorweg eine Hilfsaussage:

Satz 1.3. *Sei R ein euklidischer Ring, $a, b \in R$, $a, b \neq 0$. Wenn $b \mid a$ und $\varphi(a) = \varphi(b)$, dann sind a und b assoziiert.*

Beweis. Es ist ja $a = bp$ und $b = aq + r$ mit $r = 0$ – dann sind wir fertig – oder $\varphi(r) < \varphi(a)$. Nun ist aber $r = b - aq = b(1 - pq)$. Da b Teiler von r ist, ist $\varphi(b) \leq \varphi(r) < \varphi(a)$, wenn $r \neq 0$, und somit ist dieser Fall ausgeschlossen. □

Satz 1.4. *Sei R ein euklidischer Ring. Dann ist jede Nichteinheit $a \neq 0$ ein Produkt irreduzibler Elemente.*

Beweis. Wenn a nicht schon selbst irreduzibel ist, existieren $b, c \in R$ mit

$$a = bc,$$

wobei weder b noch c zu a assoziiert ist. Nun folgt die Behauptung mit Induktion über $\varphi(a)$, denn nach Satz 1.3 gilt $\varphi(b), \varphi(c) < \varphi(a)$. □

Ebenso bekannt, aber schwerer zu beweisen ist, daß sich jede natürliche Zahl im wesentlichen eindeutig in Primfaktoren zerlegen läßt. Es ist nützlich, zunächst den wichtigen Begriff des größten gemeinsamen Teilers einzuführen:

Definition. Seien $a, b \neq 0$ Elemente eines Integritätsbereiches R. Dann heißt $c \in R$ ein *größter gemeinsamer Teiler* von a und b, wenn c sowohl a, als auch b teilt und jeder gemeinsame Teiler von a und b wiederum c teilt.

Größte gemeinsame Teiler, sofern sie existieren, sind bis auf Assoziiertheit eindeutig bestimmt. Wir erlauben uns deshalb, von *dem* größten gemeinsamen Teiler zu sprechen, und schreiben

$$c = \mathrm{ggT}(a, b).$$

(In $\mathbb{Z}$ wählen wir stets $c > 0$.) Der Name „*größter* gemeinsame Teiler“ erklärt sich aus der Eigenschaft

$$\varphi(c) = \max\{\varphi(d) : d \mid a, b\},$$

die man üblicherweise verwendet, um den ggT in $\mathbb{Z}$ zu definieren. (Dies impliziert sofort die Existenz, aber die in der obigen Definition geforderte Eigenschaft ist dann ein zu beweisender Satz; vergleiche Aufgabe 1.14.) Die Aussage $\varphi(c) = \max\{\varphi(d) : d \mid a, b\}$ folgt aus Eigenschaft (a) in der Definition des euklidischen Ringes.

Dual zum Begriff des größten gemeinsamen Teiler ist der des kleinsten gemeinsamen Vielfachen $\mathrm{kgV}(a,b)$. Wie der ggT kann es natürlich auch für mehr als zwei Zahlen definiert werden.

Satz 1.5 (Bezout)**.** *R sei ein euklidischer Ring. Dann existiert zu $a,b \neq 0$ ein größter gemeinsamer Teiler c und es gibt $x, y \in R$ mit*

$$c = xa + yb.$$

Beweis. Sei $\mathfrak{a} := Ra + Rb$ das von a,b erzeugte Ideal. Nach Satz 1.1 ist $\mathfrak{a} = Rc$ mit einem geeigneten $c \in R$, und als Element von $Ra + Rb$ besitzt c eine Darstellung $c = xa + yb$.

Da $a,b \in Rc$, ist c ein gemeinsamer Teiler von a und b, und wenn $d \mid a,b$, so ist $Rd \supset Ra, Rb$. Es folgt $Rd \supset Ra + Rb = Rc$ und damit auch $d \mid c$. Also ist c der größte gemeinsame Teiler. □

Satz 1.5 kann auch mit dem *Euklidischen Algorithmus* bewiesen werden, einem äußerst effektiven Verfahren zur Berechnung des ggT, das überdies auch x und y mitliefert.

Sei $r_0 := a,\ r_1 := b$. Gemäß Definition ist

$$r_0 = q_1 r_1 + r_2, \qquad r_2 = 0 \quad \text{oder} \quad \varphi(r_2) < \varphi(r_1).$$

Falls $r_2 = 0$, ist offensichtlich $r_1 = \mathrm{ggT}(r_1, r_2)$. Andernfalls ist immerhin noch $\mathrm{ggT}(r_0, r_1) = \mathrm{ggT}(r_1, r_2)$, denn jeder Teiler von r_0 und r_1 ist ein Teiler von r_1 und r_2 und umgekehrt. Also fährt man fort

$$r_1 = q_2 r_2 + r_3, \ldots, r_{n-1} = q_n r_n + r_{n+1},\ r_n = q_{n+1} r_{n+1}.$$

Dieses Verfahren besitzt wegen $\varphi(r_k) < \varphi(r_{k-1})$ für $k \geq 2$ stets nach endlich vielen Schritten ab, und liefert, wie oben bereits begründet, in r_{n+1} den ggT von a und b. Um eine Darstellung $r_{n+1} = xa + yb$ zu erhalten, beachtet man

$$\begin{aligned}
r_0 = a &= x_0 a + y_0 b & x_0 &= 1, & y_0 &= 0 \\
r_1 = b &= x_1 a + y_1 b & x_1 &= 0, & y_1 &= 1 \\
r_2 = r_0 - q_1 r_1 &= x_2 a + y_2 b & x_2 &= x_0 - q_1 x_1, & y_2 &= y_0 - q_1 y_1
\end{aligned}$$

usw.

Beispiel. Zu bestimmen ist $\mathrm{ggT}(705, 423)$ in $\mathbb{Z}$.

$$\begin{aligned}
& & x_0 &= 1 & y_0 &= 0 \\
r_1 &= 705 = 1 \cdot 423 + 282 & x_1 &= 0 & y_1 &= 1 \\
r_2 &= 423 = 1 \cdot 282 + 141 & x_2 &= 1 & y_2 &= -1 \\
r_3 &= 282 = 2 \cdot 141 & x_3 &= -1 & y_3 &= 2
\end{aligned}$$

also $\mathrm{ggT}(705, 423) = 141$ und $141 = -705 + 2 \cdot 423$.

Definition. Elemente a, b eines euklidischen Ringes R heißen *teilerfremd*, wenn $\mathrm{ggT}(a,b) = 1$, mit anderen Worten, wenn $R = Ra + Rb$.

Den Schlüssel zur Eindeutigkeit der Primfaktorzerlegung liefert

Satz 1.6 (Lemma von Euklid)**.** *R sei ein euklidischer Ring. Wenn a und b teilerfremd sind und a das Produkt bc teilt, so teilt a das Element c.*

Beweis. Es existieren $x, y \in R$ mit $1 = xa + by$ und ein $d \in R$ mit $bc = ad$. Folglich

$$c = c(xa + by) = a(xc + dy). \qquad \square$$

Sei nun a irreduzibel und b ein Element, das von a nicht geteilt wird. Dann sind a und b teilerfremd und wir erhalten als Folgerung aus Satz 1.6: Wenn a irreduzibel ist und b nicht teilt, wohl aber bc, so teilt a das Element c.

Die soeben bemerkte Eigenschaft irreduzibler Elemente euklidischer Ringe benutzen wir, um den Begriff „Primelement“ zu definieren:

Definition. R sei beliebiger Integritätsbereich. Eine Nichteinheit $a \in R, a \neq 0$, heißt *Primelement* wenn gilt: Teilt a das Produkt bc, so teilt es einen der Faktoren b oder c.

Trivialerweise sind Primelemente stets irreduzibel, und wie wir gesehen haben, gilt in euklidischen Ringen auch die Umkehrung. Insbesondere sind in $\mathbb{Z}$ genau die Primzahlen p und ihre additiven Inversen $-p$ Primelemente. Bekanntlich heißt ein Ideal $\mathfrak{a}$ eines Ringes R ein *Primideal*, wenn $R/\mathfrak{a}$ Integritätsbereich ist. Man sieht sofort, daß ein Element genau dann ein Primelement ist, wenn es ein Primideal erzeugt.

Wir können nun das Hauptergebnis dieses Abschnitts formulieren.

Satz 1.7.

(a) *Jede Nichteinheit $a \neq 0$ eines euklidischen Ringes R ist ein Produkt $a = p_1 \dots p_n$ von Primelementen.*

(b) *Diese Darstellung von a als Produkt von Primelementen ist im folgenden Sinne eindeutig: Wenn $a = q_1 \dots q_m$ mit Primelementen $q_1, \dots, q_m$, so ist $m = n$, und es gibt eine Permutation σ von $\{1, \dots, n\}$ derart, daß p_i assoziiert ist zu $q_{\sigma(i)}$, $i = 1, \dots, n$.*

Beweis. (a) folgt aus Satz 1.4 und 1.6 nebst der Definition von Primelementen.

(b) Wenn a selbst ein Primelement ist, ist $m = n = 1$ und nichts zu beweisen. Andernfalls ist $n \geq 2$ und $m \geq 2$. Das Primelement p_1 teilt das Produkt $q_1 \dots q_m$, muß also wegen Satz 1.6 einen der Faktoren teilen. Nach Umordnen der q_i dürfen wir annehmen, daß $q_1 = p_1 e_1$. Dabei ist e_1 notwendig eine Einheit, q_1 also zu p_1 assoziiert. Nun gilt

$$a' := p_2 \cdots p_n = (e_1 q_2) q_3 \cdots q_m$$

und die Behauptung folgt durch Induktion. □

Definition. Ringe R, in denen sich jede Nichteinheit im Sinne des vorangegangenen Satzes eindeutig in Primfaktoren zerlegen läßt, heißen *faktoriell*.

Euklidische Ringe sind also faktoriell. Eine Normierung der Primfaktorzerlegung kann man folgendermaßen erhalten: Sei P die Menge der Primelemente von R und

$$P = \bigcup_{i \in I} P_i$$

die Zerlegung von P in die Klassen assoziierter Primelemente. Wir wählen aus jeder Klasse P_i ein Primelement p_i aus. Dann kann man 1.7 äquivalent so formulieren: Zu jedem $a \in R, a \neq 0$, existieren eine eindeutig bestimmte Einheit e und eindeutig bestimmte Zahlen $\alpha_i \in \mathbb{N}$, die mit Ausnahme endlich vieler $i \in I$ alle 0 sind, so daß gilt:

$$a = e \prod_{i \in I} p_i^{\alpha_i}.$$

In $\mathbb{Z}$ sind jeweils die Primelemente $p > 0$ und $-p$ assoziiert und wir erhalten:

Satz 1.8 (Hauptsatz der elementaren Zahlentheorie)**.** *Sei $a \in \mathbb{Z}$, $a \neq 0, 1, -1$. Dann existieren eindeutig bestimmte Primzahlen $p_1, \ldots, p_n$, $p_1 < \cdots < p_n$, und natürliche Zahlen $\alpha_i > 0$ mit*

$$a = \operatorname{sgn}(a) \prod_{i=1}^{n} p_i^{\alpha_i}.$$

Wir wollen diesen Abschnitt mit einem Satz über lineare diophantische Gleichungen abschließen:

Satz 1.9. *R sei ein euklidischer Ring. Seien $a, b, c \in R$, $a, b \neq 0$. Genau dann existieren $x, y \in R$ mit*

$$c = ax + by, \qquad (*)$$

wenn c Vielfaches des größten gemeinsamen Teilers d von a und b ist. Ist (x_0, y_0) eine Lösung der Gleichung $(*)$, so ist

$$\left\{ \left(x_0 + r\frac{b}{d},\ y_0 - r\frac{a}{d} \right) : r \in R \right\}$$

die Gesamtheit der Lösungen.

Beweis. Der erste Teil folgt unmittelbar aus $aR + bR = dR$, also daraus, daß d erzeugendes Element des von a und b erzeugten Ideals ist. Sicherlich ist jedes Paar $(x_0 + rb/d,\ y_0 - ra/d)$ Lösung von $(*)$. Ist umgekehrt $ax_0 + by_0 = ax + by$, also $(x_0 - x)a + (y_0 - y)b = 0$, so folgt

$$(x_0 - x)\frac{a}{d} + (y_0 - y)\frac{b}{d} = 0.$$

Da a/d und b/d teilerfremd sind (!), müssen nach Satz 1.6 $r, s \in R$ existieren mit $x_0 - x = -rb/d$ und $y_0 - y = sa/d$. Einsetzen liefert $r = s$. □

Satz 1.9 läßt sich auf diophantische Gleichungen in mehr als zwei Unbekannten erweitern, dabei wird aber die Formulierung des zweiten Teiles um einiges komplizierter.

Übungen.

1.10. Bestimme den größten gemeinsamen Teiler d und eine Darstellung von d als Linearkombination für (a) $146\,039, 184\,519 \in \mathbb{Z}$, (b) $23 + 92i, -10 + 11i \in \mathbb{Z}[i]$.

1.11. Definiere ggT und kgV für Elemente $a_1, \ldots, a_n$ eines euklidischen Ringes. Zeige ihre Existenz und charakterisiere sie idealtheoretisch.

1.12. Zeige: $\mathbb{Z}[i\sqrt{2}]$ ist ein euklidischer Ring.

1.13. Die Elemente u und v eines euklidischen Ringes seien teilerfremd. Zeige: Wenn $uv = a^2$, so gibt es Einheiten e und f, für die eu und fv beide Quadrate sind (d.h. es existieren b, c mit $eu = b^2$ und $fv = c^2$).

1.14. Sei R ein euklidischer Ring, $a, b \in R$, $a \neq 0$. Unter den gemeinsamen Teilern von a und b habe c den größten Wert unter φ. Zeige $c = \mathrm{ggT}(a, b)$.

1.15. Bestimme die normierte Primfaktorzerlegung von $81\,057\,226\,635\,000$.

1.16. Wie ermittelt man die Primfaktorzerlegung von $\mathrm{ggT}(m, n)$ und $\mathrm{kgV}(m, n)$ aus den Primfaktorzerlegungen von m und n?

1.17. Sei R ein kommutativer Ring. Für eine Teilmenge $A \subset R$ nennen wir

$$\mathfrak{a} = \{r_1 a_1 + \cdots + r_n a_n \; n \in \mathbb{N},\; r_i \in R,\; a_i \in A\}$$

das von A erzeugte Ideal. Falls $A = \{a_1, \ldots, a_m\}$ endlich ist, schreiben wir auch $\mathfrak{a} = Ra_1 + \cdots + Ra_m$.

(a) Zeige, daß diese Bezeichnung gerechtfertigt ist: $\mathfrak{a}$ ist ein Ideal und zwar das kleinste, das A enthält.

(b) Seien $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ Ideale. Zeige:

$$\mathfrak{a}_1 + \cdots + \mathfrak{a}_n := \{a_1 + \cdots + a_n \; a_i \in \mathfrak{a}_i\}$$

ist das von $\mathfrak{a}_1 \cup \cdots \cup \mathfrak{a}_n$ erzeugte Ideal.

(c) Das Produkt $\mathfrak{a}_1 \ldots \mathfrak{a}_n$ sei das von $\{a_1 \ldots a_n \; a_i \in \mathfrak{a}_i\}$ erzeugte Ideal. (Natürlich ist $\mathfrak{a}^2 = \mathfrak{a}\mathfrak{a}$ usw.) Zeige: Die Addition und Multiplikation von Idealen ist assoziativ und kommutativ; es gelten die Distributivgesetze.

1.18. Für $n \in \mathbb{N}$, $n \geq 1$, bezeichne $T(n)$ die Anzahl der positiven Teiler von n.

(a) Wenn $n = p_1^{\alpha_1} \dots p_r^{\alpha_r}$ die normierte Primfaktorzerlegung von n ist, so ist $T(n) = \prod_{i=1}^{r}(\alpha_i + 1)$.

(b) Wenn n und m teilerfremd sind, so ist $T(mn) = T(m)T(n)$.

(c) Bestimme die Zahl der Teiler von 100!.

1.19. Zeige: (a) Für $a, m, n \in \mathbb{N}$, $m, n \geq 1$, $a \geq 2$ gilt: $\mathrm{ggT}(a^n - 1, a^m - 1) = a^d - 1$, wobei $d = \mathrm{ggT}(n, m)$. (Man kann mit der Grundidee des euklidischen Algorithmus schließen.)

(b) Wenn $a^n - 1$ prim, so ist n prim.

(c) Wenn $a^n - 1$ prim, so ist $a = 2$.

1.20. Zeige: Wenn $a^n + 1$ prim ist für $a, n \in \mathbb{N}$, $a \geq 2$, $n \geq 1$, so ist a gerade und $n = 2^m$ mit einem $m \in \mathbb{N}$.

1.21. Bauer Grapp hat bei Bauer Piestmann 100 DM Schulden. Leider verfügt keiner der beiden über irgendwelches Bargeld, aber Bauer Grapp hat viele Kühe im Wert von je 1850 DM und Bauer Piestmann hat Schweine im Wert von je 1100 DM. Daher bezahlt Grapp mit Kühen, und Piestmann gibt ihm mit Schweinen das Wechselgeld heraus. Kann Grapp auf diese Weise überhaupt seine Schulden bezahlen, und wenn ja, wieviele Kühe wechseln dabei mindestens den Besitzer?

1.22. Einer der großen Mathematiker des Altertums war Diophant. Man weiß über seine Lebensdaten nur, daß er im dritten Jahrhundert n. Chr. gelebt hat. Eine Sammlung mathematischer Aufgaben aus dem fünften oder sechsten Jahrhundert enthält folgenden Text, der angeblich auf Diophants Grabstein eingraviert war:

„In diesem Grabe ruhet Diophant. Und diese Inschrift verkündet uns die Zahl der Jahre seines Lebens. Die Götter ließen ihn den sechsten Teil desselben die süße Knabenzeit genießen, und als noch ein Zwölftel vergangen war, entsproß seinen Wangen der erste Flaum. Wiederum ein Siebentel war vergangen, als er trat in den holden Stand der Ehe, und am fünften Jahrestage der Vermählung schenkten die Götter ihm einen Sohn. Doch als dieser die Hälfte der Lebensspanne des Vaters erreicht, sank das spät gezeugte und unglückliche Kind ins kühle Grab. Nach vier Jahren, gramgebeugt und getröstet nur von der hehren Kunde der Zahlen, folgte der Vater dem Sohne."

Wieviele Jahre lebte Diophant?

1.23. Löse das folgende Zahlenrätsel:

$$EULER = SB \cdot RL^E$$
$$GAUSS = L \cdot A \cdot LUL \cdot E^E$$
$$ABEL = A \cdot RR \cdot RL \cdot L$$

Hierbei steht jeder Buchstabe für eine Ziffer im Dezimalsystem, und in jeder Gleichung ist die rechte Seite die Primfaktorzerlegung der links stehenden Zahl.

1.24. Der Verein Vechtaer Verfechterinnen der Zahlentheorie von 1888 e.V. kauft für die bevorstehende Weihnachtsfeier 123 Stück Geflügel im Wert von 456 Gulden. Ein Hähnchen kostet $1\frac{2}{3}$, eine Wachtel $4\frac{5}{6}$ und eine Gans $7\frac{8}{9}$ Gulden. Wieviel Stück von jeder Sorte kauft der VVVdZ von 1888 e.V.?

ABSCHNITT 2

Primzahlverteilung

Bereits im Altertum war bekannt:

Satz 2.1. *Im Ring $\mathbb{Z}$ gibt es unendlich viele Primzahlen.*

Beweis. Seien $p_1 < p_2 < p_3 \ldots$ die Primzahlen in $\mathbb{Z}$. Wir betrachten

$$N := p_1 \cdots p_n + 1.$$

N ist keine Einheit, also durch eine Primzahl p teilbar. Da $p \neq p_i$ für $i = 1, \ldots, n$ ist, gibt es zu jeder Primzahl p_n eine noch größere Primzahl. □

Dieser elegante Beweis findet sich schon bei Euklid (um 300 v. Chr.). Natürlich gibt er keine Aussage über die Verteilung der Primzahlen unter den natürlichen Zahlen. Es wäre eine wesentliche Verschärfung von Satz 2.1, wenn wir wüßten, ob die Reihe $\sum 1/p$ der Reziproken der Primzahlen konvergiert oder nicht. Bekanntlich divergiert die Reihe $\sum_{n=1}^{\infty} 1/n$, während die Reihe $\sum_{n=1}^{\infty} 1/n^2$ konvergiert. In diesem Sinne gibt es also „wesentlich weniger" Quadratzahlen als natürliche Zahlen überhaupt, die Menge der Quadratzahlen ist „dünn". Das Problem der Konvergenz von $\sum 1/p$ wurde von Euler (1707–1783) gelöst. Euler verwendete dabei eine Methode, die den Keim für die sogenannte *analytische Zahlentheorie* legte.

Wir werden im folgenden Aussagen über Reihen $\sum_{n \in M} f(n)$ machen müssen, wobei M eine abzählbare Menge, und f eine Funktion auf M mit Werten in den nichtnegativen reellen Zahlen ist. Eine solche Reihe heißt *konvergent*, wenn die Teilsummen über endliche Teilmengen von M beschränkt sind. Genau dann konvergiert ja die „gewöhnliche" unendliche Reihe $\sum_{k=1}^{\infty} f(n_k)$ für irgendeine Abzählung $M = \{n_1, n_2, \ldots\}$ und jede solche Abzählung führt wegen des Umordnungssatzes für absolut konvergent Reihen auf den gleichen Grenzwert. (P bezeichne im folgenden die Menge der Primzahlen).

Satz 2.2 (Euler)**.** *Die unendliche Reihe $\sum_{p \in P} 1/p$ und das unendliche Produkt $\prod_{p \in P} (1 - 1/p)^{-1}$ divergieren.*

Euler hatte die geniale Idee, das Produkt zu betrachten und aus der Divergenz des Produkts die der Reihe herzuleiten. Sei p_r die r-te Primzahl in aufsteigender Folge. Da die Folge

$$\left(\prod_{i=1}^{r} \left(1 - \frac{1}{p_i} \right)^{-1} \right)$$

unbeschränkt ist, ist auch

$$\log\left(\prod_{i=1}^{r}\left(1-\frac{1}{p_i}\right)^{-1}\right)=\sum_{i=1}^{r}\log\left(1-\frac{1}{p_i}\right)^{-1}=-\sum_{i=1}^{r}\log\left(1-\frac{1}{p_i}\right)$$
$$=\sum_{i=1}^{r}\sum_{i=1}^{\infty}(kp_i^k)^{-1}=\frac{1}{p_1}+\cdots+\frac{1}{p_r}+\sum_{i=1}^{r}\sum_{k=2}^{\infty}(kp_i^k)^{-1}$$

unbeschränkt. Dabei haben wir die Taylorentwicklung

$$-\log(1-x)=\sum_{i=1}^{\infty}x^k/k,\qquad |x|<1,$$

ausgenutzt. Nun gilt

$$\sum_{k=2}^{\infty}\frac{1}{k}\cdot\left(\frac{1}{p_i}\right)^k<\sum_{k=2}^{\infty}\left(\frac{1}{p_i}\right)^k=\frac{1}{p_i^2}\sum_{k=0}^{\infty}\left(\frac{1}{p_i^k}\right)=\frac{1}{p_i^2}\cdot\frac{1}{1-\frac{1}{p_i}}\le\frac{2}{p_i^2}.$$

Also ist

$$\sum_{i=1}^{r}\sum_{k=2}^{\infty}(kp_i^k)^{-1}<\sum_{i=1}^{r}\sum_{k=2}^{\infty}\left(\frac{1}{p_i}\right)^k<2\sum_{n=1}^{\infty}\frac{1}{n^2}$$

beschränkt, und

$$\frac{1}{p_1}+\cdots+\frac{1}{p_r}$$

notwendig unbeschränkt.

Bevor wir die Divergenz des Produkts beweisen, gehen wir noch einen kleinen Umweg. Bekanntlich ist für jedes $s>1$ die unendliche Reihe $\sum_{n=1}^{\infty}1/n^s$ konvergent.

Definition. Die für alle $s\in\mathbb{R}$, $s>1$, definierte Funktion

$$\zeta(s):=\sum_{n=1}^{\infty}\frac{1}{n^s}$$

heißt *Riemannsche Zeta-Funktion.*

Der folgende Satz liefert die *Eulersche Produkt-Darstellung* der Zeta-Funktion:

Satz 2.3. *Für alle $s>1$ ist*

$$\zeta(s)=\sum_{n=1}^{\infty}\frac{1}{n^s}=\prod_{i=1}^{\infty}\left(1-\frac{1}{p_i^s}\right)^{-1}.$$

Beweis. Mittels der geometrischen Reihe erhält man

$$\frac{1}{1-\frac{1}{p_i}}=\sum_{k=0}^{\infty}\left(\frac{1}{p_i}\right)^k.$$

Also ist

$$\prod_{i=1}^{r}\left(1-\frac{1}{p_i^s}\right)^{-1}=\prod_{i=1}^{r}\sum_{k=0}^{\infty}\left(\frac{1}{p_i^s}\right)^k=\sum_{(k_1,\ldots,k_r)\in\mathbb{N}^r}\frac{1}{(p_1^{k_1}\ldots p_r^{k_r})^s}. \qquad (*)$$

Dabei haben wir für die rechte Gleichung die unendlichen Reihen einfach ausmultipliziert. Dies ist wegen der absoluten Konvergenz der geometrischen Reihen zulässig. Als Nenner in der ganz rechts stehenden Reihe treten genau diejenigen $n \in \mathbb{N}$ auf, die nur Primfaktoren $p \in \{p_1, \ldots, p_r\}$ besitzen. Sei $N(p_1, \ldots, p_r)$ die Menge dieser natürlichen Zahlen. Es folgt

$$\sum_{n=1}^{r}\frac{1}{n^s}\le\sum_{n\in N(p_1,\ldots,p_r)}\frac{1}{n^s}=\prod_{i=1}^{r}\left(1-\frac{1}{p_i^s}\right)^{-1}<\sum_{n=1}^{\infty}\frac{1}{n^s}$$

und daraus die behauptete Aussage über die Grenzwerte. □

Entscheidend benutzt haben wir im Beweis von 2.3 den Hauptsatz der elementaren Zahlentheorie. Wir fahren nun in Beweis von 2.2 fort. Die Gleichung $(*)$ gilt auch für $s = 1$. Wegen

$$\sum_{n\in N(p_1,\ldots,p_r)}\frac{1}{n}\ge\sum_{n=1}^{r}\frac{1}{n}$$

wächst die rechte Seite von $(*)$ unbeschränkt, womit 2.2 bewiesen ist.

Die Riemannsche Zetafunktion läßt sich zu einer meromorphen Funktion auf $\mathbb{C}$ mit einem einzigen Pol der Ordnung 1 in $s = 1$ fortsetzen. Die weitergehenden Untersuchungen über die Primzahlverteilung beruhen auf dem Studium dieser Funktion. Die Verteilung der Primzahlen wird beschrieben durch die Funktion

$$\pi : \mathbb{R} \to \mathbb{N}, \quad \pi(x) := \text{ Anzahl der Primzahlen } p \le x.$$

Es gilt der berühmte *Primzahlsatz:*

Satz 2.4 (Hadamard, de la Vallée Poussin, 1896)**.**

$$\lim_{x\to\infty}\frac{\pi(x)}{x/\log x}=1.$$

Der Primzahlsatz besagt, daß es etwa $x/\log x$ Primzahlen $\le x$ gibt, und daß der Quotient dieser Zahlen für $x \to \infty$ gegen 1 geht. Der Primzahlsatz wurde von Gauß (im Alter von ca. 16 Jahren) auf Grund numerischer Untersuchungen vermutet. Beispiele:

x	$\pi(x)$	$[x/log\ x]$
10	4	4
10^2	25	21
10^3	168	144
10^4	1229	1085
10^5	9592	8685
10^6	78498	72382
10^7	664579	620420
10^8	5761455	5428681
10^9	50847534	48254942

Der Primzahlsatz ist relativ schwierig zu beweisen. Deshalb wollen wir uns mit einer schwächeren Aussage, dem *Satz von Tschebyschew* (1852) begnügen.

Satz 2.5. *Es gibt Konstanten $C_1, C_2 > 0$ mit*

$$C_1 \frac{x}{\log x} < \pi(x) < C_2 \frac{x}{\log x}.$$

Man kann ohne allzu großen Mehraufwand zeigen, daß Satz 2.5 etwa mit $C_1 = 2/3$, $C_2 = 8/5$ gültig ist. Zum Beweis der rechten Ungleichung studieren wir die Funktion

$$\vartheta(x) := \sum_{p \le x} \log p \qquad (p \text{ bezeichnet hier Primzahlen}).$$

Satz 2.6. *Für alle $x \in \mathbb{R}$ ist $\vartheta(x) < (4 \log 2)x$.*

Beweis. Wir betrachten den Binomialkoeffizienten

$$\binom{2n}{n} = \frac{(n+1)(n+2)\dots(2n-1)2n}{1 \cdot 2 \cdots (n-1) \cdot n};$$

er ist durch alle Primzahlen p, $n < p < 2n$, teilbar. Ferner gilt

$$(1+1)^{2n} = \sum_{j=0}^{2n} \binom{2n}{j}, \quad \text{also} \quad 2^{2n} > \binom{2n}{n}.$$

Folglich

$$2^{2n} > \binom{2n}{n} > \prod_{n<p<2n} p.$$

Logarithmieren ergibt

$$2n \log 2 > \sum_{n<p<2n} \log p = \vartheta(2n) - \vartheta(n).$$

Durch Aufsummieren für $n = 1, 2, 4, 8, \ldots, 2^{m-1}$ erhalten wir

$$\vartheta(2^m) < (\log 2)(2^{m+1} - 2).$$

Für $2^{m-1} < x \leq 2^m$ ergibt sich

$$\vartheta(x) \leq \vartheta(2^m) < (\log 2)2^{m+1} = (4 \log 2)2^{m-1} < (4 \log 2)x. \qquad \square$$

Jetzt ergibt sich die rechte Ungleichung in Satz 2.5 so:

$$\begin{aligned} \pi(x) = \pi(\sqrt{x}) + \big(\pi(x) - \pi(\sqrt{x})\big) &\leq \sqrt{x} + \sum_{\sqrt{x} < p \leq x} 1 \\ &\leq \sqrt{x} + \sum_{\sqrt{x} < p \leq x} \frac{2 \log p}{\log x} \qquad \leq \sqrt{x} + \frac{2}{\log x}\vartheta(x) \\ &\leq \sqrt{x} + (8 \log 2)\frac{x}{\log x}. \end{aligned}$$

Nun ist aber $\sqrt{x} < 2x/\log x$ für $x \geq 2$. (Wir erhalten so $C_2 \leq 2 + 8 \log 2$, zu einem besseren Wert siehe oben).

Zum Beweis der linken Ungleichung betrachten wir noch einmal $\binom{2n}{n}$. Wir benötigen eine Formel dafür, mit welcher Potenz eine Primzahl p diesen Binomialkoeffizienten teilt.

Definition. Für $n \in \mathbb{Z}$, $n \neq 0$, und eine Primzahl p sei

$$v_p(n) := \max\{k : p^k \text{ teilt } n\}.$$

Satz 2.7.

$$v_p(n!) = \left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \cdots + \left[\frac{n}{p^n}\right].$$

(Dabei ist $[x] = \max\{z \in \mathbb{Z} : z \leq x\}$*).*

Beweis. Genau $[n/p]$ der Faktoren sind durch p teilbar, genau $[n/p^2]$ durch p^2 usw. und jeder Faktor wird mit der richtigen Vielfachheit gezählt. $\square$

Es folgt

$$v_p\binom{2n}{n} = v_p\left(\frac{(2n)!}{(n!)^2}\right) = \sum_{j=1}^{t_p}\left(\left[\frac{2n}{p^j}\right] - 2\left[\frac{n}{p^j}\right]\right)$$

wobei $t_p = \max\{k : p^k \leq 2n\}$, also

$$t_p = [\log_p 2n] = \left[\frac{\log 2n}{\log p}\right].$$

Da $[2x] - 2[x] = 0$ oder $= 1$, folgt

$$v_p\binom{2n}{n} \leq \left[\frac{\log 2n}{\log p}\right]$$

und damit

$$2^n \le \binom{2n}{n} \le \prod_{p<2n} p^{[\log 2n/\log p]}.$$

Logarithmieren ergibt

$$n \log 2 \le \sum_{p<2n} \left[\frac{\log 2n}{\log p}\right] \log p.$$

Der Trick besteht wieder darin, die Summe aufzuspalten: Für $p > \sqrt{2n}$ ist

$$\log p > \frac{1}{2}\log 2n, \quad \text{also } 2\log p > \log 2n \quad \text{und} \quad \left[\frac{\log 2n}{\log p}\right] = 1.$$

Es ergibt sich

$$\begin{aligned} n \log 2 &\le \sum_{p\le\sqrt{2n}} \left[\frac{\log 2n}{\log p}\right] \log p + \sum_{\sqrt{2n}<p<2n} \log p \\ &\le \sum_{p\le\sqrt{2n}} \log 2n + \vartheta(2n) \\ &\le \sqrt{2n}\log 2n + \vartheta(2n). \end{aligned}$$

Mithin

$$\vartheta(2n) \ge n\left(\log 2 - \frac{\sqrt{2n}\log 2n}{n}\right).$$

Daher existiert eine Konstante $T > 0$ derart, daß $\vartheta(2n) > Tn$ für alle hinreichend großen n. Für $2n \le x < 2n+2$ folgt

$$\vartheta(x) \ge \vartheta(2n) > Tn > T\frac{x-2}{2} > Cx$$

mit einem geeigneten $C > 0$ für alle hinreichend großen x. Dann existiert ein $C_1 > 0$ mit

$$\vartheta(x) > C_1 x \qquad \text{für alle } x \ge 2.$$

Da $\vartheta(x) = \sum_{p\le x} \log p \le \pi(x)\log x$, folgt auch die linke Ungleichung in Satz 2.5.

Aus der in Satz 2.5 enthaltenen Abschätzung von $\pi(x)$ nach oben ergibt sich, daß der relative Anteil der Primzahlen an den Zahlen $\le x$ mit wachsendem x gegen 0 geht:

Satz 2.8.

$$\lim_{x\to\infty} \frac{\pi(x)}{x} = 0.$$

Beweis. Klar wegen $\lim_{x\to\infty} 1/\log x = 0$. □

Es ist leicht einzusehen, daß es beliebig große Intervalle in $\mathbb{Z}$ gibt, die keine Primzahl enthalten; man betrachte etwa für $n \geq 3$

$$n! + 2, n! + 3, \ldots, n! + n.$$

Andererseits erlaubt es uns der Beweis von 2.5, eine Aussage darüber zu machen, daß wir in gewissen Intervallen stets eine Primzahl auffinden:

Satz 2.9. *Es gibt eine Konstante $B > 0$ derart, daß sich für jedes $n \in \mathbb{N}$ unter den ganzen Zahlen im Intervall $[n + 1, Bn]$ stets eine Primzahl befindet.*

Beweis. Wir haben im Laufe des Beweises von Satz 2.5 gesehen, daß es Konstanten $B_1, B_2 > 0$ mit

$$B_1 x < \vartheta(x) < B_2 x$$

gibt. Also ist

$$\vartheta\left(\frac{B_2}{B_1}x\right) > B_1 \cdot \frac{B_2}{B_1}x = B_2 x > \vartheta(x),$$

und im Intervall $(x, B_2/B_1 x]$ muß eine Primzahl liegen. □

Man kann durch Verfeinerung der zum Beweis von 2.5 benutzten Methoden zeigen, daß $B = 2$ gewählt werden kann und erhält das erstmals von Tschebyscheff bewiesene „Bertrandsche Postulat“:

Satz 2.10. *Für jede natürliche Zahl $n \geq 2$ liegt im Intervall $[n+1, 2n]$ mindestens eine Primzahl.*

Es ist charakteristisch für die Beweise von 2.10, daß die Aussage zunächst für $n \geq n_0$ beweist, und für kleinere n mittels einer Tabelle prüft. Für $n_0 = 64$ siehe [**Chan**].

Die größte am 21.07.2007, 13.04 Uhr bekannte Primzahl ist

$$2^{32582657} - 1.$$

Sie ist eine der *Mersenneschen Zahlen* $M_p = 2^p - 1$, p Primzahl, benannt nach M. Mersenne (1588–1648). Nur sehr wenige Mersenneschen Zahlen sind Primzahlen (die obige Zahl ist die 44.), aber man kann zumindest vermuten, daß unendlich viele von ihnen prim sind. Im Internet gibt es zahlreiche Informationsquellen für die Suche nach großen Primzahlen; siehe zum Beispiel [**Mers**].

Eine andere berühmte Zahlenfamilie bilden die *Fermat-Zahlen*

$$F_m = 2^{2^m} + 1.$$

Die Zahlen F_0, F_1, F_2, F_3, F_4 sind prim. Man weiß nicht, ob es weitere prime Fermat-Zahlen gibt, aber von $F_5, \ldots, F_{32}$ und einigen weiteren F_m ist (am 22.10.2007) bekannt, daß sie nicht prim sind. Die primen Fermat-Zahlen spielen eine wichtige Rolle bei der Frage, welche regulären n-Ecke mit Zirkel und Lineal konstruierbar sind.

Übungen.

2.11. Zeige: Es gibt unendlich viele Primzahlen der Form $4k + 3$, $k \in \mathbb{N}$.

2.12. Mittels der Divergenz von $\sum 1/p$ zeige:

(a) Für jedes $\varepsilon > 0$ ist $\pi(x(1+\varepsilon)) - \pi(x)$, $x \in \mathbb{R}$, unbeschränkt.

(b) Zu jedem $\varepsilon > 0$ gibt es unendlich viele r mit $p_r < p_{r+1} < (1+\varepsilon)p_r$. (Dabei sei $p_1, p_2, \ldots$ die Folge der Primzahlen in aufsteigender Reihenfolge.

2.13. Für eine Primzahl p und $w \in \mathbb{Q}$, $w = u/v$ mit $u, v \in \mathbb{Z}$, sei

$$v_p(w) = \begin{cases} \infty & \text{wenn } w = 0, \\ v_p(u) - v_p(v) & \text{sonst.} \end{cases}$$

(a) Zeige, daß diese Definition sinnvoll ist: $v_p(w)$ hängt nur von w, aber nicht von der gewählten Darstellung ab.

(b) Zeige: $v_p(r \cdot s) = v_p(r) + v_p(s)$ und $v_p(r+s) \geq \min(v_p(r), v_p(s))$ für alle $r, s \in \mathbb{Q}$, wobei im Fall $v_p(r) \neq v_p(s)$ in der Ungleichung stets die Gleichheit eintritt.

Eine Funktion v auf einem Körper K mit Werten in $\mathbb{N}$, die diese Eigenschaften besitzt, heißt *diskrete Bewertung* von K. Man nennt v_p die *p-adische Bewertung* von $\mathbb{Q}$.

(c) Zeige, daß durch

$$d(a,b) := \exp(-v_p(a-b))$$

eine Metrik auf $\mathbb{Q}$ definiert wird; sie heißt *p-adische Metrik*.

ABSCHNITT 3

Die Restklassenringe von $\mathbb{Z}$

Ein fundamentales Hilfsmittel der Zahlentheorie ist das Rechnen mit Restklassen.

Definition. R sei ein Ring, $\mathfrak{a}$ ein Ideal in R. Elemente $a, b \in R$ heißen *kongruent modulo* $\mathfrak{a}$, kurz $a \equiv b \bmod \mathfrak{a}$, wenn $a - b \in \mathfrak{a}$. Falls $\mathfrak{a} = Rc$ ein Hauptideal ist, schreiben wir auch $a \equiv b \mod c$, oder noch kürzer $a \equiv b \ (c)$.

Kongruenz ist bekanntlich (und trivialerweise) eine Äquivalenzrelation, daher hat man eine kanonische Abbildung

$$R \to R/\mathfrak{a}, \quad a \mapsto \overline{a},$$

von R in die Menge der Klassen kongruenter Elemente modulo $\mathfrak{a}$. Die Kongruenzklasse modulo $\mathfrak{a}$ von $a \in R$ heißt auch *Restklasse* von a. Bekanntlich gilt:

$$\overline{a} = \overline{b},\ \overline{c} = \overline{d} \quad \Longrightarrow \quad \overline{a + c} = \overline{b + d},\ \overline{ac} = \overline{bd}$$

und diese Verträglichkeit der Restklassenbildung mit Addition und Multiplikation in R erlaubt es uns, auf $R/\mathfrak{a}$ eine Ringstruktur zu definieren, indem wir einfach

$$\overline{a} + \overline{b} := \overline{a + b}, \quad \overline{a} \cdot \overline{b} := \overline{ac}$$

setzen. Diese Definition beinhaltet schon, daß die natürliche Abbildung $R \to R/\mathfrak{a}$ ein Ringhomomorphismus ist.

Es ist wichtig, daß man sich die Elemente von $R/\mathfrak{a}$ nicht als Mengen von Elementen von R vorstellt, sondern als Elemente eines Ringes mit denen man ganz „natürlich“ rechnen kann.

Wir wollen nun speziell den Fall $R = \mathbb{Z}$ betrachten. Wir wissen aus Abschnitt 1, daß jedes Ideal in $\mathbb{Z}$ von der Form $n\mathbb{Z}$ ist. Bezeichne $\mathbb{Z}_n$ den Restklassenring $\mathbb{Z}/n\mathbb{Z}$. Anwendung der Division mit Rest auf $a, b \in \mathbb{Z}$,

$$\begin{aligned} a &= q_1 n + r_1, \qquad 0 \le r_1 < |n|, \\ b &= q_2 n + r_2, \qquad 0 \le r_2 < |n|, \end{aligned}$$

liefert, daß

$$a \equiv b \ (n) \quad \Longleftrightarrow \quad r_1 = r_2.$$

Dies erklärt den Namen „Restklasse“: In der Restklasse von a modulo n liegen diejenigen Elemente, die bei Division durch n den gleichen Rest lassen.

Definition. Eine Teilmenge $\{a_1, \ldots, a_n\}$ von $\mathbb{Z}$ heißt ein *vollständiges Restsystem* modulo n, wenn $\{a_1, \ldots, a_n\}$ aus jeder Restklasse modulo n (genau) ein Element enthält.

Beispielsweise wissen wir, daß $\{0, \ldots, |n| - 1\}$ ein vollständiges Restsystem modulo n ist, und in der Regel werden wir mit diesem arbeiten. Ein nicht unwichtiger Aspekt des Rechnens mit Restklassen ist die Endlichkeit der Restklassenringe von $\mathbb{Z}$, so daß im Prinzip jedes Gleichungssystem über einem solchen Ring durch Probieren gelöst werden kann. Dies ermöglicht es sehr häufig, notwendige Bedingungen für die Lösbarkeit diophantischer Gleichungen aufzustellen.

Beispiel. Wir betrachten die Gleichung

$$x_1^2 + x_2^2 = 4x_3 + 3.$$

Ich behaupte: Diese Gleichung besitzt keine Lösung. Wir rechnen modulo 4:

$$\overline{x}_1^2 + \overline{x}_2^2 = \overline{4x}_3 + \overline{3} = \overline{0}\overline{x}_3 + \overline{3} = \overline{3}.$$

Nun nimmt aber $\overline{x}^2$ nur die Werte $\overline{0}$ und $\overline{1}$ an, $\overline{x}_1^2 + \overline{x}_2^2$ nur die Werte $\overline{0}$, $\overline{1}$, $\overline{2}$; also besitzt die Gleichung keine Lösung.

Das natürliche Hilfsmittel sind Rechnungen mit Kongruenzen bei Teilbarkeitsbetrachtungen. Wir betrachten als Beispiel Eulers Beweis, daß die Fermat-Zahl $F_5 = 2^{2^5} + 1$ nicht prim ist. Euler hat nämlich den Teiler 641 gefunden (vgl. dazu Aufgabe 3.16). Wir wollen die Behauptung $641 \mid F_5$ mit Hilfe des Rechnens im Restklassenring modulo 641 nachweisen. Da

$$641 = 640 + 1 \quad = 5 \cdot 2^7 + 1,$$
$$641 = 625 + 16 = 5^4 + 2^4,$$

folgt $5 \cdot 2^7 \equiv 640 \ (641) \equiv -1 \ (641)$. Also $5^4 \cdot 2^{28} \equiv 1 \ (641)$. Nun ist $5^4 \equiv -2^4$ (641) und $5^4 \cdot 2^{28} \equiv -2^4 \cdot 2^{28} \equiv -2^{32} \equiv 1$ (641) also $2^{32} \equiv -1$ (641) und $2^{32} + 1 \equiv 0 \ (641)$.

Wir haben in diesem Beispiel die Kongruenzschreibweise an Stelle von Gleichungen im Restklassenring verwendet. Dies ist in der zahlentheoretischen Literatur weithin üblich, vor allem dann, wenn Restklassen nach wechselnden Moduli zu bilden sind.

Als systematische Hilfsmittel wurde das Rechnen mit Kongruenzen von Gauß eingeführt. Wir werden in diesem und in den nächsten Abschnitten die Lösbarkeit von Kongruenzen studieren.

Definition. Sei $f \in \mathbb{Z}[X_1, \ldots, X_n]$ ein Polynom d-ten Grades. Eine Gleichung der Form

$$f(x_1, \ldots, x_n) \equiv 0 \mod m$$

heißt *Kongruenz d-ten Grades* in n Unbekannten.

Statt Kongruenz 1., 2., 3., 4. Grades sagen wir auch lineare, quadratische, kubische, biquadratische Kongruenz.

Bei der Lösbarkeit der Kongruenz $f(x_1, \ldots, x_n) \equiv 0 \ (m)$ fragen wir natürlich nur nach den Nullstellen $(\overline{x}_1, \ldots, \overline{x}_n) \in (\mathbb{Z}_m)^n$ des Polynoms $\overline{f} \in \mathbb{Z}_m[X_1, \ldots, X_n]$, wobei $\overline{f}$ einfach gebildet wird, indem man die Koeffizienten von f durch ihre Restklassen modulo m ersetzt. Dem entsprechend ist die *Zahl der Lösungen* der obigen Kongruenz die Zahl der Nullstellen von $\overline{f}$.

Lineare Kongruenzen bereiten uns wenig Schwierigkeiten. Wir betrachten der Kürze halber nur solche in einer Unbekannten:

Satz 3.1. *Seien a, b, $m \in \mathbb{Z}$, a, $m \neq 0$. Die Kongruenz*

$$ax \equiv b \ (m)$$

besitzt genau dann eine Lösung, wenn b ein Vielfaches des größten gemeinsamen Teilers d von a und m ist. Wenn (die Restklasse von) x_0 eine Lösung ist, so sind (die Restklassen von) $x_0 + m/d, \ldots, x_0 + (d-1)m/d$ die restlichen Lösungen.

Beweis. Genau dann gilt $ax \equiv b \ (m)$, wenn es ein $y \in \mathbb{Z}$ mit $ax - b = y \cdot m$, also $b = ax - my$ gibt. Nach 1.9 gibt es ein solches y genau dann, wenn b Vielfaches von d ist. Die x-Komponenten der Lösungen von $b = ax - my$ sind sämtlich von der Form

$$x_0 + k \cdot \frac{m}{d}.$$

Die im Satz genannten Lösungen repräsentieren daher sämtliche Lösungen modulo m, und sind überdies paarweise inkongruent modulo m. □

In anderer Sprechweise beschreibt Satz 3.1 die Lösbarkeit der Gleichung $ax = b$ in $\mathbb{Z}_m$, von der die multiplikative Struktur von $\mathbb{Z}_m$ wesentlich abhängt. Wir wollen aus Satz 3.1 Folgerungen ziehen:

Satz 3.2. *Wenn a und m teilerfremd sind, so besitzt $ax \equiv b \ (m)$ genau eine Lösung.*

Satz 3.3. *Für eine ganze Zahl $m \neq 0, \pm 1$ sind äquivalent:*

(a) *m ist eine Primzahl.*
(b) *Zu jedem $a \not\equiv 0 \ (m)$ und zu jedem b gibt es höchstens eine Lösung der Kongruenz $ax \equiv b \ (m)$.*
(c) *Zu jedem $a \not\equiv 0 \ (m)$ und zu jedem b gibt es mindestens eine Lösung der Kongruenz $ax \equiv b \ (m)$.*

Beweis. (a) $\Rightarrow$ (b), (c): Da $a \not\equiv 0 \ (m)$, sind $\mathrm{ggT}(a, m) = 1$. Also gibt es nach 3.1 genau eine Lösung.

(b) $\Rightarrow$ (a): Für einen von 1 und m verschiedenem Teiler $a > 0$ von m gilt

$$am \equiv m \ (m), \quad a \cdot \frac{m}{a} \equiv m \ (m), \quad \frac{m}{a} \not\equiv m \ (m)$$

(c) $\Rightarrow$ (a): Für einen von 1 und m verschiedenen Teiler $a > 0$ von m besitzt $ax \equiv 1$ (m) keine Lösung. □

Wir wollen Satz 3.3 noch einmal allgemeiner für euklidische Ringe formulieren. Wir betrachten dazu den Restklassenring $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z}$. Die Aussage (b) impliziert: Für jedes $a \in \mathbb{Z}_m$, $a \neq 0$, folgt aus $ax = 0$, daß $x = 0$; also ist $\mathbb{Z}_m$ nullteilerfrei. Die Aussage (c) impliziert: Für jedes $a \in \mathbb{Z}_m$, $a \neq 0$, besitzt $ax = 1$ eine Lösung; also ist $\mathbb{Z}_m$ ein Körper.

Satz 3.4. *R sei ein euklidischer Ring, $m \in R$, $m \neq 0$. Dann sind äquivalent:*

(a) *m ist Primelement,*
(b) *R/mR ist ein Integritätsbereich,*
(c) *R/mR ist ein Körper.*

Der Beweis von 3.4 verläuft wie der Beweis von 3.3, der sich ja nur auf den ersten Teil von 3.1 stützt, welcher wiederum aus 1.9 folgt.

Definition. Man nennt $a \in \mathbb{Z}$ *Einheit* modulo m, wenn $\overline{a}$ eine Einheit in $\mathbb{Z}_m$ ist.

Satz 3.1 sagt uns welche a Einheiten modulo m sind:

Satz 3.5. *Genau dann ist a eine Einheit modulo m, wenn a und m teilerfremd sind.*

Beweis. Genau dann ist $ax \equiv 1 \ (m)$ lösbar, wenn $\mathrm{ggT}(a, m) = 1$. □

Die Zahl der Einheiten modulo m zählt die *Eulersche φ-Funktion*:

$$\varphi(m) := \big|\{a : 0 \leq a < m,\ \mathrm{ggT}(a, m) = 1\}\big|.$$

Die zu m teilerfremden Restklassen bilden die Einheitengruppe des Ringes $\mathbb{Z}_m$, die wir mit $\mathbb{Z}_m^*$ bezeichnen. $\mathbb{Z}_m^*$ heißt auch *prime Restklassengruppe* modulo m. Es gilt $\varphi(m) = \mathrm{ord}\,\mathbb{Z}_m^*$. Auf $\mathbb{Z}_m^*$ können wir nun die elementaren Sätze über endliche Gruppen anwenden und erhalten

Satz 3.6 (Euler)**.** *Falls* $\mathrm{ggT}(a, m) = 1$, *so* $a^{\varphi(m)} \equiv 1 \ (m)$.

Beweis. Für $\overline{a} \in \mathbb{Z}_m^*$ ist $\overline{a}^{\mathrm{ord}\,\mathbb{Z}_m^*} = \overline{1}$. □

Satz 3.7 (Fermat)**.** *Wenn p eine Primzahl ist und $p \nmid a$, so ist $a^{p-1} \equiv 1 \ (p)$.*

Die Sätze 3.6 und 3.7 erscheinen uns als sehr einfach. Das sind sie aber nur deshalb, weil wir mit den „richtigen“ Begriffen, nämlich mit Kongruenzen, Restklassenringen und deren Einheitengruppen arbeiten. Dieser Aspekt wird auch durch den folgenden Satz beleuchtet:

Satz 3.8 (Wilson)**.** *Wenn p eine Primzahl ist, so ist*

$$(p-1)! \equiv -1 \ (p).$$

Allgemeiner gilt: In jedem endlichen Körper K ist das Produkt der von 0 verschiedenen Elemente gleich -1.

Beweis. Wir geben zwei verschiedene Argumente.

1. Die Mengen $\{1,-1\}$, $\{a,a^{-1}\}$, $a \in K$, $a \neq 1$, bilden eine Zerlegung der Einheitengruppe K^* von K (bei char $K = 2$ ist $1 = -1$). Das Produkt von 1 und -1 ist -1, und das Produkt von a und a^{-1} ist 1. Also ist das Produkt aller Elemente $\neq 0$ in K gleich -1.

Man berücksichtige dabei, daß 1 und -1 die einzigen zu sich selbst inversen Elemente sind: $x^2 - 1 = 0 \Rightarrow (x+1)(x-1) = 0 \Rightarrow x = 1$.

2. Sei q die Anzahl der Elemente von K. Dann gilt

$$x^{q-1} = 1 \quad \text{für alle } x \in K,\ x \neq 0.$$

Folglich wird $x^{q-1} - 1$ von den Linearfaktoren $x - a$, $a \in K$, $a \neq 0$, geteilt. Es muß sogar gelten

$$x^{q-1} - 1 = \prod_{a \in K^*} (x - a).$$

Multipliziert man das Produkt aus, so erhält man mittels Koeffizientenvergleich

$$\prod_{a \in K^*} a = -1. \qquad \square$$

Wenn $a \equiv b \ (m)$, so sind a und b erst recht kongruent modulo jedem Teiler d von m. Wir haben also einen natürlichen Ringhomomorphismus

$$\mathbb{Z}_m \longrightarrow \mathbb{Z}_d$$

für jeden Teiler d von m. Dieser natürliche Homomorphismus ergänzt das Diagramm

$$\begin{array}{ccc} \mathbb{Z} & \longrightarrow & \mathbb{Z}_d \\ \downarrow & \nearrow & \\ \mathbb{Z}_m & & \end{array}$$

kommutativ. Da der Ring $\mathbb{Z}_d$ weniger Elemente hat als $\mathbb{Z}_m$, ist seine Struktur in gewisser Weise einfacher. Die Struktur von $\mathbb{Z}_d$ ist vollständig von $\mathbb{Z}_m$ bestimmt. Andererseits bestimmen die Ringe $\mathbb{Z}_d$ umgekehrt die Struktur von $\mathbb{Z}_m$:

Satz 3.9 (Chinesischer Restsatz)**.** *Die Zahlen $m_1, \ldots, m_r$ seien paarweise teilerfremd. Sei $m := \prod_{i=1}^{r} m_i$. Dann ist der natürliche Homomorphismus*

$$\varphi : \mathbb{Z}_m \to \mathbb{Z}_{m_1} \oplus \cdots \oplus \mathbb{Z}_{m_r}$$
$$a \mod m \mapsto (a \mod m_1, \ldots, a \mod m_r)$$

ein Isomorphismus, mit anderen Worten: Für $x_1, \ldots, x_r \in \mathbb{Z}$ besitzt das System

$$x \equiv x_1 \ (m_1), \ldots, x \equiv x_r \ (m_r),$$

stets eine Lösung x, die modulo m eindeutig bestimmt ist.

Beweis. Wir zeigen, daß die Abbildung injektiv ist. Sei $k \in \mathbb{Z}$ mit

$$\varphi(k \mod m) = (k \mod m_1, \ldots, k \mod m_r) = (0, \ldots, 0).$$

Also gilt: $m_i \mid k$ für $i = 1, \ldots, r$. Da $m_1, \ldots, m_r$ paarweise teilerfremd sind, folgt(!) $m \mid k$. Also ist $k \equiv 0 \ (m)$. Da $\mathbb{Z}_m$ und $\mathbb{Z}_{m_1} \oplus \cdots \oplus \mathbb{Z}_{m_r}$ gleichmächtig sind, folgt aus der Injektivität von φ die Bijektivität. □

Das im Beweis des Satzes angewandte Anzahlargument ist natürlich nicht konstruktiv. Man würde schon gern die Umkehrabbildung ausrechnen, und dies geht auch. Sei

$$n_i = \prod_{j \neq i} m_j.$$

Dann ist $m = m_i n_i$, und da m_i und n_i teilerfremd sind (!), existieren a_i und b_i mit

$$1 = a_i m_i + b_i n_i.$$

Sei $e_i := b_i n_i$. Dann ist $e_i \equiv 1 \ (m_i)$ und $e_i \equiv 0 \ (m_j)$ für $j \neq i$. Zu gegebenen $x_1, \ldots, x_r$ setzt man dann

$$x := e_1 x_1 + \cdots + e_r x_r$$

und erhält wie gewünscht $x \equiv x_i \ (m_i)$, $i = 1, \ldots, r$.

Beispiel. Sei $m_1 = 3$, $m_2 = 5$, $m_3 = 7$, also $m = 105$ und $n_1 = 35$, $n_2 = 21$, $n_3 = 15$. Dann ist

$$\begin{aligned} 1 &= 12m_1 - n_1, & e_1 &= -35, \\ 1 &= -4m_2 + n_2, & e_2 &= 21, \\ 1 &= -2m_3 + n_3, & e_3 &= 15. \end{aligned}$$

Für $x_1 = 2$, $x_2 = 3$, $x_3 = 6$ ist dann

$$x = (-35) \cdot 2 + 21 \cdot 3 + 6 \cdot 15 = -70 + 63 + 90 = 83.$$

Es bietet sich natürlich sofort an, Satz 3.9 auf die normierte Primfaktorzerlegung

$$m = \operatorname{sgn}(m) \prod_{k=1}^{r} p_k^{\alpha_k}, \quad \alpha_k = v_{p_k}(m) > 0,$$

von m anzusetzen.

Satz 3.10. *Die natürliche Abbildung $\mathbb{Z}_m \to \mathbb{Z}_{p_1^{\alpha_1}} \oplus \cdots \oplus \mathbb{Z}_{p_r^{\alpha_r}}$ ist ein Isomorphismus. Eine Kongruenz*

$$f(x_1, \ldots, x_n) \equiv 0 \ (m)$$

besitzt genau dann eine Lösung modulo m, wenn sie für jedes k eine Lösung modulo $p_k^{\alpha_k}$ besitzt.

Beweis. Die erste Aussage ist lediglich eine Spezialisierung von Satz 3.9. Die Richtung „⇒“ ist trivial, und wenn für $k = 1, \ldots, r$

$$(x_1^{(k)}, \ldots, x_n^{(k)})$$

eine Lösung von $f(x_1, \ldots, x_n) \equiv 0 \ (p_k^{\alpha_k})$ ist, so wählt man einfach x, so daß $x_i \equiv x_i^{(k)} \ (p_k^{\alpha_k})$, $i = 1, \ldots, n$, und hat

$$f(x_1, \ldots, x_n) \equiv f(x_1^{(k)}, \ldots, x_n^{(k)}) \equiv 0 \ (p_k^{\alpha_k}), \quad k = 1, \ldots, r.$$

Nach Satz 3.9 folgt daraus $f(x_1, \ldots, x_n) \equiv 0 \ (m)$. □

Im nächsten Abschnitt werden wir uns mit der Struktur der primen Restklassengruppen $\mathbb{Z}_m^*$ beschäftigen. Unmittelbar aus Satz 3.9 folgt:

Satz 3.11. *Seien $m_1, \ldots, m_r$ paarweise teilerfremd und $m = m_1 \cdots m_r$. Dann ist*

$$\mathbb{Z}_m^* \cong \mathbb{Z}_{m_1}^* \times \cdots \times \mathbb{Z}_{m_r}^*$$

(wobei $\times$ hier das direkte Produkt von Gruppen bezeichnet.)

Wie wir wissen, ist

$$\operatorname{ord} \mathbb{Z}_m^* = \varphi(m) = \big|\{a : 0 \le a < m, \ a \text{ teilerfremd zu } m\}\big|.$$

Somit folgt unmittelbar aus Satz 3.11:

Satz 3.12. *Wenn m_1 und m_2 teilerfremd sind, so ist*

$$\varphi(m_1 m_2) = \varphi(m_1)\varphi(m_2).$$

Viele in der Zahlentheorie wichtigen Funktionen besitzen die in Satz 3.12 genannte Eigenschaft der Funktion φ:

Definition. Sei f eine Funktion auf $\mathbb{N}_+$ mit Werten in $\mathbb{C}$. Man nennt f *multiplikativ*, wenn $f(mn) = f(m)f(n)$ für teilerfremde $m, n \in \mathbb{N}_+$ gilt. (Wenn $f(mn) = f(m)f(n)$ sogar für beliebige $m, n \in \mathbb{N}_+$ gilt, so heißt f *voll multiplikativ*.)

Satz 3.10 zeigt, daß wir $\varphi(m)$ kennen, wenn $\varphi(p_k^{\alpha_k})$ für $k = 1, \ldots, r$ bekannt ist. Für eine Primzahl p gilt offensichtlich

$$\varphi(p^\alpha) = p^\alpha - p^{\alpha-1},$$

denn genau die $p^{\alpha-1}$ Vielfachen von p zwischen 1 und p^α sind zu p^α nicht teilerfremd.

Satz 3.13. *Für alle $m \in \mathbb{N}$, $m \neq 0$, ist*

$$\varphi(m) = m \prod_{p|m} \left(1 - \frac{1}{p}\right),$$

wobei das Produkt die Primteiler von m durchläuft.

Beweis.

$$\varphi(m) = \prod_{p|m}(p^{v_p(m)} - p^{v_p(m)-1}) = \prod_{p|m} p^{v_p(m)} \prod_{p|m}\left(1 - \frac{1}{p}\right)$$
$$= m \prod_{p|m}\left(1 - \frac{1}{p}\right). \qquad \square$$

Übungen.

3.14. Seien R, S Ringe; $\varphi : R \to S$ sei ein Ringhomomorphismus (d.h. $\varphi(a + b) = \varphi(a) + \varphi(b)$, $\varphi(ab) = \varphi(a)\varphi(b)$ für alle $a, b \in R$). Sei $I := \varphi^{-1}(0)$ der *Kern* von φ. Zeige:

(a) I ist ein Ideal.

(b) Für alle Ideale $\mathfrak{b}$ von S ist $\varphi^{-1}(\mathfrak{b})$ ein Ideal von R.

(c) Für jedes Ideal $\mathfrak{a}$ von R gilt $\varphi^{-1}(\varphi(\mathfrak{a})) = \mathfrak{a} + I$.

(d) Sei φ surjektiv. Dann ist die Zuordnung $\mathfrak{b} \mapsto \varphi^{-1}(\mathfrak{b})$ eine bijektive Abbildung zwischen der Menge der Ideale von S und der Menge I umfassenden Ideale von R.

3.15. (a) Bestimme 29^{-1} in $\mathbb{Z}_{99}^*$.

(b) Bestimme die letzten drei Ziffern von $9^{(9^9)}$ im Dezimalsystem.

(c) Bestimme alle Vielfachen der Zahl 17, deren Ziffern im Dezimalsystem alle gleich 1 sind.

3.16. (a) Eine Mersennesche Zahl $2^p - 1$, p prim, $p > 2$, hat nur Primteiler der Form $2np + 1$ mit $n \in \mathbb{N}$.

(b) Eine Fermatsche Zahl $2^{2^k} + 1$ hat nur Primfaktoren der Form $m2^{k+1} + 1$, $m \in \mathbb{N}$. (Wir werden später zeigen: Die Primfaktoren haben sogar die Gestalt $m2^{k+2} + 1$.)

3.17. Seien p, q ungerade Primzahlen, $p \neq q$. Die Zahl $p - 1$ teile $q - 1$. Zeige: Wenn $\mathrm{ggT}(n, pq) = 1$, so ist $n^{q-1} \equiv 1 \ (pq)$.

3.18. Sei p eine Primzahl. Zeige für $a, b \in \mathbb{Z}$:

(a) $a \equiv b \ (p^n) \Rightarrow a^p \equiv b^p \ (p^{n+1})$.

(b) $p \mid \binom{p}{k}$ für $k = 1, \ldots, p - 1$.

(c) $(x + y)^p = x^p + y^p$ für alle $x, y \in \mathbb{Z}_p$.

(d) Die Umkehrung von (b): Wenn $m \mid \binom{m}{k}$ für $k = 1, \ldots, m - 1$, so ist m prim.

3.19. Zeige, daß eine ungerade Zahl $p \geq 3$ genau dann prim ist, wenn

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

3.20. Bestimme die kleinsten positiven Lösungen der Systeme von Kongruenzen

$$\begin{array}{lll} \text{(a)} & x \equiv 1 \ (8) & \text{(b)} \qquad 2x \equiv 19 \ (25) \\ & x \equiv 3 \ (9) & \qquad\quad x^2 - x \equiv 3 \ (9) \\ & x \equiv 9 \ (11) & \qquad\quad x^3 \equiv 5 \ (8) \end{array}$$

Bestimme bei (b) ferner alle Lösungen mod $25 \cdot 9 \cdot 8$.

3.21. Beweise die folgende Verallgemeinerung des Chinesischen Restsatzes: Genau dann besitzt das System $x \equiv a_i \ (m_i)$, $i = 1, \ldots, n$ eine Lösung, wenn $\mathrm{ggT}(m_i, m_j) \mid (a_i - a_j)$ für alle i, j. Die Lösung ist eindeutig bestimmt mod $\mathrm{kgV}(m_1, \ldots, m_n)$.

3.22. Beweise:

(a) In einer zyklischen Gruppe G der Ordnung $n \in \mathbb{N}$ gibt es zu jedem Teiler d von n genau eine Untergruppe der Ordnung d, und diese ist ebenfalls zyklisch.

(b) Sei g ein erzeugendes Element von G. Das Element g^a, $a \in \mathbb{Z}$, erzeugt G genau dann, wenn a und n teilerfremd sind. (Wir schreiben die Verknüpfung in G als Multiplikation.)

(c) In G gibt es genau $\varphi(n)$ Elemente, die G erzeugen.

3.23. Seien G und H endliche Gruppen der Ordnungen m und n. Zeige daß für alle $x \in G \times H$ gilt: $\mathrm{ord}(x) \mid \mathrm{kgV}(m, n)$.

3.24. Die Gruppen G und H seien zyklisch von den endlichen Ordnungen m und n. Zeige daß $G \times H$ genau dann zyklisch ist, wenn m und n teilerfremd sind.

ABSCHNITT 4

Die primen Restklassengruppen

In diesem Abschnitt soll die Struktur der primen Restklassengruppen $\mathbb{Z}_m^*$ bestimmt werden. Nach 3.6 gilt für $m = p_1^{r_1} \dots p_s^{r_s}$ mit paarweise verschiedenen Primfaktoren p_i die Isomorphie

$$\mathbb{Z}_m^* \cong \mathbb{Z}_{p_1^{r_1}}^* \times \cdots \times \mathbb{Z}_{p_s^{r_s}}^*.$$

Daher genügt es, die Struktur der primen Restklassen modulo Primzahlpotenzen zu bestimmen.

Wir befassen uns zunächst mit dem Fall, in dem $m = p$ eine Primzahl ist. In der Sprache der Zahlentheorie formuliert, wird unser Ziel sein zu zeigen, daß es eine *primitive Wurzel* modulo p gibt. Nach dem Satz von Fermat ist jedes Element x von $\mathbb{Z}_p^*$ eine $(p-1)$-te Wurzel von 1 modulo p, denn $x^{p-1} = 1$ in $\mathbb{Z}_p$. Im Körper $\mathbb{C}$ der komplexen Zahlen besitzt 1 die $(p-1)$-ten Wurzeln $\zeta_i = \exp(2\pi i/(p-1))$, $i = 0, \dots, p-1$. Man sieht sofort, daß $\zeta_i = \zeta_1^i$ gilt, ζ_1 also die restlichen Wurzeln erzeugt: ζ_1 wird eine *primitive Wurzel* genannt. Dementsprechend stellt sich die Frage, ob es eine solche primitive Wurzel auch für die Gleichung $x^{p-1} = 1$ in $\mathbb{Z}_p$ gibt: existiert ein $x \in \mathbb{Z}_p$ derart, daß jedes Element von $\mathbb{Z}_p^*$ eine Potenz von x ist? Gruppentheoretisch formuliert lautet diese Frage: Ist die Gruppe $\mathbb{Z}_p^*$ zyklisch? Wie wir sehen werden, ist die Antwort auf diese Frage positiv.

Zur Vorbereitung beweisen wir zwei Sätze, die auch von eigenständigem Interesse sind.

Satz 4.1. *Für alle $n \in \mathbb{N}$, $n \geq 1$, ist $\sum_{d|n} \varphi(d) = n$.*

Beweis. Wir benutzen in diesem Beweis die Aussagen über zyklische Gruppen aus Aufgabe 3.22.

Sei d ein Teiler von n. Die zyklische Gruppe $\mathbb{Z}_n$ der Ordnung n besitzt genau eine Untergruppe H der Ordnung d, die selbst wieder zyklisch ist. Sei H etwa von $a \in H$ erzeugt. Dann sind bekanntlich die Elemente a^k, k teilerfremd zu d, $1 \leq k \leq d$, paarweise verschieden. Sie sind die Elemente der Ordnung d in H und damit in $\mathbb{Z}_n$.

Wir haben gezeigt: $\mathbb{Z}_n$ besitzt zu jedem Teiler d von n genau $\varphi(d)$ Elemente der Ordnung d. Andererseits hat $\mathbb{Z}_n$ genau n Elemente, und jedes dieser Elemente besitzt eine Ordnung, die ein Teiler von n ist. Also gilt $\sum_{d|n} \varphi(d) = n$. □

Satz 4.2. *G sei eine endliche Gruppe der Ordnung n. Zu jedem Teiler d von n existiere höchstens eine Untergruppe der Ordnung n. Dann ist G zyklisch (und besitzt daher zu jedem Teiler d von n genau eine Untergruppe der Ordnung d).*

Beweis. Für jeden Teiler d von n sei

$$i(d) := \begin{cases} 1 & \text{wenn es ein Element der Ordnung } d \text{ in } G \text{ gibt,} \\ 0 & \text{wenn es kein Element der Ordnung } d \text{ in } G \text{ gibt.} \end{cases}$$

Dann besitzt G genau

$$\sum_{d\,|\,n} i(d)\varphi(d) = n$$

Elemente. Wenn es nämlich kein Element der Ordnung d gibt, ist $i(d)\varphi(d) = 0$, und wenn es ein Element der Ordnung d gibt, erzeugt dies notwendig die einzige Untergruppe der Ordnung d, in der es dann genau $\varphi(d) = i(d)\varphi(d)$ Elemente der Ordnung d gibt. Mit Satz 4.1 folgt

$$i(d) = 1 \quad \text{für alle Teiler } d \text{ von } n,$$

insbesondere $i(n) = 1$, was zu beweisen war. □

Satz 4.3. *Sei K ein Körper und U eine endliche Untergruppe von K^*. Dann ist U zyklisch. Speziell ist $\mathbb{Z}_p^*$ zyklisch für jede Primzahl p.*

Beweis. Sei $V \subset U$ eine Untergruppe der Ordnung n. Nach dem Fermatschen Satz für endliche Gruppen (der Verallgemeinerung von Satz 3.7) ist

$$z^n = 1 \quad \text{für alle } z \in U.$$

Die Gleichung $x^n = 1$ besitzt aber im Körper K höchstens n Lösungen. Also besteht V aus allen n-ten Einheitswurzeln. Speziell gilt: U besitzt höchstens eine Untergruppe V der Ordnung n. Daher können wir den Satz 4.2 anwenden. □

Definition. Eine Zahl $a \in \mathbb{Z}$, deren Restklasse $\mathbb{Z}_m^*$ erzeugt, heißt *Primitivwurzel modulo m*.

Die Bestimmung einer Primitivwurzel erfolgt durch Ausprobieren (unter Verwendung des in Aufgabe 4.15 angegebenen Kriteriums). Beachten sollte man dabei natürlich, daß x^n keine Primitivwurzel sein kann, wenn x keine ist. Außerdem versucht man, die sich aus der Primfaktorzerlegung von $p-1$ ergebenden Folgerungen auszunutzen. Wir betrachten den Fall $p = 17$ und probieren zunächst $x = 2$. Die Potenzen von 2 haben der Reihe nach die Restklassen

$$\overline{2},\ \overline{4},\ \overline{8},\ \overline{16} = \overline{-1},$$

also ist $\overline{2}^4 = \overline{-1}$, $\overline{2}^8 = \overline{1}$, $\operatorname{ord}\overline{2} = 8$. Damit scheiden die acht Restklassen

$$\pm\overline{1},\ \pm\overline{2},\ \pm\overline{4},\ \pm\overline{8}$$

als Primitivwurzeln aus. Man kann nun schließen, daß die restlichen 8 primen Restklassen sämtlich prime Restklassen sein müssen.

Erstes Argument: $\mathbb{Z}_{17}^*$ besitzt jeweils genau eine Untergruppe U_i der Ordnungen $i = 1, 2, 4, 8, 16$ und keine weiteren Untergruppen. Da die U_i selbst wieder zyklisch sind, muß U_j in U_i enthalten sein, wenn $j \mid i$. Es gilt also

$$U_1 \subset U_2 \subset U_4 \subset U_8 \subset U_{16}.$$

U_8 wird von $\overline{2}$ erzeugt, wie oben gesehen, und für jedes Element $x \notin U_8$ muß $\langle x \rangle = U_{16}$ sein.

Zweites Argument: Als zyklische Gruppe der Ordnung 16 besitzt $\mathbb{Z}_{17}^*$ genau $\varphi(16) = 8$ erzeugende Elemente. Die genannten 8 Elemente scheiden aus; also müssen die übrigen 8 Elemente Primitivwurzeln sein.

Wir probieren aus, ob eine Primitivwurzel a modulo p auch Primitivwurzel modulo p^2 ist. Es gilt $\varphi(p^2) = p(p-1)$. Sei k die Ordnung von $\overline{a}$ in $\mathbb{Z}_{p^2}^*$. Dann ist k ein Teiler von $p(p-1)$.

Andererseits: $a^k \equiv 1 \ (p^2) \Rightarrow a^k \equiv 1 \ (p)$; k muß also auch ein Vielfaches von $p-1$ sein. Es bleiben die Fälle $k = p(p-1)$, in dem a Primitivwurzel modulo p^2 ist, und $k = p-1$, in dem a keine Primitivwurzel modulo p^2 ist. Im zweiten Fall probieren wir auch $a + p$. Da $a + p \equiv a \ (p)$, ist auch $a + p$ Primitivwurzel modulo p. Wir nehmen an, auch $a + p$ habe in $\mathbb{Z}_{p^2}^*$ die Ordnung $p - 1$:

$$a^{p-1} \equiv 1 \ (p^2) \quad \text{und} \quad (a+p)^{p-1} \equiv 1 \ (p^2).$$

Es folgt:

$$\begin{aligned} 1 \equiv (a+p)^{p-1} &\equiv a^{p-1} + p(p-1)a^{p-2} + p^2 \sum_{j=2}^{p-1} \binom{p-1}{j} a^{p-1-j} p^{j-2} \\ &\equiv a^{p-1} + p(p-1)a^{p-2} \mod p^2. \end{aligned}$$

Insgesamt: $p(p-1)a^{p-2} \equiv 0 \ (p^2)$, ein Widerspruch, weil p weder $p-1$ noch a^{p-2} teilt. Also gilt:

Satz 4.4. *Sei p eine Primzahl. Wenn a Primitivwurzel modulo p ist, so ist a oder $a + p$ Primitivwurzel modulo p^2. Insbesondere ist $\mathbb{Z}_{p^2}$ zyklisch.*

Da $\mathbb{Z}_8^*$ nicht zyklisch ist, kann es für die höheren Primzahlpotenzen nicht ganz so einfach weitergehen. Wir würden gern, soweit überhaupt möglich, aus einer Primitivwurzel modulo p^n eine Primitivwurzel modulo p^{n+1} gewinnen. Dazu ist es nützlich, sich zuerst über den Übergang von $\mathbb{Z}_{p^{n+1}}^*$ zu $\mathbb{Z}_{p^n}^*$ Gedanken zu machen. Wir erinnern uns: Da $p^n \mid p^{n+1}$, haben wir einen natürlichen Homomorphismus $\mathbb{Z}_{p^{n+1}} \to \mathbb{Z}_{p^n}$, der Elemente von $\mathbb{Z}_{p^{n+1}}^*$ auf solche von $\mathbb{Z}_{p^n}^*$ abbildet. In dieser und

ähnlichen Situationen ist es nicht schwer die Ordnung des Kerns auszurechnen: Sei $f : G \to H$ ein Homomorphismus endlicher Gruppen; dann gilt

$$\operatorname{ord}(G) = \operatorname{ord}(\text{Kern } f)\operatorname{ord}(\text{Bild } f).$$

Man hat ja eine disjunkte Zerlegung

$$G = \bigcup_{h \in \text{Bild } f} f^{-1}(h).$$

Jede der Fasern $f^{-1}(h)$ hat genau ord(Kern f) Elemente, denn wenn $f(g) = h$, so $g \cdot (\text{Kern } f) = f^{-1}(h)$.

Satz 4.5. *Sei $\pi : \mathbb{Z}^*_{p^{n+1}} \to \mathbb{Z}^*_{p^n}$ der natürliche Homomorphismus, $x \in \mathbb{Z}^*_{p^{n+1}}$.*

(a) *π ist surjektiv,* Kern π *hat die Ordnung p und ist daher zyklisch.*

(b) *Wenn x die Ordnung $(p-1)p^m$ hat, so hat $\pi(x)$ die Ordnung $(p-1)p^{m-1}$ oder $(p-1)p^m$.*

Beweis. (a) Seien $a_1, \ldots, a_k, k = \varphi(p^n) = p^{n-1}(p-1)$, die zu p^n teilerfremden Zahlen $\leq p^n$. Dann sind die Restklassen von $a_1, \ldots, a_k$ in $\mathbb{Z}_{p^{n+1}}$ ebenfalls prime Restklassen und $a_i \bmod p^{n+1}$ wird von π auf a_i modulo p^n abgebildet. Der Rest folgt aus gruppentheoretischen Sätzen.

(b) Im folgenden bezeichnen wir mit $\langle x \rangle$ die von x erzeugte Untergruppe. Es gilt $\pi(\langle x \rangle) = \langle \pi(x) \rangle$ und damit

$$\operatorname{ord} \pi(x) = \operatorname{ord}\langle \pi(x) \rangle = \frac{\operatorname{ord}\langle x \rangle}{\operatorname{ord}\{y \in \langle x \rangle \mid \pi(y) = 1\}}.$$

Da $\{y \in \langle x \rangle \mid \pi(y) = 1\} = \langle x \rangle \cap (\text{Kern } \pi)$ eine Untergruppe von Kern π ist, gibt es nur zwei Möglichkeiten:

$$\begin{aligned} &\operatorname{ord}((\text{Kern } \pi) \cap \langle x \rangle) = p, \quad \text{damit } \operatorname{ord} \pi(x) = p^{m-1}(p-1), \\ &\operatorname{ord}((\text{Kern } \pi) \cap \langle x \rangle) = 1, \quad \text{damit } \operatorname{ord} \pi(x) = p^m(p-1). \end{aligned}$$

□

Die erste Folgerung aus 4.5:

Satz 4.6. *$\mathbb{Z}^*_{2^n}$ ist nicht zyklisch für $n \geq 3$.*

Beweis. $\mathbb{Z}^*_{2^n}$ hat für $n \geq 3$ mindestens zwei Elemente der Ordnung 2: $-\overline{1}$ und das von $\overline{1}$ verschiedene Element von Kern π, denn $-\overline{1} \notin \text{Kern } \pi$. Dabei haben wir π wie in 4.5 gewählt. □

Sei $a \in \mathbb{Z}$, $p \nmid a$. Trotz der negativen Aussage von 4.6 wollen wir untersuchen, wann in 4.5 (b) der Fall $\operatorname{ord}(a \bmod p^n) = (p-1)p^{m-1}$, $\operatorname{ord}(a \bmod p^{n+1}) = (p-1)p^m$ eintritt, die Ordnung der Restklasse von a mit wachsenden Exponenten von p steigt. Es ist klar, daß dies ohne weitere Voraussetzungen nicht gilt. Die Voraussetzung, die wir benötigen, ist aber überraschend schwach. Der folgende

Satz zeigt, daß die Ordnung immer weiter steigt, sobald sie nur ein einziges Mal angestiegen ist.

Satz 4.7. *Sei p prim, $a \in \mathbb{Z}$, $p \nmid a$. Es sei $n \geq 2$, wenn $p \geq 3$, und $n \geq 3$, wenn $p = 2$. Dann gilt:*

$$\operatorname{ord}(a \mod p^{n-1}) = (p-1)p^{m-2}, \qquad \operatorname{ord}(a \mod p^n) = (p-1)p^{m-1}$$
$$\Longrightarrow \qquad \operatorname{ord}(a \mod p^{n+1}) = (p-1)p^m.$$

Beweis. Nach Voraussetzung ist

$$a^{(p-1)p^{m-2}} \equiv 1 \ (p^{n-1}), \quad a^{(p-1)p^{m-2}} \not\equiv 1 \ (p^n).$$

Folglich $a^{(p-1)p^{m-2}} = 1 + bp^{n-1}$ mit $p \nmid b$. Nach 4.5 (b) kommen nur in Frage

$$\operatorname{ord}(a \mod p^{n+1}) = (p-1)p^{m-1} \quad \text{oder} \quad \operatorname{ord}(a \mod p^{n+1}) = (p-1)p^m.$$

Der erste Fall ist auszuschließen. Es gilt

$$\begin{aligned} a^{(p-1)p^{m-1}} &= (a^{(p-1)p^{m-2}})^p = (1 + bp^{n-1})^p \\ &= 1 + bpp^{n-1} + \sum_{k=2}^{p-1} \binom{p}{k} b^k p^{(n-1)k} + b^p p^{(n-1)p} \\ &\equiv 1 + bp^n \mod p^{n+1}, \end{aligned}$$

denn für $1 < k < p$ teilt p den Binomialkoeffizienten $\binom{p}{k}$ und

$$(n-1)k + 1 \geq (n-1)2 + 1 = (n+1) + (n-2) \geq n+1$$

für $n \geq 2$ und $(n-1)p \geq n+1$ für $p \geq 3, n \geq 2$ oder $p = 2, n \geq 3$. Wäre nun

$$a^{(p-1)p^{m-1}} \equiv 1 \ (p^{n+1}),$$

so würde $bp^n \equiv 0 \ (p^{n+1})$ folgen, im Widerspruch zu $p \nmid b$. □

Da wir für die ungeraden Primzahlen nur $n \geq 2$ in 4.7 voraussetzen müssen, erhalten wir per Induktion sehr leicht:

Satz 4.8. *Sei $p \geq 3$ prim und $a \in \mathbb{Z}$ eine Primitivwurzel modulo p^2. Dann ist a Primitivwurzel modulo p^n für alle $n \geq 2$. Insbesondere ist $\mathbb{Z}_{p^n}^*$ zyklisch.*

Beweis. Für eine Primitivwurzel a modulo p^2 gilt

$$\operatorname{ord}(a \mod p) = p-1, \quad \operatorname{ord}(a \mod p^2) = (p-1)p.$$

Aus 4.7 folgt $\operatorname{ord}(a \mod p^3) = (p-1)p^2$ und dann mit Induktion

$$\operatorname{ord}(a \mod p^n) = (p-1)p^{n-1}$$

für alle $n \geq 1$. □

Im Fall $p = 2$ greift Satz 4.7 erst von $n = 3$ an. Da $\operatorname{ord}(5 \bmod 2^2) = 1 = p-1$ und $\operatorname{ord}(5 \bmod 2^3) = 2 = (p-1)p$, folgt Teil (a) des nächsten Satzes mittels 4.7.

Satz 4.9.

(a) $\operatorname{ord}(5 \mod 2^n) = 2^{n-2}$ *für alle* $n \geq 3$.

(b) *Jedes Element* $x \in \mathbb{Z}_{2^n}^*$ *läßt sich auf genau eine Weise in der Form*

$$x = (-\overline{1})^u \overline{5}^k, \quad u \in \{0, 1\}, \quad k \in \{0, \ldots, 2^{n-2} - 1\}$$

darstellen.

Beweis. Es ist nur noch (b) zu beweisen. Wir betrachten die Menge der Produkte $(-1)^k \overline{5}^u$, k, u wie oben. Wenn

$$(-\overline{1})^u \overline{5}^k = (-\overline{1})^v \overline{5}^\ell,$$

so folgt im Fall $u = v$, daß $\overline{5}^k = \overline{5}^\ell$, also $k = \ell$ gemäß Teil (a), daß im Fall $u \neq v$,

$$-\overline{1} = (-\overline{1})^{u-v} = \overline{5}^{\ell-k}.$$

Als Kongruenz geschrieben lautet diese Gleichung

$$-1 \equiv 5^{\ell-k} \ (2^n).$$

Da $n \geq 2$, folgt $-1 \equiv 5$ (4), ein Widerspruch. Es gibt also $2 \cdot 2^{n-2} = 2^{n-1}$ Elemente der Form $(-\overline{1})^u \overline{5}^k$, u, k wie oben, und $\mathbb{Z}_{2^n}^*$ hat genau 2^{n-1} Elemente. □

Wir formulieren Satz 4.9 noch einmal gruppentheoretisch.

Satz 4.10. *Für alle* $n \geq 3$ *ist* $\mathbb{Z}_{2^n}^*$ *das* direkte Produkt *der von* $\overline{5}$ *erzeugten Untergruppe* U *und der von* $-\overline{1}$ *erzeugten Untergruppe* V*, d.h. jedes Element von* $\mathbb{Z}_{2^n}^*$ *läßt sich auf genau eine Weise in der Form* $u \cdot v$, $u \in U$, $v \in V$, *darstellen. Die Abbildung* $\tau : U \times V \to \mathbb{Z}_{2^n}^*$, $\tau(u, v) = u \cdot v$, *ist ein Isomorphismus von Gruppen.*

Beweis. Einzig zu zeigen bleibt, daß τ ein Homomorphismus ist. Dies aber ist trivial, weil $\mathbb{Z}_{2^n}^*$ abelsch ist. □

Satz 4.10 ist auch für $n = 1, 2$ richtig, dann aber besteht U nur aus dem neutralen Element.

Wir fassen die bisherigen Ergebnisse zusammen:

Satz 4.11. *Genau dann ist* $\mathbb{Z}_m^*$ *eine zyklische Gruppe, wenn* $m = 2, 4, p^n$ *oder* $2p^n$ *mit einer ungeraden Primzahl* p *und* $n \geq 1$.

Beweis. „⇒“ Für $m = 2, 4$ ist dies trivial, für $m = p^n$ Aussage von Satz 4.11. Ist $m = 2p^n$, so ist

$$\mathbb{Z}_m^* \cong \mathbb{Z}_2^* \times \mathbb{Z}_{p^n}^* \cong \mathbb{Z}_{p^n}^*,$$

da $\mathbb{Z}_2^*$ nur aus dem neutralen Element besteht.

„⇐“ Sei $m = p_1^{r_1} \dots p_s^{r_s}$ die normierte Primfaktorzerlegung von m. Dann ist

$$\mathbb{Z}_m^* \cong \mathbb{Z}_{p_1^{r_1}}^* \times \cdots \times \mathbb{Z}_{p_s^{r_s}}^* .$$

Falls $s = 1$ ist, m also Primzahlpotenz, muß $m = 2$, 4 oder Potenz einer ungeraden Primzahl sein gemäß 4.6. Sei $s \geq 2$. Wenn m zwei ungerade Primfaktoren hat, kommen in der Produkt-Darstellung von $\mathbb{Z}_m^*$ zwei Faktoren gerader Ordnung vor, und $\mathbb{Z}_m^*$ kann nicht zyklisch sein (vergleiche Aufgabe 3.24). Gleiches gilt im Fall $m = 2^{r_1} \cdot p^{r_2}$, wenn $r_1 \geq 2$. □

Übungen.

4.12. (a) Bestimme Primitivwurzeln modulo n für $n = 41, 49, 50$.

(b) Bestimme sämtliche Primitivwurzeln modulo 41.

4.13. Zeige, daß $x^2 \equiv -1 \ (p)$ genau dann lösbar ist, wenn $p \equiv 1 \ (4)$, und daß $x^4 \equiv -1 \ (p)$ genau dann lösbar ist, wenn $p \equiv 1 \ (8)$ (Dabei ist $p > 2$ eine Primzahl).

4.14. Sei p eine Primzahl und z eine Primitivwurzel modulo p. Zeige:

(a) Genau dann ist z^k quadratischer Rest modulo p, wenn k gerade ist. (Man sagt, daß a quadratischer Rest modulo p ist, wenn die Restklasse von a modulo p ein Quadrat in $\mathbb{Z}_p$ ist.)

(b) Sei $p - 1 = 2^t q$, q ungerade. Dann ist a quadratischer Rest modulo p genau dann, wenn $\operatorname{ord}(a \mod p) = 2^u r$, r ungerade, $u < t$.

4.15. Sei p eine Primzahl und $a \in \mathbb{Z}$ teilerfremd zu p. Zeige, daß a genau dann Primitivwurzel modulo p ist, wenn $a^{(p-1)/q} \not\equiv 1 \ (p)$ für jeden Primteiler q von $p - 1$ ist.

4.16. Sei p eine ungerade Primzahl. Zeige, daß es genau $p^{k-2}(p-1)\,\varphi(p-1)$ Primitivwurzeln modulo p^k gibt.

ABSCHNITT 5

Primzahltests, Kryptographie und Faktorisierung

In diesem Abschnitt sprechen wir zunächst über *Primzahltests*, also über Verfahren, mit denen man prüfen kann, ob eine Zahl m Primzahl ist. Allerdings sind diese Tests nicht in allen Fällen aussagekräftig, sie lassen bei negativem Ausgang stets den Schluß zu, daß m keine Primzahl ist, machen bei positivem Ausgang aber keine definitive Aussage über m. Allerdings schlüpfen nur wenige zusammengesetzte Zahlen durch die Maschen der Tests.

Der naive Primzahltest ist das *Probedividieren:* man prüft m auf Teilbarkeit durch alle potentiellen Teiler. Dabei braucht man nicht alle Zahlen $\leq m$ auszuprobieren: m ist genau dann Primzahl, wenn für alle (Primzahlen) n mit $2 \leq n \leq \sqrt{m}$ gilt: $n \nmid m$. Soweit vorhanden, kann man die potentiellen Teiler p einer Primzahltabelle entnehmen. Ist der Bereich der Primzahltabelle erschöpft, kann man die Nichtprimzahlen unter den potentiellen Teilern nur noch sehr grob aussondern: man prüft auf Teilbarkeit durch ungerade Zahlen (Ersparnisfaktor: 1/2), besser noch auf Teilbarkeit durch zu 6 teilerfremde Zahlen (Faktor 1/3), noch besser auf Teilbarkeit durch zu 30 teilerfremde Zahlen (Faktor 8/30) usw. Der Rechenaufwand wird jedoch bei allem Bemühen proportional zu

$$c \cdot \sqrt{n}$$

sein, wobei c eine (möglicherweise sehr kleine) Konstante ist. Ist m eine 100-stellige Primzahl, $c = 10^{-9}$ (dies dürfte schwer zu erreichen sein) und hat man einen Computer, der 10^9 Divisionen pro Sekunde ausführen kann, so benötigt man etwa 10^{32} Sekunden $\approx 3 \cdot 10^{24}$ Jahre, um zu erkennen, daß m Primzahl ist. Das Verfahren ist also für Zahlen dieser Größenordnung völlig impraktikabel, allerdings liefert es uns auch mehr als die Information, ob m Primzahl ist. Es liefert stets den kleinsten Primfaktor von m, bei iterierter Anwendung sogar die Primfaktorzerlegung.

Daß m zusammengesetzt ist, kann man aber häufig sehr schnell erkennen, und zwar ohne Bestimmung eines echten Teilers. Die folgenden Überlegungen beruhen alle darauf, daß uns die Struktur von $\mathbb{Z}_p^*$ für Primzahlen p bekannt ist. Zunächst brauchen wir nur auszunutzen, daß ord $\mathbb{Z}_p^* = p-1$ ist. Der Satz von Fermat liefert den

Fermat-Test: Ist $a^{m-1} \not\equiv 1 \ (m)$ für ein $a \in \mathbb{Z}$, $a \not\equiv 0 \ (m)$, so ist m keine Primzahl.

Man führt den Fermat-Test durch, indem man $a = 2, 3, 5, \ldots$ wählt und prüft, ob $a^{m-1} \equiv 1 \ (m)$. Um die Potenzen (modulo m) schnell zu berechnen, stellt man $m-1$ im Dualsystem dar,

$$m-1 = \sum_{k=0}^{u} \delta_k 2^k, \quad \delta_k \in \{0,1\}, \quad a_u \neq 0,$$

und bestimmt a^{m-1} im Fall $a \neq 1$ so: Sei $a_u := a$ und mit rückwärts laufender Rekursion

$$a_k := \begin{cases} a_{k+1}^2 \cdot a \ (m) & \text{wenn } \delta_k = 1 \\ a_{k+1}^2 \ (m) & \text{wenn } \delta_k = 0. \end{cases}$$

Dann ist $a_0 \equiv a^{m-1} \ (m)$. Auf diese Weise brauchen wir höchstens $2\log_2(m-1)$ Multiplikationen modulo m.

Beispiel. Sei etwa $m = 1903$, $a = 2$,

$$1902 = 11101101110_2.$$

Dann ist $a_{11} = 1$, $a_{10} = 2$, $a_9 = 8$, $a_8 = 128$, $a_7 = 1160$, $a_6 = 358$, $a_5 = 1326$, $a_4 = 1807$, $a_3 = 1305$, $a_2 = 1583$, $a_1 = 1179$, $a_0 = 851$.

Es gilt $2^{1902} \equiv 851 \ (1903)$, also ist 1903 keine Primzahl. Die Primfaktorzerlegung von 1903 ist $1903 = 11 \cdot 173$.

Definition. m heißt *quasiprim zur Basis* a, wenn $a^{m-1} \equiv 1 \ (m)$.

Mißlich ist nur, daß es zusammengesetzte Zahlen gibt, die auf diese Weise kaum als nicht prim erkannt werden können. Es gilt natürlich die Umkehrung des Fermatschen Satzes: $a^{m-1} \equiv 1 \ (m)$ für *alle* a, $1 \leq a \leq m-1 \Rightarrow a$ prim. Wenn man aber zu viele a probieren muß, geht die Effektivität des Tests verloren.

Zum Beispiel ist $341 = 11 \cdot 31$ quasiprim zur Basis 2:

$$\left.\begin{array}{l} 2^{340} \equiv 1 \ (11), \quad \text{da } 2^{10} \equiv 1 \ (11) \\ 2^{340} \equiv 1 \ (31), \quad \text{da } 2^{5} \equiv 1 \ (31) \end{array}\right\} \Rightarrow 2^{340} \equiv 1 \ (341).$$

Da $3^{10} \not\equiv 1 \ (31)$, ist $3^{340} \not\equiv 1 \ (341)$, 341 also nicht quasiprim zur Basis 3.

Die Wirksamkeit des Fermat-Testes wird ernsthaft erschüttert durch die Existenz der Carmichael-Zahlen:

Definition. m heißt *Carmichael-Zahl*, wenn m nicht prim, aber $a^{m-1} \equiv 1 \ (m)$ für alle zu m teilerfremden $a \in \mathbb{Z}$.

Die kleinste Carmichael-Zahl ist $561 = 3 \cdot 11 \cdot 17$. Für a mit $(a, 561) = 1$ gilt

$$\left.\begin{array}{l} a^{560} \equiv a^0 \equiv 1 \ (3) \\ a^{560} \equiv a^0 \equiv 1 \ (11) \\ a^{560} \equiv a^0 \equiv 1 \ (17) \end{array}\right\} \Rightarrow a^{560} \equiv 1 \ (561).$$

Nicht prime Quasiprimzahlen oder gar Carmichael-Zahlen sind ziemlich selten. Zum Beispiel gibt es 882 206 716 Primzahlen $< 20 \cdot 10^9$, aber nur 19 865 Quasiprimzahlen zur Basis 2 in diesem Bereich, die nicht prim sind.

Man kann andererseits leicht zeigen, daß es unendlich viele nicht prime Quasiprimzahlen zur Basis 2 gibt (Aufgabe 5.7). Es ist seit 1994 bekannt, daß es unendlich viele Carmichael-Zahlen gibt; die Anzahl der Carmichael-Zahlen $\leq x$ ist von der Größenordnung $x^{1/10}$.

Wenn die Primfaktorzerlegung von $m-1$ bekannt ist, kann man den Fermat-Test zum $(p-1)$-*Primzahltest* verschärfen und wirklich entscheiden, ob m prim ist:

Satz 5.1. *Wenn es ein $a \in \mathbb{Z}$ gibt mit $a^{m-1} \equiv 1 \ (m)$, aber $a^{(m-1)/p} \not\equiv 1 \ (m)$ für jeden Primteiler p von $m-1$, so ist m prim.*

Beweis. Sei k die Ordnung von a in $\mathbb{Z}_m^*$. Da $a^{m-1} \equiv 1 \ (m)$, ist k ein Teiler von $m-1$. Wäre k ein echter Teiler von $m-1$, so würde es eine der Zahlen $(m-1)/p$ teilen. Mithin ist $\mathbb{Z}_m^*$ (zyklisch) von der Ordnung $m-1$ und m daher prim. (Vergleiche auch Aufgabe 4.15.) □

Unter den Voraussetzungen von Satz 5.1 ist a Primitivwurzel modulo m. Eine solche ist i.a. schnell gefunden.

Eine andere, ebenfalls sehr wirksame Verschärfung des Fermat-Testes ist der *Rabin-Test*. Er beruht auf folgendem Satz:

Satz 5.2. *Sei $p > 2$ eine Primzahl, $p \nmid a$. Sei $p - 1 = 2^h \cdot q$, q ungerade. dann gilt*

$$a^q \equiv 1 \ (p) \quad \textit{oder} \quad a^{q2^i} \equiv -1 \ (p) \quad \textit{für ein } i,\ 0 \leq i \leq h-1. \qquad (*)$$

Beweis. $\mathbb{Z}_p^*$ ist zyklisch und hat genau ein Element der Ordnung 2, nämlich -1. Sei i die kleinste Zahl $-1 \leq i \leq h-1$, für die

$$a^{q2^{i+1}} \equiv 1 \ (p).$$

Die Zahl i ist wohldefiniert, weil $a^{p-1} = a^{q2^h} \equiv 1 \ (p)$. Im Fall $i = -1$ ist $a^q \equiv 1 \ (p)$. Im Fall $i \geq 0$ ist $(a^{q2^i})^2 = a^{q2^{i+1}} \equiv 1 \ (p)$. Da $a^{q2^i} \not\equiv 1 \ (p)$, also die Restklasse von der Ordnung 2 in $\mathbb{Z}_p^*$ dargestellt, muß $a^{q2^i} \equiv -1 \ (p)$ sein. □

Obwohl der Rechenaufwand zur Überprüfung der Bedingung $(*)$ nicht viel größer ist als der für den Fermat-Test, ist der Rabin-Test wesentlich wirksamer: Es gibt keine zusammengesetzte Zahl $< 25 \cdot 10^9$, die den Rabin-Test zu den Basen 2, 3, 5, 7 und 11 übersteht.

Insofern ist folgende Definition gerechtfertigt:

Definition. Eine ungerade Zahl m heißt *stark quasiprim* zur Basis a, falls m (an Stelle von p) die Bedingung $(*)$ in Satz 5.2 erfüllt.

Daß es Analoga zu den Carmichael-Zahlen für den Rabin-Test nicht gibt, zeigt der folgende Satz:

Satz 5.3. *Sein $m \in \mathbb{N}$, $m \geq 2$ ungerade, $m-1 = 2^h \cdot q$, q ungerade. Sei*

$$B := \{x \in \mathbb{Z}_m^* : x^q = 1 \text{ oder } x^{q2^i} = -1 \text{ für ein } i,\ 0 \leq i \leq h-1\}.$$

Wenn m nicht prim ist, so ist $|B| \leq \varphi(m)/4$, mit Ausnahme von $m = 9$, $|B| = 2$.

Satz 5.3 läßt sich so interpretieren: Wenn wir a zufällig im Bereich $1 \leq a \leq m-1$ wählen, so ist die Wahrscheinlichkeit, daß a zu B gehört, höchstens 1/4, also die Wahrscheinlichkeit dafür, mittels der Basis a im Rabin-Test m nicht als zusammengesetzt zu erkennen, höchstens 1/4. Führt man den Rabin-Test mit k unabhängig voneinander gewählten Basen aus, ist die Fehlerquote des Rabin-Tests höchstens $(1/4)^k$.

Beweis. Wir setzen

$$F := \{x \in \mathbb{Z}_m^* : x^{m-1} = 1\}$$

Dies ist die Menge der Restklassen a modulo m, für die m quasiprim zur Basis a ist. Als Kern einer Potenzabbildung auf $\mathbb{Z}_m^*$ ist F eine Untergruppe.

Sei zunächst $m = p^e$ Potenz einer ungeraden Primzahl p, $e \geq 2$. Dann ist $\mathbb{Z}_m^*$ zyklisch von der Ordnung $\varphi(m) = (p-1)p^{e-1}$. Für $x \in F$ gilt $\operatorname{ord}(x) \mid p^e - 1$ und $\operatorname{ord}(x) \mid (p-1)p^{e-1}$, also $\operatorname{ord}(x) \mid p-1$. Damit ist F die Untergruppe der Ordnung $p-1$ von $\mathbb{Z}_m^*$, und $|F| = \varphi(m)/p^{e-1}$. Mit Ausnahme von $p = 3$, $e = 2$ gilt $p^{e-1} > 4$. (Im Fall $m = 9$ ist $F = B = \{\pm 1\}$.)

Wir können jetzt annehmen, daß m mehr als einen Primfaktor hat. Sei $m = p_1^{e_1} \cdots p_s^{e_s}$ die Primfaktorzerlegung von m mit paarweisen verschiedenen p_i. Es ist

$$a^{m-1} \equiv 1 \ (m) \quad \Longleftrightarrow \quad a^{m-1} \equiv 1 \ (p_i^{e_i}),\ i = 1, \ldots, s.$$

Mit der Zerlegung $\mathbb{Z}_m^* \cong \mathbb{Z}_{p_1^{e_1}}^* \times \cdots \times \mathbb{Z}_{p_s^{e_s}}^*$ folgt daraus

$$F \cong F_1 \times \cdots \times F_r, \qquad F_i := \{x \in \mathbb{Z}_{p_i^{e_i}}^* : x^{m-1} = 1\}.$$

Beachte, daß $-1 \in F_i$ für alle i, so daß $|F_i|$ stets gerade ist. Wir schreiben

$$|F_i| = 2^{t_i} u_i, \qquad u_i \text{ ungerade.}$$

Wir können annehmen, daß $t_1 \leq \cdots \leq t_s$.

Für jedes $x \in F$ gilt

$$\operatorname{ord}(x) \mid \operatorname{kgV}(|F_1|, \ldots, |F_s|) =: 2^t \cdot u, \quad t = t_s,\ u = \operatorname{kgV}(u_1, \ldots, u_s). \qquad (*)$$

Wie wir gerade beobachtet haben, ist $t \geq 1$. Man kann also den Endomorphismus $\sigma : F \to F$ mittels

$$\sigma(x) := x^{u \cdot 2^{t-1}}$$

definieren. Wegen $(*)$ ist $(\sigma(x))^2 = 1$ für alle $x \in F$.

Sei nun $x \in B$. Wenn $x^q = 1$, hat x ungerade Ordnung, und aus $x^{u2^t} = 1$ folgt schon $x^u = 1$, erst recht $\sigma(x) = x^{u\cdot 2^{t-1}} = 1$. Wenn $x^{q2^i} = -1$ ist, dann ist $-1 \in \langle x\rangle$. Weil $\langle x\rangle$ zyklisch ist, $\sigma(x) \in \langle x\rangle$ und $\sigma(x)^2 = 1$, muß $\sigma(x) = \pm 1$ sein.

(i) Wir betrachten zunächst den Fall, daß alle t_i übereinstimmen. Für $y \in F_i$ ist dann $y^{u\cdot 2^{t-1}} = -1$, wenn y ungerade Ordnung hat, und sonst 1. Bei Anwendung des Isomorphismus $\pi : F \to F_1 \times \cdots \times F_r$ sehen wir, daß $\sigma(x) = \pm 1$ genau dann gilt, wenn

$$\pi(x) \in E_1 \times \cdots \times E_s \cup O_1 \times \cdots \times O_s$$

ist. Dabei bezeichne E_i die Elemente gerader Ordnung in F_i, O_i die Elemente ungerader Ordnung. Da $|E_i| = |O_i| = |F_i|/2$ für alle i, gilt

$$|B| \le |E_1 \times \cdots \times E_s| \cup |O_1 \times \cdots \times O_s| = 2\frac{|F|}{2^s} = \frac{|F|}{2^{s-1}}$$

Nach Voraussetzung über m ist $s \ge 2$. Wenn m keine Carmichael-Zahl ist, folgt

$$|B| \le \frac{|F|}{2} \le \frac{\varphi(m)}{4}.$$

Wenn aber m eine Carmichael-Zahl ist, muß $s \ge 3$ sein (Aufgabe 5.8), und es ergibt sich

$$|B| \le \frac{|F|}{4} = \frac{\varphi(m)}{4}.$$

(ii) Es gilt $t_1 < t$. Da $t_1 \ge 1$, ist dann $t \ge 2$. Die Ordnung $u_1 2^{t_1}$ von F_1 teilt $u2^{t-1}$. Daher ist $y^{u\cdot 2^{t-1}} = 1$ für alle $y \in F_1$. Der Fall $\sigma(x) = -1$ kann also gar nicht eintreten. Mithin $\sigma(x) = 1$ für alle $x \in B$. Sei $U = \operatorname{Kern}\sigma$. Wie gerade gesehen, ist $B \subset U$. Außerdem ist U eine echte Untergruppe von F, wie wir jetzt zeigen.

Sei w eine Primitivwurzel modulo $p_s^{e_s}$. Wir wählen

$$a \equiv 1 \ (p_i^{e_i}),\ i = 1, \ldots, s-1, \qquad a \equiv w \ (p_s^{e_s})$$

und betrachten die Restklasse x_1 von a. Offensichtlich ist $x_1 \in F$. Es gilt $\sigma(x_1) \ne 1$ für die Restklasse x_1 von a, denn $a^{u2^{t-1}} \equiv -1 \ (p_s^{e_s})$. Beachte, daß ebenso $\sigma(x_1) \ne -1$, denn $a^{u2^{t-1}} \equiv 1 \ (p_i^{e_i})$, $i = 1, \ldots, r$. Insgesamt: $x_1 \in F$, $x_1 \notin U$.

Da $t \ge 2$, können wir zusätzlich den Endomorphismus $\tau : F \to F$ betrachten,

$$\tau(x) := x^{u\cdot 2^{t-2}}.$$

Wegen $\tau(x)^2 = \sigma(x) = 1$ für $x \in B$, folgt mit dem gleichen Argument wie oben, daß $\tau(x) = \pm 1$ für alle $x \in B$. Sei V die Untergruppe aller $z \in U$ mit $\tau(z) = \pm 1$. Wenn wir noch zeigen, daß V eine echte Untergruppe von U ist, dann ergibt sich

$$|B| \le |V| \le \frac{|U|}{2} \le \frac{|F|}{4} \le \frac{\varphi(m)}{4}.$$

Wir wählen nun a wie oben, betrachten aber die Restklasse $x_2 = x_1^2$ von a^2. Für diese gilt

$$\tau(x_2) = \sigma(x_1) \neq \pm 1, \qquad \text{aber} \qquad \sigma(x_2) = \sigma(x_1^2) = \sigma(x_1)^2 = 1.$$

Mit x_2 haben wir also ein Element gefunden, daß zwar in U, aber nicht in V liegt. □

Mathematiker haben seit etwa zwanzig Jahren revolutionär neue *Kryptosysteme* entwickelt. Mehrere dieser Verfahren beruhen auf zahlentheoretischen Sätzen und Algorithmen.

Die klassischen Anwendungsbereiche von Kryptosystemen sind Militär, Diplomatie und – nicht weit davon entfernt – die Geheimdienste. Die neuen Verschlüsselungsverfahren werden natürlich auch von diesen Institutionen mit Freuden eingesetzt, entwickelt worden sind sie aber für eine andere Aufgabe: die Geheimhaltung von Daten in öffentlich zugänglichen Datenbanken und bei der Übertragung von Nachrichten in den riesigen elektronischen Kommunikationsnetzen, die uns alle schon jetzt einbeziehen. Solche Datenbanken und Kommunikationsnetze sind höchst unsicher: Unbefugte können leicht „mithören“ oder sich unter dem Deckmantel einer falschen Identität in fremde Datenbestände einschleichen. Daher ist die Verwendung von Kryptosystemen zu ihrem Schutz unentbehrlich.

Wir wollen die Ausgangssituation bei der Verwendung eines Kryptosystems in der Sprache der Mathematik beschreiben. Der Absender B möchte eine Nachricht N, genannt *Klartext* an den Empfänger A senden; um sie schützen, verschlüsselt er sie und sendet das so erhaltene *Chiffrat* C an A. Der Empfänger A entschlüsselt C, um den Klartext N zurückzugewinnen. Das verwendete Kryptosystem besteht also im wesentlichen aus zwei Funktionen, der Verschlüsselung f, $f(N) = C$, und der Entschlüsselung f^{-1}, $f^{-1}(C) = N$. Wenn wir die Menge der Klartexte mit $\mathscr{N}$ bezeichnen,

$$\mathscr{N} = \{\, N : N \text{ Klartext} \,\},$$

und diejenige der Chiffrate mit $\mathscr{C}$,

$$\mathscr{C} = \{\, C : C \text{ Chiffrat} \,\},$$

geht es „nur“ um zwei Abbildungen

$$f : \mathscr{N} \to \mathscr{C}, \qquad f^{-1} : \mathscr{C} \to \mathscr{N}, \qquad f^{-1}(f(N)) = N \quad \text{für alle } N \in \mathscr{N}.$$

Natürlich spielen auch die „Alphabete“, in denen Klartexte und ihre Chiffrate notiert werden, und die technische Realisation der Übermittlung eine wichtige Rolle. In unserer Diskussion können wir sie aber vernachlässigen.

Bevor ein Kryptosystem benutzt werden kann, muß der Empfänger A dem Absender B den Schlüssel f, mit dem B Nachrichten an A verschlüsseln soll, mitteilen. Alle klassischen Kryptosysteme, haben die Eigenschaft, daß mit der Verschlüsselung f auch die Entschlüsselung f^{-1} bekannt ist oder zumindest sehr schnell ermittelt werden kann. Daher muß A den Schlüssel f auf einem sicheren Kanal mitteilen, und f muß gegenüber Dritten geheimgehalten werden. Der Zwang zur Geheimhaltung von f ist nur deshalb gegeben, weil die zum unbefugten Mithören notwendige Entschlüsselung f^{-1} sofort aus f gewonnen werden kann.

Die Notwendigkeit eines zweiten, sicheren Kanals für den Schlüsselaustausch macht die Verwendung eines klassischen Kryptosystems in Kommunikationsnetzen mit mehreren zehntausend Teilnehmern unmöglich. Den Ausweg bieten die vor ca. zwanzig Jahren erfundenen

Kryptosysteme mit öffentlichem Schlüssel,

in der englischsprachigen Literatur *public key cryptosystems* genannt. Sie beruhen auf der Verwendung von *Falltürfunktionen* für die Verschlüsselung. Falltüren zeichnen sich ja dadurch aus, daß man sie in einer Richtung sehr leicht, in der Gegenrichtung aber nur sehr schwer durchschreiten kann – es sei denn, man kennt den geheimen „Knopf“. Wir wollen also eine Funktion

$$f : \mathscr{N} \to \mathscr{C}$$

eine Falltürfunktion nennen, wenn sich für jedes $N \in \mathscr{N}$ das Chiffrat $f(N)$ und für jedes $C \in \mathscr{C}$ der Klartext $f^{-1}(\mathscr{C})$ sehr leicht berechnen lassen, es hingegen praktisch nicht möglich ist, trotz der Kenntnis von f die Entschlüsselung f^{-1} zu bestimmen. Wir können hier die vagen Begriffe „leicht“ und „praktisch nicht möglich“ nicht präzisieren. (Das dafür zuständige mathematische Gebiet ist die *Komplexitätstheorie.*)

Ein Kryptosystem mit öffentlichem Schlüssel funktioniert nun so:

(1) Jeder Teilnehmer A konstruiert eine Falltürfunktion f_A; die in die Konstruktion einfließende, nur ihm bekannte Zusatzinformation liefert f_A^{-1}.
(2) Er gibt f_A bekannt, hält aber die Entschlüsselung f_A^{-1} natürlich geheim.
(3) Will nun ein anderer Teilnehmer B eine Nachricht N an A senden, so verschlüsselt er sie mittels f_A und sendet das Chiffrat $C = f_A(N)$ an A.
(4) Wenn A das Chiffrat C empfängt, wendet er die nur ihm bekannte Entschlüsselung f_A^{-1} an und erhält $N = f_A^{-1}(f_A(N))$ zurück.

Da es nicht möglich ist, f_A^{-1} aus f_A zu bestimmen, kann f_A öffentlich bekannt sein und die Notwendigkeit eines sicheren Kanals zum Austausch der Schlüssel entfällt.

Man kann natürlich einwenden, daß es prinzipiell immer möglich ist, zu einer bekannten Funktion die Umkehrfunktion zu bestimmen. Wenn wir auch nicht mathematisch präzisiert haben, was „praktisch nicht möglich“ heißt: Die Ermittlung von f_A^{-1} muß einen Aufwand erfordern, der den durch das Brechen des Kryptosystems zu erzielenden Gewinn bei weitem überwiegt, oder eine so lange Zeit benötigen, daß nach ihrem Ablauf jegliches Interesse an der zu erlangenden Information erloschen ist.

Bei der Verwendung von Kryptosystemen mit öffentlichem Schlüssel muß indes einer Gefahr vorgebeugt werden: der Vorspiegelung einer falschen Identität. Niemand hindert ja einen dritten Teilnehmer C daran, sich für B auszugeben und unter Verwendung von f_A verschlüsselte Nachrichten an A zu senden, die A nicht von Mitteilungen unterscheiden kann, die wirklich von B stammen. Jedes Chiffrat muß also mit einer „Unterschrift“ versehen sein, die den Absender eindeutig identifiziert. Glücklicherweise lassen sich solche Unterschriften gerade in diesen Systemen exzellent realisieren; wir kommen darauf noch zurück.

Im folgenden beschreiben wir das RSA-Kryptosystem. Dafür brauchen wir noch eine zahlentheoretische Aussage:

Satz 5.4. *Genau dann ist $m \in \mathbb{Z}$, $m > 1$, Produkt paarweise verschiedener Primzahlen wenn für alle Zahlen a gilt:*

$$a^{\varphi(m)+1} \equiv a \ \ (m).$$

Den Beweis überlassen wir dem Leser zur Übung. Die im folgenden wichtige Implikation „$\Rightarrow$“ folgt leicht aus dem Satz von Fermat und dem Chinesischen Restsatz.

Das RSA-Kryptosystem arbeitet nach folgendem Schema:

(1) Jeder Teilnehmer A wählt zufällig zwei große Primzahlen p und q mit etwa 100 Dezimalstellen.

(2) Er bildet das Produkt $m_A = pq$ und $\varphi(m_A) = (p-1)(q-1)$.

(3) Er wählt eine zu $\varphi(m)$ teilerfremde Zahl e_A, den *Verschlüsselungsexponenten*, zwischen 1 und $\varphi(m)$.

(4) Er bestimmt als *Entschlüsselungsexponenten* die Zahl d_A, $1 \leq d_A \leq \varphi(m_A)$ mit $e_A d_A \equiv 1 \ \ (\varphi(m_A))$.

(5) Er gibt m_A und e_A öffentlich bekannt, hält aber d_A, p, q und $\varphi(m_A)$ geheim.

(6) Wenn B eine Nachricht an A senden will, so stellt er diese zunächst als eine Folge von Zahlen $n_1, \ldots, n_r$ zwischen 0 und $m_A - 1$ dar. Dann bestimmt er zu jeder dieser Zahlen n_i das Chiffrat

$$c_i \equiv n_i^{e_A} \ \ (m_A)$$

und sendet die Folge $c_1, \ldots, c_r$ an A. (Jede der Restklassen wird dabei natürlich durch c_i mit $0 \leq c_i \leq m_A - 1$ repräsentiert.)

(7) A empfängt $c_1, \ldots, c_r$ und bildet die Potenzen

$$c_i^{d_A} \equiv n_i \ \ (m_A)$$

und erhält die ursprüngliche Nachricht $n_1, \ldots, n_r$ zurück.

Nach Wahl von d_A gilt ja

$$e_A d_A \equiv 1 \ \ (\varphi(m_A))$$

also

$$e_A d_A = t\,\varphi(m_A) + 1$$

mit einer Zahl $t \geq 0$. Damit ist mit $m = m_A$

$$\begin{aligned} c_i^{d_A} \equiv n_i^{e_A d_A} \equiv n_i^{t\varphi(m)+1} \equiv (n_i^{\varphi(m)})^t n_i \equiv (n_i^{\varphi(m)})^{(t-1)} n_i^{\varphi(m)} n_i \\ \equiv n_i^{\varphi(m)(t-1)} n_i \equiv \cdots \equiv n_i \ \ (m) \end{aligned}$$

nach Satz 5.4.

Um ein RSA-Kryptosystem zu brechen, muß man die Zahl d_A bestimmen, und hierfür ist bisher kein effektiveres Verfahren bekannt als die Bestimmung der Primfaktoren p und q von m_A. Für Zahlen der genannten Größenordnung erfordert dies jedoch mit den besten bekannten Methoden und schnellsten verfügbaren Computern immer noch astronomische Rechenzeiten. Daß man sich andererseits relativ schnell große Primzahlen verschaffen kann, die man für das Verfahren ja braucht, haben die oben besprochenen Tests zumindest plausibel gemacht. Bei der Auswahl dieser Primzahlen hat man noch gewisse Vorsichtsmaßnahmen zu treffen, um m_A möglichst immun gegen die bekannten Faktorisierungsverfahren zu machen. Die Zahlen p und q sollten zum Beispiel nicht zu nahe benachbart sein, und $p-1$ und $q-1$ sollten beide einen großen Primfaktor enthalten. Den Grund für diese Forderung werden wir unten kennenlernen.

Schließlich wollen wir noch das Problem der „Unterschrift“ diskutieren. Es läßt sich zum Beispiel so lösen: Wenn B eine Nachricht an A senden will, so sendet er mit jedem Chiffrat c_i zusätzlich $c_i^{d_B} \ \ (m_B)$. Der Empfänger A kennt ja e_B, kann dann $c_i^{d_B e_B}$ bilden und erhält so auf einem zweiten Wege c_i. Da ein dritter Teilnehmer C die nur B bekannte Zahl d_B nicht kennt, kann er die zu c_i gehörende Unterschrift $c_i^{d_B}$ nicht erzeugen.

Ein Problem bei der Anwendung des RSA-Kryptosystems ist der im Vergleich zu herkömmlichen Verfahren vergleichsweise hohe Rechenaufwand, der verhindert, daß Chiffrierung und Dechiffrierung im Tempo der Nachrichtenübertragung möglich sind. Daher benutzt man das RSA-Kryptosystem häufig nur zum Austausch von Schlüsseln für weniger aufwendige Verfahren oder dann, wenn eine zeitliche Verzögerung der Nachrichtenübermittlung unwesentlich ist.

Die Zahlentheorie galt über Jahrhunderte als „unschuldige“, weil anwendungsfreie mathematische Disziplin. Sehr klar kommt dies in den Worten G. H. Hardy's zum Ausdruck (*A mathematician's apology*, (1940)):

> But here I must deal with a misconception. It is sometimes suggested that pure mathematicians glory in the uselessness of their work, and make it a boast that it has no practical applications. The imputation is usually based on an incautious saying attributed to Gauss, to the effect that, if mathematics is the queen of the sciences, then the theory of numbers is, because of its supreme uselessness, the queen of mathematics – I have never been able to find an exact quotation. I am sure that Gauss's saying (if indeed it be his) has been rather crudely misinterpreted. If the theory of numbers could be employed for any practical and obviously honourable purpose, if it could be turned directly to the furtherance of human happiness or the relief of human suffering, as physiology and even chemistry can, then surely neither Gauss nor any other mathematician would have been so foolish as to decry or regret such applications. But science works for evil as well as for good (and particularly, of course, in time of war); and both Gauss and lesser mathematicians may be justified in rejoicing that there is one science at any rate, and that their own, whose very remoteness from ordinary human activities should keep it gentle and clean.

Mit der Anwendung in der Kryptographie hat die Zahlentheorie ihre Unschuld verloren.

Im letzten Teil dieses Abschnitts besprechen wir zwei raffinierte, von J. M. Pollard erfundene Faktorisierungsverfahren. Sie zeigen, daß man die Probedivision bereits mit sehr geringem Aufwand erheblich übertreffen kann. Vorweg sei gesagt, daß man natürlich immer versuchen wird, die im jeweils benutzten Computerprogramm tabellierten Primzahlen (z.B. $< 10^6$) per Probedivision aus der zu faktorisierenden Zahl herauszuziehen.

Sei N eine zu zerlegende Zahl und p ein unbekannter Primfaktor – daß N zusammengesetzt ist, sollte man tunlichst mit einem Primzahltest nachweisen. Wir betrachten eine „Zufallsfolge“

$$y_0, y_1, \ldots, y_k, \ldots$$

in $\mathbb{Z}_p$. Da diese Menge endlich ist, muß irgendwann eine Wiederholung eintreten. Man kann mit Mitteln der Wahrscheinlichkeitsrechnung leicht zeigen, daß dies mit Wahrscheinlichkeit $1/2$ schon bei $k \approx 1.2\sqrt{p}$ der Fall ist.

Sei etwa $x_i \equiv x_j \ (p)$. Dann gilt $p \mid \mathrm{ggT}(x_i - x_j, N) \mid N$ und wir haben einen nichttrivialen Teiler von N gefunden, falls $x_i \neq x_j$.

Da man p nicht kennt, hat es den Anschein, als müßte man alle Zahlen $\mathrm{ggT}(x_i - x_j, N)$ ausrechnen, und wenn man erst bei $k \approx 1.2\sqrt{p}$ mit der ersten Wiederholung modulo p rechnen kann, muß man ca. p größte gemeinsame Teiler bestimmen, so daß überhaupt kein Vorteil gegenüber der Probedivision zu erkennen ist. Der Trick besteht darin, die Folge der x_i in einer ganz bestimmten Weise zu wählen.

Sei $F \in \mathbb{Z}[X]$ ein Polynom. Wir starten mit einem zufällig gewählten $x_0 \in \mathbb{Z}_N$ und betrachten dann die *Bahn* von x_0 unter der von F auf $\mathbb{Z}_N$ induzierten Abbildung:

$$x_0, x_1 = F(x_0), \ldots, x_n = F(x_{n-1}), \ldots$$

Weil F ein Polynom ist, induziert es ebenso eine Abbildung von $\mathbb{Z}_p$ in sich selbst und die Bahn der Restklasse y_0 entsteht gerade aus den Restklassen y_i der x_i. Da $\mathbb{Z}_p$ endlich ist, können in der Bahn von y_0 nur endlich viele Elemente vorkommen. Sei n der kleinste Index, für den $y_n = y_j$ mit einem $j > n$ ist, und dann k wiederum der kleinste Index, für den $y_n = y_k$. Dann besteht die Bahn offensichtlich aus der *Vorperiode* $y_0, \ldots, y_{n-1}$ und einem Zykel der Länge $\pi = k - n$, der bei y_i beginnt und periodisch mit der *Periode* π durchlaufen wird. Die Form der Bahn erinnert an den griechischen Buchstaben ρ und daher heißt das gerade diskutierte Faktorisierungsverfahren *Pollardsche ρ-Methode*.

Die entscheidende Aussage für das Aufspüren einer Wiederholung modulo p in der Bahn von x_0 ist

Satz 5.5. *Mit den eingeführten Bezeichnungen sei m das kleinste Vielfache $\geq n$ von π. Dann ist $y_m = y_{2m}$ (und m ist der kleinste unter den Indizes i mit $y_i = y_{2i}$).*

Zu beweisen ist wirklich nur die in Klammern gesetzte Aussage, und dies sei dem Leser zur Übung überlassen. Man kann also Wiederholungen modulo p einfach durch Vergleich von x_i und x_{2i} aufspüren. Wenn die Periode sehr groß ist, müssen sehr viele x_i gespeichert werden. Mit einem Trick von R. W. Floyd spart man aber Speicherplatz auf Kosten von Rechenzeit, indem man gleichzeitig mit zwei Folgen rechnet:

$$z_0 = x_0 \quad \text{und} \quad x_i = F(x_{i-1}),\ z_i = F(F(z_{i-1})) \quad \text{für } i > 0.$$

Einzig festzulegen ist noch F; in der Praxis haben sich die Polynome $F = X^2 + a$ mit $a \not\equiv 0, -2 \ (N)$ gut bewährt. Insgesamt geht man also so vor:

(1) Man wählt zufällig eine Zahl x_0, $0 \leq x_0 < N$, und setzt $z_0 = x_0$.

(2) Für $i > 0$ setzt man der Reihe nach

$$x_i = F(x_{i-1}),\ z_i = F(F(z_{i-1})),\ d_i = \mathrm{ggT}(x_i - z_i, N)$$

und bricht ab, sobald man einen echten Teiler d_i von N gefunden hat.

Wenn das Verfahren nicht zum Erfolg führt, liegt einer der folgenden Gründe vor: Erstens können natürlich alle Primfaktoren von N so groß sein, daß man innerhalb der Laufzeit keine Chance hat, einen von ihnen zu finden, und zweitens kann der Fall $x_i = z_i$ eintreten. Im zweiten Fall kann man zumindest versuchen, mit einem anderen Polynom F zum Ziel zu kommen.

Da die zu erwartende Rechenzeit proportional zu $\sqrt{p}$ ist, kann man mit der ρ-Methode im Vergleich zur Probedivision Primfaktoren etwa der doppelten Stellenzahl finden. Man kann Rechenzeit sparen, indem man nicht bei jedem Schritt den ggT berechnet, sondern immer eine gewisse Anzahl Differenzen modulo N aufmultipliziert, und erst den „akkumulierten“ ggT bestimmt. Ferner läßt sich der Floydsche Trick weiter verfeinern.

Das zweite Verfahren ist die Pollardsche $(p-1)$-*Methode*. Mit ihr kann man sehr gut Primfaktoren p finden, für die $p-1$ selbst nur „kleine“ Primpotenzfaktoren besitzt. Dabei sei wie oben p ein Primfaktor der zusammengesetzten Zahl N. Wir wählen zunächst eine Schranke B und wählen für jede Primzahl $q \leq B$ die Zahl $\mu(q, B)$ als den größten Exponenten i, für den $q^i \leq B$. Sei dann

$$k = \prod_q q^{\mu(q,B)}.$$

Wenn dann p eine Primzahl ist, für die die Primpotenzfaktoren von $p-1$ sämtlich $\leq B$ sind, so gilt

$$x^k \equiv 1 \ (p)$$

nach dem Fermatschen Satz, denn $p-1$ teilt k. Für die praktische Durchführung bildet man k nicht als ein einziges Produkt, sondern setzt es aus Teilprodukten k_i zusammen, die man zu einer aufsteigenden Folge

$$1 = B_0 < \cdots < B_m = B$$

folgendermaßen berechnet: Man setzt $k_0 = 1$, und für $i > 0$ ist k_i ist das Produkt aller Primzahlen q mit $B_{i-1} < q \leq B_i$ und aller *ganzen* Zahlen n mit $\sqrt{B_{i-1}} < n \leq \sqrt{B_i}$. Das Produkt der k_i ist dann ein Vielfaches der oben definierten Zahl k.

Ausgehend von einer zufällig gewählten Zahl x_0, $0 < x_0 < N$, setzt man dann sukzessive $x_i \equiv x_{i-1}^{k_i} \ (N)$ und prüft, ob $\mathrm{ggT}(x_i - 1, N)$ ein nichttrivialer Teiler von N ist.

Das $(p-1)$-Verfahren ist mit Sicherheit erfolgreich, falls es einen Primfaktor p von N gibt, der die eingangs gemachte Annahme über die Primfaktoren von $p-1$ erfüllt, aber auch schon dann, wenn man glücklicherweise x_0 so trifft, daß seine Ordnung in $\mathbb{Z}_p^*$ nur Primfaktoren $\leq B$ hat. Mit Hilfe des Primzahlsatzes kann man leicht zeigen, daß der Rechenaufwand beim $(p-1)$-Verfahren etwa proportional zu B ist. Wenn man es sehr oft anwenden will, kann man die von N unabhängigen Zahlen k_i natürlich abspeichern und erhebliche Zeit sparen. Es gibt

zum $(p-1)$-Verfahren noch eine „big prime"-Variante, die auch dann noch erfolgreich sein kann, falls $p-1$ außer kleinen Primfaktoren nur noch einen weiteren mit der Vielfachheit 1 enthält.

Die Grundidee des $(p-1)$-Verfahrens wird auch bei anderen, neueren Faktorisierungsmethoden verwendet: Man rechnet in einer „p nahestehenden" endlichen abelschen Gruppe; beim $(p-1)$-Verfahren ist dies $\mathbb{Z}_p^*$.

Übungen.

5.6. Sei p eine Primzahl. Zeige, daß die Mersennesche Zahl M_p quasiprim zur Basis 2 ist.

5.7. Sei $a \in \mathbb{Z}$, $a > 1$, und p eine der ungeraden Primzahlen, die $a(a^2-1)$ nicht teilen. Zeige:

$$m = \frac{a^p - 1}{a-1}\,\frac{a^p + 1}{a+1}$$

ist quasiprim zur Basis a, aber nicht prim. Insbesondere gibt es unendlich viele zusammengesetzte Zahlen, die quasiprim zur Basis a sind.

Anleitung: Zeige zunächst, daß $2p \mid m-1$.

5.8. (a) Zeige: Genau dann ist m eine Carmichael-Zahl, wenn m das Produkt paarweise verschiedener Primzahlen $p_1, \ldots, p_n$ und $p_i - 1 \mid m-1$ für $i = 1, \ldots, n$.

(b) Zeige: Eine Carmichael-Zahl hat mindestens 3 Primfaktoren.

(b) Zeige: Wenn $p > 3$ prim, derart daß auch $2p-1$ und $3p-2$ prim sind, so ist $m = p(2p-1)(3p-2)$ Carmichael-Zahl.

(c) Finde mittels (b) zwei weitere Carmichael-Zahlen.

5.9. Beweise in Verallgemeinerung von Satz 5.1 folgende Aussage: Sei $m \in \mathbb{Z}$, $m \geq 2$. Es gelte $m-1 = r \cdot s$ und zu jedem Primfaktor q von r existiere ein a mit

$$a^{m-1} \equiv 1 \ \ (m) \quad \text{und} \quad \ggT(a^{(m-1)/q} - 1, m) = 1.$$

Dann gilt: Jeder Primteiler p von m erfüllt $p \equiv 1 \ \ (r)$; wenn $r \geq s$, ist m prim.

Diese Aussage zeigt, wie man mittels partieller Faktorisierung von $m-1$ die Menge der potentiellen Teiler von m erheblich einschränken oder sogar zeigen kann, daß m prim ist.

ABSCHNITT 6

Quadratische Reste und quadratische Reziprozität

In diesem Abschnitt beschäftigen wir uns mit quadratischen Kongruenzen

$$a_2x^2 + a_1x + a_0 \equiv 0 \ (m).$$

Diese lassen sich sehr einfach auf Kongruenzen der Form

$$x^2 \equiv a \ (m')$$

reduzieren:

$$\begin{aligned} a_2x^2 + a_1x + a_0 \equiv 0 \ (m) &\iff 4a_2^2x^2 + 4a_2a_1x + 4a_2a_0 \equiv 0 \ (4a_2m) \\ &\iff (2a_2x + a_1)^2 \equiv a_1^2 - 4a_2a_0 \ (4a_2m) \end{aligned}$$

Insofern genügt es, den Spezialfall

$$x^2 \equiv a \ (m)$$

zu betrachten. (Natürlich muß man nach Bestimmung der Lösungen von $y^2 \equiv a_1 - 4a_2a_0 \ (4a_2m)$ noch die lineare Kongruenz $2a_2x + a_1 \equiv y \ (4a_2m)$ lösen.)

Definition. Die Zahl a heißt *quadratischer Rest* bzw. *quadratischer Nichtrest* modulo m je nachdem, ob die Kongruenz

$$x^2 \equiv a \ (m)$$

eine Lösung besitzt oder nicht.

Mit anderen Worten: a ist quadratischer Rest genau dann, wenn $\bar{a}$ ein Quadrat in $\mathbb{Z}_m$ ist.

Mittels des chinesischen Restsatzes können wir das Problem, zu entscheiden ob a quadratischer Rest modulo m ist, auf Primzahlpotenzen m reduzieren: Sei $m = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$ die normierte Primfaktorzerlegung von m. Unmittelbar aus 3.9 folgt: a ist quadratischer Rest modulo m genau dann, wenn a quadratischer Rest modulo $p_i^{\alpha_i}$ ist für $i = 1, \ldots, r$. Der folgende Satz reduziert das Problem dann letztendlich auf den Fall einer Primzahl.

Satz 6.1. *Sei zunächst p eine ungerade Primzahl.*

(a) *Falls $p \nmid a$, so ist a genau dann quadratischer Rest modulo p^α, wenn a quadratischer Rest modulo p ist.*

(b) *Falls $a = p^k b$ mit $p \nmid b$, $k < \alpha$, so ist a genau dann quadratischer Rest modulo p^α, wenn k gerade und b quadratischer Rest modulo p^α ist.*

Sei nun $p = 2$.

(c) *Genau dann ist* a *quadratischer Rest modulo* 4, *wenn* $a \equiv 0, 1 \ (4)$.

(d) *Wenn* $\alpha > 2$ *und* $2 \nmid a$, *so ist* a *genau dann quadratischer Rest modulo* 2^α, *wenn* $a \equiv 1 \ (8)$.

(e) *Wenn* $2 \mid a$, *so gilt* (b) *entsprechend.*

Wir übergehen den Beweis, den wir (zumindest teilweise) einer Übungsaufgabe überlassen (vgl. auch [**Gaus**], Art. 101ff.).

Ein nützliches Hilfsmittel bei der Behandlung der Kongruenz $x^2 \equiv a \ (p)$, p ungerade Primzahl (der Fall $p = 2$ ist trivial), ist das Legendre-Symbol: Für $a \in \mathbb{Z}$, $p \nmid a$, sei

$$\left(\frac{a}{p}\right) := \begin{cases} 1 & \text{wenn } a \text{ quadratischer Rest,} \\ -1 & \text{wenn } a \text{ quadratischer Nichtrest.} \end{cases}$$

Offensichtlich ist $(a/p) = (b/p)$, wenn $a \equiv b \ (p)$. Daher kann man auch $\mathbb{Z}_p^*$ als natürlichen Definitionsbereich des Legendresymbols ansehen. Im folgenden kann a in (a/p) also sowohl eine ganze Zahl, als auch eine Restklasse modulo p bedeuten. Ferner bezeichnet p stets eine ungerade Primzahl.

Die unmittelbar klaren Eigenschaften des Legendre-Symbols beschreibt

Satz 6.2.

(a) *Die Abbildung* $\varphi : \mathbb{Z}_p^* \to \mathbb{Z}_p^*$, $\varphi(x) = x^2$, *ist ein Homomorphismus mit Kern* $\{1, -1\}$.

(b) $\operatorname{Bild}\varphi$, *die Menge der Quadrate in* $\mathbb{Z}_p^*$, *ist eine Untergruppe vom Index* 2; *es gibt also genau so viel Quadrate wie Nichtquadrate in* $\mathbb{Z}_p^*$.

(c) *Die Abbildung* $a \mapsto (a/p)$ *von* $\mathbb{Z}_p^*$ *in* $\{+1, -1\}$ *ist ein Gruppenhomomorphismus; es gilt*

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

Das Produkt ab *ist genau dann ein Quadrat, wenn* a *und* b *beide Quadrate oder beide Nichtquadrate sind.*

Beweis. (a) ist trivial, (b) folgt unmittelbar aus (a) und (c) ergibt sich aus der Identifizierung

$$\mathbb{Z}_p^* / \operatorname{Bild}\varphi \cong \mathbb{Z}_2 \cong \{+1, -1\} \quad (= \mathbb{Z}^*) \qquad \square$$

Ein erstes Hilfsmittel zur Bestimmung von (a/p) ist das *Eulersche Kriterium*.

Satz 6.3. *Für* $a \in \mathbb{Z}$, $p \nmid a$, *ist* $a^{(p-1)/2} \equiv (a/p) \ (p)$.

Beweis. Nach dem Satz von Fermat ist $(a^{(p-1)/2})^2 = a^{p-1} \equiv 1 \ (p)$, also

$$a^{(p-1)/2} \equiv 1 \ (p) \quad \text{oder} \quad a^{(p-1)/2} \equiv -1 \ (p).$$

Falls $a = b^2$, so $a^{(p-1)/2} = b^{(p-1)/2} \equiv 1 \ (p)$. Sei nun umgekehrt $a^{(p-1)/2} \equiv 1 \ (p)$. Nach Satz 4.3 ist $\mathbb{Z}_p^*$ zyklisch mit einem erzeugenden Element c. Sei $k \in \mathbb{Z}$ so gewählt, daß $\bar{a} = c^k$. Dann ist

$$\bar{a}^{(p-1)/2} = c^{k(p-1)/2} = 1.$$

Da c die Ordnung $p-1$ hat, muß $k(p-1)/2$ ein Vielfaches von $p-1$ und k somit gerade sein: $\bar{a} = (c^{k/2})^2$ ist ein Quadrat in $\mathbb{Z}_p^*$. □

Mit Hilfe des Eulerschen Kriteriums können wir sofort den „quadratischen Charakter“ von -1 bestimmen. Durch Einsetzen von $a = -1$ erhalten wir

Satz 6.4.

$$\left(\frac{-1}{p}\right) = (-1)^{(p-1)} = \begin{cases} 1 & \textit{wenn } p \equiv 1 \ (4), \\ -1 & \textit{wenn } p \equiv 3 \ (4). \end{cases}$$

In einer Übungsaufgabe wurde gezeigt, daß es unendlich viele Primzahlen der Form $4k+3$ gibt. Mit Hilfe von 6.4 können wir zeigen, daß es auch unendlich viele Primzahlen der Form $4k+1$ gibt. Seien $p_1, \ldots, p_m$ Primzahlen der Form $4k+1$ und p ein Primteiler von

$$(2p_1 \ldots p_m)^2 + 1.$$

Dann ist offensichtlich $(-1/p) = 1$, also p eine von $p_1, \ldots, p_m$ verschiedene Primzahl der Form $4k+1$.

Das nächste Kriterium stammt von Gauß. Es ist ein sehr wirksames Hilfsmittel, wie wir noch sehen werden. Zu gegebenem p betrachten wir die Menge

$$S = \left\{-\frac{p-1}{2}, -\frac{p-1}{2}+1, \ldots, -1, 1, \ldots, \frac{p-1}{2}-1, \frac{p-1}{2}\right\}.$$

Wir nennen S die Menge der *Minimalreste modulo* p. Zu jedem $a \in \mathbb{Z}$, $p \nmid a$, gibt es genau ein $s \in S$ mit $a \equiv s \ (p)$.

Sei $a \in \mathbb{Z}$ gegeben, $p \nmid a$. Für $n = 1, \ldots, (p-1)/2$ seien ε_n und s_n definiert durch

$$na \equiv \varepsilon_n s_n \ (p), \quad s_n \in S, s_n > 0 \quad \text{und} \quad \varepsilon_n \in \{1, -1\}.$$

Dann gilt:

Satz 6.5 (Gaußsches Lemma)**.**

$$\left(\frac{a}{p}\right) = \varepsilon_1 \ldots \varepsilon_t, \quad t = \frac{p-1}{2}.$$

Beweis. Wir rechnen der bequemeren Schreibweise halber direkt in $\mathbb{Z}_p^*$. Dann ist $S = \mathbb{Z}_p^*$. Da $\mathbb{Z}_p^*$ eine Gruppe ist, ist die Multiplikation mit $a \in \mathbb{Z}_p^*$ eine Bijektion:

$$\begin{aligned} a \cdot S &= \Big\{a\Big(-\frac{p-1}{2}\Big), \ldots, a\frac{p-1}{2}\Big\} \\ &= \Big\{-\varepsilon_t s_t, \ldots, -\varepsilon_1 s_1, \varepsilon_1 s_1, \ldots, \varepsilon_t s_t\Big\}. \end{aligned}$$

Deshalb muß $\{s_1, \ldots, s_t\} = \{1, \ldots, t\}$ sein und wir erhalten modulo p

$$t! \cdot a^t = \prod_{n=1}^{t} n \cdot \prod_{n=1}^{t} a = \prod_{n=1}^{t} an \equiv \prod_{n=1}^{t} \varepsilon_n s_n \equiv \gamma(a) \cdot \prod_{n=1}^{t} s_n \equiv \gamma(a) \cdot t!$$

mit $\gamma(a) = \varepsilon_1 \ldots \varepsilon_t$. Da $t! \not\equiv 0 \ (p)$, können wir $t!$ kürzen, so daß $a^t = a^{(p-1)/2} = \gamma(a)$. Daraus folgt die Behauptung mit dem Eulerschen Kriterium. □

Mit Hilfe des Gaußschen Lemmas bestimmen wir den quadratischen Charakter von 2:

Satz 6.6.

$$\Big(\frac{2}{p}\Big) = (-1)^{(p^2-1)/8} = \begin{cases} 1, & \text{wenn } p \equiv \pm 1 \ (8), \\ -1, & \text{wenn } p \equiv \pm 3 \ (8). \end{cases}$$

Beweis. Für $n = 1, \ldots, (p-1)/2$ ist der Repräsentant $s \in S$ von $2n$ modulo p genau dann negativ, wenn $2n > (p-1)/2$. Also müssen wir die Anzahl m der ganzen Zahlen n mit $\frac{p-1}{4} < n \leq \frac{p-1}{2}$ bestimmen. Im Fall $p = 8k+1$ ist $p - 1/4 \in \mathbb{Z}$ und

$$m = \Big|\Big\{\frac{p-1}{4} + 1, \ldots, \frac{p-1}{2}\Big\}\Big| = \frac{p-1}{2} - \frac{p-1}{4} = \frac{p-1}{4}$$

ist gerade. Im Fall $p = 8k+7$ ist $p - 1/4 \notin \mathbb{Z}$ und

$$m = \Big|\Big\{\frac{p-1}{4} + \frac{1}{2}, \ldots, \frac{p-1}{2}\Big\}\Big| = \frac{p-1}{2} - \frac{p-1}{4} + \frac{1}{2} = \frac{p+1}{4}$$

ist ebenfalls gerade. Die anderen beiden Fälle überlassen wir dem Leser. □

Wenn man die Primfaktorzerlegung $a = q_1 \ldots q_r$ von a kennt, kann man (a/p) mittels 6.2(c) ausrechnen:

$$\Big(\frac{a}{p}\Big) = \Big(\frac{q_1}{p}\Big) \ldots \Big(\frac{q_r}{p}\Big).$$

Den Wert von $(2/p)$ kennen wir aus 6.6, also bleiben noch die Legendre-Symbole (q/p) zu bestimmen, bei denen q eine ungerade Primzahl ist.

Wir erhalten aber nicht eine Aussage wie etwa 6.6, die explizit sagt, wann $(q/p) = 1$ ist. Stattdessen erhalten wir eine Aussage, die (p/q) und (q/p) auf eine unerwartete und höchst elegante Weise in Beziehung setzt:

Satz 6.7 (Quadratisches Reziprozitätsgesetz)**.** *Für ungerade Primzahlen p und q gilt*

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

in äquivalenter Formulierung:

$$\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right) \quad \text{wenn } p \equiv 1 \ (4) \text{ oder } q \equiv 1 \ (4),$$
$$\left(\frac{p}{q}\right) = -\left(\frac{q}{p}\right) \quad \text{wenn } p \equiv 3 \ (4) \text{ und } q \equiv 3 \ (4).$$

Die Aussagen 6.6 und 6.4 heißen auch *Ergänzungssätze zum quadratischen Reziprozitätsgesetz.*

Satz 6.7 wurde erstmals von Gauß 1796 bewiesen. Entdeckt hatte das Gesetz schon Euler (ca. 1745), und Legendre hatte gewisse Spezialfälle bewiesen. Gauß war äußerst stolz auf seine Leistung. Er hat insgesamt acht verschiedene Beweise geliefert, davon sechs veröffentlicht und zwei in seinem Nachlaß. Satz 6.5, das Gaußsche Lemma, ist die Grundlage des dritten von Gauß veröffentlichten Beweises, von dem wir eine vereinfachte Fassung wiedergeben, die auf Kronecker zurückgeht. Wir leiten zunächst einen Hilfssatz her, in dem wir folgende Bezeichnung benutzen: Für $a \in \mathbb{Z}$ sei

$$S(a,p) = \sum_{n=1}^{t} \left[\frac{an}{p}\right], \qquad t = \frac{p-1}{2}.$$

Satz 6.8. *Sei $a \in \mathbb{Z}$, $p \nmid a$, $t := (p-1)/2$. Dann gilt $(a/p) = (-1)^{S(2a,p)}$. Für ungerades a ist*

$$\left(\frac{a}{p}\right) = (-1)^{S(a,p)}.$$

Beweis. Seien ε_n, s_n wie in 6.5 definiert. Für die erste Behauptung genügt es,

$$\varepsilon_n = (-1)^{\left[\frac{2an}{p}\right]}.$$

zu zeigen. Es gilt

$$\left[\frac{2an}{p}\right] = \left[2\left[\frac{an}{p}\right] + 2\left(\frac{an}{p} - \left[\frac{an}{p}\right]\right)\right] = 2\left[\frac{an}{p}\right] + \left[2\left(\frac{an}{p} - \left[\frac{an}{p}\right]\right)\right]$$

denn $[x+y] = x + [y]$ für $x \in \mathbb{Z}$. Also

$$\left[\frac{2an}{p}\right] \text{ gerade} \iff \left[2\left(\frac{an}{p} - \left[\frac{an}{p}\right]\right)\right] = 0 \iff \frac{an}{p} - \left[\frac{an}{p}\right] < \frac{1}{2}$$
$$\iff an - p\left[\frac{an}{p}\right] < \frac{1}{2}p.$$

Nun ist $an - p[\frac{an}{p}]$ der Rest von an bei Division durch p und deshalb ist

$$\Big[\frac{2an}{p}\Big] \text{ gerade } \iff \varepsilon_n = 1.$$

Sei nun a ungerade. Dann ist $a + p$ gerade. Folglich

$$\Big(\frac{2}{p}\Big)\Big(\frac{a}{p}\Big) = \Big(\frac{2a}{p}\Big) = \Big(\frac{2(a+p)}{p}\Big) = \left(\frac{4\frac{a+p}{2}}{p}\right) = \Big(\frac{4}{p}\Big)\Big(\frac{(a+p)/2}{p}\Big)$$
$$= \Big(\frac{(a+p)/2}{p}\Big) = (-1)^{S(a+p,p)}$$

nach der schon bewiesenen ersten Behauptung. Es gilt

$$S(a+p,p) = \sum_{n=1}^{t}\Big[\frac{an}{p}\Big] + n = S(a,p) + \frac{t(t+1)}{2}.$$

Hieraus ergibt sich noch einmal $(2/p)$, denn bei $a = 1$ ist $[an/p] = 0$ für $n = 1,\dots,t$ und $(a/p) = 1$, also

$$\Big(\frac{2}{p}\Big) = (-1)^{t(t+1)/2} = (-1)^{(p^2-1)/8}.$$

Für beliebiges ungerades a ist nach Kürzung durch $(2/p)$

$$\Big(\frac{a}{p}\Big) = (-1)^{S(a,p)}. \qquad \square$$

Beweis des quadratischen Reziprozitätsgesetzes. Sei $t := \frac{p-1}{2}$, $u := \frac{q-1}{2}$. Nach 6.8 ist

$$\Big(\frac{p}{q}\Big)\Big(\frac{q}{p}\Big) = (-1)^{S(p,q)+S(q,p)}.$$

Daher genügt es zu zeigen:

$$tu = S(p,q) + S(q,p).$$

Sei

$$Q := \{qn - pm : 1 \le n \le t,\ 1 \le m \le u\}.$$

Es gilt $|Q| = t \cdot u$ und $0 \notin Q$. Sei $N := \{z \in Q : z < 0\}$, $P := \{z \in Q : z > 0\}$. Wir bestimmen die Anzahl der Elemente von P:

$$qn - pm > 0 \iff \frac{qn}{p} > m \iff 1 \le m \le \Big[\frac{qn}{p}\Big],$$

(da $\frac{qn}{p} \notin \mathbb{Z}$). Also gibt es für festes n genau $[\frac{qn}{p}]$ Zahlen $qn - pm > 0$ in Q, und somit ist

$$|P| = S(q,p).$$

Genauso zeigt man: $|N| = S(p,q)$. Da $|N| + |P| = |Q|$, folgt die Behauptung. $\square$

Neben seiner außergewöhnlichen theoretischen Bedeutung ist das Quadratische Reziprozitätsgesetz ein nützliches Hilfsmittel bei der konkreten Rechnung. Wir betrachten ein (von Gauß gegebenes)

Beispiel. Ist 453 quadratischer Rest modulo 1236? Die Primfaktorzerlegung liefert $1236 = 3 \cdot 4 \cdot 103$. Offensichtlich ist 453 quadratischer Rest modulo 3 und modulo 4. Da $453 \equiv 41 \ (103)$, bleibt $(41/103)$ zu bestimmen. Es gilt

$$\begin{aligned}\Big(\frac{41}{103}\Big) = \Big(\frac{103}{41}\Big) = \Big(\frac{21}{41}\Big) = \Big(\frac{3}{41}\Big) \cdot \Big(\frac{7}{41}\Big) = \Big(\frac{41}{3}\Big)\Big(\frac{41}{7}\Big) \\ = \Big(\frac{2}{3}\Big)\Big(\frac{-1}{7}\Big) = (-1) \cdot (-1) = 1\end{aligned}$$

Insgesamt: 453 ist quadratischer Rest modulo 1236.

Im folgenden Satz betrachten wir das Legendre-Symbol (a/p) als Funktion von p. Eine der überraschenden Folgerungen des Reziprozitätsgesetzes ist, daß (a/p) in einem gewissen Sinn auch als Funktion von p multiplikativ ist. Dazu erweitern wir das Legendre-Symbol durch $(a/p) := 0$, falls $p \mid a$.

Satz 6.9. *Sei $a \in \mathbb{N}_+$ und p, q, p' ungerade Primzahlen.*

(a) *Wenn $p \equiv \pm q \ (4a)$, so $(a/p) = (a/q)$.*
(b) *Wenn $pp' \equiv q \ (4a)$, so $(a/p)(a/p') = (a/q)$.*

Beweis. (a) Wegen 6.2(c) genügt es, (a) für Primzahlen a zu beweisen. Für $a = 2$ folgt (a) dann aus 6.6. Sei nun a eine ungerade Primzahl.

Wenn $p \equiv q \ (4a)$, so $p \equiv q \ (a)$ und $(p/a) = (q/a)$. Der Fall $p = a$ ist trivial. Sei $p \neq a$. Mit dem quadratischen Reziprozitätsgesetz folgt

$$\Big(\frac{a}{p}\Big) = (-1)^{\frac{a-1}{2}\frac{p-1}{2}}\Big(\frac{p}{a}\Big) = (-1)^{\frac{a-1}{2}\frac{p-1}{2}}\Big(\frac{q}{a}\Big) = (-1)^{\frac{a-1}{2}\frac{q-1}{2}}\Big(\frac{q}{a}\Big) = \Big(\frac{a}{q}\Big)$$

denn $p \equiv q \ (4a)$ impliziert $p \equiv q \ (4)$ und damit $\frac{q-1}{2} \equiv \frac{p-1}{2} \ (2)$.

Der Fall $p \equiv -q \ (4a)$ erledigt sich analog: Dann ist

$$\Big(\frac{p}{a}\Big) = \Big(\frac{-1}{a}\Big) \cdot \Big(\frac{q}{a}\Big) = (-1)^{\frac{a-1}{2}}\Big(\frac{q}{a}\Big)$$

und

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

Auch Teil (b) braucht natürlich nur für Primzahlen a bewiesen zu werden. Für $a = 2$ ist die Abbildung $\mathbb{Z}_8^* \to \{+1, -1\}, \ \pm\bar{1} \to 1, \ \pm 5 \to -1$, ein Homomorphismus (nämlich der natürliche $\mathbb{Z}_8^* \to \mathbb{Z}_8^*/\{-\bar{1}, \bar{1}\}$); daher folgt in diesem Fall die Behauptung aus 6.6.

Sei a eine ungerade Primzahl. Wenn $q = a$, so $p = a$ oder $p' = a$. Sei nun $p, p' \neq a$. Dann ist auch $q \neq a$, und wir haben

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

denn $\frac{p+p'-2}{2} \equiv \frac{q-1}{2}$ (2), wenn $pp' \equiv q$ (4). □

Unter Zuhilfenahme von 6.4 kann man aus Teil (a) von 6.9 umgekehrt das quadratische Reziprozitätsgesetz ableiten (vgl. etwa [**IrRo**], p. 61). Obwohl die Eleganz des quadratischen Reziprozitätsgesetzes in der Formulierung von 6.9(a) nicht zum Ausdruck kommt, ist diese Fassung im Sinne weitestreichender Verallgemeinerungen die „richtige".

Der zweite Teil von Satz 6.9 legt es nahe, den Definitionsbereich des Legendre-Symbols zu erweitern, zumal wenn man weiß, daß jede teilerfremde Restklasse modulo $4a$ eine Primzahl enthält, was wir freilich nicht bewiesen haben: Es sollte $(a/p)(a/p') = (a/pp')$ gelten. Dementsprechend definieren wir das *Jacobi-Symbol*.

Definition. Sei $a \in \mathbb{Z}$ und b eine positive ungerade Zahl, $b = p_1 \dots p_r$, p_i prim für $i = 1, \dots, r$. Dann sei

$$\left(\frac{a}{b}\right) := \left(\frac{a}{p_1}\right) \cdots \left(\frac{a}{p_r}\right).$$

Warnung: (a/b) kann 1 sein, ohne daß a quadratischer Rest modulo b ist, denn a ist quadratischer Rest modulo b genau dann, wenn a quadratischer Rest modulo p_i für alle $i = 1, \dots, r$ (vgl. 6.1).

Unmittelbar aus der Definition des Jacobi-Symbols ergibt sich:

Satz 6.10.

(a) $(a_1/b) = (a_2/b)$, *wenn* $a_1 \equiv a_2$ (b),
(b) $(a_1a_2/b) = (a_1/b)(a_2/b)$,
(c) $(a/b_1b_2) = (a/b_1)(a/b_2)$.

Für das Jacobi-Symbol gelten zu 6.4, 6.6, 6.7 und 6.9 analoge Aussagen. Sie lassen sich ohne große Mühe aus diesen Sätzen herleiten. Wir überlassen den Beweis (teilweise) einer Übungsaufgabe.

Satz 6.11. *Seien a, b positive ungerade Zahlen. Dann gilt:*

(a) $(-1/b) = (-1)^{(b-1)/2}$,
(b) $(2/b) = (-1)^{(b^2-1)/8}$,
(c) $(a/b)(b/a) = (-1)^{(a-1)(b-1)/4}$, *wenn a und b teilerfremd.*

(d) *Sei c eine beliebige ganze Zahl. Wenn dann $a \equiv b\ (4c)$, so $(c/a) = (c/b)$.*

Der Modul $4c$ in 6.11 bzw. $4a$ in 6.9 kann häufig durch einen kleineren ersetzt werden, vgl. etwa [**Hass**], § 9.

Übungen.

6.12. Bestimme alle Lösungen der quadratischen Kongruenzen

(a) $3x^2 + 4x - 11 \equiv 0\ (36)$ (b) $7x^2 + 6x + 8 \equiv 0\ (36)$.

Natürlich kann man dies sehr einfach per Computerprogramm lösen. Es ist aber instruktiver das zu Beginn dieses Abschnitts entwickelte Lösungsverfahren von Hand durchzuführen.

6.13. Beweise Satz 6.1(a) und (d).

6.14. Bestimme $\left(\frac{547}{3389}\right)$ und $\left(\frac{1063}{1999}\right)$.

(a) Zeige: $\left(\frac{10}{p}\right) = 1 \iff p \equiv 1, 3, 9, 13, 27, 31, 37, 39\ (40)$.

(b) Bestimme in ähnlicher Weise die Primzahlen p, für die $\left(\frac{15}{p}\right) = 1$.

6.15. Sei p eine ungerade Primzahl, für die $q = 2p + 1$ ebenfalls prim ist. Zeige:

(a) Falls $p \equiv 1\ (4)$, so ist 2 Primitivwurzel modulo q.

(b) Falls $p \equiv 3\ (4)$, so ist 2 keine Primitivwurzel modulo q.

Zeige ferner:

(c) Wenn $p \equiv 1\ (4)$, so teilt q nicht die Mersennesche Zahl $2^p - 1$.

(d) Wenn $p \equiv 3\ (4)$, so teilt q die Mersennesche Zahl $2^p - 1$.

Finde mittels (d) Mersennesche Zahlen, die nicht prim sind.

6.16. Sei $F_m = 2^{2^m} + 1$ die m-te Fermat-Zahl.

(a) Zeige: F_m prim $\iff 3^{(F_m-1)/2} \equiv -1\ (F_m)$.

(b) Beweise die (dort bereits erwähnte) Verschärfung von Aufgabe 3.16: Für $m \geq 2$ haben alle Primteiler p von F_m die Form $p = n \cdot 2^{m+2} + 1$.

Hinweis: $\left(\frac{2}{p}\right) = 1$!

6.17. Zeige: Es gibt unendlich viele Primzahlen $p \equiv 3\ (8)$ und unendlich viele Primzahlen $p \equiv 5\ (8)$.

6.18. Beweise Satz 5.11.

6.19. Eine ungerade Zahl q heißt *quadratisch quasiprim zur Basis* a, wenn

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

(a) Erkläre diese Bezeichnung und zeige: Ist q stark quasiprim zur Basis a, so ist q auch quadratisch quasiprim zur Basis a. (Nicht ganz einfach.)

Somit ist der Rabin-Test besser als der Solovay-Strassen-Test, der prüft, ob q quadratisch quasiprim zur Basis a ist. Es gilt aber: Es gibt keine „quadratisch-Carmichael-Zahlen“:

(b) Zeige: Ist q quadratisch quasiprim zur Basis a für jedes zu q teilerfremde a, so ist q Carmichael-Zahl.

(c) Zeige: Es gibt keine zusammengesetzten Zahlen q, die quadratisch quasiprim zur Basis a für jedes zu q teilerfremde a sind.

ABSCHNITT 7

Ganze Gaußsche Zahlen und Summen von Quadraten

Wir wissen bereits, daß der Ring $\mathbb{Z}[i]$, genannt Ring der ganzen Gaußschen Zahlen, ein euklidischer Ring ist (Abschnitt 1). Wir wollen in diesem Abschnitt den Ring $\mathbb{Z}[i]$ näher untersuchen und die dabei gewonnenen Kenntnisse auf gewisse diophantische Gleichungen anwenden.

Bereits Fermat hat den folgenden, von Euler bewiesenen Satz formuliert:

Satz 7.1. *Sei $p > 2$ eine Primzahl. Dann besitzt die diophantische Gleichung*

$$x^2 + y^2 = p$$

genau dann eine Lösung, wenn $p \equiv 1 \ (4)$.

Man sieht leicht, daß die Bedingung $p \equiv 1 \ (4)$ notwendig ist: Falls $p \equiv 3 \ (4)$ und $x^2 + y^2 = p$, so $x^2 + y^2 \equiv 3 \ (4)$, was unmöglich ist. Sei nun $p \equiv 1 \ (4)$. Genau dann ist $x^2 + y^2 = p$, wenn

$$p = (x + iy)(x - iy).$$

Also hängt die Darstellbarkeit von p als Summe zweier Quadrate eng zusammen mit der multiplikativen Zerlegung von p in $\mathbb{Z}[i]$. Wir beobachten außerdem, daß $x - iy$ die zu $x + iy$ konjugiert komplexe Zahl ist. Die zu $z \in \mathbb{C}$ konjugiert komplexe Zahl bezeichnen wir im folgenden mit z'. Die komplexe Konjugation überführt offenbar $\mathbb{Z}[i]$ in sich und ist daher ein Automorphismus dieses Rings. Wir notieren zunächst:

Satz 7.2. *Die Einheiten in $\mathbb{Z}[i]$ sind $1, -1, i, -i$.*

Beweis. Falls $uv = 1$, so $N(u) \cdot N(v) = 1$. Da die Norm nur ganzzahlige nichtnegative Werte annimmt, müssen $N(u), N(v) = 1$ sein. Daraus folgt $u = 1, -1, i$ oder $-i$. □

Satz 7.3. *Sei $\pi \in \mathbb{Z}[i]$, $\pi \neq \pm 1 + \pm i$, ein Primelement, das nicht zu einer Primzahl $q \in \mathbb{Z}$ assoziiert ist. Dann ist π' ein nicht zu π assoziiertes Primelement.*

Beweis. Offensichtlich gilt: $a \mid b \iff a' \mid b'$. Daraus folgt unmittelbar, daß π' Primelement ist. Sei $\pi = a + ib$. Da die Assoziierten von π keine (Prim-)Zahlen in $\mathbb{Z}$ sind, müssen $a, b \neq 0$ sein. Dann können die Gleichungen $\pi = \pi'$, $\pi = -\pi'$ nicht erfüllt sein. Wegen $\pi \neq 1 + i$, $1 - i$ kann aber auch keine der Gleichungen $\pi = i\pi'$, $\pi' = -i\pi'$ gelten. □

Wir kommen zurück zum Ausgangsproblem und nehmen an, $p \neq 2$ sei kein Primelement in $\mathbb{Z}[i]$. Sei

$$p = \pi_1 \cdots \pi_r$$

eine Primfaktorzerlegung von p in $\mathbb{Z}[i]$. Dann ist

$$p^2 = N(p) = N(\pi_1) \cdots N(\pi_r),$$

und wegen $N(\pi_i) > 1$ muß $r = 2$, $N(\pi_1) = N(\pi_2) = p$ gelten, und da auch $p = \pi_1 \pi_1'$ ist, folgt $\pi_2 = \pi_1'$. Also ist die Primfaktorzerlegung von p gegeben durch $p = \pi\pi'$, $N(\pi) = p$.

Für $p \neq 2$ ist π nicht assoziiert zu $1 + i$, folglich π gemäß 7.3 nicht assoziiert zu π'. Für $p = 2$ ergibt sich $2 = (1 + i)(1 - i) = -i(1 + i)^2$: Somit ist 2 kein Primelement in $\mathbb{Z}[i]$.

Da umgekehrt aus der Darstellbarkeit $p = x^2 + y^2 = (x + iy)(x - iy)$ folgt, daß p kein Primelement in $\mathbb{Z}[i]$ ist, haben wir bewiesen:

Satz 7.4. *Genau dann ist die diophantische Gleichung $p = x^2 + y^2$ für eine Primzahl p lösbar, wenn p kein Primelement in $\mathbb{Z}[i]$ ist.*

Natürlich bleibt immer noch zu zeigen, daß Primzahlen $p \equiv 1 \ (4)$ keine Primelemente in $\mathbb{Z}[i]$ sind. Dies aber ist nicht schwierig. Dazu braucht man keineswegs eine Zerlegung $p = uv$ anzugeben. Es genügt ja schon eine Gleichung

$$pt = uv,$$

bei der p weder u noch v teilt. Da $p \equiv 1 \ (4)$ ist $(\frac{-1}{p}) = 1$, es gibt also ein $z \in \mathbb{Z}$ mit $z^2 \equiv -1 \ (p)$ und somit $z^2 + 1 \equiv 0 \ (p)$. Folglich gibt es ein $t \in \mathbb{Z}$ mit

$$(z + i)(z - i) = tp.$$

Offensichtlich teilt p weder $z + i$ noch $z - i$. Damit ist 7.1 vollständig bewiesen.

Die Brücke zwischen der Lösbarkeit der Gleichung $x^2 + y^2 = p$ für Primzahlen p und der Teilbarkeitstheorie in $\mathbb{Z}[i]$, wurde durch die Äquivalenz

$$x^2 + y^2 = p \iff N(x + iy) = p$$

geliefert. Diese Äquivalenz hilft uns auch bei der Erweiterung auf zusammengesetzte Zahlen:

Satz 7.5. *Wenn sowohl m, als auch n Summen von zwei Quadraten sind, ist auch $m \cdot n$ Summe von zwei Quadraten.*

Beweis. Wenn $m = N(u)$ und $n = N(v)$, so $m \cdot n = N(u)N(v) = N(uv)$. □

Damit erhalten wir eine vollständige Übersicht über die ganzen Zahlen, die sich als Summe von zwei Quadraten darstellen lassen:

Satz 7.6. *Sei $n \in \mathbb{N}_+$, $n = a^2 b$ mit $a, b \in \mathbb{N}_+$, b quadratfrei. Die Zahl n ist Summe von zwei Quadraten genau dann, wenn b nur von 2 oder Primzahlen $p \equiv 1 \ (4)$ geteilt wird.*

Beweis. „⇐“: Mit 7.1 und 7.6 genügt die Feststellung, daß $2 = 1 + 1$ Summe zweier Quadrate ist.

„⇒“: Zu zeigen ist, daß $v_q(n)$ gerade für jede Primzahl $q \equiv 3 \ (4)$, die n teilt. Sei $n = x^2 + y^2$,

$$n = (x + iy)(x - iy) = zz'.$$

Nach 7.1 und 7.4 ist q ein Primelement in $\mathbb{Z}[i]$. Wenn q die Zahl n teilt, so muß es damit z oder z' teilen, also beide. □

Wir haben nebenbei eine Übersicht über das Zerlegungsverhalten der Primzahlen $p \in \mathbb{Z}$ in $\mathbb{Z}[i]$ gewonnen:

Satz 7.7.

(i) *Wenn $p = 2$, so $p = -i(1 + i)^2$, $1 + i$ prim in $\mathbb{Z}[i]$.*
(ii) *Wenn $p \equiv 1 \ (4)$, so $p = \pi\pi'$, π prim in $\mathbb{Z}[i]$, $N(\pi) = p$.*
(iii) *Wenn $p \equiv 3 \ (4)$, so ist p prim in $\mathbb{Z}[i]$.*

Eine merkwürdige Koinzidenz: Die Fälle (i), (ii), (iii) entsprechen den Fällen

$$\left(\frac{-4}{p}\right) = \begin{cases} 0 \\ 1 \\ -1 \end{cases}.$$

Unsere Vorgehensweise beim Studium der diophantischen Gleichung $x^2 + y^2 = n$ folgt einem Programm, das man allgemein auf Gleichungen der Form $ax^2 + bxy + cy^2 = n$ anwenden kann. In jedem Fall haben wir die Einsicht gewonnen, daß es zur Untersuchung zahlentheoretischer Probleme, die nur die ganzen Zahlen betreffen, zweckmäßig sein kann, den Kreis der betrachteten Objekte über den Ring $\mathbb{Z}$ hinaus auszudehnen.

Zunächst kommen wir auf die pythagoräischen Tripel zurück. Wir lösen die Aufgabe, die primitiven pythagoräischen Tripel

$$x^2 + y^2 = z^2, \qquad \mathrm{ggT}(x, y, z) = 1$$

zu bestimmen, noch einmal. Da $2 \nmid x$ oder $2 \nmid y$, folgt aus $x^2 + y^2 \equiv z^2 \ (4)$ sofort, daß z ungerade.

Sei $w = x + iy$ und

$$N(w) = (x + iy)(x - iy) = z^2.$$

Sei π ein Primteiler von w. Falls π auch $w' = x - iy$ teilt, teilt π auch $w + w' = 2x$ und $w - w' = 2iy$, also auch $2y$. Folglich: $\pi \mid \mathrm{ggT}(2x, 2y) = 2$ und somit ist π assoziiert zu $1 + i$. (Auch in $\mathbb{Z}[i]$ sind x und y teilerfremd.)

Es folgt $N(\pi) = 2 \mid N(w) = z^2$ im Widerspruch dazu, daß z ungerade ist. Insgesamt: $\pi \nmid w'$.

Wir erhalten

$$v_\pi(w) = v_\pi(z^2) = 2v_\pi(z).$$

Also ist $v_\pi(w)$ gerade für jedes Primelement π, das w teilt:

$$w = \varepsilon v^2, \quad \varepsilon \in \{1, -1, i, -i\}.$$

Mit $v = m+in$ folgt $w = \varepsilon(m^2-n^2+2mni)$. Dabei dürfen wir $m > 0$ annehmen.

Wenn wir voraussetzen, daß $x, y, z > 0$ und daß y gerade ist, muß auch $n > 0$ sein und $\varepsilon = 1$, und wir erhalten

$$x = m^2 - n^2, \quad y = 2mn, \quad z = m^2 + n^2.$$

Teilerfremdheit von x, y, z impliziert, daß m und n teilerfremd und nicht beide ungerade, und umgekehrt sind unter diesen Bedingungen auch x, y, z teilerfremd. Ferner dürfen wir uns natürlich auf $m, n > 0$ beschränken:

Satz 7.8. *Die primitiven pythagoräischen Tripel (x, y, z) mit $x, y, z > 0$ und $2 \mid y$, sind genau die Tripel*

$$(m^2 - n^2, \quad 2mn, \quad m^2 + n^2)$$

mit teilerfremden $m, n \in \mathbb{N}_+$, $m > n$, $m \not\equiv n$ (2).

Diese Beschreibung der pythagoräischen Tripel erlaubt es uns, das Fermatsche Problem für den einzigen elementaren Fall, nämlich den Exponenten 4 zu lösen:

Satz 7.9. *Die diophantische Gleichung $x^4 + y^4 = z^2$ besitzt keine Lösung mit $xyz \neq 0$, erst recht besitzt $x^4 + y^4 = z^4$ keine nichttriviale Lösung.*

Beweis. Wir nehmen an, es gäbe eine nichttriviale Lösung von $x^4 + y^4 = z^2$ und betrachten eine Lösung (x, y, z), für die z minimal ist. Dann ist notwendig $\mathrm{ggT}(x, y, z) = 1$, also auch $\mathrm{ggT}(x^2, y^2, z) = 1$. Mittels 7.8 finden wir (nach eventueller Vertauschung von x und y) m, n mit

$$x^2 = m^2 - n^2, \quad y^2 = 2mn, \quad z = m^2 + n^2.$$

Dabei teilt 2 genau eine der Zahlen m und n. Dies muß n sein, denn andernfalls

$$x^2 \equiv m^2 - n^2 \equiv -1 \ (4)$$

was nicht möglich ist. Also ist $n = 2n'$. Wir betrachten nun $x^2 + n^2 = m^2$. Da m, n teilerfremd, ist (x, n, m) ein primitives pythagoräisches Tripel:

$$x = u^2 - v^2, \quad n = 2uv, \quad m = u^2 + v^2.$$

Nun kommt der entscheidende Punkt: Wir zeigen, daß m, u, v Quadrate sind. Da m, n' teilerfremd und $y^2 = 4mn'$, müssen m, n' Quadrate sein. Da $n' = uv$ und

u, v teilerfremd, müssen auch u und v Quadrate sein. Mit $m = (z')^2$, $u = (y')^2$, $v = (x')^2$ ergibt sich

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

mit $z' < z$, im Widerspruch zur Wahl von (x, y, z). □

Der vorangegangene Beweis illustriert sehr schön die von Fermat erfundene Technik des *unendlichen Abstiegs*, bei der aus einer Lösung einer diophantischen Gleichung eine neue, kleinere konstruiert wird. Die Unmöglichkeit des unendlichen Abstiegs zeigt dann, daß es keine Lösung gibt.

Nachdem wir das Problem der Darstellung von Zahlen durch Summen von zwei Quadraten gelöst haben, sollen ohne Beweis die entsprechenden Aussagen für drei und vier Quadrate mitgeteilt werden:

Satz 7.10 (Gauß)**.** *Eine positive Zahl n ist genau dann Summe von drei Quadraten, wenn sie nicht von der Form* $4^a(8b - 1)$ *ist.*

Satz 7.11 (Lagrange)**.** *Jede positive Zahl n ist Summe von vier Quadraten.*

Man kann 7.11 in gewisser Weise analog zu 7.6 beweisen. Statt im Ring der Gaußschen Zahlen arbeitet man mit den Hurwitzschen Quaternionen (vgl. etwa [**Fors**]).

Übungen.

7.12. (a) Zeige: In $\mathbb{Z}[i]$ sind prim genau die Elemente (i) π mit $N(\pi) = p$, p Primzahl, $p \equiv 1\ (4)$, (ii) $q \in \mathbb{Z}$, q Primzahl, $q \equiv 3\ (4)$, und die dazu assoziierten Elemente, (iii) $1 + i$ und die dazu assoziierten Elemente.

(b) Stelle eine Tabelle der Primelemente π in $\mathbb{Z}[i]$ mit $N(\pi) \leq 50$ auf.

7.13. Sei $n \in \mathbb{Z}$, $Q(n) = \{z \in \mathbb{Z}[i]\, N(z) = n\}$ und $q(n)$ die Anzahl der Elemente von $Q(n)$.

(a) Zeige: Für $n, m \in \mathbb{Z}$, $\mathrm{ggT}(m, n) = 1$ ist $q(mn) = (1/4)q(m)q(n)$. Betrachte dazu die Abbildung $\psi\ Q(m) \times Q(n) \to Q(m \cdot n)$, $\psi(w, z) = wz$ und zeige: Jedes Element aus $Q(m \cdot n)$ hat genau 4 Urbilder.

(b) Sei $p \in \mathbb{Z}$ prim. Zeige:

$$q(p^\alpha) = \begin{cases} 4 & \text{wenn } p \equiv 3\ (4),\ \alpha \equiv 0\ (2) \\ 0 & \text{wenn } p \equiv 3\ (4),\ \alpha \equiv 1\ (2) \\ 4 & \text{wenn } p = 2 \\ 4(\alpha + 1) & \text{wenn } p \equiv 1\ (4). \end{cases}$$

7.14. Entwickle die Theorie der diophantischen Gleichung $x^2 + 2y^2 = n$ analog zu den Aussagen über $x^2 + y^2 = n$. Nutze dabei, daß $\mathbb{Z}[i\sqrt{2}]$ euklidisch.

Beginne damit, die möglichen Werte von $x^2 + 2y^2 = n$ modulo 8 zu bestimmen. Formuliere und beweise das Zerlegungsgesetz. Bestimme die $n \in \mathbb{N}$, die in der Form $x^2 + 2y^2 = n$ darstellbar sind.

7.15. Zeige: Die diophantische Gleichung $y^2 + 2 = x^3$ hat nur die Lösungen $x = 3$, $y = \pm 5$.

Hinweis: Wenn $y^2+2 = x^3$ für $x, y \in \mathbb{N}$, so sind $y+i\sqrt{2}$ und $y-i\sqrt{2}$ teilerfremd in $\mathbb{Z}[i\sqrt{2}]$.

ABSCHNITT 8

Algebraische und ganz-algebraische Zahlen

In Kapitel 7 zeigte sich, daß es sehr vorteilhaft sein kann, von $\mathbb{Z}$ zu einem größeren Zahlbereich überzugehen, selbst wenn man an der Lösung eines nur $\mathbb{Z}$ betreffenden Problems interessiert ist. Diese Erweiterungen von $\mathbb{Z}$ sind in der Regel dadurch gekennzeichnet, daß über ihnen gewisse Polynome, die über $\mathbb{Z}$ irreduzibel sind, in Linearfaktoren zerfallen, wie etwa $x^2 + y^2 = (x + iy)(x - iy)$ über $\mathbb{Z}[i]$. Solche Bereiche sollen in diesem Abschnitt systematisch eingeführt werden.

Nach dem Fundamentalsatz der Algebra zerfällt jedes Polynom $a_n X^n + \cdots + a_1 X + a_0$ mit $a_i \in \mathbb{C}$ in ein Produkt von Linearfaktoren $X - b$, $b \in \mathbb{C}$. Speziell gilt dies für Polynome mit rationalen oder ganzen Koeffizienten. Insofern genügt es, sich auf Teilmengen und Elemente von $\mathbb{C}$ zu beschränken.

Definition. Eine komplexe Zahl a heißt *algebraisch*, wenn es ein Polynom

$$f = X^n + c_{n-1}X^{n-1} + \cdots + c_1 X + c_0, \quad c_i \in \mathbb{Q},$$

mit $f(a) = 0$ gibt.

Man gelangt offenbar zu dem gleichen Begriff, wenn man nur Polynome mit Koeffizienten aus $\mathbb{Z}$ zuläßt, aber nicht verlangt daß der Leitkoeffizient 1 ist.

Beispiele algebraischer Zahlen: $\sqrt[n]{r}$ für jedes $r \in \mathbb{Q}$, $r > 0$, $n \in \mathbb{N}_+$, und i, $-i$, $i\sqrt[n]{r}$, die Nullstellen des Polynoms $X^2 + X + 1 = 0$, also $x = -1/2 \pm i\sqrt{3}/2$, natürlich alle rationalen Zahlen. Es gibt aber auch algebraische Zahlen, die sich nicht aus rationalen Zahlen mittels der arithmetischen Operationen und des Ziehens n-ter Wurzeln, $n \in \mathbb{N}_+$, erhalten lassen. Dies gilt z.B. für die Nullstellen von $X^5 - 2X^4 + 2$, was mit Hilfe der Galoistheorie bewiesen werden kann.

Nicht algebraische komplexe Zahlen nennt man *transzendent*, berühmte Beispiele für transzendente Zahlen sind e und π.

Für $a \in \mathbb{C}$ betrachten wir die Teilmenge

$$\mathbb{Q}[a] := \{f(a) : f \in \mathbb{Q}[X]\}$$

von $\mathbb{C}$, also die Menge aller $\mathbb{Q}$-Linearkombinationen der Potenzen von a. Sie ist ein Teilring von $\mathbb{C}$, insbesondere ein Integritätsbereich, und entsteht als homomorphes Bild von $\mathbb{Q}[X]$ mittels der Substitution $X \mapsto a$. Also

$$\mathbb{Q}[a] = \mathbb{Q}[X]/\mathfrak{a}, \quad \mathfrak{a} = \{f \in \mathbb{Q}[X] : f(a) = 0\}. \qquad (*)$$

Sicher ist $\mathfrak{a}$ ein Primideal. Genau dann ist $\mathfrak{a} \neq 0$, wenn a algebraisch ist. Da $\mathbb{Q}[X]$ euklidisch ist, wird $\mathfrak{a}$ in diesem Fall von einem Primelement $p \in \mathbb{Q}[X]$ erzeugt:

$$\mathfrak{a} = \mathbb{Q}[X] \cdot p, \quad p = \alpha_n X^n + \cdots + \alpha_0,\ \alpha_i \in \mathbb{Q},\ \alpha_n \neq 0.$$

Nach Multiplikation mit $1/\alpha_n$ dürfen wir $\alpha_n = 1$ annehmen:

Satz und Definition 8.1. *Sei $a \in \mathbb{C}$ eine algebraische Zahl. Dann gibt es genau ein irreduzibles normiertes Polynom $p \in \mathbb{Q}[X]$ mit $p(a) = 0$. Dieses Polynom heißt* Minimalpolynom *von a, sein Grad heißt auch* Grad von *a (oder von $\mathbb{Q}[a]$).*

Die Behauptung hinsichtlich „genau ein“ ist klar: Jedes zweite solche Polynom muß zu p assoziiert sein, sich also von p um einen konstanten Faktor $\neq 0$ unterscheiden. Da die Leitkoeffizienten beide 1 sind, ist der Faktor gerade 1.

Beispiele. (a) Sei $p = X^2 + \alpha X + \beta$, $\alpha, \beta \in \mathbb{Q}$, ohne Nullstellen in $\mathbb{Q}$. Dann ist p irreduzibel, und die Nullstellen

$$-\alpha/2 \pm \sqrt{\alpha^2/4 - \beta}$$

sind algebraische Zahlen des Grades 2.

(b) Das Polynom $X^3 - 2$ ist irreduzibel über $\mathbb{Q}$. (Wäre es reduzibel, so müßte es ein $a \in \mathbb{Q}$ mit $a^3 = 2$ geben). Also ist $\sqrt[3]{2}$ eine algebraische Zahl des Grades 3.

Da $\mathbb{Q}[X]$ euklidisch ist, ist $p \in \mathbb{Q}[X]$ genau dann irreduzibel, wenn der Restklassenring $\mathbb{Q}[X]/p\mathbb{Q}[X]$ ein Körper ist (vgl. 3.4). Deshalb folgt aus der Isomorphie $(*)$, daß $\mathbb{Q}[a]$ ein Körper ist, wofür wir gleich noch einen zweiten Beweis geben.

Zunächst eine Charakterisierung der algebraischen Zahlen mit Hilfe der linearen Algebra:

Satz 8.2. *Für eine komplexe Zahl a sind äquivalent:*

(a) *a ist algebraisch.*
(b) *Der $\mathbb{Q}$-Vektorraum $\mathbb{Q}[a]$ ist endlich-dimensional.*
(c) *Es existiert ein endlich-dimensionaler $\mathbb{Q}$-Vektorraum $V \subset \mathbb{C}$, $V \neq 0$, mit $aV \subset V$.*

Beweis. (a) $\Rightarrow$ (b): Sei $p = X_n + \alpha_{n-1}X^{n-1} + \cdots + \alpha_1 X + \alpha_0$ das Minimalpolynom von a. Da $p(a) = 0$, ist

$$a^n = -(\alpha_{n-1}a^{n-1} + \cdots + \alpha_1 a + \alpha_0 \cdot 1) \qquad (**)$$

Linearkombination von $a^{n-1}, \ldots, 1$. Durch sukzessive Multiplikation mit a und Einsetzen der rechten Seite von $(**)$ für a^n erhält man aus $(**)$, daß jede Potenz von a Linearkombination von $1, \ldots, a^{n-1}$ ist. Also erzeugen die Potenzen $1, \ldots, a^{n-1}$ den $\mathbb{Q}$-Vektorraum $\mathbb{Q}[a]$.

(b) $\Rightarrow$ (c): Wir wählen einfach $V = \mathbb{Q}[a]$.

(c) $\Rightarrow$ (a): Sei $v_1, \ldots, v_n$ ein Erzeugendensystem von V. (Es kommt nicht darauf an, daß $v_1, \ldots, v_n$ linear unabhängig sind.) Für jedes $v \in V$ ist av eine $\mathbb{Q}$-Linearkombination von $v_1, \ldots, v_n$. Speziell gibt es also rationale Zahlen γ_{ij} mit

$$av_j = \gamma_{1j} v_1 + \cdots + \gamma_{nj} v_n, \qquad j = 1, \ldots, n.$$

Man bringt nun av_j auf die rechte Seite und erhält dann die Matrizengleichung

$$(\Gamma - aE_n) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = 0$$

(Dabei sei $\Gamma = (\gamma_{ij})$ und E_n die n-reihige Einheitsmatrix.) Da nach Voraussetzung $v_j \neq 0$ für mindestens ein j, ist a Eigenwert und somit Nullstelle des charakteristischen Polynoms von Γ. Dies ist ein Polynom des Grades n mit Leitkoeffizient ± 1. Somit ist a algebraisch.

Für spätere Verwendung halten wir fest: *Falls die γ_{ij} sogar ganze Zahlen sind, ist a Nullstelle eines normierten Polynoms über $\mathbb{Z}$.* □

Nun ergibt sich leicht:

Satz und Definition 8.3. *Sei $a \in \mathbb{C}$ eine algebraische Zahl des Grades n. Dann ist $\mathbb{Q}[a]$ ein $\mathbb{Q}$-Vektorraum der Dimension n und sogar ein Körper. Die Potenzen $1 = a^0, \ldots, a^{n-1}$ bilden eine $\mathbb{Q}$-Basis von $\mathbb{Q}[a]$. Der Körper $\mathbb{Q}[a]$ heißt ein* algebraischer Zahlkörper. *Alle Elemente von $\mathbb{Q}[a]$ sind algebraisch.*

Beweis. Daß die Potenzen $1, a, \ldots, a^{n-1}$ den $\mathbb{Q}$-Vektorraum $\mathbb{Q}[a]$ erzeugen, haben wir oben bereits bewiesen. Sie sind auch linear unabhängig: Falls

$$\beta_{n-1} a^{n-1} + \cdots + \beta_0 = 0,$$

so muß $q := \beta_{n-1} X^{n-1} + \cdots + \beta_0$ ein Vielfaches des Minimalpolynoms p sein, was nur für $q = 0$ möglich ist.

Sei $b \in \mathbb{Q}[a]$, $b \neq 0$. Da $b\mathbb{Q}[a] \subset \mathbb{Q}[a]$ folgt mit Satz 8.2, daß b algebraisch ist.

Sei nun $b \neq 0$. Die Multiplikation mit b bildet $\mathbb{Q}[a]$ injektiv und $\mathbb{Q}$-linear auf sich selbst ab. Da $\mathbb{Q}[a]$ endlich-dimensional ist, ist diese Abbildung auch surjektiv; es existiert also ein $\tilde{b} \in \mathbb{Q}[a]$ mit $b\tilde{b} = 1$. Folglich ist der Ring $\mathbb{Q}[a]$ sogar ein Körper. □

Die Argumente des vorangegangenen Beweises verwenden wir noch einmal zum Beweis des nächsten Satzes.

Satz 8.4. *Summe, Produkt und Quotient algebraischer Zahlen sind algebraisch.*

Beweis. Seien a, b algebraisch von den Graden m und n. Wir betrachten den von allen Potenzen $a^i b^j$, $i, j \geq 0$, aufgespannten $\mathbb{Q}$-Untervektorraum K von $\mathbb{C}$. Da

nun aber a^m Linearkombination von $1, \dots, a^{m-1}$ und b^n Linearkombination von $1, \dots, b^{n-1}$ ist, sieht man sofort, daß dieser Untervektorraum schon von den endlich vielen Potenzen $a^i b^j$ mit $i < m$ und $j < n$ erzeugt wird. Satz 8.2 impliziert, daß $a + b$ und ab algebraisch sind, denn $(a + b)K \subset K$ und $abK \subset K$.

Wie im Beweis von Satz 8.3 folgt nun, daß K ein Körper ist, und ebenso folgt, daß alle Element von K algebraische Zahlen sind. Zu ihnen gehört a/b. □

Aus Satz 8.4 ergibt sich sofort, daß die algebraischen Zahlen einen Teilkörper von $\mathbb{C}$ bilden, den man üblicherweise mit $\overline{\mathbb{Q}}$ bezeichnet.

Unser Interesse wird speziell den quadratischen Zahlkörpern gelten; das sind die algebraischen Zahlkörper des Grades 2. Sei $p = X^2 + \alpha X + \beta$ irreduzibel, $a = -\alpha/2 + \sqrt{\alpha^2/4 - \beta}$. Dann ist

$$\mathbb{Q}[a] = \mathbb{Q}\left[\sqrt{\alpha^2/4 - \beta}\right],$$

also von der Form $\mathbb{Q}[\sqrt{r}]$ mit einem $r \in \mathbb{Q}$. (Dabei sei $\sqrt{r} > 0$, falls $r > 0$, $\sqrt{r} := i\sqrt{|r|}$, falls $r < 0$.) Nun ist

$$\sqrt{r} = \frac{1}{t}\sqrt{st}, \quad \text{wenn } r = \frac{s}{t},\ s, t \in \mathbb{Z},\ t > 0,$$

und daher

$$\mathbb{Q}[\sqrt{r}] = \mathbb{Q}[\sqrt{u}], \quad u = st \in \mathbb{Z}.$$

Wenn wir nun $u = \tilde{u}^2 d$ mit quadratfreiem $d \in \mathbb{Z}$ schreiben, so erhalten wir insgesamt: $\mathbb{Q}[a] = \mathbb{Q}[\sqrt{d}]$, $d \in \mathbb{Z}$, d quadratfrei, d.h. $d = -1$ oder $d = \pm p_1 \cdots p_k$ mit paarweise verschiedenen Primzahlen p_j, $d \neq 1$.

Wenn die quadratfreien Zahlen $d, d' \in \mathbb{Z}$ verschieden sind, so gilt andererseits

$$\mathbb{Q}[\sqrt{d}] \neq \mathbb{Q}[\sqrt{d'}].$$

Dies sieht man so ein: Wir dürfen annehmen, daß $d \neq -1$, und unter der Annahme $\mathbb{Q}[\sqrt{d'}] = \mathbb{Q}[\sqrt{d}]$ besitzt $\sqrt{d'}$ eine Darstellung

$$\sqrt{d'} = \alpha + \beta\sqrt{d} \quad \text{mit } \alpha, \beta \in \mathbb{Q}.$$

Quadrieren ergibt $d' = (\alpha^2 + \beta^2 d) + 2\alpha\beta\sqrt{d}$. Da $1, \sqrt{d}$ linear unabhängig sind, folgt $\alpha = 0$ oder $\beta = 0$. Im Fall $\beta = 0$, ist $d' = \alpha^2$, im Fall $\alpha = 0$ erhalten wir $d' = \beta^2 d$, wobei $\beta^2 \neq 1$. Jede der Alternativen ist ein Widerspruch zur Annahme, daß d' quadratfrei ist.

Satz 8.5. *Die quadratischen Zahlkörper sind die Körper $\mathbb{Q}[\sqrt{d}]$, $d \in \mathbb{Z}$ quadratfrei, $d \neq 1$. Diese Körper sind paarweise verschieden.*

Ein quadratischer Körper heißt *reell*, wenn $d > 0$, und *imaginär*, wenn $d < 0$.

Es war in Abschnitt 7 wesentlich, daß wir im Ring $\mathbb{Z}[i]$ gearbeitet haben und nicht im Körper $\mathbb{Q}[i]$. Dieser und verwandte Ringe werden von ganz-algebraischen Zahlen gebildet:

Definition. Die komplexe Zahl a heißt *ganz-algebraisch*, wenn es ein Polynom $f = X^n + a_{n-1}X^{n-1} + \cdots + a_0$, $a_i \in \mathbb{Z}$, mit $f(a) = 0$ gibt. Ein solches Polynom heißt eine *Ganzheitsgleichung* für a.

Die bisher angeführten nichtrationalen algebraischen Zahlen sind sogar ganze algebraische Zahlen. Es ist aber auch sehr leicht, algebraische Zahlen anzugeben, die nicht ganz-algebraisch sind:

Satz 8.6. *Eine Zahl $a \in \mathbb{Q}$ ist ganz-algebraisch genau dann, wenn $a \in \mathbb{Z}$.*

Beweis. Die Implikation „$\Leftarrow$“ ist trivial. Zum Beweis der anderen Implikation sei $r \in \mathbb{Q}$ ganz-algebraisch, etwa

$$r^n + a_{n-1}r^{n-1} + \cdots + a_r r + a_0 = 0$$

mit $a_i \in \mathbb{Z}$. Wir schreiben $r = s/t$ mit $r, s \in \mathbb{Z}$, $\mathrm{ggT}(r, s) = 1$.

Einsetzen und Multiplikation mit t^n zeigt: $s^n \in t\mathbb{Z}$, also $\mathrm{ggT}(s^n, t) = t$. Da andererseits $\mathrm{ggT}(s^n, t) = 1$, folgt $t = 1$. □

Es ist in der Zahlentheorie üblich, ganz-algebraische Zahlen kurz *ganz* zu nennen. Die hier getroffene Erweiterung des Begriffes „ganz“ fangen wir dadurch auf, daß wir die Elemente von $\mathbb{Z}$ künftig *ganz-rational* nennen (vgl. 8.6).

Glücklicherweise braucht man zwischen dem Minimalpolynom und dem „ganzen Minimalpolynom“ nicht zu unterscheiden:

Satz 8.7. *Sei a eine ganz-algebraische Zahl. Dann liegt das Minimalpolynom p von a in $\mathbb{Z}[X]$ und jedes Polynom $q \in \mathbb{Z}[X]$ mit $q(a) = 0$ ist Vielfaches von p (nicht nur in $\mathbb{Q}[X]$, sondern bereits) in $\mathbb{Z}[X]$.*

Beweis. Jedes normierte Polynom $f \in \mathbb{Z}[X]$ läßt sich, wie man mittels Induktion über den Grad sofort feststellt, als Produkt irreduzibler normierter Polynome in $\mathbb{Z}[X]$ darstellen. Es gibt also ein solches Polynom in $\mathbb{Z}[X]$, unter dessen Nullstellen auch a ist. Gemäß dem folgenden Satz hat man bereits das Minimalpolynom gefunden. Der letzte Teil ergibt sich dann einfach aus der Division mit Rest (für Polynome). □

Satz 8.8. *Ein normiertes Polynom $p \in \mathbb{Z}[X]$, das in $\mathbb{Z}[X]$ irreduzibel ist, ist auch in $\mathbb{Q}[X]$ irreduzibel.*

Satz 8.8 firmiert auch unter der Bezeichnung „Gaußsches Lemma“ und sollte aus der Algebra bekannt sein. Satz 8.8 gilt allgemeiner für primitive Polynome in $\mathbb{Z}[X]$, das sind solche, deren Koeffizienten den größten gemeinsamen Teiler 1 haben. Besonders nützlich ist 8.8 auch bei der Untersuchung konkreter Polynome auf Irreduzibilität, denn diese ist über $\mathbb{Z}$ i.a. viel leichter nachzuweisen als über $\mathbb{Q}$. In Analogie zu $\mathbb{Q}[a]$ setzen wir

$$\mathbb{Z}[a] = \{f(a) : f \in \mathbb{Z}[X]\},$$

mit anderen Worten: $\mathbb{Z}[a]$ wird von allen $\mathbb{Z}$-Linearkombinationen der Potenzen von a gebildet.

Für die Charakterisierung der ganz-algebraischen Zahlen in Analogie zu Satz 8.2 führen wir in Analogie zur „gewöhnlichen“linearen Algebra einige Begriffe ein:

Definition. Eine abelsche Gruppe M nennt man einen $\mathbb{Z}$-*Modul*. Die Verknüpfung wird in $\mathbb{Z}$-Moduln üblicherweise additiv geschrieben.

Man sagt, daß $x_1, \ldots, x_n$ den $\mathbb{Z}$-Modul M *erzeugen* oder ein *Erzeugendensystem* von M bilden , falls jedes Element $y \in M$ eine Darstellung $y = a_1x_1 + \cdots + a_nx_n$ mit $a_1, \ldots, a_n \in \mathbb{Z}$ besitzt.

Mann nennt $x_1, \ldots, x_n$ *linear unabhängig*, falls die Darstellung von $0 \in M$ nur mit $a_1, \ldots, a_n = 0$ möglich ist.

Eine linear unabhängiges Erzeugendensystem von M nennt man eine *Basis*.

In Analogie zu Satz 8.2 und seinem Beweis, den wir nicht zu wiederholen brauchen, erhalten wir nun:

Satz 8.9. *Für eine komplexe Zahl a sind äquivalent:*

(a) *a ist ganz-algebraisch.*
(b) *Der $\mathbb{Z}$-Modul $\mathbb{Z}[a]$ ist endlich erzeugt.*
(c) *Es existiert ein endlich erzeugter $\mathbb{Z}$-Modul $M \subset \mathbb{C}$, $M \neq 0$, mit $aM \subset M$.*

Zumindest Teile der Sätze 8.3 und 8.4 können wir auf die ganz-algebraischen Zahlen übertragen. Wiederum brauchen wir die Beweise nicht zu wiederholen. (Die einzige Eigenschaft endlich-dimensionaler $\mathbb{Q}$-Vektorräumen, die bei 8.3 und 8.4 ausgenutzt wurde, aber auf endlich erzeugte $\mathbb{Z}$-Moduln nicht zutrifft, ist die Surjektivität einer injektiven linearen Abbildung.)

Satz 8.10. *Sei a ganz-algebraisch vom Grad n. Dann bilden $1, \ldots, a^{n-1}$ eine $\mathbb{Z}$-Basis von $\mathbb{Z}[a]$. Alle Elemente von $\mathbb{Z}[a]$ sind ganz-algebraisch.*

Satz 8.11. *Summe und Produkt ganz-algebraischer Zahlen sind ganz-algebraisch.*

Die ganz-algebraischen Zahlen bilden also einen Teilring von $\mathbb{C}$, und Gleiches gilt für die ganz-algebraischen Zahlen innerhalb eines algebraischen Zahlkörpers. Diese Ringe sind die zentralen Objekte der algebraischen Zahlentheorie. Für die quadratischen Zahlkörper werden wir sie im nächsten Abschnitt mit direkten Methoden bestimmen.

Übungen.

8.12. Bestimme das Minimalpolynom von $\sqrt{2} + \sqrt{3}$.

8.13. Sei $f = X_3 + a_2X^2 + a_1X + a_0 \in \mathbb{Q}[X]$ irreduzibel und a Nullstelle von f. Bestimme das Minimalpolynom von a^2.

8.14. Seien $d, d' \in \mathbb{Z}$, d, d' quadratfrei, $d \neq d'$. Zeige: Die Körper $\mathbb{Q}[\sqrt{d}]$ und $\mathbb{Q}[\sqrt{d'}]$ sind nicht isomorph. (Dies ist eine geringfügige Verschärfung von Satz 8.5.)

8.15. Sei $a \in \mathbb{C}$ algebraisch. Zeige, daß na ganz ist für ein geeignetes $n \in \mathbb{Z}$, $n \neq 0$.

ABSCHNITT 9

Die ganzen Elemente quadratischer Zahlkörper

Die Bestimmung der ganzen Zahlen eines algebraischen Zahlkörpers ist i.a. eine schwierige Aufgabe. Im Spezialfall der quadratischen Körper ist sie aber sehr einfach.

Sei a ganz-algebraisch. Dann liegt die Vermutung nahe, die ganzen Elemente von $\mathbb{Q}[a]$ seien gerade durch $\mathbb{Z}[a]$ gegeben. Dies ist aber nicht richtig, nicht einmal im Fall der quadratischen Zahlkörper. Zum Beispiel gilt $\mathbb{Q}[2i] = \mathbb{Q}[i]$, aber $\mathbb{Z}[2i] \subsetneq \mathbb{Z}[i]$, und die Elemente von $\mathbb{Z}[i]$ sind sämtlich ganz. Dies ist zwar ein „billiges Gegenbeispiel“, weil $2i = \sqrt{-4}$ und -4 nicht quadratfrei ist. Aber selbst wenn $d \in \mathbb{Z}$ quadratfrei ist, erfaßt $\mathbb{Z}[\sqrt{d}]$ oft nicht alle ganzen Elemente von $\mathbb{Q}[\sqrt{d}]$, wie wir gleich sehen werden.

Sowohl $\sqrt{d}$ als auch $-\sqrt{d}$ haben $X^2 - d$ zum Minimalpolynom. Insofern sind sie, wenn man von $\mathbb{Q}$ ausgeht, in ihren algebraischen Eigenschaften nicht zu unterscheiden. Dies äußert sich darin, daß es einen Automorphismus des Körpers $\mathbb{Q}[\sqrt{d}]$ gibt, der $\sqrt{d}$ auf $-\sqrt{d}$ abbildet (und umgekehrt):

Satz 9.1. *Seien $d \in \mathbb{Q}$, $\sqrt{d} \notin \mathbb{Q}$. Dann ist die Abbildung $' : \mathbb{Q}[\sqrt{d}] \to \mathbb{Q}[\sqrt{d}]$*

$$(a + b\sqrt{d})' := a - b\sqrt{d}$$

ein Automorphismus von $\mathbb{Q}[\sqrt{d}]$, der zu sich selbst invers ist. Genau dann gilt $c' = c$, wenn $c \in \mathbb{Q}$.

Dieser Automorphismus heißt *Konjugation*. Er ist im imaginär-quadratischen Fall lediglich die Einschränkung der komplexen Konjugation auf $\mathbb{Q}[\sqrt{d}]$, im reell-quadratischen Fall aber keineswegs, dann bei $K \subset \mathbb{R}$ ist die Einschränkung der komplexen Konjugation auf K die Identität.

Der Beweis von 9.1 ist trivial. Ferner ist leicht zu sehen, daß die Konjugation der einzige von der Identität verschiedene Automorphismus von $\mathbb{Q}[\sqrt{d}]$ ist: ein solcher muß $\mathbb{Q}$ invariant lassen und $\sqrt{d}$ auf sich selbst oder $-\sqrt{d}$ abbilden.

Mit Hilfe der Konjugation definiert man zwei wichtige Funktionen

$$N : \mathbb{Q}[\sqrt{d}] \to \mathbb{Q}, \qquad N(c) := cc',$$
$$S : \mathbb{Q}[\sqrt{d}] \to \mathbb{Q}, \qquad S(c) := c + c'$$

N heißt *Norm* und S *Spur*. Daß die Norm und die Spur tatsächlich Werte in $\mathbb{Q}$ annehmen, folgt aus 9.1:

$$(cc')' = c'c'' = c'c = cc', \quad \text{also } cc' \in \mathbb{Q},$$

und analog ergibt sich dies für die Spur. Man kann es natürlich auch explizit ausrechnen:

$$N(a + b\sqrt{d}) = (a + b\sqrt{d})(a - b\sqrt{d}) = a^2 - b^2 d,$$
$$S(a + b\sqrt{d}) = 2a.$$

Wichtige Eigenschaften von Norm und Spur enthält der nächste Satz.

Satz 9.2.

(a) $N(uv) = N(u)N(v)$*: Die Norm ist multiplikativ.*
(b) $S(u+v) = S(u)+S(v)$, $S(rv) = rS(v)$ *für* $r \in \mathbb{Q}$*: Die Spur ist* $\mathbb{Q}$*-linear.*

Auch dies rechnet man unmittelbar nach.

Sei nun $a \in \mathbb{Q}[\sqrt{d}]$, $a \notin \mathbb{Q}$. Der von den Potenzen a^i, $i \in \mathbb{N}$, erzeugte $\mathbb{Q}$-Untervektorraum hat mindestens die Dimension 2, weil er $\mathbb{Q}$ enthält, aber $a \notin \mathbb{Q}$ ist. Andererseits hat er höchstens die Dimension 2, weil er Unterraum des zweidimensionalen $\mathbb{Q}$-Vektorraums $\mathbb{Q}[\sqrt{d}]$ ist. Das Minimalpolynom $X^2 + \alpha X + \beta$ von a hat daher genau den Grad 2. Dann ist $a^2 + \alpha a + \beta = 0$ und ebenso

$$0 = (a^2 + \alpha a + \beta)' = a'^2 + \alpha a' + \beta.$$

Die zweite Nullstelle ist also a' und wir erhalten

$$(X^2 + \alpha X + \beta) = (X - a)(X - a') = X^2 - (a + a')X + aa'.$$

Diese Übertragung zeigt erstens:

Satz 9.3. *Für* $a \in \mathbb{Q}[\sqrt{d}]$, $a \notin \mathbb{Q}$ *ist* $X^2 - S(a)X + N(a)$ *das Minimalpolynom von* a.

Und zweitens (in Verbindung mit 8.7):

Satz 9.4.

(a) *Mit* a *ist auch* a' *ganz.*
(b) a *ist genau dann ganz, wenn Norm und Spur von* a *ganz sind.*

Satz 9.4 eröffnet uns nun die Möglichkeit, die ganzen Elemente der quadratischen Zahlkörper vollständig zu bestimmen. Für $r, s, \in \mathbb{Q}$ ist

$$a = r + s\sqrt{d} \quad \text{ganz} \iff 2r \in \mathbb{Z} \quad \text{und} \quad r^2 - s^2 d \in \mathbb{Z}$$

Also gilt $r = r'/2$, $r' \in \mathbb{Z}$, und $4s^2 d \in \mathbb{Z}$. Mittels Primfaktorzerlegung von Zähler und Nenner von s folgt

$$s = \frac{s'}{2}, \quad s' \in \mathbb{Z}.$$

Einsetzen ergibt

$$r^2 - s^2 d = \frac{r'^2 - s'^2 d}{4} \in \mathbb{Z} \iff r'^2 - s'^2 d \equiv 0 \ (4).$$

Da d quadratfrei ist, ist $d \not\equiv 0 \ (4)$. Für die einzelnen Fälle gilt:

$$\begin{aligned}
&d \equiv 1 \ (4)\text{: } r'^2 - s'^2 = (r' - s')(r' + s') \equiv 0 \ (4) && \iff r' \equiv s' \ (2)\\
&d \equiv 2 \ (4)\text{: } r'^2 - 2s'^2 \equiv 0 \ (4) && \iff r', s' \equiv 0 \ (2)\\
&d \equiv 3 \ (4)\text{: } r'^2 - 3s'^2 \equiv r'^2 + s'^2 \equiv 0 \ (4) && \iff r', s' \equiv 0 \ (2)
\end{aligned}$$

Wenn umgekehrt r' und s' diese Bedingungen erfüllen, ist $r + s\sqrt{d}$ ganz. Damit haben wir die ganzen Elemente in $\mathbb{Q}[\sqrt{d}]$ bestimmt.

Satz 9.5.

(a) *Falls* $d \equiv 2, 3 \ (4)$*, so sind die ganzen Elemente von* $\mathbb{Q}[\sqrt{d}]$ *gegeben durch*

$$r + s\sqrt{d}, \quad r, s \in \mathbb{Z}.$$

(b) *Falls* $d \equiv 1 \ (4)$*, so sind die ganzen Elemente von* $\mathbb{Q}[\sqrt{d}]$ *gegeben durch*

$$\frac{r}{2} + \frac{s}{2}\sqrt{d}, \quad r, s \in \mathbb{Z}, \ r \equiv s \ (2).$$

Im Fall $d \equiv 1 \ (4)$ ergibt eine einfache Umformung:

$$\frac{r}{2} + \frac{s}{2}\sqrt{d} = \frac{r - s}{2} + s\left(\frac{1 + \sqrt{d}}{2}\right), \qquad \frac{r - s}{2} \in \mathbb{Z}.$$

Falls umgekehrt $a = u + v\left(\frac{1+\sqrt{d}}{2}\right)$, $u, v \in \mathbb{Z}$, so

$$a = \frac{2u + v}{2} + \frac{v}{2}\sqrt{d} \quad \text{und} \quad 2u + v \equiv v \ (2)$$

Insgesamt:

Satz 9.6. *Sei* $d \in \mathbb{Z}$, $d \neq 1$ *quadratfrei und*

$$\omega_d := \begin{cases} \sqrt{d} & \text{falls } d \equiv 2, 3 \ (4) \\ \frac{1+\sqrt{d}}{2} & \text{falls } d \equiv 1 \ (4) \end{cases}$$

Dann sind genau die Elemente von $\mathbb{Z}[\omega_d]$ *ganz in* $\mathbb{Q}[\sqrt{d}]$.

Es sei angemerkt, daß sich nicht in allen algebraischen Zahlkörpern K Elemente a finden lassen, für die $\mathbb{Z}[a]$ gerade die Teilmenge der ganzen Elemente ist. Bereits in Zahlkörpern des Grades 3 ist dies im allgemeinen nicht richtig.

Zur Vereinfachung der Schreibweise setzen wir

$$A_d = \mathbb{Z}[\omega_d].$$

Im Abschnitt 7 haben wir den Ring $A_{-1} = \mathbb{Z}[i]$ studiert. Die Hoffnung, alle Ringe A_d würden eine ähnlich einfache Struktur besitzen, zerstört sich rasch. Für $d = -5$ etwa ist $\omega_d = i\sqrt{5}$, und wir erhalten in A_{-5} zwei Zerlegungen der Zahl 6:

$$6 = 2 \cdot 3 = (1 + \omega_d)(1 - \omega_d).$$

Es gilt

$$N(2) = 4,\ N(3) = 9,\ N(1 + \omega_d) = 6,\ N(1 - \omega_d) = 6$$

Kein $z \in A_{-5}$, $z = u + vi\sqrt{5}$ mit $v \neq 0$ kann 2 teilen, denn $N(z) \geq 5$. Also muß 2 irreduzibel sein in A_{-5}. Aber 2 teilt weder $1 + \omega_d$ noch $1 - \omega_d$, ist also nicht prim. Die Existenz eines irreduziblen, nicht primen Elementes in A_{-5} zeigt: In A_{-5} gilt nicht der Satz von der eindeutigen Primfaktorzerlegung. Damit können wir nicht erwarten, daß die diophantische Gleichung $x^2 + 5y^2 = n$ eine ähnliche elegante Theorie aufweist wie $x^2 + y^2 = n$.

Als ein weiteres Beispiel zur Anwendung der Theorie der Ringe A_d wollen wir das quadratische Reziprozitätsgesetz betrachten. Der in Abschnitt 6 gegebene Beweis ist zwar eine Verifikation, eine Erklärung für die Gültigkeit des quadratischen Reziprozitätsgesetzes liefert er nicht. Im Rahmen der Theorie der quadratischen Körper kann man das Reziprozitätsgesetz wirklich begreifen. Dies wollen wir hier zunächst unter einschränkenden Ausnahmen demonstrieren. Seien p, q ungerade Primzahlen, $p \neq q$. Wir nehmen an, $A := A_q$ sei faktoriell. Sei zunächst $q \equiv 3$ (4) und quadratischer Rest modulo p. Wir wählen $a \in \mathbb{Z}$ mit

$$a^2 \equiv q \ \ (p).$$

Dann existiert ein $t \in \mathbb{Z}$ mit $a^2 - q = tp$. Die linke Seite können wir in A zerlegen,

$$(a + \sqrt{q})(a - \sqrt{q}) = tp.$$

Da p weder $a + \sqrt{q}$ noch $a - \sqrt{q}$ (in A) teilt, ist p kein Primelement, also gemäß der Annahme, A sei faktoriell, zerlegbar:

$$p = uv, \quad u, v \in A_q,$$

wobei u und v keine Einheiten sind. Deshalb gilt $N(u), N(v) \neq \pm 1$. (Wir diskutieren die Einheiten der Ringe A_d im nächsten Abschnitt.) Es folgt $N(p) = p^2 = N(u)N(v)$. Da notwendig $N(u) = \pm p$ ist, erhalten wir mit $a = \alpha + \beta\sqrt{q}$, $\alpha, \beta \in \mathbb{Z}$,

$$\pm p = N(u) = \alpha^2 - \beta^2 q$$

Es folgt $\pm p \equiv \alpha^2 + \beta^2$ (4), also muß gelten

$$\alpha^2 - \beta^2 q = \begin{cases} p & \text{wenn} \quad p \equiv 1 \ (4), \\ -p & \text{wenn} \quad p \equiv 3 \ (4). \end{cases}$$

Bei $p \equiv 1 \ (4)$ folgt $p \equiv \alpha^2 \ (q)$, p ist quadratischer Rest modulo q. Bei $p \equiv 3 \ (4)$ folgt: $-p$ ist quadratischer Rest modulo q, mithin p quadratischer Nichtrest, denn $(-1/q) = -1$.

Ähnlich diskutiert man den Fall $q \equiv 1 \ (4)$, $(q/p) = 1$. Es ergibt sich

$$\pm 4p \equiv \alpha^2 - \beta^2 q,$$

und man kann sofort schließen, daß p quadratischer Rest ist modulo q.

Zum Beweis des quadratischen Reziprozitätsgesetzes braucht man nur noch den Fall $p \equiv q \equiv 3 \ (4)$, $(p/q) = (q/p) = -1$ anzuschließen: Bei $(q/p) = -1$ ist dann $(-q/p) = 1$. Man zieht nun den Ring A_{-q} heran und schließt analog $4p = \alpha^2 + \beta^2 q$, woraus wieder $(p/q) = 1$ folgt.

Die einschränkende Annahme in diesem Beweis besteht in der Annahme, die Ringe A_q und A_{-q} seien faktoriell. Dies trifft leider fast nie zu. (Siehe dazu auch die Aufgaben zu diesem Abschnitt.) Dennoch kann man das quadratische Reziprozitätsgesetz aus dem Zerlegungsverhalten von p in A_q und A_{-q} (und dem von p und q in A_{pq} im Fall $p \equiv q \equiv 3 \ (4)$) herleiten. Dies freilich erfordert eine ganz neue Theorie, die wir im Folgenden entwickeln werden.

Übungen.

9.7. Sei $d \in \mathbb{Z}$ quadratfrei und $A = A_d$. Sei $p \in \mathbb{Z}$ eine Primzahl. Zeige:

(a) Ein Element $a \in A$ mit $N(a) = \pm p$ ist unzerlegbar.

(b) Genau dann ist p in A zerlegbar, wenn es ein $a \in A$ gibt mit $N(a) = \pm p$.

9.8. Sei $d < 0$. Für $d = -1$ ist $\sqrt{d}$ Einheit in $A := A_d$. Zeige, daß für $d < -1$ gilt:

(a) $\sqrt{d}$ ist unzerlegbar.

(b) Falls $|d|$ keine Primzahl ist, ist $\sqrt{d}$ kein Primelement in A, folglich A nicht faktoriell.

9.9. Sei $d \in \mathbb{Z}$ quadratfrei. Zeige:

(a) 2 ist kein Primelement in $\mathbb{Z}[\sqrt{d}]$. ($2 \mid d(d-1)\,!$)

(b) Falls $d \equiv 1 \ (4)$, ist $\mathbb{Z}[\sqrt{d}]$ nicht faktoriell. (Achtung: A_d kann durchaus faktoriell sein.)

(c) Falls $d < 0$, ist $\mathbb{Z}[\sqrt{d}]$ faktoriell genau dann, wenn $d = -1$ oder $d = -2$. Insbesondere ist im Fall $d < 0$, $d \equiv 2, 3 \ (4)$ der Ring A_d nur für $d = -1$ und $d = -2$ faktoriell.

9.10. Sei $d > 0$, $d = pr$, wobei p Primzahl, r teilerfremd zu p. Ferner sei vorausgesetzt, daß weder p noch $-p$ quadratischer Rest mod r. Zeige:

(a) p ist unzerlegbar in $A = A_d$.

(b) p ist kein Primelement in A.

Gib Beispiele an, auf die die Voraussetzungen dieser Aufgabe zutreffen.

ABSCHNITT 10

Die Einheiten quadratischer Zahlkörper

Zum genaueren Studium der Ringe A_d ist es erforderlich, deren Einheiten zu bestimmen. Nur wenn man die Einheiten kennt, kann man z.B. die Assoziierten eines Elements überblicken. Die Einheiten von A_d nennt man auch (mißverständlich) die *Einheiten von* $\mathbb{Q}[\sqrt{d}]$.

Zur Untersuchung von Zerlegungsproblemen in A_d zieht man immer die Norm heran: Für $a, b, c \in A_d$ mit $ab = c$ ist $N(a)N(b) = N(c)$. Also hat man die Teiler von c unter den Elementen zu suchen, deren Norm $N(c)$ teilt. Speziell für $c = 1$ ergibt sich:

Satz 10.1.

(a) *Genau dann ist $z \in A_d$ eine Einheit, wenn $N(z) = \pm 1$.*

(b) *Für $d \equiv 2, 3 \ (4)$, $r, s \in \mathbb{Z}$, ist*

$$r + s\sqrt{d} \text{ Einheit in } A_d \iff r^2 - ds^2 = \pm 1.$$

(c) *Für $d \equiv 1 \ (4)$, $r, s \in \mathbb{Z}$, ist*

$$\frac{r}{2} + \frac{s}{2}\sqrt{d} \text{ Einheit in } A_d \iff r^2 - ds^2 = \pm 4.$$

Beweis. (a) Falls $wz = 1$, ist $N(w)N(z) = 1$, also $N(z) = \pm 1$. Umgekehrt: Falls $N(z) = \pm 1$, ist $zz' = 1$ oder $z(-z') = 1$, also z Einheit.

(b) folgt unmittelbar aus (a) und ebenso (c), wenn man beachtet, daß $r^2 - ds^2 = \pm 4$ bereits $r \equiv s \ (2)$ erzwingt. □

Um die Einheiten in A_d zu bestimmen, haben wir also für $d \equiv 2, 3 \ (4)$ die Lösungsgesamtheit der Gleichungen

$$x^2 - dy^2 = \pm 1, \quad x, y \in \mathbb{Z},$$

im Fall $d \equiv 1 \ (4)$ die Lösungsgesamtheit der Gleichungen

$$x^2 - dy^2 = \pm 4, \quad x, y \in \mathbb{Z},$$

zu bestimmen. Diese Gleichungen heißen *Pellsche Gleichungen*.

Im Fall $d < 0$ scheidet $N(z) < 0$ von vornherein aus. Dieser Fall ist sehr einfach:

Satz 10.2. *Die Einheiten von A_d, $d < 0$, sind:*

(a) $1, i, -1, -i$ *für* $d = -1$,

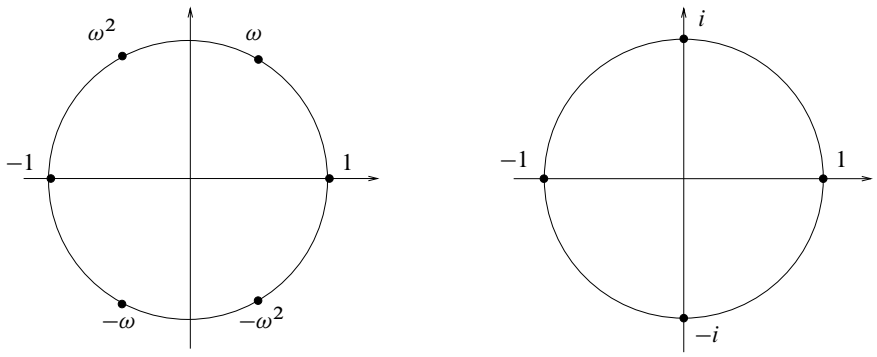

ABBILDUNG 1. Einheiten in A_{-3} und A_{-1}

(b) $1, \omega_{-3}, \omega_{-3}^2, -1, -\omega_{-3}, -\omega_{-3}^2$ *für* $d = -3$,
(c) $1, -1$ *sonst.*

Beweis. In den Fällen $d \equiv 2, 3 \pmod 4$ ergibt sich

$$x^2 - dy^2 = x^2 + |d|y^2 = 1 \iff \begin{cases} x \text{ oder } y = \pm 1, x \cdot y = 0 & \text{falls } d = -1, \\ x = \pm 1, y = 0 & \text{sonst .} \end{cases}$$

Sei $d \equiv 1 \pmod 4$. Dann ist

$$x^2 - dy^2 = x^2 + |d|y^2 = 4 \iff \begin{cases} x = \pm 2, y = 0 & \text{für } d < -4 \\ \left.\begin{matrix} x = \pm 2, y = 0 \\ x = \pm 1, y = \pm 1 \end{matrix}\right\} & \text{für } d = -3. \end{cases}$$

Beachte im Fall $d = -3$: $\omega_{-3}^2 = \omega_{-3} - 1$. □

Im Fall $d = -3$ sind die Einheiten gerade die 6-ten Einheitswurzeln. Im Fall $d = -1$ die 4-ten Einheitswurzeln, wie in Abbildung 1 illustriert.

Im reell-quadratischen Fall liegen die Dinge komplizierter. Vor allem gibt es stets unendlich viele Einheiten. (Die imaginär-quadratischen sind die überhaupt die einzigen Zahlkörper mit endlicher Einheitengruppe.) Mit ε ist auch ε^n für jedes n Einheit in A_d, und für $\varepsilon \in \mathbb{R}$, $\varepsilon \neq \pm 1, 0$ ist die Folge $(|\varepsilon_n|)$ stets streng monoton wachsend oder streng monoton fallend. Also folgt die Existenz unendlich vieler Einheiten bereits aus

Satz 10.3. *Sei $d > 0$. Dann gibt es in A_d eine von ± 1 verschiedene Einheit.*

Den Beweis von Satz 10.3 schieben wir auf. Wir wollen aus ihm zunächst die Struktur der Einheitengruppe vollständig herleiten. Dazu zeigen wir als nächstes

Satz und Definition 10.4. *Unter den Einheiten $\varepsilon \in A_d$ mit $\varepsilon > 1$ gibt es eine kleinste. Diese Einheit heißt* Fundamentaleinheit.

Beweis. Der schwierigste Punkt des Beweises besteht darin zu zeigen, daß es überhaupt eine Einheit $\varepsilon > 1$ in A_d gibt. Diese Aufgabe haben wir im wesentlichen in den Beweis von 10.3 verschoben. Sei also $\tilde{\varepsilon} \neq \pm 1$ eine nach 10.3 existierende Einheit. Dann ist eine der Einheiten

$$\tilde{\varepsilon},\ \frac{1}{\tilde{\varepsilon}},\ -\tilde{\varepsilon},\ -\frac{1}{\tilde{\varepsilon}}$$

größer als 1. Nach Umbenennung dürfen wir annehmen: $\tilde{\varepsilon} > 1$. Die Menge

$$E := \{\varepsilon \text{ Einheit in } A_d : 1 < \varepsilon \leq \tilde{\varepsilon}\}$$

ist nicht leer. Es genügt zu zeigen, daß sie nur endlich viele Elemente enthält.

Sei $\varepsilon \in E$. Dann ist $N(\varepsilon) = \pm 1$, also

$$\varepsilon' = \pm\frac{1}{\varepsilon} \quad \text{und} \quad |\varepsilon'| < 1.$$

Daraus folgt: $|S(\varepsilon')| = |\varepsilon + \varepsilon'| \leq |\varepsilon| + |\varepsilon'| \leq \tilde{\varepsilon} - 1$.

Da $S(\varepsilon') \in \mathbb{Z}$, kommen nur endlich viele Werte für $S(\varepsilon')$ in Frage. Da überdies $N(\varepsilon) = \pm 1$ gilt, kommen nur die endlich vielen Polynome

$$X^2 \pm mX \pm 1, \qquad m \in \mathbb{Z}, \quad 0 \leq m \leq \tilde{\varepsilon} + 1,$$

als Minimalpolynom in Frage. Diese Polynome besitzen nur endlich viele Nullstellen, unter denen sich alle Elemente von E befinden müssen. □

Sei nun ε die Fundamentaleinheit von A_d und u eine weitere Einheit, $u \geq 1$. Da $\varepsilon^n \to \infty$ mit wachsendem n, liegt u in einem der Intervalle

$$[\varepsilon^n, \varepsilon^{n+1}), \quad n \in \mathbb{Z}.$$

Es folgt $1 \leq u/\varepsilon^n < \varepsilon$, und da auch u/ε^n eine Einheit in A_d ist, muß nach 10.4 $u/\varepsilon_n = 1$, also $u = \varepsilon^n$ sein. Ferner liegt für jede Einheit $u \in A_d$ eines der Elemente $u, -u, 1/u, -1/u$ im Intervall $[1, \infty)$ und es folgt

Satz 10.5. *Sei ε die Fundamentaleinheit von A_d. Dann sind die Einheiten von A_d die (paarweise verschiedenen) Elemente $\pm\varepsilon^n$, $n \in \mathbb{Z}$.*

Zu zeigen bleibt Satz 10.3. Zunächst sieht es so aus, als müßten wir in den Fällen $d \equiv 2, 3\ (4)$ und $d \equiv 1\ (4)$ unterschiedlich argumentieren. Wir beweisen aber eine hinreichend starke Aussage, die alle Fälle simultan erfaßt.

Satz 10.6. *Sei $D \in \mathbb{Z}$, $D > 0$, D kein Quadrat. Dann besitzt die diophantische Gleichung*

$$x^2 - Dy^2 = 1$$

stets eine von $(x, y) = (\pm 1, 0)$ verschiedene Lösung.

Eine solche Lösung für $D = d$ liefert uns ein Element $x + y\sqrt{d} \in \mathbb{Z}[\sqrt{d}] \subset A_d$ mit $N(x + y\sqrt{d}) = 1$, $y \neq 0$, und damit eine Einheit $\neq \pm 1$ in A_d.

Der Beweis von 10.6 vollzieht sich in drei Schritten. Nehmen wir an, wir suchten eine rationale Zahl u/v, die die irrationale Zahl $\sqrt{D}$ möglichst gut approximiert. Dazu müssen wir den Betrag der Differenz $u/v - \sqrt{D}$ oder, was auf das gleiche hinausläuft, den Betrag der Differenz $u^2/v^2 - D$ oder den von $u^2 - Dv^2$ möglichst klein machen. Satz 10.6 besagt: Wir können sogar für geeignete u, v den bestmöglichen Wert $u^2 - Dv^2 = 1$ erreichen. Er macht also eine Aussage über die rationale Approximation von $\sqrt{D}$. Der erste Schritt des Beweises von 10.6 ist daher eine Aussage über die rationale Approximation von Irrationalzahlen.

Satz 10.7. *Sei $\delta \in \mathbb{R}$ irrational. Dann existieren zu jedem $n \in \mathbb{N}_+$ ganze Zahlen a, b mit $0 < b \leq n$ und*

$$|b\delta - a| < \frac{1}{n}, \quad \text{äquivalent: } \left|\delta - \frac{a}{b}\right| < \frac{1}{bn}.$$

Beweis. Sei $\delta_k := k\delta - [k\delta]$. Dann gilt $0 \leq \delta_k < 1$. Ferner sind die Zahlen δ_k paarweise verschieden, weil δ irrational ist. Das Intervall $[0, 1)$ ist die disjunkte Vereinigung der n Intervalle

$$I_j = \left[\frac{j}{n}, \frac{j+1}{n}\right), \quad j = 0, \ldots, n-1.$$

Von den $n + 1$ Zahlen δ_k, $k = 0, \ldots, n$, müssen mindestens zwei in dem gleichen Intervall liegen: Es gibt k, l, $0 \leq l < k \leq n$, mit

$$|\delta_k - \delta_l| < \frac{1}{n}.$$

Also

$$|k\delta - [k\delta] - l\delta + [l\delta]| = |(k-l)\delta - ([k\delta] - [l\delta])| < \frac{1}{n}. \qquad \square$$

Im zweiten Schritt wählen wir zu jedem $n \in \mathbb{N}_+$ Zahlen a_n, b_n gemäß 10.7 zu $\delta = \sqrt{D}$ und setzen $\alpha_n := a_n + \sqrt{D}b_n$. Dann ist $\alpha_n \in \mathbb{Q}[\sqrt{D}]$ und

$$\begin{aligned} |N(\alpha_n)| &= |(a_n + \sqrt{D}b_n)(a_n - \sqrt{D}b_n)| = |a_n + \sqrt{D}b_n||a_n - \sqrt{D}b_n| \\ &\leq \frac{1}{n}|a_n + \sqrt{D}b_n| = \frac{1}{n}|a_n - \sqrt{D}b_n + 2\sqrt{D}b_n| \\ &\leq \frac{1}{n}\big(|a_n - \sqrt{D}b_n| + 2\sqrt{D}b_n\big) \leq \frac{1}{n^2} + 2\sqrt{D} \leq 1 + 2\sqrt{D}. \end{aligned}$$

Da $N(\alpha_n)$ stets ganzzahlig ist und nur endlich viele Werte für $N(\alpha_n)$ in Frage kommen, nämlich die ganzen Zahlen $\neq 0$ im Intervall $[-(1 + 2\sqrt{D}), 1 + 2\sqrt{D}]$, existiert eine ganze Zahl m mit

$$N(\alpha_n) = m \text{ für unendlich viele } n \in \mathbb{N}.$$

Der dritte Schritt: Sei $M := \{n \in \mathbb{N} : N(\alpha_n) = m\}$. Die Abbildung $M \to \mathbb{Z}_m \times \mathbb{Z}_m$, $\alpha_n \mapsto (\overline{a}_n, \overline{b}_n)$ nimmt nur endlich viele Werte an. Es gibt also n, n' mit

$$a_n \equiv a_{n'} \ (m), \quad b_n \equiv b_{n'} \ (m).$$

Sei $a := a_n$, $b := b_n$, $\tilde{a} := a_{n'}$, $\tilde{b} := b_{n'}$. Dann ist

$$N\left(\frac{a + b\sqrt{D}}{\tilde{a} + \tilde{b}\sqrt{D}}\right) = \frac{N(a + b\sqrt{D})}{N(\tilde{a} + \tilde{b}\sqrt{D})} = \frac{m}{m} = 1$$

und

$$\frac{a + b\sqrt{D}}{\tilde{a} + \tilde{b}\sqrt{D}} = \frac{(a + b\sqrt{D})(\tilde{a} - \tilde{b}\sqrt{D})}{m} = \frac{(a\tilde{a} - b\tilde{b}D) + (b\tilde{a} - a\tilde{b})\sqrt{D}}{m}.$$

Weiter ist

$$a\tilde{a} - b\tilde{b}D \equiv a^2 - b^2 D \equiv m \equiv 0 \ (m)$$
$$b\tilde{a} - a\tilde{b} \equiv ab - ab \equiv 0 \ (m).$$

Also sind $x := (a\tilde{a} - b\tilde{b}D)/m$ und $y := (b\tilde{a} - a\tilde{b})/m$ ganze Zahlen mit $N(x + y\sqrt{D}) = x^2 - Dy^2 = 1$. Wir müssen noch sicherstellen, daß $x + y\sqrt{D} \neq \pm 1$. Falls $x + y\sqrt{D} = \pm 1$, folgt

$$a_n + b_n\sqrt{D} = \pm(a_{n'} + b_{n'}\sqrt{D}).$$

Da die Folge (a_k/b_k) gegen $\sqrt{D}$ konvergiert, müssen die Nenner b_{k_i} in jeder Teilfolge (a_{k_i}/b_{k_i}) unbeschränkt wachsen. Also kommen unter den Paaren (a_k, b_k), $k \in M$, $a_k \equiv a_n \ (m)$, $b_k \equiv b_n \ (m)$, unendlich viele paarweise verschiedene vor, so daß wir bei der Wahl von n' vermeiden können, daß $a_n + b_n\sqrt{D} = \pm(a'_n + b'_n\sqrt{D})$. Damit ist Satz 10.6 bewiesen.

Bevor wir die mit den Einheiten zusammenhängenden Phänomene weiter diskutieren, brauchen wir zur Produktion von Beispielen erst einmal ein Verfahren zur Bestimmung der Fundamentaleinheit.

Satz 10.8. *Sei $d \in \mathbb{Z}$, d quadratfrei, $d > 1$. Sei*

$$\varepsilon = a + b\sqrt{d} \quad (d \equiv 2, 3 \ (4)) \qquad \textit{bzw.} \qquad \varepsilon = \frac{a + b\sqrt{d}}{2} \quad (d \equiv 1 \ (4))$$

die Fundamentaleinheit von A_d. Dann gilt $a, b > 0$ und b ist die kleinste natürliche Zahl $\neq 0$, für die eine der diophantischen Gleichungen

$$a^2 - db^2 = \begin{cases} \pm 1 & \textit{wenn } d \equiv 2, 3 \ (4), \\ \pm 4 & \textit{wenn } d \equiv 1 \ (4) \end{cases}$$

eine Lösung besitzt.

Beweis. Sei $\alpha > 1$ Einheit in A_d. Dann ist α die größte von den Einheiten

$$\alpha, \frac{1}{\alpha}, -\frac{1}{\alpha}, -\alpha$$

und wenn $\alpha = u + v\sqrt{d}$, so sind diese vier Einheiten durch

$$\pm|u| \pm |v|\sqrt{d}$$

gegeben. Daraus folgt unmittelbar, daß $a, b > 0$. Die Einheiten in A_d, die größer als 1 sind, sind die Potenzen

$$\varepsilon^n = a_n + b_n\sqrt{d} \qquad \text{bzw.} \qquad \varepsilon^n = \frac{a_n + b_n\sqrt{d}}{2}.$$

Im Fall $d \equiv 2, 3 \pmod 4$ hat man

$$\begin{aligned}\varepsilon^n = \varepsilon \cdot \varepsilon^{n-1} &= (a + b\sqrt{d})(a_{n-1} + b_{n-1}\sqrt{d}) \\ &= (aa_{n-1} + bb_{n-1}d) + (ab_{n-1} + a_{n-1}b)\sqrt{d}.\end{aligned}$$

Im Fall $d \equiv 1 \pmod 4$ hat man

$$\varepsilon^n = \varepsilon \cdot \varepsilon^{n-1} = \frac{1}{4}((aa_{n-1} + bb_{n-1}d) + (ab_{n-1} + a_{n-1}b)\sqrt{d}).$$

Im ersten Fall folgt unmittelbar: $b_n = ab_{n-1} + ba_{n-1} > b_{n-1}$. Im zweiten Fall ergibt sich:

$$b_n = \frac{1}{2}(ab_{n-1} + bn_{n-1}) > b_{n-1}, \qquad \text{sobald } a \geq 2.$$

Im Fall $a = 1$ ist notwendig $d = 5$, $b = 1$, was die Behauptung in diesem Fall beweist: Eine kleinere Einheit $\alpha > 1$ als

$$\omega_5 = \frac{1 + \sqrt{5}}{2}$$

kann es in A_5 nicht geben.

In allen anderen Fällen folgt die Behauptung aus der strengen Monotonie der Folge (b_n). (Beachte, daß im Fall $d \equiv 1 \pmod 4$ (4) bei $a^2 - db^2 = \pm 4$ notwendig $a \equiv b \pmod 2$ ist, so daß tatsächlich $\frac{1}{2}(a + b\sqrt{d}) \in A_d$.) □

Beispiele. (a) $d = 5$. Wie bereits gesehen, ist $\varepsilon = 1/2(1 + \sqrt{5})$, $N(\varepsilon) = -1$.
(b) $d = 2 : \varepsilon = 1 + \sqrt{2}$, $N(\varepsilon) = -1$.
(c) $d = 3 : \varepsilon = 2 + \sqrt{3}$, $N(\varepsilon) = 1$.
(d) $d = 7$: Wir probieren der Reihe nach die Werte $b = 1, 2, \ldots$

b	$db^2 + 1$	$db^2 - 1$
1	8	6
2	29	27
3	64	

also $\varepsilon = 8 + 3\sqrt{7}$, $N(\varepsilon) = 1$.

Dieses Verfahren ist zwar sehr einfach, kann aber schon für mäßig große d sehr langwierig sein. Z.B.:

$$d = 31: \quad \varepsilon = 1520 + 237\sqrt{31}$$
$$d = 94: \quad \varepsilon = 2143295 + 221064\sqrt{94}.$$

Es gibt schnellere Verfahren, die die von uns nicht diskutierte Kettenbruchentwicklung von $\sqrt{d}$ benutzen.

Wie wir oben gesehen haben, gibt es quadratische Körper mit $N(\varepsilon) = 1$ und andere mit $N(\varepsilon) = -1$. Bisher ist kein Kriterium bekannt, das auf einfache Weise eine Entscheidung über die Norm der Fundamentaleinheit gestattet. Wir geben zwei Aussagen an, von denen die erste einfach zu beweisen ist:

Satz 10.9. *Sei $d \in \mathbb{Z}$, d quadratfrei, $d > 1$. Falls d einen Primteiler der Form $p \equiv 3 \ (4)$ hat, besitzt die Fundamentaleinheit in A_d die Norm 1.*

Beweis. Wenn $a^2 - db^2 = -1$, so ist -1 quadratischer Rest modulo d und dann auch modulo aller Primteiler von d, was den Fall $d \equiv 2, 3 \ (4)$ erledigt. Im Fall $d \equiv 1 \ (4)$ betrachtet man entsprechend $a^2 - db^2 = -4$. Wieder folgt: -1 ist quadratischer Rest modulo d. □

Satz 10.10. *Wenn $p \in \mathbb{N}$ prim, $p \equiv 1 \ (4)$, so ist -1 die Norm der Fundamentaleinheit in A_p.*

Beweis. Es genügt zu zeigen: Es gibt $u, v \in \mathbb{Z}$ mit $v^2 - pu^2 = -1$. Wir wählen nun $r \in \mathbb{N}$ minimal unter der Bedingung $r^2 - ps^2 = 1$ für ein geeignetes $s \in \mathbb{Z}$. (Dies ist möglich nach 10.6.) Offensichtlich gilt $r \equiv 1 \ (2)$, $s \equiv 0 \ (2)$. Also hat man in $\mathbb{Z}$ die Gleichung

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

Da $(r+1)/2$, $(r-1)/2$ teilerfremd sind, muß einer der folgenden Fälle eintreten:

$$\frac{r+1}{2} = u^2, \ \frac{r-1}{2} = pv^2, \quad \text{oder} \quad \frac{r+1}{2} = pu^2, \ \frac{r-1}{2} = v^2$$

mit $u, v \in \mathbb{Z}$. Im ersten Fall ist $u^2 - pv^2 = 1$ im Widerspruch zur Wahl von r. Im zweiten Fall ergibt sich $v^2 - pu^2 = -1$, was zu erreichen war. □

Mit Satz 10.5 überblickt man natürlich auch die Struktur der Einheitengruppen der Unterringe R von A_d, speziell von $\mathbb{Z}[\sqrt{d}]$ für $d \equiv 1 \ (4)$. Dabei beachte man, daß alle Elemente von R, die Einheiten in A_d sind, auch Einheiten in R sind: mit $\varepsilon \in A_d$ gehören auch ε' und folglich ε^{-1} zu R,

Diese Überlegung und Satz 10.6 zeigen, daß R im Fall $R \neq \mathbb{Z}$ eine von ± 1 verschiedene Einheit enthält. Daher ist die Untergruppe der positiven Einheiten von R eine von $\{1\}$ verschiedene Untergruppe der von der Fundamentaleinheit ε

erzeugten Gruppe der positiven Einheiten von A_d, selbst also wieder unendlich zyklisch und erzeugt von ε^n mit $n = \min\{m : \varepsilon^m \in R\}$. Falls $d \equiv 1\ (4)$ und $R = \mathbb{Z}[\sqrt{d}]$, so läßt sich leicht zeigen, daß $n = 1$ oder $n = 3$ ist (siehe Aufgabe 10.13).

Übungen.

10.11. Bestimme die Fundamentaleinheit von $\mathbb{Q}[\sqrt{19}]$.

10.12. Eine reelle Zahl a heiße eine *quadratische Einheit*, wenn es ein quadratfreies $d \in \mathbb{Z}$ gibt, für das a Einheit in A_d ist.

(a) Zeige: $(1 + \sqrt{5})/2$ ist die kleinste quadratische Einheit > 1.

(b) Bestimme die beiden nächstgrößeren quadratischen Einheiten.

10.13. Sei $d > 0$, $d \equiv 1\ (4)$. Sei η eine Einheit in A_d. Zeige: $\eta^3 \in \mathbb{Z}[\sqrt{d}]$.

10.14. Die Zahlen $\frac{1}{2}y(y+1)$, $y \in \mathbb{Z}$, $y \geq 1$, heißen *Dreieckszahlen*. Bezeichne Δ_n die in aufsteigender Reihenfolge n-te Zahl, die sowohl Dreiecks- als auch Quadratzahl ist. Bestimme eine Rekursionsformel für Δ_n.

Hinweis: Nach geeigneter Umformung führt die Gleichung $x^2 = \frac{1}{2}y(y+1)$ auf eine Gleichung der Form $u^2 - dv^2 = 1$, $d \not\equiv 1\ (4)$. (d ist natürlich zu finden.) Sei $\varepsilon > 1$ erzeugendes Element der Gruppe der positiven Einheiten mit Norm 1 in A_d. (Wenn die Fundamentaleinheit Norm 1 hat, ist ε die Fundamentaleinheit, sonst ihr Quadrat.) Aus $\varepsilon^{n+1} = \varepsilon\varepsilon^n$ ergibt sich (letzten Endes) die gesuchte Rekursionsformel.

10.15. Es ist offensichtlich, daß man in imaginär-quadratischen Zahlkörpern in endlich vielen Schritten testen kann, ob ein ganzes Element mit gegebener Norm n existiert. (Bei $d \equiv 2, 3\ (4)$ braucht man höchstens $\sqrt{n}/d$ Schritte, bei $d \equiv 1\ (4)$ höchstens $4\sqrt{n}/d$ Schritte.) Aber auch für reell-quadratische Zahlkörper ist dies möglich. Sei $d > 0$, $A = A_d$ und η die Fundamentaleinheit von A. Ein Element $a \in A$ ist *in spezieller Lage*, wenn $a > 0$ und $1 \leq |a/a'| < \eta^2$. Zeige:

(a) Zu jedem Element $a \in A$, $a \neq 0$, gibt es genau ein assoziiertes Element $\tilde{a}$ in spezieller Lage.

(b) Falls a in spezieller Lage ist, gilt $|S(a)| < \sqrt{|N(a)|}(\eta + 1)$, a erfüllt also eine Gleichung $a^2 - sa + N(a) = 0$ mit $|s| < \sqrt{|N(a)|}(\eta + 1)$.

ABSCHNITT 11

Euklidische quadratische Körper

Die gesamte Zahlentheorie in $\mathbb{Z}$ basiert auf der Division mit Rest: Die Aussagen über die Primfaktorzerlegung und die Existenz und Darstellbarkeit größter gemeinsamer Teiler folgen aus ihr. Nun war von Anfang an klar, daß die Zahlentheorie in Bereichen, die eine Division mit Rest zulassen, in vieler Hinsicht ähnlich zu der in $\mathbb{Z}$ sein muß, und wir hatten solche Ringe euklidisch genannt. Wir schränken diese Definition zunächst etwas ein:

Definition. Sei $d \in \mathbb{Z}$, d quadratfrei. A_d oder (mißverständlich) $\mathbb{Q}[\sqrt{d}]$ heißt *N-euklidisch*, wenn A_d mit der Funktion $a \mapsto |N(a)|$ ein euklidischer Ring ist.

Beispiele sind $A_1 = \mathbb{Z}$, $A_{-1} = \mathbb{Z}[i]$, $A_{-2} = \mathbb{Z}[i\sqrt{2}]$. Wie im letzten Abschnitt erweist sich auch hier der Fall $d < 0$ als der weitaus einfachere. Sei zunächst d beliebig.

Seien $a, b \in A_d$, $a = qb + r$, $q, r \in A_d$, $b \neq 0$. Es gilt $|N(r)| < |N(b)|$ genau dann, wenn

$$\left| N\left(\frac{a}{b} - q\right) \right| = \left| N\left(\frac{r}{b}\right) \right| = \frac{|N(r)|}{|N(b)|} < 1.$$

Umgekehrt läßt sich jedes Element von $\mathbb{Q}[\sqrt{d}]$ in der Form a/b, $a, b \in A_d$, darstellen und wir erhalten:

Satz 11.1. *Genau dann ist A_d N-euklidisch, wenn zu jedem $c \in \mathbb{Q}[\sqrt{d}]$ ein $q \in A_d$ mit $|N(c - q)| < 1$ existiert.*

Die Bedingung von 11.1 läßt sich im Fall $d < 0$ sehr einfach überprüfen, weil $N(c) = |c|^2$ ist. Für $d \equiv 2, 3$ (4) bilden die Elemente von $A_d = \mathbb{Z}[\sqrt{d}]$ ein Gitter in der komplexen Ebene, wie in Abbildung 1 angedeutet.

Die Punkte mit dem größten Minimalabstand an den Gitterpunkten sind gerade die Mittelpunkte der „Grundmaschen“. Die Bedingung von Satz 11.1 ist genau dann erfüllt, wenn die Diagonale einer Grundmasche eine Länge < 2 hat, also wenn

$$\sqrt{1 + |d|} < 2.$$

Dies ist genau dann erfüllt, wenn $d = -1, -2$. (Allgemeiner haben wir gezeigt: $\mathbb{Z}[\sqrt{d}]$ ist genau dann euklidisch, wenn $d = -1, -2$. (Vgl. auch Aufgabe 9.9.)

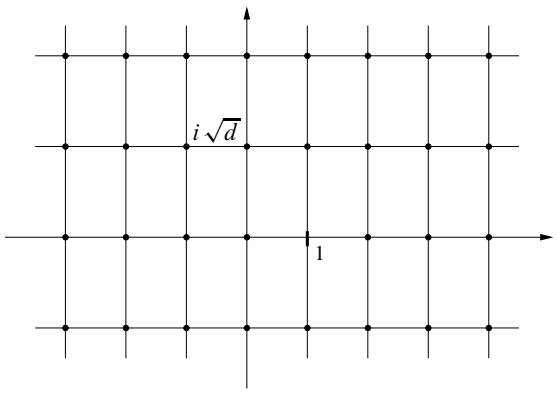

ABBILDUNG 1

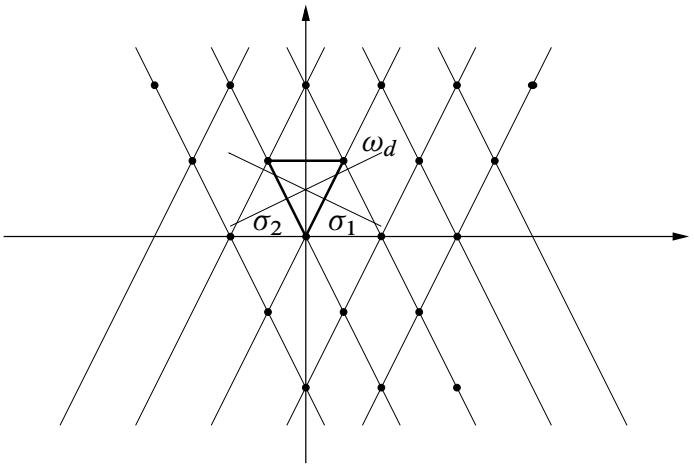

ABBILDUNG 2

Im Fall $d \equiv 1\ (4)$, $A_d = \mathbb{Z}[\omega_d]$, $\omega_d = \frac{1}{2}(1 + i\sqrt{|d|})$, ist die Situation geringfügig komplizierter. Die Punkte von A_d bilden ein Gitter wie in Abbildung 2 angedeutet.

Hier ist die Grundmasche eine Raute. Um die Punkte mit dem maximalen Minimalabstand zu finden, betrachten wir eine Rautenhälfte: ein gleichschenkliges Dreieck mit Basislänge 1 und Schenkellänge

$$\frac{1}{2}\sqrt{1 + |d|}.$$

Der Punkt mit dem maximalen Minimalabstand ist der Umkreismittelpunkt dieses Dreiecks. Er ist gerade der Schnittpunkt der Mittelsenkrechten auf den Dreiecksseiten. Die Gleichung der Mittelsenkrechten σ_1 ist in Punkt-Steigungs-Form (im

x-y-Koordinatensystem)

$$y - \frac{1}{4}\sqrt{|d|} = -\frac{1}{\sqrt{|d|}}\left(x - \frac{1}{4}\right), \quad \text{also } y = \frac{1}{\sqrt{|d|}}\left(x - \frac{1}{4}(|d|) + 1\right),$$

der Schnittpunkt mit der imaginären Achse ist $(0, \frac{1}{4}(\sqrt{|d|} + \frac{1}{\sqrt{|d|}}))$, und der Umkreisradius ist

$$\frac{1}{4}\left(\sqrt{|d|} + \frac{1}{\sqrt{|d|}}\right)$$

Es gilt

$$\begin{aligned}\frac{1}{4}\left(\sqrt{|d|} + \frac{1}{\sqrt{|d|}}\right) < 1 &\iff \frac{1}{16}\left(|d| + 2 + \frac{1}{|d|}\right) < 1 \\ &\iff |d| \le 13 \quad (\text{für } d \in \mathbb{Z}).\end{aligned}$$

Wir erhalten damit die Äquivalenz (a) $\iff$ (b) in

Satz 11.2. *Für $d < 0$ sind äquivalent:*

(a) $d = -1, -2, -3, -7, -11$,
(b) *A_d ist N-euklidisch,*
(c) *A_d ist euklidisch.*

Beweis. Die vielleicht überraschende Verschärfung (c) $\Rightarrow$ (a) von (b) $\Rightarrow$ (a) ist nicht schwer einzusehen. Wir benötigen allerdings ein Argument, das erst im folgenden Abschnitt bereitgestellt wird.

Sei A_d euklidisch bezüglich einer Funktion φ. Unter allen Nichteinheiten $\neq 0$ wählen wir $z \in A_d$ so, daß $\varphi(z)$ minimal ist. Sei $w \in A_d$. Dann ist

$$w = qz + r \qquad \text{mit } r = 0 \text{ oder } \varphi(r) < \varphi(z). \tag{$*$}$$

Nach Wahl von z folgt: $r = 0$ oder r ist Einheit. Sei $\mathfrak{a} = A_d z$ das von z erzeugte Ideal in A_d und $\pi : A_d \to A_d/\mathfrak{a}$ der natürliche Epimorphismus. Die Aussage $(*)$ impliziert: $\pi(w) = \pi(r)$ mit einer Einheit r oder $r = 0$.

Wir nehmen an, daß d keiner der genannten Werte ist und leiten einen Widerspruch ab. Die Einheiten in A_d sind dann $+1$ und -1. Also

$$\pi(w) \in \{\pi(0), \pi(1), \pi(-1)\}$$

für alle $w \in A_d$ und somit $|A_d/\mathfrak{a}| \le 3$. Wie wir im Abschnitt 12 zeigen werden, gilt

$$|A_d/\mathfrak{a}| = N(z),$$

woraus $N(z) \le 3$ folgt. Im Fall $d \equiv 2, 3 \ (4)$, $|d| > 3$, ist

$$N(z) = N(a + b\sqrt{d}) = a^2 + b^2|d| \le 3 \iff |a| = 1,\ b = 0 \quad (z \neq 0)$$

Im Fall $d \equiv 1 \ (4)$, $|d| > 12$, ist

$$N(z) = N\left(\frac{a + b\sqrt{d}}{2}\right) = \frac{1}{4}a^2 + \frac{b^2}{4}|d| \leq 3 \iff |a| = 2,\ b = 0.$$

Stets ist $z = \pm 1$ – im Widerspruch zur Wahl von z. □

Eine schöne Anwendung der Euklidizität von A_{-3} ist der Beweis, daß die Fermat-Gleichung $x^3 + y^3 = z^3$ nur die triviale Lösung für $x, y, z \in \mathbb{Z}$ hat; siehe dazu [**AdGo**].

Der reell-quadratische Fall ist erheblich komplizierter. Wir erwähnen hier ohne Beweis:

Satz 11.3 (Chatland-Davenport, 1950)**.** *Für $d > 0$ ist A_d genau dann N-euklidisch, wenn*

$$d = 2, 3, 5, 6, 7, 11, 13, 17, 19, 21, 29, 33, 37, 41, 57, 73$$

In [**HaWr**], Theorem 248, findet man einen einheitlichen Beweis für $d = 2, 3, 5, 6, 7, 11, 13, 17, 21, 29$. Ferner liefert [**HaWr**], Theorem 249 (Beweis) auf einfache Weise die Information, daß A_d nicht N-euklidisch ist für

$$d \geq \begin{cases} 50 & \text{wenn } d \equiv 2 \ (4), \\ 90 & \text{wenn } d \equiv 3 \ (4). \end{cases}$$

Siehe dazu auch [**Leut**], Abschnitt 8.4. Die Sätze 11.2 und 11.3 zeigen, daß N-Euklidizität eine ganz selten auftretende Eigenschaft in quadratischen Körpern ist. Es gibt für $d > 0$ Ringe A_d gibt, die zwar nicht bezüglich der Norm, wohl aber bezüglich einer anderen geeigneten Funktion euklidisch sind. Bekannt ist dies für A_{69} (D. Clark, *A quadratic field which is Euclidean but not norm-Euclidean*, Manuscr. Math. **83** 327–330 (1994)). Der Kandidat mit kleinstmöglichem d ist A_{14} (siehe Aufgabe 11.10).

Wir wissen, daß euklidische Ringe faktoriell sind, d. h. daß sich in ihnen jede Nichteinheit $a \neq 0$ (auf genau eine Weise) als Produkt von Primelementen schreiben läßt. Es gibt bereits unter den Ringen A_d faktorielle Ringe, die nicht euklidisch sind. Für $d < 0$ sind diese Ringe vollständig bekannt:

Satz 11.4. *Außer den in* 11.2 *genannten Fällen ist A_d bei $d < 0$ faktoriell nur noch für*

$$d = -19, -43, -67, -163.$$

Der Beweis von 11.4 wurde 1967 von Stark abgeschlossen. Zwar sind die $d > 0$ mit faktoriellen A_d immer noch nicht vollständig bekannt – bereits Gauß hat vermutet, daß es unendlich viele solcher d gibt –, aber es ist dennoch gerechtfertigt zu sagen, daß auch die faktoriellen A_d eine „dünne“ Menge bilden. Insofern ist

es notwendig, eine völlig neue Teilbarkeitstheorie zu entwickeln, die allen A_d gerecht wird. Die Gegenstände dieser Teilbarkeitstheorie sind nicht mehr Elemente, sondern Ideale. Nachdem wir die Teilbarkeitstheorie für Ideale entwickelt haben, kommen wir noch einmal auf die Teilbarkeitstheorie für Elemente zurück.

Interessanterweise kann man für die in Satz 11.4 genannten Ringe auch elementar nachweisen, daß sie faktoriell sind. Wir führen dies für $d = -19$ aus. Die anderen Fälle erfordern einen höheren Aufwand. Wenn die Norm in diesen Ringen auch nicht mehr die Euklidizität sichert, so reicht sie dennoch aus, diese Ringe zu Hauptidealringen zu machen.

Es ist aus der Algebra bekannt, daß alle nullteilerfreien Hauptidealringe faktoriell sind. Weil wir dies in Abschnitt 1 aber nicht explizit bewiesen haben, erklären wir an dieser Stelle wenigstens, weshalb die Hauptidealringe unter den A_d faktoriell sind:

Satz 11.5. *Die Hauptidealringe unter den A_d sind faktoriell.*

Beweis. Zunächst einmal sieht man mittels Induktion über $|N(a)|$, daß sich jede Nichteinheit $a \neq 0$ als Produkt irreduzibler Elemente schreiben läßt. (Im Fall eines beliebigen Hauptidealringes muß man dies mit Hilfe des Zornschen Lemmas beweisen.) Außer dieser Tatsache haben wir beim Beweis, daß euklidische Ringe faktoriell sind, nur noch die Existenz von $\mathrm{ggT}(a, b)$ und seine Darstellung als Linearkombination von a und b benutzt. Das hat man aber auch in Hauptidealringen R: Für $a, b \in R$ ist jedes erzeugende Element von $Ra + Rb$ ein ggT. □

Man benutzt nun folgendes Kriterium.

Satz 11.6. *Ein Integritätsbereich R ist genau dann ein Hauptidealring, wenn es eine Funktion $\varphi : R \to \mathbb{N}$ gibt mit folgenden Eigenschaften: Zu jedem $b \in R$, $b \neq 0$, und jedem $a \in R$, $b \nmid a$, existieren $u, v \in R$, so daß*

$$\varphi(0) < \varphi(au - vb) < \varphi(b)$$

Beweis. „⇒“ Jeder Hauptidealring ist faktoriell. Wir setzen $\varphi(0) = 0$, $\varphi(r) = 1$, wenn r Einheit und $\varphi(r) = n + 1$, wenn in der Primfaktorzerlegung von r genau n Faktoren vorkommen. Man wählt nun u, v so, daß $d = ua + (-v)b$ das Ideal $Ra + Rb$ erzeugt: Da d ein ggT von a und b ist und sogar ein echter Teiler von b, folgt $\varphi(0) < \varphi(d) < \varphi(b)$.

„⇐“ Sei $\mathfrak{a} \subset R$ ein Ideal, $\mathfrak{a} \neq 0$. Dann wählt man $b \in \mathfrak{a}$ so, daß

$$\varphi(b) = \min\{\varphi(c) : c \in \mathfrak{a}, c \neq 0\}.$$

Sei $a \in I$ beliebig. Da auch $ua - vb \in I$ für alle $u, v \in R$, erzwingt die Wahl von b, daß der Fall $b \nmid a$ nicht eintreten kann. Es folgt $I = Rb$. □

Wir zeigen nun, daß die Norm in $A = A_{-19}$ die Bedingungen von 11.6 erfüllt. Wegen der Multiplikativität der Norm kann man die Bedingung in 11.6 so umformulieren:

$$0 < N\left(u\frac{a}{b} - v\right) < 1$$

Sei $\omega = \omega_{-19}$ und

$$\frac{a}{b} = \left(p' + \frac{p}{q}\right) + \left(r' + \frac{r}{s}\right)\omega$$

mit $p', r', p, q, r, s \in \mathbb{Z}$, $0 \le p < q$, $0 \le r < s$, p/q, r/s gekürzt. Dabei sei noch $q = 1$, falls $p = 0$, und $s = 1$, falls $r = 0$.

Wenn wir $u', v' \in A_{-19}$ gefunden haben, für die

$$0 < N\left(u'\left(\frac{p}{q} + \frac{r}{s}\omega\right) - v'\right) < 1,$$

sind wir fertig: Wähle $u = v'$, $v = v' - u'(p' + r'\omega)$. Wir dürfen also annehmen: $a/b = p/q + (r/s)\omega$.

1. Wenn $r = 0$ wählen wir $u = 1$, $v = 0$. Im folgenden sei $r > 0$.
2. Wenn $q \nmid s$, wählen wir $u = s$. Dann ist

$$N\left(u\frac{a}{b}\right) = N\left(\frac{sp}{q} + r\omega\right) = N\left(m + \frac{w}{q} + r\omega\right) \quad m \in \mathbb{Z},\ 0 \le w < q.$$

Da $sp \not\equiv 0\ (q)$, ist $w \neq 0$, und wir können $v = m + r\omega$ wählen. Es folgt

$$N\left(u\frac{a}{b} - v\right) = N\left(\frac{w}{q}\right) \in (0, 1).$$

3. Sei q ein Teiler von s, $s > 2$. Es existiert ein m mit $mr \equiv 1\ (s)$. Mit $u = m$ gilt

$$u\frac{a}{b} = \frac{mp}{q} + \frac{mr}{s}\omega = \left(l + \frac{r_1}{r_2}\right) + \left(k + \frac{1}{s}\right)\omega$$

$$\text{mit } l, k, r_1, r_2 \in \mathbb{Z},\ \left|\frac{r_1}{r_2}\right| \le \frac{1}{2}.$$

Wir wählen $v = l + k\omega$. Damit gilt:

$$\begin{aligned} 0 \neq N\left(u\frac{a}{b} - v\right) &= N\left(\frac{r_1}{r_2} + \frac{1}{s}\frac{1 + \sqrt{-19}}{2}\right) \\ &= \left(\frac{r_1}{r_2} + \frac{1}{2s}\right)^2 + \frac{19}{4s^2} = \left(\frac{r_1}{r_2}\right)^2 + \frac{r_1}{r_2 s} + \frac{20}{4s^2} \\ &\le \frac{1}{4} + \frac{1}{6} + \frac{20}{36} = \frac{35}{36}. \end{aligned}$$

4. Sei nun $s = 2$, $q \mid s$. Dann ist (a) $a/b = \omega/2$ oder (b) $a/b = (1+\omega)/2$. Wir wählen im Fall (a) $u = 1+\omega$, $v = -2+\omega$, und im Fall (b) $u = \omega$, $v = -2+\omega$. Dann ist

$$N\left(u\frac{a}{b} - v\right) = N\left(-\frac{1}{2}\right) = \frac{1}{4}.$$

In den Fällen $d = -43, -67, -163$ geht man prinzipiell genauso vor. Der wesentliche Unterschied ist, daß unter 4. erheblich mehr Fälle auftreten.

Bemerkung 11.7. Wir werden später zeigen, daß A_d genau dann faktoriell ist, wenn es Hauptidealring ist. Wenn A_d Hauptidealring ist, folgt leicht, daß $|N|$ die Bedingung von Satz 11.6 erfüllt: Wähle $au - vb = \mathrm{ggT}(a,b)$. Insgesamt folgt: A_d ist faktoriell genau dann, wenn $|N|$ die Bedingung von Satz 11.6 erfüllt.

Übungen.

11.8. Zeige, daß A_2, A_3, A_5 und A_{13} euklidisch sind.

11.9. Führe den oben gegebenen Beweis für die Hauptidealringeigenschaft von A_{-19} für A_{-43} durch. (Es treten zwar mehr Fälle in Schritt 4. auf, sie sollten sich aber noch übersehen lassen.)

11.10. Zeige, daß A_{14} nicht N-euklidisch ist und zwar in folgenden Schritten:

(a) Zeige, daß keine der sechs Gleichungen $x^2 - 14y^2 = \pm 1, \pm 2 \pm 3$ eine Lösung (x, y) besitzt, bei der sowohl x als auch y ungerade ist. Rechne dazu modulo 8 und modulo 3.

(b) Zeige, daß kein $a \in A_{14}$ existiert mit $N\left(\frac{1+\sqrt{14}}{2} - a\right) < 1$.

(Dennoch ist A_{14} faktoriell, und es wird vermutet, daß A_{14} bezüglich einer geeigneten Funktion euklidisch ist.)

ABSCHNITT 12

Ideale in quadratischen Körpern

Wir erinnern zunächst an die Definition des Ideals. Dabei kann R ein beliebiger kommutativer Ring mit Einselement sein: Eine Teilmenge $\mathfrak{a} \subset R$ heißt Ideal, wenn für alle $a, b \in \mathfrak{a}$, $r \in R$ gilt: $a + b \in \mathfrak{a}$, $ra \in \mathfrak{a}$.

Es folgt sofort, daß $\mathfrak{a}$ bezüglich der Addition eine Gruppe ist. Die wichtigsten Verknüpfungen zwischen Idealen sind Addition, Multiplikation und Durchschnitt:

Satz 12.1. *Seien $\mathfrak{a}$, b Ideale in R.*

(a) *Dann sind die Teilmengen*

$$\mathfrak{a} + \mathfrak{b} := \{a + b : a \in \mathfrak{a},\ b \in \mathfrak{a}\}$$
$$\mathfrak{a} \cdot \mathfrak{b} := \{a_1 b_1 + \cdots + a_n b_n : a_i \in \mathfrak{a},\ b_i \in b,\ i = 1, \ldots, n,\ n \in \mathbb{N}_+\}$$
$$\mathfrak{a} \cap \mathfrak{b}$$

Ideale in R.

(b) *Addition und Multiplikation von Idealen sind assoziativ und kommutativ:*

$$\mathfrak{a} + (\mathfrak{b} + \mathfrak{c}) = (\mathfrak{a} + \mathfrak{b}) + \mathfrak{c} \qquad \mathfrak{a}(\mathfrak{b}\mathfrak{c}) = (\mathfrak{a}\mathfrak{b})\mathfrak{c}$$
$$\mathfrak{a} + \mathfrak{b} = \mathfrak{b} + \mathfrak{a} \qquad \mathfrak{a}\mathfrak{b} = \mathfrak{b}\mathfrak{a}.$$

Ferner gilt das Distributivgesetz $\mathfrak{a}(\mathfrak{b} + \mathfrak{c}) = \mathfrak{a}\mathfrak{b} + \mathfrak{a}\mathfrak{c}$.

(b) *Addition und Multiplikation sind mit der Inklusion verträglich: Falls $\mathfrak{a} \subset \mathfrak{b}$, so*

$$\mathfrak{a} + \mathfrak{c} \subset \mathfrak{b} + \mathfrak{c} \quad \textit{und} \quad \mathfrak{a}\mathfrak{c} \subset \mathfrak{b}\mathfrak{c}.$$

Ideale sind diejenigen Teilmengen von R, nach denen man Restklassenringe bilden kann. Für $r \in R$ sei

$$\overline{r} := \{r + a : a \in \mathfrak{a}\}$$

die *Restklasse* von r nach $\mathfrak{a}$, und $R/\mathfrak{a}$ sei die Menge dieser Restklassen. Es gilt bekanntlich: $R/\mathfrak{a}$ ist ein Ring derart, daß die natürliche Abbildung $R \to R/\mathfrak{a}$, $r \to \overline{r}$, ein Ringhomomorphismus ist.

Ausführlich Gebrauch haben wir von dieser Aussage in den ersten Abschnitten gemacht, als wir die Restklassenringe $\mathbb{Z}/n\mathbb{Z}$ von $\mathbb{Z}$ untersucht haben. Die Ideale in $\mathbb{Z}$ waren sämtlich von der Form $n\mathbb{Z}$, also Hauptideale. Nun interessieren wir uns für die Ideale in den Ringen A_d, die wir – mißverständlich – auch Ideale von

$\mathbb{Q}[\sqrt{d}]$ nennen. Wir wollen uns zunächst überlegen, wie sich die Ideale in A_d erzeugen lassen:

Definition. Sei $\mathfrak{a}$ ein Ideal im Ring R. Wir sagen, die Elemente $a_1, \ldots, a_n \in \mathfrak{a}$ *erzeugen* $\mathfrak{a}$ (oder bilden ein *Erzeugendensystem von* $\mathfrak{a}$), wenn sich jedes Element $a \in \mathfrak{a}$ in der Form

$$a = r_1 a_1 + \cdots + r_n a_n$$

mit $r_i \in R$, $i = 1, \ldots, n$, darstellen läßt.

Beispiele. (a) Das Ideal $2\mathbb{Z}$ in $\mathbb{Z}$ wird von 2 erzeugt, aber auch $6, 8$ ist ein Erzeugendensystem.

(b) Eine Einheit e erzeugt das *Einsideal* $R = 1R = eR$.

(c) In $A = A_{-5}$ betrachten wir das von 2, $1 + \sqrt{-5}$ erzeugte Ideal $\mathfrak{a}$. Zunächst ist $\mathfrak{a}$ nicht das Einsideal. Aus einer Darstellung

$$1 = a2 + b(1 + \sqrt{-5}) \qquad \text{mit } a, b \in A$$

würde nämlich folgen, daß $1 + \sqrt{-5}$ eine Einheit modulo dem von 2 erzeugten Ideal ist. Dies aber ist ausgeschlossen, denn

$$(1 + \sqrt{-5})^2 = 1 + 2\sqrt{-5} - 5 = -4 + 2\sqrt{-5} = 2 \cdot (-2 + \sqrt{-5}) \in 2A.$$

Also $\mathfrak{a} \neq A$. Die Restklassen von A modulo 2, also die Elemente von $A/2A$, werden repräsentiert durch

$$0, 1, \sqrt{-5}, 1 + \sqrt{-5},$$

die Elemente von $A/\mathfrak{a}$ durch 0 und 1: $A/\mathfrak{a}$ ist ein Körper, $\mathfrak{a}$ ein Primideal, sogar ein maximales Ideal. Jedoch ist $\mathfrak{a}$ kein Hauptideal, denn für $c \in \mathfrak{a}$ mit $2 = ac$, $1 + \sqrt{5} = bc$ müßte gelten

$$N(c) \mid N(2) = 4 \quad \text{und} \quad N(c) \mid N(1 + \sqrt{5}) = 6,$$

also $N(c) = 2$, denn $N(c) = 1$ ist wegen $\mathfrak{a} \neq A$ ausgeschlossen. Es gibt aber in A kein Element der Norm 2.

Wir können folgern: A_{-5} ist kein Hauptidealring. Wir wissen bereits, daß A_{-5} nicht faktoriell ist. Später wird sich zeigen, daß die Eigenschaften „faktoriell“ und „Hauptidealring“ für die Ringe A_d äquivalent sind.

Andererseits besitzt jedes Ideal in einem der Ringe A_d ein Erzeugendensystem aus höchstens zwei Elementen. Dies ergibt sich als Folgerung aus einer wesentlich schärferen Aussage, bei der wir nur Linearkombinationen mit Koeffizienten aus $\mathbb{Z}$ betrachten. (Wegen $\mathbb{Z} \subset A_d$ ist eine $\mathbb{Z}$-Basis eines Ideals $\mathfrak{a}$ in A_d natürlich stets ein Erzeugendensystem, die Umkehrung dagegen ist falsch: 2 ist zwar ein Erzeugendensystem von $2A_{-1}$, aber keine $\mathbb{Z}$-Basis.)

Satz 12.2. *Sei $A = A_d$ für ein quadratfreies $d \in \mathbb{Z}$, $\mathfrak{a} \neq 0$ ein Ideal in A. Dann gilt:*

(a) *$\mathfrak{a}$ besitzt eine $\mathbb{Z}$-Basis a_1, a_2 mit $a_1 \in \mathbb{Z}$, wobei $\mathfrak{a} \cap \mathbb{Z} = \mathbb{Z}a_1$.*

(b) *Jede $\mathbb{Z}$-Basis von $\mathfrak{a}$ besitzt genau zwei Elemente.*

Beweis. (a) Sei $a \in \mathfrak{a}$, $a \neq 0$. Dann ist $aa' = N(a) \in \mathbb{Z} \cap \mathfrak{a}$. Also ist $\mathbb{Z} \cap \mathfrak{a} \neq 0$, und $\mathbb{Z} \cap \mathfrak{a}$ ist ein Ideal in $\mathbb{Z}$: Es gilt $\mathbb{Z} \cap \mathfrak{a} = a_1\mathbb{Z}$, $a_1 \neq 0$.

Das Element $a_2 = \alpha_2 + \beta_2\omega_d \in \mathfrak{a}$, $\alpha_2, \beta_2 \in \mathbb{Z}$, wählen wir so, daß

$$|\beta_2| = \min\{|\beta| : \alpha + \beta\omega_d \in \mathfrak{a},\ \beta \neq 0\}.$$

Da z.B. $a_1\omega_d \in \mathfrak{a} \setminus \mathbb{Z}$, ist β_2 sinnvoll definiert.

Damit gilt $a_1 \in \mathbb{Q}$, $a_2 \notin \mathbb{Q}$, $a_1 \neq 0$. Die Elemente a_1, a_2 müssen als Elemente des $\mathbb{Q}$-Vektorraums $\mathbb{Q}[\sqrt{d}]$ linear unabhängig sein, bilden also eine Basis dieses Vektorraums. Zu jedem Element $a \in \mathbb{Q}[\sqrt{d}]$, speziell zu jedem $a \in \mathfrak{a}$, existieren eindeutig bestimmte $q_1, q_2 \in \mathbb{Q}$ mit

$$a = q_1a_1 + q_2a_2.$$

Zu zeigen bleibt, daß $q_1, q_2 \in \mathbb{Z}$ für $a \in \mathfrak{a}$. Sei $a = \alpha + \beta\omega_d$ und $\beta = t_2\beta_2 + u_2$, $t_2, u_2 \in \mathbb{Z}$, $0 \leq u_2 <| \beta_2 |$. Dann ist

$$a - t_2a_2 = (\alpha - t_2\alpha_2) + u_2\omega_d \in \mathfrak{a}.$$

Nach Wahl von a_2 muß $u_2 = 0$ gelten: $a - t_2a_2 \in \mathbb{Z} \cap \mathfrak{a}$ und $a - t_2a_2 = t_1a_1$ mit $t_1 \in \mathbb{Z}$:

$$a = t_1a_1 + t_2a_2.$$

(b) Die Elemente einer $\mathbb{Z}$-Basis von $\mathfrak{a}$ sind linear unabhängig im $\mathbb{Q}$-Vektorraum $\mathbb{Q}[\sqrt{d}]$. Daher kann eine $\mathbb{Z}$-Basis höchstens zwei Elemente haben. Sie kann andererseits auch nicht aus nur einem Element bestehen, sonst wären ja die Elemente a_1, a_2 gemäß (a) als Elemente eines 1-dimensionalen $\mathbb{Q}$-Vektorraums linear unabhängig über $\mathbb{Q}$. □

Wenn a_1, a_2 eine $\mathbb{Z}$-Basis des Ideals $\mathfrak{a}$ ist, schreiben wir $\mathfrak{a} = \mathbb{Z}a_1 + \mathbb{Z}a_2$. Wir erhalten als unmittelbare Folgerung aus 12.2:

Satz 12.3. *Ein Ideal in einem der Ringe A_d besitzt ein Erzeugendensystem aus höchstens zwei Elementen.*

Es ist nützlich, die Beschreibung der im Beweis von Satz 12.2 gefundenen $\mathbb{Z}$-Basis noch etwas zu präzisieren:

Satz 12.4. *Sei $\mathfrak{a}$ ein Ideal in A_d mit $\mathbb{Z}$-Basis $a_1 = \alpha_1$, $a_2 = \alpha_2 + \beta_2\omega_d$. Dann gilt : $\beta_2 \mid \alpha_1, \alpha_2$. Daher ist $\mathfrak{a} = \beta_2\mathfrak{b}$, wobei $\mathfrak{b}$ die $\mathbb{Z}$-Basis $b_1 = \gamma_1$, $b_2 = \gamma_2 + \omega_d$ besitzt, $\gamma_i = \alpha_i/\beta_2$, $i = 1, 2$.*

Beweis. Mit α_1 gehört auch $\alpha_1\omega_d$ zu $\mathfrak{a}$. Folglich existieren $\varepsilon_1, \varepsilon_2 \in \mathbb{Z}$ mit

$$\alpha_1\omega_d = \varepsilon_1\alpha_1 + \varepsilon_2(\alpha_2 + \beta_2\omega_d).$$

Der Koeffizientenvergleich ergibt

$$\alpha_1 = \varepsilon_2\beta_2 \qquad \text{und} \qquad \alpha_2 = -\varepsilon_1\beta_2.$$

Damit ist die Teilbarkeitsaussage bewiesen. Wenn wir nun $\mathfrak{b} = \mathbb{Z}\gamma_1 + \mathbb{Z}(\gamma_2 + \omega_d)$ setzen, folgt erstens, daß $\mathfrak{a} = \beta_2\mathfrak{b}$ und zweitens, daß $\mathfrak{b} = (1/\beta_2)\mathfrak{a}$ auch wirklich ein Ideal ist. □

Wenn ein Erzeugendensystem $b_1, \ldots, b_n$ eines Ideals $\mathfrak{a} \subset A_d$ gegeben ist, kann man eine $\mathbb{Z}$-Basis folgendermaßen systematisch konstruieren: Für $i = 1, \ldots, n$ sei $b_{n+i} := \omega_d b_i$; dann bilden $b_1, \ldots, b_{2n}$ ein $\mathbb{Z}$-Erzeugendensystem von $\mathfrak{a}$. Sei

$$b_j = \gamma_j + \delta_j\omega_d, \quad j = 1, \ldots, 2n, \quad \gamma_j,\ \delta_j \in \mathbb{Z},$$

und

$$\beta_2 = \mathrm{ggT}(\delta_1, \ldots, \delta_{2n}).$$

Der größte gemeinsame Teiler besitzt eine Darstellung

$$\beta_2 = \varepsilon_1\delta_1 + \cdots + \varepsilon_{2n}\delta_{2n} \quad \text{mit } \varepsilon_j \in \mathbb{Z}.$$

Sei dann

$$a_2 = \varepsilon_1 b_1 + \cdots + \varepsilon_{2n} b_{2n} = \alpha_2 + \beta_2\omega_d.$$

Dann subtrahiert man nacheinander:

$$b'_j := b_j - \frac{\delta_j}{\beta_2} a_2 \in \mathbb{Z} \cap \mathfrak{a}.$$

Schließlich ist $a_1 := \mathrm{ggT}(b'_1, \ldots, b'_{2n})$. Man prüft nun leicht nach, daß a_1 und a_2 linear unabhängig und $b_1, \ldots, b_n$ $\mathbb{Z}$-Linearkombinationen von ihnen sind.

Beispiel. Sei $d = -3$, und $\mathfrak{a}$ werde von $4 + 6\omega$, $8 + 3\omega$, $6 + 9\omega$ erzeugt, $\omega = \omega_{-3} = (1 + \sqrt{3}/2)$. Dann ist $\omega^2 = \omega - 1$. Es gilt

$$\begin{aligned} b_1 &= 4 + 6\omega, & b_4 &= b_1\omega = -6 + 10\omega \\ b_2 &= 8 + 3\omega, & b_5 &= b_2\omega = -3 + 11\omega \\ b_3 &= 6 + 9\omega, & b_6 &= b_3\omega = -9 + 15\omega \end{aligned}$$

Es folgt $\beta_2 = 1$ .Mit $\varepsilon_2 = 2$, $\varepsilon_5 = -1$, $\varepsilon_i = 0$ sonst ergibt sich $a_2 = 11 + \omega$ und

$$b'_1 = -62, \quad b'_2 = -25.$$

Damit bereits muß $a_1 = 1$ sein: 1, $11 + \omega$ ist eine Basis von $\mathfrak{a}$, und wir sehen daß $\mathfrak{a} = A_{-3}$. (Man kann nun natürlich a_2 noch abändern durch Subtraktion von Vielfachen von a_1.)

Mittels einer $\mathbb{Z}$-Basis kann man eine wichtige Invariante eines Ideals bestimmen. Dazu betrachten wir zwei Basen a_1, a_2 und b_1, b_2. Dann existieren $\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}$ und $\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22} \in \mathbb{Z}$ mit

$$a_i = \alpha_{ij} b_1 + \alpha_{i2} b_2 \quad \text{und} \quad b_i = \beta_{i1} a_1 + \beta_{i2} a_2, \qquad i = 1, 2.$$

In Matrizenschreibweise mit $A = \left(\begin{smallmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & a_{22} \end{smallmatrix}\right)$, $B = \left(\begin{smallmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{smallmatrix}\right)$ folgt

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = A \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = B \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \text{ und } \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = AB \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}.$$

Da a_1, a_2 über $\mathbb{Q}$ linear unabhängig sind, folgt $AB = E_2$ und damit

$$\det AB = \det E_2 = 1, \quad \text{also} \quad \det A = \det B = \pm 1,$$

denn $\det A, \det B \in \mathbb{Z}$. Mit den Konjugierten a_1', a_2', b_1', b_2' ergibt sich

$$\begin{pmatrix} a_1 & a_1' \\ a_2 & a_2' \end{pmatrix} = A \begin{pmatrix} b_1 & b_1' \\ b_2 & b_2' \end{pmatrix}.$$

Folglich ist

$$\det \begin{pmatrix} a_1 & a_1' \\ a_2 & a_2' \end{pmatrix} = \pm \det \begin{pmatrix} b_1 & b_1' \\ b_2 & b_2' \end{pmatrix}$$

und

$$\left(\det \begin{pmatrix} a_1 & a_1' \\ a_2 & a_2' \end{pmatrix} \right)^2 = \left(\det \begin{pmatrix} b_1 & b_1' \\ b_2 & b_2' \end{pmatrix} \right)^2.$$

Definition. Sei $\mathfrak{a}$ ein Ideal in A_d, a_1, a_2 eine $\mathbb{Z}$-Basis von $\mathfrak{a}$. Dann heißt

$$\Delta(\mathfrak{a}) := \det \begin{pmatrix} a_1 & a_1' \\ a_2 & a_2' \end{pmatrix}^2$$

die *Diskriminante* von $\mathfrak{a}$.

Diese Definition ist sinnvoll, denn wie wir zuvor gesehen haben, ist die Diskriminante unabhängig von der Wahl der Basis. Die Diskriminante von A_d heißt auch *Diskriminante von* $\mathbb{Q}[\sqrt{d}]$. Es gilt

$$\Delta(A_d) = \det \begin{pmatrix} 1 & 1 \\ \omega_d & \omega_d{}' \end{pmatrix}^2 = (\omega_d{}' - \omega_d)^2 = \begin{cases} 4d & \text{für } d \equiv 2, 3 \ (4) \\ d & \text{für } d \equiv 1 \ (4) \end{cases}$$

Es ist häufig günstiger, die quadratischen Körper nach ihren Diskriminanten zu ordnen, anstatt nach den sie definierenden quadratfreien Zahlen $d \in \mathbb{Z}$.

Wir wenden uns nun den Restklassenringen der Ringe A_d zu:

Satz 12.5. *Sei* $\mathfrak{a} \neq 0$ *ein Ideal in* $A = A_d$.

(a) *Dann besitzt* $A/\mathfrak{a}$ *nur endlich viele Elemente. Genauer gilt:*

(b) *Wenn a_1, a_2 mit $a_1 = \alpha_1 \in \mathbb{Z}$ und $a_2 = \alpha_2 + \beta_2\omega_d$ eine $\mathbb{Z}$-Basis von $\mathfrak{a}$ ist, so besitzt $A/\mathfrak{a}$ genau $|\alpha_1||\beta_2|$ Elemente, die von*

$$\alpha + \beta\omega_d, \quad 0 \le \alpha < |\alpha_1|, \quad 0 \le \beta < |\beta_2|$$

repräsentiert werden.

Beweis. (a) folgt direkt aus (b). Zu (b): Sei $c = \gamma + \delta\omega \in A$, $\omega := \omega_d$. Mittels Division mit Rest in $\mathbb{Z}$ wählt man $\zeta \in \mathbb{Z}$ so, daß

$$\beta = \zeta\beta_2 + \vartheta \quad \text{mit} \quad 0 \le \zeta < |\beta_2|$$

und setzt $c' = c - \zeta a_2 = \gamma' + \delta'\omega$. Anschließend wählt man wiederum mittels Division mit Rest $\varepsilon \in \mathbb{Z}$ so, daß

$$\gamma' = \varepsilon\alpha_1 + \eta, \quad 0 \le \eta < |\alpha_1|.$$

Dann gilt

$$c = (\varepsilon a_1 + \zeta a_2) + (\eta + \vartheta\omega).$$

Da der erste Summand in $\mathfrak{a}$ liegt, der zweite von der in (b) genannten Form ist, bleibt nur zu zeigen, daß die in (b) genannten Elemente paarweise verschiedene Restklassen repräsentieren: Aus

$$(\alpha + \beta\omega) - (\alpha' + \beta'\omega) \in \mathfrak{a}$$

folgt, daß $\beta - \beta'$ Vielfaches von β_2 ist. Bei $0 \le \beta, \beta' < |\beta_2|$ ist dies nur für $\beta = \beta'$ möglich:

$$\alpha - \alpha' \in \mathbb{Z} \cap \mathfrak{a} = \mathbb{Z}a_1$$

impliziert bei $0 \le \alpha, \alpha' < |a_1|$, daß $\alpha = \alpha'$. □

Definition. Die Anzahl der Elemente von $A_d/\mathfrak{a}$ heißt *Norm* des Ideals $\mathfrak{a}$, bezeichnet durch $N(\mathfrak{a})$.

Die Norm kann man mittels einer beliebigen $\mathbb{Z}$-Basis b_1, b_2 berechnen. Mit $b_i = \gamma_i + \delta_i\omega$ ist

$$N(\mathfrak{a}) = \left|\det \begin{pmatrix} \gamma_1 & \delta_1 \\ \gamma_2 & \delta_2 \end{pmatrix}\right|.$$

Mit den vor der Definition der Diskriminante eingeführten Bezeichnungen ist

$$\begin{pmatrix} \alpha_1 & 0 \\ \alpha_2 & \beta_2 \end{pmatrix} = A \begin{pmatrix} \gamma_1 & \delta_1 \\ \gamma_2 & \delta_2 \end{pmatrix}$$

und wegen $|\det A| = 1$ folgt

$$\left|\det \begin{pmatrix} \gamma_1 & \delta_1 \\ \gamma_2 & \delta_2 \end{pmatrix}\right| = |\alpha_1|\,|\beta_2|.$$

Der soeben eingeführte Normbegriff ist mit dem Begriff „Norm eines Elements“ verträglich:

Satz 12.6. *Sei $a \in A_d$ und $\mathfrak{a}$ das von a erzeugte Ideal. Dann ist*

$$N(\mathfrak{a}) = |N(a)|.$$

Beweis. Sei $\omega = \omega_d$. Die Elemente $a, a\omega$ bilden eine $\mathbb{Z}$-Basis von $\mathfrak{a}$. Mit $a = \alpha + \beta\omega$ ist

$$a\omega = \beta\frac{d-1}{4} + (\alpha+\beta)\omega, \quad N(\mathfrak{a}) = \left|\det\begin{pmatrix} \alpha & \beta \\ \frac{\beta(d-1)}{4} & \alpha+\beta \end{pmatrix}\right| = |N(a)|$$

für $d \equiv 1 \ (4)$ und im Fall $d \equiv 2, 3 \ (4)$ ist

$$a\omega = \beta d + a\omega, \quad N(\mathfrak{a}) = \left|\det\begin{pmatrix} \alpha & \beta \\ \beta d & \alpha \end{pmatrix}\right| = |N(a)|. \qquad \square$$

Wir beschließen, diesen Abschnitt mit einer auf den ersten Blick unscheinbaren, aber äußerst konsequenzenreichen Aussage. Mit $\mathfrak{a}$ ist auch

$$\mathfrak{a}' := \{a' : a \in \mathfrak{a}\}$$

ein Ideal in A_d. Es gilt:

Satz 12.7. *Das Ideal $\mathfrak{a}\mathfrak{a}'$ wird von einer ganzrationalen Zahl erzeugt:*

$$\mathfrak{a}\mathfrak{a}' = qA_d \qquad \textit{mit} \quad q \in \mathbb{Z}.$$

Beweis. Sei a_1, a_2 ein Erzeugendensystem von $\mathfrak{a}$. Dann ist a_1', a_2' ein Erzeugendensystem von $\mathfrak{a}'$ und

$$a_1a_1', \ a_2a_1', \ a_1'a_2, \ a_2a_2'$$

sicherlich ein Erzeugendensystem von $\mathfrak{a}\mathfrak{a}'$. Wegen $(a_1a_2' + a_2a_1')' = a_1a_2' + a_2a_1'$ ist $a_1a_2' + a_2a_1' \in \mathbb{Z}$, ebenso

$$a_1a_1' = N(a_1), \quad a_2a_2' = N(a_2) \in \mathbb{Z}.$$

Sei $g := \mathrm{ggT}\big(N(a_1), N(a_2), a_1a_2' + a_1a_1'\big)$ (in $\mathbb{Z}$). Natürlich liegt g in $\mathfrak{a}\mathfrak{a}'$, denn es ist Linearkombination von a_1a_1', a_2a_2', $a_1a_2' + a_2a_1'$. Umgekehrt gilt sofort $a_1a_1' \in gA_d$, $a_2a_2' \in gA_d$. Ferner ist

$$N\left(\frac{a_1a_2'}{g}\right) = \frac{N(a_1a_2')}{N(g)} = \frac{N(a_1)N(a_2)}{N(g)} \in \mathbb{Z} \quad \text{und auch}$$

$$S\left(\frac{a_1a_2'}{g}\right) = \frac{1}{g}S(a_1a_2') = \frac{1}{g}(a_1a_2' + a_2a_1') \in \mathbb{Z}.$$

Also ist $(a_1a_2'/g) \in A_d$: Es existiert ein $c \in A_d$ mit $a_1a_2' = cg$, und genau so folgt $a_2a_1' \in gA_d$. $\square$

Wir können Satz 12.7 noch verschärfen:

Satz 12.8. *$\mathfrak{a}\mathfrak{a}'$ wird von $N(\mathfrak{a})$ erzeugt.*

Beweis. Wir setzen $A = A_d$, $\omega = \omega_d$. Nach Satz 12.4 kann man $\mathfrak{a} = \beta\mathfrak{b}$ mit $\beta \in \mathbb{Z}$, $\beta > 0$, schreiben, wobei $\mathfrak{b}$ eine $\mathbb{Z}$-Basis $b_1 = \gamma$, $b_2 = \delta + \omega_d$ besitzt, $\gamma, \delta \in \mathbb{Z}$. Aus Satz 12.5 folgt nun unmittelbar, daß $N(\mathfrak{a}) = \beta^2 N(\mathfrak{b})$. Da ferner $\mathfrak{a}\mathfrak{a}' = \beta^2\mathfrak{b}\mathfrak{b}'$ gilt, genügt es, den Fall $\mathfrak{a} = \mathfrak{b}$ zu betrachten. Wir dürfen also annehmen, daß $\mathfrak{a}$ die Basis $a_1 = \gamma, a_2 = \delta + \omega$ hat.

Sei $g \in \mathbb{Z}$ das durch Satz 12.7 gegebene erzeugende Element von $\mathfrak{a}\mathfrak{a}'$. Dann gilt einerseits $\gamma \mid g$, denn

$$g \in \mathfrak{a}\mathfrak{a}' \cap \mathbb{Z} \subset \mathfrak{a} \cap \mathbb{Z} = \mathbb{Z}\gamma,$$

andererseits $g \mid \gamma$, denn

$$\gamma a_2 \in \mathfrak{a}'\mathfrak{a} = Ag, \qquad \text{also} \quad \frac{1}{g}\gamma a_2 = \frac{\gamma\delta}{g} + \frac{\gamma}{g}\omega \in A.$$

Mithin $\gamma/g \in \mathbb{Z}$, und es folgt $g = \pm\gamma$. □

Wenn man versucht, 12.8 direkt zu beweisen, muß man (mit den Bezeichnungen im Beweis von 12.8) zeigen, daß

$$\gamma = \mathrm{ggT}(\gamma^2, (\delta + \omega)(\delta + \omega)', \gamma(2\delta + \omega + \omega')).$$

Dies sei dem Leser zur Übung empfohlen.

Eine wichtige Folgerung aus 12.8:

Satz 12.9. *Für Ideale* $\mathfrak{a}, \mathfrak{b} \subset A$ *ist* $N(\mathfrak{a}\mathfrak{b}) = N(\mathfrak{a})N(\mathfrak{b})$.

Beweis. Da $\mathfrak{a}\mathfrak{a}' = AN(\mathfrak{a})$ und $\mathfrak{b}\mathfrak{b}' = AN(\mathfrak{b})$, gilt

$$AN(\mathfrak{a}\mathfrak{b}) = (\mathfrak{a}\mathfrak{b})(\mathfrak{a}'\mathfrak{b}') = a\mathfrak{a}'\mathfrak{b}\mathfrak{b}' = AN(\mathfrak{a})N(\mathfrak{b}).$$

Da $N(\mathfrak{a}b), N(\mathfrak{a})N(b) \in \mathbb{N}_+$, folgt $N(\mathfrak{a}\mathfrak{b}) = N(\mathfrak{a})N(\mathfrak{b})$. □

Übungen.

12.10. Sei $d = -5$. Die Ideale $\mathfrak{a}, \mathfrak{b}$ seien gegeben durch

$$\mathfrak{a} = A12 + A(1 + \sqrt{-5}) \qquad \text{und} \qquad \mathfrak{b} = A3 + A(1 - \sqrt{-5}).$$

(a) Bestimme $\mathbb{Z}$-Basen von $\mathfrak{a}, \mathfrak{b}, \mathfrak{a} + \mathfrak{b}, \mathfrak{a}\mathfrak{b}, \mathfrak{a} \cap \mathfrak{b}$.

(b) Ist eines dieser fünf Ideale ein Hauptideal?

(c) Bestimme ihre Diskriminanten und Normen.

12.11. Sei $A = A_d, \omega = \omega_d$.

(a) Zeige: Eine Untergruppe U von A (hinsichtlich der Addition) ist genau dann ein Ideal, wenn für jedes $u \in U$ auch $\omega u \in U$.

(b) Versuche, die in Satz 12.4 angegebenen Bedingungen so zu erweitern, daß sich an ihnen ablesen läßt, ob gegebene Elemente $a_1 = \alpha_1, a_2 = \alpha_2 + \beta_2\omega$ eine $\mathbb{Z}$-Basis

des von ihnen erzeugten Ideals bilden. Nach (a) ist also zu untersuchen, unter welchen Bedingungen an $\alpha_1, \alpha_2, \beta_2$ die Elemente ωa_1 und ωa_2 Linearkombinationen von a_1 und a_2 mit Koeffizienten aus $\mathbb{Z}$ sind.)

12.12. Zeige: Elemente $a_1, a_2 \in \mathbb{Q}[\sqrt{d}]$ sind genau dann linear unabhängig über $\mathbb{Q}$ (oder $\mathbb{Z}$ – das ist gleichgültig), wenn

$$\det \begin{pmatrix} a_1 & a_1' \\ a_2 & a_2' \end{pmatrix} \neq 0.$$

12.13. Sei $d \equiv 1 \ (4)$ und $R = \mathbb{Z}[\sqrt{d}]$. Zeige:

(a) $2, 1 + \sqrt{d}$ ist $\mathbb{Z}$-Basis des von ihnen erzeugten Ideals $\mathfrak{p}$ in R.

(b) $|R/\mathfrak{p}| = 2$ und $\mathfrak{p}$ ist ein Primideal.

(c) $4, 2 + 2\sqrt{d}$ ist $\mathbb{Z}$-Basis von $\mathfrak{p}^2$.

(d) Satz 12.9 läßt sich nicht auf R verallgemeinern.

ABSCHNITT 13

Teilbarkeitstheorie für Ideale

Für die Teilbarkeitstheorie der Ideale erweitert man zweckmäßigerweise den Idealbegriff noch etwas:

Definition. Eine Teilmenge $\mathfrak{b}$ von $\mathbb{Q}[\sqrt{d}]$ heißt ein *gebrochenes Ideal*, wenn es ein Ideal $\mathfrak{a}$ in A und ein $b \in A \setminus \{0\}$ gibt, derart daß

$$\mathfrak{b} = \frac{\mathfrak{a}}{b} := \left\{\frac{a}{b} : a \in \mathfrak{a}\right\}.$$

Ein gebrochenes Ideal erhält man also, indem man ein Ideal hernimmt und alle Brüche a/b, $a \in \mathfrak{a}$, mit einem festen Nenner b betrachtet.

Die Aussagen von Satz 12.1 übertragen sich unmittelbar auf gebrochene Ideale. Dabei berücksichtigt man, daß man, wenn nötig, den Nenner erweitern kann:

$$\frac{\mathfrak{a}}{b} = \frac{c\mathfrak{a}}{cb} \quad \text{für alle } c \in A,\ c \neq 0.$$

Es ist ferner zweckmäßig – und davon haben wir soeben schon Gebrauch gemacht – folgende abkürzende Schreibweise zu verwenden:

$$c\mathfrak{a} := (Ac)\mathfrak{a}.$$

Dies ist auch gerechtfertigt, weil in der Tat $(Ac)\mathfrak{a} = \{ca : a \in \mathfrak{a}\}$.

Offensichtlich ist das Einsideal A neutral bezüglich der Multiplikation von gebrochenen Idealen. Mit Hilfe von Satz 12.7 können wir nun zeigen, daß jedes gebrochene Ideal $\mathfrak{b} \neq 0$ ein Inverses bezüglich der Multiplikation besitzt:

Satz 13.1. *Die vom Nullideal verschiedenen gebrochenen Ideale bilden bezüglich der Multiplikation eine Gruppe. Die Bildung des Inversen ist mit der Inklusion verträglich: Wenn* $\mathfrak{a} \subset \mathfrak{b}$, *so* $\mathfrak{a}^{-1} \supset \mathfrak{b}^{-1}$.

Beweis. Sei zunächst $\mathfrak{a} \subset A$ ein Ideal. Nach Satz 12.7 ist

$$\mathfrak{a}\mathfrak{a}' = Aa$$

ein Hauptideal (sogar mit $a \in \mathbb{Z}$). Nun gilt trivialerweise

$$Aa \cdot \frac{A}{a} = A$$

und es folgt $\mathfrak{a}(\mathfrak{a}' \cdot A/a) = A$, $\mathfrak{a}^{-1} = \mathfrak{a}' \cdot A/a$.

Sei nun $\mathfrak{b} = \mathfrak{a}/b$ ein gebrochenes Ideal. Dann ist

$$\mathfrak{b} \cdot (b\mathfrak{a}^{-1}) = A.$$

Wenn $\mathfrak{a} \subset \mathfrak{b}$, so $A = \mathfrak{a}\mathfrak{a}^{-1} \subset \mathfrak{b}\mathfrak{a}^{-1}$ und $\mathfrak{b}^{-1} \subset \mathfrak{a}^{-1}$. □

Aus Satz 13.1 ergibt sich unmittelbar die Kürzungsregel für gebrochene Ideale:

$$\mathfrak{c}\mathfrak{a} = \mathfrak{c}\mathfrak{b},\ \mathfrak{c} \neq 0 \quad \Rightarrow \quad \mathfrak{a} = \mathfrak{b}.$$

In natürlicher Weise definiert man für Ideale $\mathfrak{a}, \mathfrak{b} \subset A$

$$\mathfrak{a} \mid \mathfrak{b} \iff \text{es existiert ein Ideal } \mathfrak{c} \subset A \text{ mit } \mathfrak{b} = \mathfrak{a}\mathfrak{c}.$$

Im Fall eines Hauptideals Ab schreiben wir auch $\mathfrak{a} \mid b$ statt $\mathfrak{a} \mid Ab$.

Die Teilbarkeit ist zu einer sehr viel elementareren Beziehung äquivalent:

Satz 13.2. *Genau dann gilt* $\mathfrak{a} \mid \mathfrak{b}$, *wenn* $\mathfrak{a} \supset \mathfrak{b}$.

Beweis. Trivialerweise ist $\mathfrak{a}\mathfrak{c} \subset \mathfrak{a}$, also $\mathfrak{b} \subset \mathfrak{a}$, falls $\mathfrak{a} \mid \mathfrak{b}$. Zur Umkehrung: Wegen

$$\mathfrak{b} = \mathfrak{a} \cdot (\mathfrak{a}^{-1}\mathfrak{b})$$

genügt es zu zeigen, daß $\mathfrak{a}^{-1}\mathfrak{b} \subset A$, falls $\mathfrak{a} \supset \mathfrak{b}$. Bei $\mathfrak{a} \supset \mathfrak{b}$ ist aber $\mathfrak{b}^{-1} \supset \mathfrak{a}^{-1}$ und daher

$$\mathfrak{a}^{-1}\mathfrak{b} \subset \mathfrak{b}^{-1}\mathfrak{b} = A.$$ □

Die Krönung der Teilbarkeitstheorie für Ideale ist die Aussage, daß sich jedes (gebrochene) Ideal $\neq 0$ eindeutig in Primfaktoren zerlegen läßt, daß also der Hauptsatz der elementaren Zahlentheorie auch in quadratischen Körpern gilt, wenn man ihn als eine Aussage über *Ideale* begreift. Wir gehen in zwei Schritten vor. Zunächst zeigen wir, daß sich jedes Ideal $\neq 0$ als Produkt irreduzibler Ideale schreiben läßt: Ein Ideal $\mathfrak{a} \subset A$ heißt *irreduzibel*, wenn eine Darstellung $\mathfrak{a} = \mathfrak{b}\mathfrak{c}$ mit Idealen $\mathfrak{b}, \mathfrak{c} \subset A$ nur mit $\mathfrak{b} = A$ oder $\mathfrak{c} = A$ möglich ist.

Satz 13.3. *Sei* $\mathfrak{a} \neq 0$ *ein Ideal in A. Dann gibt es irreduzible Ideale* $\mathfrak{p}_1, \ldots, \mathfrak{p}_n$ *mit* $\mathfrak{a} = \mathfrak{p}_1 \cdots \mathfrak{p}_n$.

Beweis. Wir führen eine Induktion über $N(\mathfrak{a})$. Falls $N(\mathfrak{a}) = 1$, ist $\mathfrak{a} = A$ trivialerweise irreduzibel. Sei nun $N(\mathfrak{a}) > 1$. Wenn $\mathfrak{a}$ nicht schon selbst irreduzibel ist, existieren Ideale $\mathfrak{b}, \mathfrak{c} \subset A$ mit

$$\mathfrak{a} = \mathfrak{b}\mathfrak{c}, \quad \mathfrak{b} \neq A,\ \mathfrak{c} \neq A.$$

Es folgt unmittelbar $N(\mathfrak{b}), N(\mathfrak{c}) < N(\mathfrak{a})$, und $\mathfrak{b}$ und $\mathfrak{c}$ sind nach Induktionsvoraussetzung zerlegbar. □

Wir müssen nun noch zeigen, daß irreduzible Ideale prim sind. Dabei gehen wir genauso vor wie in Abschnitt 1, als wir gezeigt haben, daß irreduzible Elemente in euklidischen Ringen Primelemente sind.

Definition. Wir nennen das Ideal $\mathfrak{a}+\mathfrak{b}$ den *größten gemeinsamen Teiler* der Ideale $\mathfrak{a}$ und $\mathfrak{b}$. Falls $\mathfrak{a}+\mathfrak{b}=R$ ist, heißen $\mathfrak{a}$ und $\mathfrak{b}$ *teilerfremd*.

In Analogie zum Euklidischen Lemma gilt:

Satz 13.4. *Seien $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$ Ideale in A, derart, daß $\mathfrak{a}$ und $\mathfrak{b}$ teilerfremd sind und $\mathfrak{a}$ das Produkt $\mathfrak{b}\mathfrak{c}$ teilt. Dann teilt $\mathfrak{a}$ das Ideal $\mathfrak{c}$.*

Beweis. $\mathfrak{a}+\mathfrak{b}=A$ und $\mathfrak{a} \mid (\mathfrak{a}+\mathfrak{b})\mathfrak{c}$, also $\mathfrak{a} \mid \mathfrak{c}$. □

Da ein irreduzibles Ideal $\mathfrak{p}$ als Teiler nur sich selbst und das Einsideal besitzt, ist es zu jedem Ideal $\mathfrak{b} \not\subset \mathfrak{p}$ teilerfremd: Falls $\mathfrak{b} \not\subset \mathfrak{p}$, so ist $\mathfrak{b}+\mathfrak{p} \underset{\neq}{\supset} \mathfrak{p}$, also nach 13.2 ein Teiler von $\mathfrak{p}$, mithin $\mathfrak{b}+\mathfrak{p}=R$.

Satz 13.5. *Ein irreduzibles Ideal $\mathfrak{p}$ in A, $\mathfrak{p} \neq A$, ist prim.*

Beweis. Aus 13.2 folgt unmittelbar, daß ein irreduzibles Ideal maximal ist; erst recht ist es prim. □

Die Umkehrung von 13.5 ist in beliebigen Ringen richtig: Ein Primideal ist stets irreduzibel.

Aus 13.3 und 13.5 können wir nun schließen:

Satz 13.6. *Jedes Ideal $\mathfrak{a} \neq 0$ in A läßt sich auf genau eine Weise als Produkt von Primidealen darstellen: Es existieren (bis auf die Reihenfolge) eindeutig bestimmte Primideale $\mathfrak{p}_1, \ldots, \mathfrak{p}_n$, $\mathfrak{p}_i \neq \mathfrak{p}_j$ für $i \neq j$, und eindeutig bestimmte Zahlen $r_1, \ldots, r_n \in \mathbb{N}_+$ mit*

$$\mathfrak{a} = \mathfrak{p}_1^{r_1} \cdots \mathfrak{p}_n^{r_n}.$$

(Im Fall $\mathfrak{a}=A$ hat man rechts das leere Produkt, das man definitionsgemäß gleich A setzt.)

Beweis. Die Existenz der Darstellung folgt aus 13.5 und 13.3, die Eindeutigkeit schließt man mittels 13.4 und der Kürzungsregel ebenso wie im Beweis des Hauptsatzes der elementaren Zahlentheorie. □

Satz 13.6 charakterisiert die Ringe A_d insofern, als daß seine Aussage auf keinen Ring R, der echt zwischen $\mathbb{Z}$ und A_d liegt verallgemeinerbar ist. (Vergleiche dazu auch Aufgabe 13.10.)

Im nächsten Abschnitt untersuchen wir die Primideale in den Ringen A_d näher. Für die Verwendung in dem unten diskutierten Beispiel notieren wir schon hier:

Satz 13.7. *Wenn $N(\mathfrak{a})$ eine Primzahl ist, so ist $\mathfrak{a}$ ein Primideal.*

Beweis. Ein Ring, dessen Elementezahl prim ist, ist ein Körper, nämlich isomorph zu $\mathbb{Z}_p$. □

Beispiel. $d = -5$, $A = A_{-5}$, $\mathfrak{a} = A6$. Es gilt $\mathfrak{a} = A2 \cdot A3$, aber auch $(1+\sqrt{-5}) \cdot (1-\sqrt{-5}) \in \mathfrak{a}$. Da $1+\sqrt{-5}$, $1-\sqrt{-5} \notin A2, A3$, ist weder $A2$ noch $A3$ Primideal. Ist andererseits $\mathfrak{p} \supset A2$ ein Primideal, gilt $1+\sqrt{-5} \in \mathfrak{p}$ oder $1-\sqrt{-5} \in \mathfrak{p}$. Da $(1+\sqrt{-5}) - (1-\sqrt{-5}) \in A2$, folgt $1+\sqrt{-5} \in \mathfrak{p}$ und $1-\sqrt{-5} \in \mathfrak{p}$.

Wir betrachten das Ideal $A2 + A(1+\sqrt{-5})$. Es besitzt das $\mathbb{Z}$-Erzeugendensystem 2, $1+\sqrt{-5}$, $-5+\sqrt{-5}$, $2\sqrt{-5}$. Daraus ergibt sich als $\mathbb{Z}$-Basis: 2, $1+\sqrt{-5}$, und die Norm des Ideals ist 2: $\mathfrak{p} = A2 + A(1+\sqrt{-5})$ ist Primideal. Da $(1+\sqrt{-5})' = 1-\sqrt{-5} \in \mathfrak{p}$, folgt $\mathfrak{p} = \mathfrak{p}'$ und $A2 = AN(\mathfrak{p}) = \mathfrak{p}\mathfrak{p}' = \mathfrak{p}^2$.

Ebenso naheliegend ist es, daß Ideal $\mathfrak{q} = A3 + A(1+\sqrt{-5})$ zu untersuchen: Wie soeben erhält man die $\mathbb{Z}$-Basis 3, $1+\sqrt{-5}$, also $N(\mathfrak{q}) = 3$. Somit ist $\mathfrak{q}$ Primideal. Da $2 \notin \mathfrak{q}$, ist auch $1-\sqrt{-5} = 2-(1+\sqrt{-5}) \notin \mathfrak{q}$, somit $\mathfrak{q} \neq \mathfrak{q}'$. Es gilt $\mathfrak{q}\mathfrak{q}' = AN(\mathfrak{q}) = A3$.

Zusammengefaßt: $\mathfrak{a} = \mathfrak{p}^2\mathfrak{q}\mathfrak{q}'$ ist die Primzerlegung von $\mathfrak{a}$.

Wie spiegelt sich die Zerlegung $(1+\sqrt{-5})(1-\sqrt{-5}) = 6$ in der Primfaktorzerlegung wider? Da $N(1+\sqrt{-5}) = 6$, muß $A(1+\sqrt{-5}) = \mathfrak{p}\mathfrak{q}$ oder $A(1+\sqrt{-5}) = \mathfrak{p}\mathfrak{q}'$ gelten. Da $1+\sqrt{-5} \notin \mathfrak{q}'$ kommt nur $A(1+\sqrt{-5}) = \mathfrak{p}\mathfrak{q}$ in Frage. Folglich

$$\mathfrak{a} = (\mathfrak{p}\mathfrak{q})(\mathfrak{p}\mathfrak{q}') = A(1+\sqrt{-5})A(1-\sqrt{-5}). -$$

An diesem Beispiel sehen wir, daß eine Primzahl $p \in \mathbb{Z}$ in A in der Form

$$Ap = \mathfrak{p}^2 \quad \text{oder} \quad A\mathfrak{p} = \mathfrak{p}\mathfrak{q},\ \mathfrak{p} \neq \mathfrak{q},$$

zerfallen kann. Das Beispiel $d = -1$, $p \equiv 3\ (4)$ zeigt, daß auch der Fall

$$Ap = \mathfrak{p}$$

auftritt. Wir werden später beweisen, daß diese drei Fälle alle möglichen Zerlegungstypen darstellen, die für Primzahlen $p \in \mathbb{Z}$ in A_d auftreten können.

Aus den Primzerlegungen zweier Ideale $\mathfrak{a}$, $\mathfrak{b}$ kann man leicht die Primzerlegung von Produkt, Durchschnitt und Summe errechnen:

Satz 13.8. *Seien $\mathfrak{a}$, $\mathfrak{b} \subset A$ Ideale, $\mathfrak{p}_1, \ldots, \mathfrak{p}_n$ paarweise verschiedene Primideale, r_i, $s_i \in \mathbb{N}$, $i = 1, \ldots, n$ derart, daß*

$$\mathfrak{a} = \mathfrak{p}_1^{r_1} \cdots \mathfrak{p}_n^{r_n} \quad \textit{und } \mathfrak{b} = \mathfrak{p}_1^{s_1} \cdots \mathfrak{p}_n^{s_n}.$$

Dann gilt:

$$\begin{aligned} \mathfrak{a}\mathfrak{b} &= \mathfrak{p}_1^{r_1+s_1} \cdots \mathfrak{p}_n^{r_n+s_n}, \\ \mathfrak{a}+\mathfrak{b} &= \mathfrak{p}_1^{m_1} \cdots \mathfrak{p}_n^{m_n}, \qquad m_i = \min(r_i, s_i), \\ \mathfrak{a}\cap\mathfrak{b} &= \mathfrak{p}_1^{M_1} \cdots \mathfrak{p}_n^{M_n}, \qquad M_i = \max(r_i, s_i). \end{aligned}$$

Ferner ist $(\mathfrak{a}\cap\mathfrak{b})(\mathfrak{a}+\mathfrak{b}) = \mathfrak{a}\mathfrak{b}$.

Den einfachen Beweis lassen wir aus. Es ist zweckmäßig, mit den $\mathfrak{p}$-Exponenten

$$v_{\mathfrak{p}}(\mathfrak{a}) := \max\{n : \mathfrak{p}^n \mid \mathfrak{a}\} = \max\{n : \mathfrak{p}^n \supset \mathfrak{a}\}$$

zu arbeiten. Mit dieser Bezeichnung ergibt sich die Darstellung

$$\mathfrak{a} = \prod \mathfrak{p}^{v_{\mathfrak{p}}(\mathfrak{a})}.$$

Weiterhin erlaubt es uns 13.6, die Aussagen über die Existenz von Erzeugendensystemen von Idealen zu verschärfen:

Satz 13.9. *Seien $\mathfrak{a} \subset \mathfrak{b} \subset A$ Ideale $\mathfrak{a} \neq 0$. Dann existiert ein $b \in \mathfrak{b}$ mit*

$$\mathfrak{b} = \mathfrak{a} + Ab.$$

Speziell gilt: Zu jedem $a \in \mathfrak{b}$, $a \neq 0$ existiert ein $b \in \mathfrak{b}$ mit $\mathfrak{b} = Aa + Ab$.

Beweis. Sei $\mathfrak{a} = \mathfrak{p}_1^{r_1} \cdots \mathfrak{p}_n^{r_n}$ die Primzerlegung von $\mathfrak{a}$. Dann ist $\mathfrak{b} = \mathfrak{p}_1^{s_1} \cdots \mathfrak{p}_n^{s_n}$ mit $s_i \leq r_i$. Wir setzen

$$\begin{aligned}
\mathfrak{b}_1 &:= \mathfrak{p}_1^{s_1} \cdot \mathfrak{p}_2^{s_2+1} \cdots \mathfrak{p}_n^{s_n+1} \\
\mathfrak{b}_2 &:= \mathfrak{p}_1^{s_1+1} \mathfrak{p}_2^{s_2} \mathfrak{p}_3^{s_3+1} \cdots \mathfrak{p}_n^{s_n+1} \\
&\vdots \\
\mathfrak{b}_n &:= \mathfrak{p}_1^{s_1+1} \cdots \mathfrak{p}_{n-1}^{s_{n-1}+1} \cdot \mathfrak{p}_n^{s_n}
\end{aligned}$$

und wählen b_i so, daß

$$b_i \in \mathfrak{b}_i, \; b_i \notin \mathfrak{b}_i \mathfrak{p}_i.$$

Sei $b := b_1 + \cdots + b_n$. Sei $\mathfrak{q} \supset \mathfrak{a} + Ab$ ein Primideal. Dann gilt $\mathfrak{q} = \mathfrak{p}_i$ für ein $i \in \{1, \ldots, n\}$. Da $b_i \notin \mathfrak{p}_i^{s_i+1}$, ist

$$v_{\mathfrak{p}_i}(\mathfrak{a} + Ab) < s_i + 1.$$

Andererseits ist $\mathfrak{a} + Ab \subset \mathfrak{b}$ und daher $v_{\mathfrak{p}_i}(\mathfrak{a} + Ab) \geq v_{\mathfrak{p}_i}(b) = s_i$:

$$v_{\mathfrak{p}_i}(\mathfrak{a} + Ab) = s_i.$$

Es folgt: $\mathfrak{a} + Ab = \mathfrak{p}_1^{s_1} \cdots \mathfrak{p}_n^{s_n} = \mathfrak{b}$. □

Die Idealtheorie ist weitgehend eine Schöpfung des Braunschweiger Mathematikers R. Dedekind (1831–1916). Der Name „Ideal“ rührt her von den von E. Kummer (1810–1893) eingeführten „idealen Faktoren“, mit deren Hilfe man Elemente in Primfaktoren zerlegen konnte, für die das eigentlich, d.h. im Ring A, nicht möglich war. Die „idealen Faktoren“ Kummers sind ganz-algebraische Zahlen b, die möglicherweise in einem echten Oberring B von A liegen. Der Durchschnitt $A \cap Bb$ ist dann ein Ideal.

Übungen.

13.10. Sei $d \equiv 1$ (4). Wir betrachten $R = \mathbb{Z}[\sqrt{d}]$, das gleiche Primideal $\mathfrak{p}$ wie in Aufgabe 12.13 und das Ideal $\mathfrak{a} = R2$. Zeige: $\mathfrak{p}^2 \subsetneq \mathfrak{a} \subsetneq \mathfrak{p}$.

Schließe, daß der Satz von der eindeutigen Zerlegung in Primideale in R nicht gilt.

13.11. (Verallgemeinerter Chinesischer Restsatz) Seien $\mathfrak{a}, \mathfrak{b}$ Ideale eines kommutativen Ringes R mit $\mathfrak{a} + \mathfrak{b} = R$. (In den Ringen der Zahlentheorie heißen solche Ideale teilerfremd, im allgemeinen sollte man sie *komaximal* nennen.) Zeige:

(a) $\mathfrak{a} \cap \mathfrak{b} = \mathfrak{a}\mathfrak{b}$.

(b) $R/\mathfrak{a}\mathfrak{b} \cong R/\mathfrak{a} \times R/\mathfrak{b}$.

Verallgemeinere auf Ideale $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$.

13.12. (Verallgemeinerte Eulersche Φ-Funktion) Für $\mathfrak{a} \subset A = A_d$, $\mathfrak{a} \neq 0$, sei $\Phi(\mathfrak{a})$ die Anzahl der Einheiten des Restklassenrings $A/\mathfrak{a}$. Zeige:

(a) Für teilerfremde Ideale $\mathfrak{a}, \mathfrak{b}$ ist $\Phi(\mathfrak{a}\mathfrak{b}) = \Phi(\mathfrak{a})\Phi(\mathfrak{b})$.

(b) Für ein Primideal $\mathfrak{p}$ ist $\Phi(\mathfrak{p}) = N(\mathfrak{p}) - 1$.

(c) Für ein Primideal $\mathfrak{p}$ und $m \in \mathbb{N}$, $m \geq 1$, ist

$$\Phi(\mathfrak{p}^m) = N(\mathfrak{p})^m \left(1 - \frac{1}{N(\mathfrak{p})}\right).$$

(d) Für jedes Ideal $\mathfrak{a} \subset A$, $\mathfrak{a} \neq 0$, ist

$$\Phi(\mathfrak{a}) = N(\mathfrak{a}) \prod_{\mathfrak{p}|\mathfrak{a}} \left(1 - \frac{1}{N(\mathfrak{p})}\right).$$

Hinweis zu (c): Die Restklasse eines Elementes $a \in A$ ist genau dann Einheit mod $\mathfrak{p}^m$, wenn $a \notin \mathfrak{p}$ (weshalb?).

ABSCHNITT 14

Ganzrationale Primzahlen und Primideale

Eine erste Beziehung zwischen Primzahlen $p \in \mathbb{Z}$ und Primidealen $\mathfrak{p} \subset A_d =: A$ gibt uns

Satz 14.1. *Zu jedem Primideal $\mathfrak{p} \subset A$ gibt es genau eine Primzahl $p \in \mathbb{Z}$ mit $\mathfrak{p} \mid Ap$. Es gilt $N(\mathfrak{p}) = p$ oder $N(\mathfrak{p}) = p^2$.*

Beweis. Das Ideal $\mathfrak{p} \cap \mathbb{Z}$ in $\mathbb{Z}$ ist ein Primideal. Also ist $\mathfrak{p} \cap \mathbb{Z} = \mathbb{Z}p$ mit einer eindeutig bestimmten Primzahl p. Da $p \in \mathfrak{p}$, gilt $\mathfrak{p} \mid Ap$. Jede von p verschiedene Primzahl $q \in \mathbb{Z}$ ist zu p auch in A teilerfremd: $\mathbb{Z}p + \mathbb{Z}q = \mathbb{Z}$ impliziert $A = Ap + Aq$. Wenn nun $\mathfrak{p} \mid Aq$, so $\mathfrak{p} \mid Ap + Aq$ und $\mathfrak{p} \mid A$, was nicht möglich ist.

Da $\mathfrak{p} \mid Ap$, folgt $N(\mathfrak{p}) \mid N(Ap) = p^2$, was nur die Möglichkeiten $N(\mathfrak{p}) = p^2$ oder $N(\mathfrak{p}) = p$ zuläßt. □

Nach 12.2 besitzt $\mathfrak{p}$ eine $\mathbb{Z}$-Basis der Form p, π. Da $\pi\pi' \in \mathfrak{p} \cap \mathbb{Z} = \mathbb{Z}p$, gilt $p \mid N(\pi)$:

Satz 14.2. *Seien $\mathfrak{p}$, p wie in* 14.1. *Dann besitzt $\mathfrak{p}$ eine $\mathbb{Z}$-Basis p, π mit $p \mid N(\pi)$.*

Aus 14.1 können wir sofort alle Zerlegungstypen ganzrationaler Primzahlen in A bestimmen:

Satz und Definition 14.3. *Sei $p \in \mathbb{Z}$ Primzahl. Dann gibt es für die Primzerlegung von Ap nur folgende Möglichkeiten:*

(i) $Ap = \mathfrak{p}^2$,	$N(\mathfrak{p}) = p$	*(p ist* verzweigt*)*
(ii) $Ap = \mathfrak{p}\mathfrak{p}'$, $\mathfrak{p} \neq \mathfrak{p}'$,	$N(\mathfrak{p}) = N(\mathfrak{p}') = p$	*(p ist* zerlegt*)*
(iii) $Ap = \mathfrak{p}$,	$N(\mathfrak{p}) = p^2$	*(p ist* träge*)*

Beweis. Sei $Ap = \mathfrak{p}_1^{r_1} \cdots \mathfrak{p}_n^{r_n}$ die Primzerlegung von Ap. Dann ist

$$p^2 = N(Ap) = N(\mathfrak{p}_1)^{r_1} \cdots N(\mathfrak{p}_n)^{r_n},$$

und hierfür sind nur folgende Fälle möglich:

(i) $n = 1,\ N(\mathfrak{p}_1) = p,\ r_1 = 2$;
(ii) $n = 2,\ N(\mathfrak{p}_1) = N(\mathfrak{p}_2) = p,\ r_1 = r_2 = 1$;
(iii) $n = 1,\ N(\mathfrak{p}_1) = p^2,\ r_1 = 1$.

Es bleibt nur noch zu zeigen, daß im Fall (ii) $\mathfrak{p}_2 = \mathfrak{p}_1'$. Nach 12.7 ist $\mathfrak{p}_1\mathfrak{p}_1' = Aa$ mit $a \in \mathbb{Z}$. Da $N(\mathfrak{p}_1\mathfrak{p}_1') = N(Aa) = N(a) = p^2$, folgt $a = p$ und $\mathfrak{p}_2 = \mathfrak{p}_1'$ wegen der Eindeutigkeit der Zerlegung. □

Satz 14.3 ist noch unbefriedigend, weil er offen läßt, welcher Fall auf eine gegebene Primzahl zutrifft. Aber auch dieses Problem besitzt eine vollständige Lösung, bei der wir überdies konkrete Formeln für die Primideale $\mathfrak{p}$ erhalten. Zunächst betrachten wir Primzahlen $p > 2$:

Satz 14.4. *Die Primzahl $p > 2$ ist genau dann träge in A_d, wenn $p \nmid d$ und $(d/p) = -1$, zerlegt, wenn $p \nmid d$ und $(d/p) = 1$, und verzweigt, wenn $p \mid d$.*

Beweis. Sei $A = A_d$. Wir bestimmen zunächst diejenigen d, für die p nicht träge ist, und nehmen dazu an, p sei nicht träge.

Sei zunächst $d \equiv 2, 3 \ (4)$. Dann gilt $Ap = \mathfrak{p}\mathfrak{p}'$ oder $Ap = \mathfrak{p}^2$ mit $N(\mathfrak{p}) = p$, $\mathfrak{p} = \mathbb{Z}p + \mathbb{Z}\pi$, $p \mid N(\pi)$ gemäß den vorangegangenen Sätzen. Sei $\pi = a + b\sqrt{d}$. Es ist $N(\mathfrak{p}) = p|b| = p$. Wir dürfen also $b = 1$ annehmen. Die Bedingung $p \mid N(\pi)$ impliziert $p \mid a^2 - d$. Also ist d quadratischer Rest mod p (wobei wir auch zulassen, daß $p \mid d$).

Sei nun $d \equiv 1 \ (4)$. Wieder ist $Ap = \mathfrak{p}\mathfrak{p}'$ oder $Ap = \mathfrak{p}^2$ mit $N(\mathfrak{p}) = p$, $\mathfrak{p} = \mathbb{Z}p + \mathbb{Z}\pi$, $p \mid N(\pi)$. Sei $\pi = (a + b\sqrt{d})/2$ mit $a \equiv b \ (2)$. Dann ist $\pi = (a - b)/2 + b\omega$. Wie oben dürfen wir $b = 1$ annehmen. Die Bedingung $p \mid N(\pi)$ impliziert

$$p \mid \frac{a^2 - d}{4}.$$

Also ist d wiederum quadratischer Rest mod p.

Umgekehrt setzen wir jetzt voraus, d sei quadratischer Rest modulo p und betrachten zunächst $d \equiv 2, 3 \ (4)$. Wir wählen ein $a \in \mathbb{Z}$ mit $a^2 \equiv d \ (p)$ und setzen $\mathfrak{p} = \mathbb{Z}p + \mathbb{Z}(a + \sqrt{d})$. Wenn $\mathfrak{p}$ ein Ideal ist, dann eines der Norm p, so daß wir ein Primideal $\mathfrak{p} \neq Ap$ gefunden habe, das p teilt.

Es gilt $\sqrt{d}\,p = -ap + p(a + \sqrt{d})$ und $\sqrt{d}(a + \sqrt{d}) = -bp + a(a + \sqrt{d})$, wobei wir $b = (a^2 - d)/p$ gewählt haben. Daher ist $\mathfrak{p}$ wie behauptet ein Ideal.

Sei nun $d \equiv 1 \ (4)$. In diesem Fall wählt man $\mathfrak{p} = \mathbb{Z}p + \mathbb{Z}(a + \sqrt{d})/2$, wobei $a^2 \equiv d \ (p)$, $a \equiv 1 \ (2)$. Wie im Fall $d \equiv 2, 3 \ (4)$ ist zu zeigen, daß $\omega_d\mathfrak{p} \subset \mathfrak{p}$, und wiederum kann man dies unmittelbar nachrechnen.

Nun ist noch zu ermitteln, wann p verzweigt oder zerlegt ist. Dazu müssen wir prüfen, ob das Primideal $\mathfrak{p}$ mit $\mathfrak{p}'$ übereinstimmt. Jeweils ist zu entscheiden, ob $y' \in \mathfrak{p}$ für das von p verschiedene Basiselement y von $\mathfrak{p}$.

Sei zunächst $d \equiv 2, 3 \ (4)$. Es gilt $a - \sqrt{d} \in \mathfrak{p}$ genau dann, wenn $p \mid 2a$, denn die einzig mögliche Darstellung von $a - \sqrt{d}$ als $\mathbb{Q}$-Linearkombination von p und

$a + \sqrt{d}$ ist von der Form

$$a - \sqrt{d} = bp - (a + \sqrt{d}), \qquad \text{wobei } bp = 2a.$$

Da $a^2 \equiv d \ (p)$, ist $b \in \mathbb{Z}$ äquivalent zu $p \mid 2a$ und dies wiederum äquivalent zu $p \mid d$.

Im Fall $d \equiv 1 \ (4)$ gilt analog, daß $\mathfrak{p} = \mathfrak{p}'$ genau dann, wenn $p \mid d$, denn nur dann ist $(a - \sqrt{d})/2$ eine $\mathbb{Z}$-Linearkombination von p und $(a + \sqrt{d})/2$. □

Nun zur Primzahl 2.

Satz 14.5. *Wenn $d \equiv 5 \ (8)$, ist 2 träge, und wenn $d \equiv 1 \ (8)$, ist 2 zerlegt in A_d. Falls $d \equiv 2, 3 \ (4)$, ist 2 verzweigt.*

Beweis. Sei $A = A_d$. Im Fall $d \equiv 2 \ (4)$ ist $\mathbb{Z}2 + \mathbb{Z}\sqrt{d}$ ein Primideal der Norm 2, im Fall $d \equiv 3 \ (4)$ gilt dies für $\mathbb{Z}2 + \mathbb{Z}(1 + \sqrt{d})$, und in beiden Fällen ist offensichtlich $\mathfrak{p} = \mathfrak{p}'$.

Sei nun $d \equiv 1 \ (4)$. Ein Primideal der Norm 2 ist von der Form $\mathfrak{p} = \mathbb{Z}2 + \mathbb{Z}(a + \sqrt{d})/2$, $a \equiv 1 \ (2)$. Die einzige Möglichkeit, $\omega_d(a + \sqrt{d})/2$ als $\mathbb{Q}$-Linearkombination der Basiselemente darzustellen, ist

$$\frac{\omega_d(a + \sqrt{d})}{2} = \frac{a + d + (a+1)\sqrt{d}}{4} = \frac{d - a^2}{8} \cdot 2 + \frac{a+1}{2} \cdot \frac{a + \sqrt{d}}{2}.$$

Der Koeffizient von 2 ist in $\mathbb{Z}$ genau dann, wenn $d \equiv 1 \ (8)$, während der andere Koeffizient stets zu $\mathbb{Z}$ gehört.

Dies zeigt: 2 ist träge im Fall $d \equiv 5 \ (8)$, während im Fall $d \equiv 1 \ (8)$ ein Primideal der Norm 2 durch $\mathfrak{p} = \mathbb{Z}2 + \mathbb{Z}\omega_d$ gegeben ist. Man prüft unmittelbar nach, daß es von $\mathfrak{p}'$ verschieden ist. □

Wir fassen die beiden vorangegangenen Sätze noch einmal tabellarisch zusammen, wobei wir der Übersicht halber auf die Angabe der $\mathbb{Z}$-Basen der Primideale zugunsten von Erzeugendensystemen verzichten (der Leser möge verifizieren, daß sie korrekt ermittelt sind):

Satz 14.6. *Sei $d \in \mathbb{Z}$ quadratfrei, $A = A_d$ und $p \in \mathbb{Z}$ prim. Dann ist die Primfaktorzerlegung von Ap gegeben durch:*

$p = 2$:

(i)	$A2$		$d \equiv 5 \ (8)$
(ii)	$A2 = \mathfrak{p}\mathfrak{p}'$,	$\mathfrak{p} = A2 + A(1 + \sqrt{d})/2$	$d \equiv 1 \ (8)$
(iii)	$A2 = \mathfrak{p}^2$,	$\mathfrak{p} = A2 + A(a + \sqrt{d})$	$d \equiv 2, 3 \ (4)$, $a \equiv d \ (2)$

$p > 2$:

(i)	Ap		$(d/p) = -1$
(ii)	$Ap = \mathfrak{p}\mathfrak{p}'$,	$\mathfrak{p} = Ap + A(a + \sqrt{d})$	$(d/p) = 1$, $a^2 \equiv d \ (p)$
(iii)	$Ap = \mathfrak{p}^2$,	$\mathfrak{p} = Ap + A\sqrt{d}$	$p \mid d$.

Beispiel. (a) $d = -1$: Es ist $d \equiv 3 \ (4)$, und wir erhalten

$$A2 = \mathfrak{p}^2, \ \mathfrak{p} = A2 + A(1+i),$$
$$Ap = Ap, \ p > 2, \ p \equiv 3 \ (4),$$
$$Ap = \mathfrak{p}\mathfrak{p}', \ \mathfrak{p} = Ap + A(a+i), \ a^2 \equiv -1 \ (p), \ p > 2, \ p \equiv 1 \ (4).$$

Natürlich wissen wir schon, daß alle auftretenden Primideale Hauptideale sind, denn die Elemente von A_{-1} lassen sich in Produkte von Primelementen zerlegen. Diese zusätzliche Information kann man Satz 14.6 nicht ohne weiteres entnehmen.

(b) $d = -5$. Wieder ist $d \equiv 3 \ (4)$.

$$A2 = \mathfrak{p}^2, \ \mathfrak{p} = A2 + A(1+\sqrt{-5}),$$
$$A3 = \mathfrak{p}\mathfrak{p}', \ \mathfrak{p} = A3 + A(1+\sqrt{-5}),$$

denn $1^2 \equiv -5 \ (3)$. Dieses Beispiel haben wir in Abschnitt 12 betrachtet und hier – wie vorhergesagt – die Primzerlegung ohne Mühe erhalten.

Die offensichtliche Analogie zwischen den Fällen $p = 2$ und $p > 2$ in Satz 14.6 legt eine weitere Vereinheitlichung nahe. Eine besonders elegante Formulierung erhält man, wenn man mit der Diskriminante Δ

$$\Delta = \begin{cases} d & d \equiv 1 \ (4), \\ 4d & d \equiv 2, 3 \ (4) \end{cases}$$

von A_d arbeitet. Dazu erweitert man das Legendre-Symbol auf dem Fall des „Nenners“ 2:

$$\left(\frac{a}{2}\right) = \begin{cases} 0 & \text{wenn } 4 \mid a, \\ 1 & \text{wenn } a \equiv 1 \ (8), \\ -1 & \text{wenn } a \equiv 5 \ (8). \end{cases}$$

In allen anderen Fällen ist $(a/2)$ nicht definiert.

Satz 14.7 (Zerlegungsgesetz in quadratischen Körpern)**.** *Eine Primzahl $p \in \mathbb{Z}$ ist in A_d genau dann*

(i) *verzweigt, wenn* $(\Delta/p) = 0$,
(ii) *zerlegt, wenn* $(\Delta/p) = 1$,
(iii) *träge, wenn* $(\Delta/p) = -1$.

Mit diesen Ergebnissen lassen sich beliebige Ideale $\mathfrak{a}$ in A schnell in Primfaktoren zerlegen. Man bestimmt zunächst eine Basis $a_1 = \alpha_1, a_2 = \alpha_2 + \beta_2\omega$ von $\mathfrak{a}$. Dann ist

$$a_1 = \gamma\beta_2, \quad a_2 = \delta\beta_2 + \beta_2\omega, \qquad \gamma, \delta \in \mathbb{Z},$$

also $\mathfrak{a} = \beta_2\tilde{\mathfrak{a}}$, $\tilde{\mathfrak{a}}$ gegeben durch die Basis $\tilde{a}_1 = \gamma, \tilde{a}_2 = \delta + \omega$. Die ganzrationale Zahl β_2 läßt sich mit 14.6 sofort zerlegen, so daß wir nun nur noch ein Ideal zu betrachten haben, das von keiner ganzrationalen Zahl geteilt wird. Das hat folgenden

Vorteil: $\tilde{\mathfrak{a}}$ kann nicht zugleich von einem Primideal $\mathfrak{p}$ und seinem Konjugierten $\mathfrak{p}'$ geteilt werden, wenn $\mathfrak{p} \neq \mathfrak{p}'$, und nicht von $\mathfrak{p}^2$, wenn $\mathfrak{p} = \mathfrak{p}'$.

Man zerlegt jetzt $|\gamma| = N(\tilde{\mathfrak{a}})$ und hat damit gemäß 12.8 die Primfaktorzerlegung von $\tilde{\mathfrak{a}}\tilde{\mathfrak{a}}'$ gefunden. Sei

$$A\gamma = \pi \mathfrak{p}^{v_{\mathfrak{p}}(\gamma)}.$$

Dann ist

$$v_{\mathfrak{p}}(\tilde{\mathfrak{a}}) = \begin{cases} 1 & \text{wenn } v_{\mathfrak{p}}(\gamma) > 0,\ \mathfrak{p} = \mathfrak{p}' \\ v_{\mathfrak{p}}(\gamma) & \text{wenn } \mathfrak{p} \mid \tilde{\mathfrak{a}},\ \mathfrak{p} \neq \mathfrak{p}' \\ 0 & \text{wenn } \mathfrak{p}' \mid \tilde{\mathfrak{a}},\ \mathfrak{p} \neq \mathfrak{p}' \end{cases}$$

Beispiel. $d = -5$, $\mathfrak{a} = A30 + A(5 + \sqrt{-5})$; $\mathfrak{a}$ wird von keiner ganzrationalen Zahl geteilt, $N(\mathfrak{a}) = 30$.

$$30 = 2 \cdot 3 \cdot 5, \quad A30 = \mathfrak{p}^2 \mathfrak{q}_1 \mathfrak{q}_2 \mathfrak{r}^2$$

mit

$$\mathfrak{p} = A2 + A(1 + \sqrt{-5}) = \mathfrak{p}', \quad \mathfrak{r} = A5 + A\sqrt{-5} = \mathfrak{r}'$$
$$\mathfrak{q}_1 = A3 + A(1 + \sqrt{-5}), \qquad \mathfrak{q}_2 = \mathfrak{q}_1'.$$

Man sieht sofort: $\mathfrak{q}_2 \mid \mathfrak{a}$, und damit $\mathfrak{a} = \mathfrak{p}\mathfrak{q}_2\mathfrak{r}$.

Sei d quadratfrei, p eine ungerade Primzahl mit $(\Delta/p) = (d/p) = -1$. Dann ist p träge in $A = A_d$, und A/Ap ist ein Körper mit p^2 Elementen. Seine Einheitengruppe ist nach Satz 4.3 zyklisch von der Ordnung $p^2 - 1$. Dies kann man ähnlich wie die Struktur von $\mathbb{Z}_p^*$ für Primzahltests und Faktorisierungsverfahren ausnutzen. Voraussetzung für die Wirksamkeit von $(p-1)$-Primzahltest und $(p-1)$-Faktorisierungsverfahren (siehe Abschnitt 5) ist, daß $p-1$ in „kleine“ Primfaktoren zerfällt. Da $p^2 - 1 = (p-1)(p+1)$, sieht es zunächst so aus, als wäre durch den Übergang zu $(A/Ap)^*$ nichts gewonnen. Die entscheidende Idee ist nun, statt der vollen Gruppe die (eindeutig bestimmte) Untergruppe der Ordnung $p+1$ ins Spiel zu bringen. Dann ist nur die Primfaktorzerlegung von $p+1$ ausschlaggebend, und dies eröffnet neue Chancen. Wir werden die $(p+1)$-Verfahren nicht vollständig diskutieren (siehe dazu [**Fors**]), sondern uns auf den Lucas-Test für Mersennesche Primzahlen beschränken: für eine Mersennesche Zahl $M = 2^q - 1$, q prim, besitzt $M + 1 = 2^q$ in der Tat nur kleine Primfaktoren.

Zunächst ist die Untergruppe der Ordnung $p+1$ zu beschreiben. Wir machen uns dazu klar, daß A/Ap zum Körper $\mathbb{Z}_p$ in der gleichen Beziehung steht wie $\mathbb{Q}$ zu seiner quadratischen Erweiterung $\mathbb{Q}[\sqrt{d}]$:

(a) Zunächst können wir $\mathbb{Z}_p$ als Unterring von A/Ap betrachten, indem wir a mod p mit der Restklasse von a in A/Ap identifizieren. Offensichtlich ist A/Ap ein Vektorraum der Dimension 2 über $\mathbb{Z}_p$, in dem die Restklassen von 1 und ω_d eine Basis bilden (vergleiche Satz 12.5). Wir können als zweites Basiselement

aber auch stets die Restklasse von $\sqrt{d}$ wählen: Falls die Restklasse von $\sqrt{d}$ in $\mathbb{Z}_p \subset A/Ap$ liegt, muß d quadratischer Rest modulo p sein, was aber durch die Voraussetzung $(d/p) = -1$ ausgeschlossen ist.

Dies rechtfertigt die Schreibweisen $A/Ap = \mathbb{Z}_p[\sqrt{d}]$ und $a + b\sqrt{d}$ für die Elemente von A/Ap, wobei a und b durch $\mathbb{Z}_p$ laufen.

(b) Die Konjugation auf A induziert die Konjugation $(a + b\sqrt{d})' = a - b\sqrt{d}$ auf $\mathbb{Z}_p[\sqrt{d}]$. Diese ist wieder ein involutorischer Automorphismus. Norm und Spur können nun entweder von A_d „geerbt“ oder direkt definiert werden: Für $x, y \in \mathbb{Z}_p[\sqrt{d}]$ ist $N(x) = xx' \in \mathbb{Z}_p$, $S(x) = x + x' \in \mathbb{Z}_p$, und es gilt $N(xy) = N(x)N(y)$, $S(x + y) = S(x) + S(y)$.

(c) Da jeder Automorphismus α des Körpers $\mathbb{Z}_p[\sqrt{d}]$ auf dem Teilkörper $\mathbb{Z}_p$ als identische Abbildung wirkt (dies folgt aus $\alpha(1) = 1$), ist er durch $\alpha(\sqrt{d})$ schon eindeutig bestimmt. Da andererseits nur $\sqrt{d}$ und $-\sqrt{d}$ für $\alpha(\sqrt{d})$ in Frage kommen, gilt: Die Identität und die Konjugation sind die einzigen Automorphismen von $\mathbb{Z}_p[\sqrt{d}]$.

Nach diesen Vorbereitungen können wir leicht die Untergruppe der Ordnung $p + 1$ von $\mathbb{Z}_p[\sqrt{d}]^*$ bestimmen:

Satz 14.8. *Der Kern des Gruppenhomomorphismus $N : \mathbb{Z}_p[\sqrt{d}]^* \to \mathbb{Z}_p^*$ ist die (eindeutig bestimmte und notwendig zyklische) Untergruppe der Ordnung $p + 1$ von $\mathbb{Z}_p[\sqrt{d}]^*$.*

Beweis. Als Einheitengruppe eines endlichen Körpers ist $\mathbb{Z}_p[\sqrt{d}]^*$ zyklisch; folglich besitzt sie zu jedem Teiler q ihrer Ordnung genau eine Untergruppe der Ordnung q, und diese ist ebenfalls zyklisch.

Die Abbildung $F(x) = x^p$ ist ein Automorphismus des Körpers $\mathbb{Z}_p[\sqrt{d}]$, weil alle Binomialkoeffizienten $\binom{p}{k}$, $0 < k < p$, durch p teilbar sind. Da die Gleichung $x^p - x = 0$ in $\mathbb{Z}_p[\sqrt{d}]$ höchstens p Lösungen hat, aber alle Elemente von $\mathbb{Z}_p$ Lösungen sind, bleiben genau die Elemente von $\mathbb{Z}_p$ unter F fest. Also ist F *nicht* die Identität. Wie oben gezeigt, muß dann

$$x^p = x' \qquad \text{für alle } x \in \mathbb{Z}_p[\sqrt{d}]$$

gelten.

Es folgt $N(x) = xx' = x^{p+1}$ für alle $x \in \mathbb{Z}_p[\sqrt{d}]$. Der Kern von N besteht daher genau aus den Elementen $x \in \mathbb{Z}_p[\sqrt{d}]^*$, deren Ordnung $p + 1$ teilt. Diese bilden die Untergruppe der Ordnung $p + 1$. □

Wir können nun den Lucas-Test für Mersennesche Primzahlen begründen:

Satz 14.9. *Sei p eine ungerade Primzahl. Dann sind äquivalent:*

(a) *Die Mersennesche Zahl $M_p = 2^p - 1$ ist eine Primzahl.*

(b) *Für die durch* $s_0 = 4$, $s_{n+1} = s_n^2 - 2$ *definierte Folge gilt:* $M_p \mid s_{p-2}$.

Beweis. Zunächst ist das Geheimnis der rekursiv definierten Folge (s_n) zu lüften. Sei $\varepsilon = 2 + \sqrt{3} \in A = A_3$. Dann gilt $N(\varepsilon) = 1$ und $\varepsilon' = \varepsilon^{-1}$. (Daß ε die Fundamentaleinheit ist, ist unwesentlich.) Wir behaupten:

$$s_n = S\left(\varepsilon^{2^n}\right) \qquad \text{für alle } n.$$

Dies überprüft man mit Induktion. Für $n = 0$ ist $S\left(\varepsilon^{2^0}\right) = S(\varepsilon) = 4$. Für $n > 0$ ergibt sich

$$s_{n+1} = s_n^2 - 2 = \left(\varepsilon^{2^n} + \varepsilon^{-2^n}\right)^2 - 2 = \varepsilon^{2^{n+1}} + \varepsilon^{-2^{n+1}} = S\left(\varepsilon^{2^{n+1}}\right).$$

Wir beginnen mit der einfacheren Implikation (b) $\Rightarrow$ (a), die wir schon viel früher hätten beweisen können: Die Voraussetzung $s_{p-2} \equiv 0 \mod M_p$ ist äquivalent mit

$$\varepsilon^{2^{p-2}} + \varepsilon^{-2^{p-2}} \equiv 0 \mod M_p.$$

Multiplikation mit $\varepsilon^{2^{p-2}}$ liefert

$$\varepsilon^{2^{p-1}} \equiv -1 \mod M_p \qquad \text{und somit} \qquad \varepsilon^{2^p} \equiv 1 \mod M_p.$$

Diese Kongruenzen gelten erst recht modulo eines jeden Primteilers q von M_p. Daher besitzt die Restklasse von ε in der Einheitengruppe von A/Aq die Ordnung 2^p, denn diese Ordnung teilt 2^p, aber nicht 2^{p-1}.

Wenn die Zahl M_p nicht prim ist, besitzt sie einen Primteiler $q \leq \sqrt{M_p}$. Die Restklasse von ε hat in der Einheitengruppe des Ringes A/Aq die Ordnung $2^p > q^2$. Andererseits hat $(A/Aq)^*$ weniger als q^2 Elemente. Dies ist unmöglich.

Nun zu (a) $\Rightarrow$ (b): Es gilt $M_p \equiv 1$ (3) und daher $(M_p/3) = 1$. Da aber $M_p \equiv 3$ (4), folgt $(3/M_p) = -1$. Die Primzahl M_p ist träge in A. Daher können wir die obigen Überlegungen auf $(\mathbb{Z}/M_p\mathbb{Z})[\sqrt{3}]$ anwenden.

Die Norm der Restklasse von ε ist die Restklasse der Norm. Folglich liegt ε in der Untergruppe U der Ordnung $M_p + 1 = 2^p$ von $(Z/M_p\mathbb{Z})[\sqrt{3}]$. Wir behaupten, daß diese sogar von der Restklasse von ε erzeugt wird. Da die Ordnung von U eine Potenz von 2 ist, sind die erzeugenden Elemente genau die Nichtquadrate in U.

Zum Zwecke eines Widerspruchs nehmen wir an, die Restklasse von ε sei ein Quadrat in U. In A bedeutet dies: Es existieren $u, v \in \mathbb{Z}$ mit $N(u + v\sqrt{3}) = u^2 - 3v^2 \equiv 1 \mod M_p$ und

$$2 + \sqrt{3} \equiv (u + v\sqrt{3})^2 = (u^2 + 3v^2) + 2uv\sqrt{3} \mod M_p.$$

Daraus folgt

$$2 \equiv 2u^2 - 1 \mod M_p, \qquad \text{also} \qquad 2u^2 \equiv 3 \mod M_p$$

in $\mathbb{Z}$. Wegen $M_p \equiv -1 \pmod 8$ ist 2 quadratischer Rest modulo M_p. Andererseits ist $(3/M_p) = -1$. Widerspruch!

Da $\left(\varepsilon^{2^{p-1}}\right)^2 \equiv 1$, aber $\varepsilon^{2^{p-1}} \not\equiv 1 \mod M_p$, muß -1 die Restklasse von $\varepsilon^{2^{p-1}}$ sein, denn die quadratische Gleichung $x^2 = 1$ hat nur zwei Lösungen im Körper $(Z/M_p\mathbb{Z})[\sqrt{3}]$. Mithin folgt

$$S\left(\varepsilon^{2^{p-1}}\right) \equiv -2 \mod M_p.$$

Dies impliziert nach der Rekursionsformel $s_{p-2}^2 \equiv 0 \mod M_p$, und damit $s_{p-2} \equiv 0 \mod M_p$. □

Lucas hat 1876 mit Hilfe seines Tests gezeigt, daß M_{127} eine Primzahl ist. Der Lucas-Test erfordert „nur“ $p-2$ Multiplikationen von Zahlen der Größenordnung M_p. Dies erklärt, weshalb die größten bekannten Primzahlen Mersennesche Zahlen sind. Dennoch ist der Rechenaufwand in deren Bereich bereits enorm. Man kann zu seiner Reduktion raffinierte, auf der sogenannten schnellen Fouriertransformation beruhende Multiplikationsalgorithmen verwenden (siehe [**Fors**]).

Übungen.

14.10. Zerlege das von 18 erzeugte Ideal in A_{-17} in ein Produkt von Primidealen und erkläre, wie sich die Darstellungen $18 = 2 \cdot 3 \cdot 3$ und $18 = (1 + \sqrt{-17})(1 - \sqrt{-17})$ aus der Primzerlegung ergeben.

14.11. Wieviele Ideale der Norm 252 gibt es in A_{-5}?

14.12. Bestimme alle Ideale in A_{-29}, die das Element 30 enthalten.

14.13. Zerlege (a) das Ideal $A(41 + 5\sqrt{3})$ in $A = A_3$ in ein Produkt von Primidealen, (b) das Element $41 + 5\sqrt{3}$ in A_3 in ein Produkt von Primelementen.

ABSCHNITT 15

Die Endlichkeit der Klassenzahl

Wir wissen, daß die Ringe A_d i.a. nicht faktoriell sind. In diesem Abschnitt diskutieren wir ein Hilfsmittel, mit dem man die Abweichung eines Ringes A_d vom Faktoriellsein messen kann. Mit seiner Hilfe kann man bei vielen Anwendungen auch dort weiterkommen, wo man auf den ersten Blick meint, ohne die Zerlegung von Elementen in Primfaktoren nicht auszukommen. Dies trifft zum Beispiel zu, wenn man das quadratische Reziprozitätsgesetz im Rahmen der Theorie der quadratischen Zahlkörper beweisen will, ferner bei Anwendungen auf die Theorie der quadratischen Formen.

Wir haben in Satz 11.5 gesehen, daß die Hauptidealringe unter den A_d faktoriell sind. Es gilt auch die Umkehrung:

Satz 15.1. *Sei $d \in \mathbb{Z}$ quadratfrei. Der Ring A_d ist genau dann Hauptidealring, wenn er faktoriell ist.*

Beweis. Sei $A := A_d$ faktoriell. Jedes Ideal in A ist Produkt von Primidealen, und ein Produkt von Hauptidealen ist stets ein Hauptideal. Daher genügt es zu zeigen, daß die Primideale in A Hauptideale sind. Sei $\mathfrak{p} \subset A$ Primideal. Dann ist $\mathfrak{p} = Ap$, $p \in \mathbb{Z}$ prim, also „direkt“ Hauptideal oder

$$\mathfrak{p}\mathfrak{p}' = Ap, \qquad p \in \mathbb{Z} \text{ Primzahl.}$$

Das Element p zerfällt nach Voraussetzung in Primfaktoren: $p = \pi_1 \dots \pi_n$. (Natürlich ist $n = 1$ oder $n = 2$, aber das ist hier unwichtig.) Folglich: $Ap = (A\pi_1) \dots (A\pi_n) = \mathfrak{p}\mathfrak{p}'$. Wegen der Eindeutigkeit der Zerlegung in Primideale folgt: Es existiert ein i mit $\mathfrak{p} = A\pi_i$. □

Die gebrochenen Hauptideale, das sind die gebrochenen Ideale der Form Aa, $a \in \mathbb{Q}[\sqrt{d}]$, bilden eine Untergruppe H der Gruppe I aller gebrochenen Ideale. Satz 15.1 besagt: Genau dann ist A_d faktoriell, wenn diese beiden Gruppen übereinstimmen. Die Abweichung der Untergruppe H von der Gruppe I wird gemessen durch die Faktorgruppe I/H.

Definition. Die Faktorgruppe $C(A) := I/H$ heißt *Klassengruppe* (oder auch Idealklassen- oder Divisorenklassengruppe) des Ringes A (oder auch des quadratischen Zahlkörpers $\mathbb{Q}[\sqrt{d}]$).

Nochmals 15.1: Genau dann ist A faktoriell, wenn $C(A)$ nur aus dem neutralen Element besteht, mit anderen Worten: wenn $C(A)$ die Ordnung 1 hat.

Wir sagen, Ideale $\mathfrak{a}, \mathfrak{b} \neq 0$ *gehören zur gleichen Klasse*, $\mathfrak{a} \sim \mathfrak{b}$, wenn ihre Restklasse in $C(A)$ übereinstimmt.

Satz 15.2. *Genau dann gehören $\mathfrak{a}$ und $\mathfrak{b}$ zur gleichen Klasse, wenn es Elemente $a, b \in A$, $a, b \neq 0$, mit $a\mathfrak{a} = b\mathfrak{b}$ gibt.*

Beweis. Nach Definition von $C(A)$ gilt $\mathfrak{a} \sim \mathfrak{b}$ genau dann, wenn $\mathfrak{a}\mathfrak{b}^{-1} \in H$, wenn also ein $c \in \mathbb{Q}[\sqrt{d}]$ existiert mit

$$\mathfrak{a}\mathfrak{b}^{-1} = Ac.$$

Mit $c = b/a$, $a, b \in A$, folgt $a\mathfrak{a} = b\mathfrak{b}$. Umgekehrt: Falls $a\mathfrak{a} = b\mathfrak{b}$, so $\mathfrak{a}\mathfrak{b}^{-1} = Ab/a$, und $\mathfrak{a} \sim \mathfrak{b}$. □

Ob das Hilfsmittel Klassengruppe wirklich brauchbar ist, hängt entscheidend davon ab, welche Aussagen man über sie machen kann. Wir wollen hier eine fundamentale Eigenschaft beweisen: Die Klassengruppe ist endlich.

Dazu beweisen wir zunächst, daß es nur endlich viele Ideale $\mathfrak{a} \subset A$ gibt, deren Norm unterhalb einer gegebenen Schranke liegt.

Satz 15.3. *Sei r eine reelle Zahl. Dann gibt es nur endlich viele Ideale $\mathfrak{a} \subset A$ mit $N(\mathfrak{a}) \leq r$.*

Beweis. Sei etwa $N(\mathfrak{a}) = n$. Dann ist $nA \subset \mathfrak{a} \subset A$. A/nA ist endlich. Es existieren nur endlich viele Ideale zwischen nA und A. □

Zunächst ist klar: Jede Idealklasse enthält ein „ganzes“ Ideal $\mathfrak{a} \subset A$, denn sie enthält ein gebrochenes Ideal $\mathfrak{b}$, und zu diesem gibt es per Definition ein $b \in A$ mit $\mathfrak{b} = \mathfrak{a}/b$, $\mathfrak{a}$ Ideal in A. Satz 15.3 impliziert dann: Wenn es ein $r \in \mathbb{R}$ gibt derart, daß jede Idealklasse ein Ideal $\mathfrak{a} \subset A$ mit $N(\mathfrak{a}) \leq r$ enthält, dann gibt es auch nur endlich viele Idealklassen.

Die Verwirklichung dieses Beweisplans beruht auf der Möglichkeit, in einem Ideal Elemente „kleiner“ Norm zu finden, und diese wiederum auf dem Minkowskischen Gitterpunktsatz, den wir zunächst besprechen. (Elementare, das Schubfachprinzip benutzende Überlegungen führen zwar prinzipiell auch zum Ziel, münden aber in effektiv wesentlich schlechtere Schranken, vgl. [**Hass**], pp. 375–377.)

Ein (ebenes) *Gitter* ist eine Teilmenge der Form $\mathbb{Z}a_1 + \mathbb{Z}a_2$ von $\mathbb{R}^2$, $a_1, a_2 \in \mathbb{R}^2$ linear unabhängig. Die *Grundmasche* $\mathfrak{G}$ von L ist die konvexe Hülle der Menge $\{0, a_1, a_2, a_1 + a_2\}$. Die Teilmengen $a + \mathfrak{G}$, $a \in L$, heißen Maschen des Gitters. Ihre Vereinigung ist die ganze Ebene; sie haben paarweise keine inneren Punkte gemeinsam.

Die Grundmasche ist natürlich abhängig von der gewählten Basis, nicht hingegen ihr Flächeninhalt – vgl. dazu die Argumentation zur Basisunabhängigkeit der

Diskriminante. Überraschenderweise geht nur der Flächeninhalt in den folgenden Satz ein, bei dessen Beweis wir elementare Aussagen der Maßtheorie benutzen.

Satz 15.4 (Minkowski)**.** *Sei $L = \mathbb{Z}a_1 + \mathbb{Z}a_2$ ein Gitter in $\mathbb{R}^2$ mit Grundmasche $\mathfrak{G}$. Sei ferner K eine zu 0 zentralsymmetrische, kompakte konvexe Teilmenge von $\mathbb{R}^2$. Wenn*

$$V(K) \geq 4V(\mathfrak{G}),$$

so enthält K einen von 0 verschiedenen Punkt des Gitters.

Beweis. Wir bilden die Durchschnitte $K_i = K \cap \mathfrak{H}_i$ mit den Maschen $\mathfrak{H}_i$ des Gitters $2L = \mathbb{Z}2a_1 + \mathbb{Z}2a_2$. Seine Grundmasche $\mathfrak{H} = \mathfrak{H}_1$ hat das Volumen $4V(\mathfrak{G}) \leq V(K)$. Bis auf endlich viele Ausnahmen sind die K_i leer, da K beschränkt ist. Sämtliche K_i sind kompakt.

Mittels Parallelverschiebung von $\mathfrak{H}_i$ auf $\mathfrak{H}$ „transportiert“ man nun K_i in $\mathfrak{H}$ hinein und erhält als Bild K_i'. Wir nehmen an, die K_i' wären paarweise disjunkt, insbesondere etwa

$$K_1' \cap \bigcup_{i \neq 1} K_i' = \emptyset.$$

Da $K \cap \mathfrak{H}_i \neq \emptyset$ für die vier den Nullpunkt enthaltenden Maschen $\tilde{\mathfrak{H}}_1, \ldots, \tilde{\mathfrak{H}}_4$, ist $K_1' \neq \emptyset$, $\bigcup_{i \neq 1} K_i' \neq \emptyset$. Ferner sind beide Mengen kompakt. Daher haben sie positiven Minimalabstand ε. Wir wählen einen Punkt $x_0 \in K_1'$ und einen Punkt $x_1 \in \bigcup_{i \neq 1} K_i'$, und zwar beide als innere Punkte von $\mathfrak{H}$ (dies ist möglich, weil sowohl $V(K_1)$ als auch $V(\bigcup_{i \neq 1} K_i) > 0$: der Nullpunkt muß innerer Punkt von K sein.) Die Verbindungsstrecke S verläuft im Inneren der konvexen Menge $\mathfrak{H}$. Da K_1' konvex und kompakt ist, ist $S \cap K_1'$ ebenfalls eine Strecke. Ihrem von x_0 eventuell verschiedenen Endpunkt folgt eine Strecke S' mindestens der Länge ε, deren Inneres keinen Punkt von $K_1' \cup \bigcup_{i \neq 1} K_i'$ enthält. Der offene Kreis mit Radius $\varepsilon/2$ um den Mittelpunkt von S' ist ebenfalls disjunkt zu $K_1' \cup \bigcup_{i \neq 1} K_1'$ und ein eventuell kleinerer Kreis mit Radius δ um denselben Mittelpunkt ganz in $\mathfrak{H}$ enthalten. Damit folgt:

$$V(\mathfrak{H}) \geq V\left(\bigcup_i K_i'\right) + \pi\delta^2 > V\left(\bigcup_i K_i'\right) = \sum_i V(K_i') \geq V(\mathfrak{H}).$$

Widerspruch!

Also gibt es $i \neq j$ mit $K_i' \cap K_j' \neq \emptyset$. Sei y das Urbild von $x \in K_i' \cap K_j'$ in K_i und z das Urbild in K_j. Dann gilt $y \neq z$, weil die Verschiebungsvektoren

$$x - y \in 2L \quad \text{und} \quad x - z \in 2L$$

verschieden sind, ferner $-(x - y) + (x - z) = y - z \in 2L$. Damit folgt $1/2(y - z) \in L$. Da K punktsymmetrisch zu 0, gilt auch $-z \in K$, und weil K konvex,

schließlich $1/2(y + (-z)) \in K$. Insgesamt:

$$\frac{1}{2}(y - z) \in L \cap K. \qquad \square$$

Der Beweis zeigt, daß sich der Satz direkt auf Gitter im $\mathbb{R}^n$ verallgemeinern läßt; man muß nur den Faktor 4 durch 2^n ersetzen. Als Folgerung ergibt sich:

Satz 15.5. *Sei Δ die Diskriminante von $\mathbb{Q}[\sqrt{d}]$. Zu jedem Ideal $\mathfrak{a} \subset A_d$, $\mathfrak{a} \neq 0$, gibt es ein $a \in \mathfrak{a}$, $a \neq 0$ mit*

$$|N(a)| \leq \begin{cases} \frac{2}{\pi}\sqrt{|\Delta|}N(\mathfrak{a}) & \text{für } d < 0, \\ \frac{1}{2}\sqrt{|\Delta|}N(\mathfrak{a}) & \text{für } d > 0. \end{cases}$$

Beweis. Wir betten $\mathbb{Q}[\sqrt{d}]$ in $\mathbb{R}^2$ ein:

(i) im Fall $d < 0$ mittels seiner natürlichen Einbettung φ in $\mathbb{C} \cong \mathbb{R}^2$,

(ii) im Fall $d > 0$ mittels der $\mathbb{Q}$-linearen Abbildung

$$\varphi : r + s\sqrt{d} \mapsto (r, s\sqrt{d}).$$

Sei $a_1 = \alpha_1$, $a_2 = \alpha_2 + \beta_2\omega$, $\omega = \omega_d$, eine Basis von $\mathfrak{a}$. Die Elemente $\varphi(a)$, $a \in \mathfrak{a}$, bilden das von $\varphi(a_1)$ und $\varphi(a_2)$ aufgespannte Gitter L. Also

$$\varphi(a_1) = (\alpha_1, 0), \quad \varphi(a_2) = (\alpha_2, \beta_2\sqrt{|d|}) \qquad d \equiv 2, 3 \ (4)$$

$$\varphi(a_1) = (\alpha_1, 0), \quad \varphi(a_2) = (\alpha_2 + \beta_2/2,\ \beta_2\sqrt{|d|}/2) \quad d \equiv 1 \ (4).$$

Damit ergibt sich

$$V(\mathfrak{G}) = \left|\det \begin{pmatrix} \alpha_1 & 0 \\ \alpha_2 & \beta_2\sqrt{|d|} \end{pmatrix}\right| = \left|\alpha_1\beta_2\sqrt{|d|}\right|$$

im Fall $d \equiv 2, 3 \ (4)$ und analog $V(\mathfrak{G}) = \left|(1/2)\alpha_1\beta_2\sqrt{|d|}\right|$, falls $d \equiv 1 \ (4)$.

Somit gilt in beiden Fällen

$$V(\mathfrak{G}) = \frac{1}{2}\sqrt{|\Delta|}N(\mathfrak{a}), \qquad 4V(\mathfrak{G}) = 2\sqrt{|\Delta|}N(\mathfrak{a}),$$

denn $\Delta = 4d$ im Fall $d \equiv 2, 3 \ (4)$. Sei zunächst $d < 0$. Der abgeschlossene Kreis K um 0 mit Radius

$$\sqrt{\frac{2}{\pi}\sqrt{|\Delta|}N(\mathfrak{a})}$$

enthält einen Gitterpunkt $\varphi(a)$, $a \in \mathfrak{a}$. Es gilt $N(a) \leq 2/\pi\sqrt{|\Delta|}N(\mathfrak{a})$, wie behauptet.

Im Fall $d > 0$ ist die Norm *nicht* das Quadrat des euklidischen Abstands. Die Punkte $x \in \mathbb{R}^2$ mit $|N(\varphi^{-1}(x))| \leq c$ bilden vielmehr das Komplement der von vier Hyperbelästen eingeschlossenen konvexen Gebiete; siehe Abbildung 1. Dieses Gebiet ist nicht konvex, aber für die Punkte (x, y) mit $|x| \leq \sqrt{c}$, $|y| \leq \sqrt{c}$

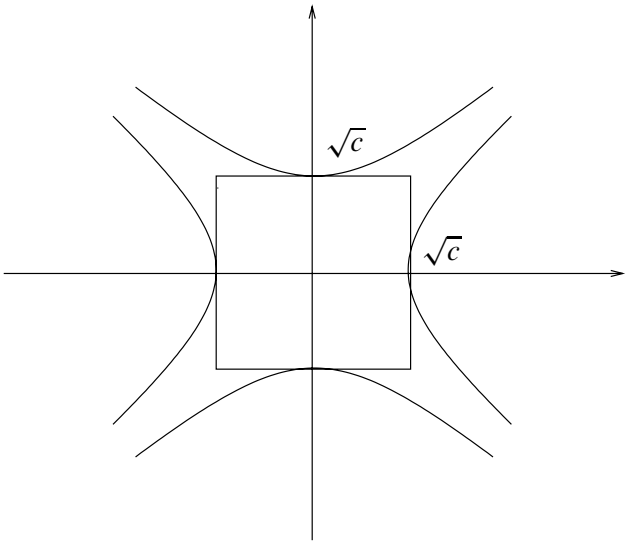

ABBILDUNG 1

gilt $|N(\varphi^{-1}(x, y))| = |x^2 - y^2| \leq c$. Wir wählen daher c so, daß

$$4c = 4\left(\frac{1}{2}\sqrt{|\Delta|} \cdot N(\mathfrak{a})\right).$$

Dann enthält das achsensymmetrische Quadrat mit Mittelpunkt 0 und Kantenlänge $2\sqrt{c}$ mindestens einen Gitterpunkt. □

Die entscheidende Folgerung aus 15.5

Satz 15.6. *Jede Idealklasse enthält ein Ideal* $\mathfrak{a} \subset A_d$ *mit*

$$N(\mathfrak{a}) \leq \begin{cases} \frac{2}{\pi}\sqrt{|\Delta|} & \text{für } d < 0,\ \Delta = \Delta(A_d), \\ \frac{1}{2}\sqrt{|\Delta|} & \text{für } d > 0,\ \Delta = \Delta(A_d). \end{cases}$$

Beweis. Sei $\mathfrak{C}$ die betrachtete Idealklasse. Die inverse Klasse $\mathfrak{C}^{-1}$ enthält ein Ideal $\tilde{\mathfrak{a}} \subset A$. In $\tilde{\mathfrak{a}}$ wählen wir ein Element gemäß 15.5 und erhalten und $(Aa)\tilde{\mathfrak{a}}^{-1} \subset A$ weil $\tilde{\mathfrak{a}} \mid Aa$ und

$$N\left((Aa)\tilde{\mathfrak{a}}^{-1}\right) = \frac{|N(a)|}{N(\tilde{\mathfrak{a}})} \leq \begin{cases} \frac{2}{|\pi|}\sqrt{|\Delta|} & \text{für } d < 0, \\ \frac{1}{2}\sqrt{|\Delta|} & \text{für } d > 0. \end{cases}$$ □

Wie bereits diskutiert, ergibt sich als Folgerung der fundamentale Satz über die Endlichkeit der Klassenzahl:

Satz 15.7. *Die Klassengruppe* $C(A)$ *ist endlich.*

Definition. Die Ordnung der Klassengruppe heißt *Klassenzahl* von A (oder von $\mathbb{Q}[\sqrt{d}]$).

Häufig benötigt man folgende Aussage:

Satz 15.8. *Sei h die Klassenzahl von A und $\mathfrak{a}$ ein Ideal in A. Dann ist $\mathfrak{a}^h$ ein Hauptideal.*

Beweis. Sei $\mathfrak{C}$ die Klasse von $\mathfrak{a}$. Dann gehört $\mathfrak{a}^h$ zur Klasse $\mathfrak{C}^h$. Da aber h die Ordnung der Gruppe $C(A)$ ist, ist $\mathfrak{C}^h$ das neutrale Element in $C(A)$ nach dem Satz von Fermat für endliche Gruppen. Das neutrale Element ist gerade die Klasse der Hauptideale: $\mathfrak{a}^h \in H$. □

Sei r die in 15.5 gegebene Schranke. Um alle Idealklassen zu finden, bestimmt man alle Primideale $\mathfrak{p}_i$ mit $N(\mathfrak{p}) \leq r$ und bildet alle möglichen Produkte $\mathfrak{p}_1^{n_1} \cdots \mathfrak{p}_m^{n_m}$ mit $\prod N(\mathfrak{p}_i)^{n_i} \leq r$. Dann muß man herausfinden, welche dieser Produkte zur gleichen Klasse gehören. Aus diesen Daten kann man dann die Struktur der Klassengruppe ablesen. Diese Prozedur klingt sehr einfach, birgt aber einige Detailprobleme in sich, insbesondere im Fall $d > 0$. Diese wollen wir hier nicht im einzelnen diskutieren, sondern nur einige Beispiele betrachten:

Beispiele. (a) $d = -1$, $\Delta = -4$, $(2/\pi)\sqrt{|\Delta|} \leq 4/3$: Jede Idealklasse enthält ein Ideal der Norm 1, also A: Es gibt nur eine Idealklasse, nämlich die Hauptideale, und A ist folglich faktoriell, was wir natürlich schon wissen.

(b) $d = -2$, $\Delta = -8$, $(2/\pi)\sqrt{8} < 2$: A_{-2} ist faktoriell.

(c) $d = -3$, $\Delta = -3$, $(2/\pi)\sqrt{3} < 2$: A_{-3} ist faktoriell.

(d) $d = -5$, $\Delta = -20$, $(2/\pi)\sqrt{20} < 3$: Ein Ideal der Norm 2 ist notwendig ein Primideal, nämlich das Ideal $\mathfrak{p} = A2 + A(1 + \sqrt{-5})$. Da $\mathfrak{p}' = \mathfrak{p}$, ist $\mathfrak{p}$ das einzige Ideal der Norm 2. Andererseits wissen wir schon, daß $\mathfrak{p}$ kein Hauptideal ist. Daher gibt es zwei Idealklassen, die der Hauptideale und die von $\mathfrak{p}$.

(e) $d = -19$, $\Delta = -19$, $(2/\pi)\sqrt{19} < 3$: Wieder sind die Primideale der Norm 2 zu bestimmen. Dazu betrachtet man die Zerlegung von 2 in A_{-19}. Nach Satz 14.7 ist 2 träge, denn $-19 \equiv 5(8)$. Es gibt also kein Ideal der Norm 2. Folglich ist A_{-19} faktoriell, das einfachste Beispiel eines nichteuklidischen Hauptidealringes, das wir ja schon in Abschnitt 11 ausführlich analysiert haben.

(f) Sei $d = -14$, $\Delta = -56$, $(2/\pi)\sqrt{56} < 5$. Wir betrachten die Zerlegung von 2 und 3:

$$(56/2) = 0,\ 2 \text{ ist verzweigt, } A2 = \mathfrak{p}^2,\ \mathfrak{p} = A2 + A\sqrt{-14},\ N(\mathfrak{p}) = 2,$$
$$(-56/3) = (1/3) = 1,\ 3 \text{ ist zerlegt, } A3 = \mathfrak{q}_1\mathfrak{q}_2,\ \mathfrak{q}_1 = A3 + A(1 + \sqrt{-14}),$$
$$\mathfrak{q}_2 = A3 + A(1 - \sqrt{-14}) = \mathfrak{q}_1'.$$

Sei $\mathfrak{C}_1$ die Klasse von $\mathfrak{p}$, $\mathfrak{C}_2$ die Klasse von $\mathfrak{p}_1$, $\mathfrak{C}_3$ die Klasse von $\mathfrak{q}_2$, H die Klasse der Hauptideale. Wir wissen nun schon: $\mathfrak{C}_1^2 = H$, $\mathfrak{C}_2\mathfrak{C}_3 = H$, $\mathfrak{C}_1 = \mathfrak{C}_1^{-1}$, $\mathfrak{C}_3 = \mathfrak{C}_2^{-1}$.

Die einzigen Ideale mit Norm < 5 sind A, $\mathfrak{p}$, $\mathfrak{p}^2$, $\mathfrak{q}_1$, $\mathfrak{q}_2$. Da $\mathfrak{p}^2$ ein Hauptideal ist, gibt es höchstens 4 Idealklassen: H, $\mathfrak{C}_1$, $\mathfrak{C}_2$, $\mathfrak{C}_3$.

Es ist $N(\mathfrak{p}\mathfrak{q}_1) = N(\mathfrak{p}\mathfrak{q}_2) = 6$, und da es keine Elemente der Norm 6 gibt, sind $\mathfrak{p}\mathfrak{q}_1$ und $\mathfrak{p}\mathfrak{q}_2$ keine Hauptideale:

$$\mathfrak{C}_1\mathfrak{C}_2 = \mathfrak{C}_1\mathfrak{C}_3^{-1} \neq H: \quad \mathfrak{C}_1 \neq \mathfrak{C}_3$$
$$\mathfrak{C}_1\mathfrak{C}_3 = \mathfrak{C}_1\mathfrak{C}_2^{-1} \neq H: \quad \mathfrak{C}_1 \neq \mathfrak{C}_2.$$

$N(\mathfrak{q}_1^2) = 9$ zeigt, daß $\mathfrak{q}_1^2$ ebenfalls kein Hauptideal ist, denn die einzigen Elemente mit Norm 9 sind ± 3 und $A3 = \mathfrak{q}_1\mathfrak{q}_2 \neq \mathfrak{q}_1^2$:

$$\mathfrak{C}_2^2 = \mathfrak{C}_2\mathfrak{C}_3^{-1} \neq H: \quad \mathfrak{C}_2 \neq \mathfrak{C}_3.$$

Es gibt also genau vier Idealklassen: H, $\mathfrak{C}_1$, $\mathfrak{C}_2$, $\mathfrak{C}_3$.

Bis auf Isomorphie existieren zwei Gruppen der Ordnung 4, $\mathbb{Z}_4$ und $\mathbb{Z}_2 \times \mathbb{Z}_2$. Da in $\mathbb{Z}_2 \times \mathbb{Z}_2$ jedes Element zu sich selbst invers ist, aber $\mathfrak{C}_2 \neq \mathfrak{C}_2^{-1}$, muß $C(A)$ isomorph zu $\mathbb{Z}_4$ sein. Da $\mathfrak{C}_1 = \mathfrak{C}_1^{-1}$, ist $\mathfrak{C}_1$ das Element der Ordnung 2, und wir haben als endgültiges Ergebnis:

$$H,\ \mathfrak{C}_2,\ \mathfrak{C}_2^2 = \mathfrak{C}_1,\ \mathfrak{C}_2^3 = \mathfrak{C}_3$$

sind die vier Idealklassen von A.

Im Fall $d > 0$ liegt die Hauptschwierigkeit in dem Problem zu entscheiden, ob es Elemente einer gegebenen Norm gibt. Dies ist zwar stets in endlich vielen Schritten entscheidbar, aber etwas langwierig (siehe Aufgabe 10.15). Deshalb wollen wir den Fall $d > 0$ hier übergehen.

In den schon in den Sätzen 11.2 und 11.4 erwähnten Fällen

$$d = -7, -11, -19, -43, -67, -163$$

kann man wie oben in den Beispielen (a), (b), (c) und (e) leicht sehen, daß A_d faktoriell ist. Allgemein gilt das folgende einfache Kriterium:

Satz 15.9. *Sei $d < 0$ und Δ die Diskriminante von $\mathbb{Q}[\sqrt{d}]$. Für $\Delta \leq -10$ gilt: A_d ist genau dann faktoriell, wenn $(\Delta/p) = -1$ für alle Primzahlen $p < (2/\pi)\sqrt{|\Delta|}$.*

Beweis. Hinreichend ist das Kriterium für alle $\Delta < 0$: Falls $(\Delta/p) = -1$ für alle Primzahlen $p < (2/\pi)\sqrt{|\Delta|}$, sind alle diese Primzahlen träge. Folglich gibt es keine Primideale der Norm p, $p < (2/\pi)\sqrt{|\Delta|}$, und die einzigen Primideale mit Norm $< (2/\pi)\sqrt{|\Delta|}$ sind die Hauptideale Ap, $p \in \mathbb{Z}$, p prim, $p < \sqrt{(2/\pi)\sqrt{|\Delta|}}$.

Sei umgekehrt A_d faktoriell. Dann sind alle Primideale Hauptideale. Die Norm eines Elementes $a \in A_d$, $a \notin \mathbb{Z}$, beträgt mindestens $|\Delta|/4$:

$$N(\alpha + \beta\sqrt{d}) = \alpha^2 + \frac{|\Delta|}{4}\beta^2 \quad d \equiv 2, 3,\ (4),$$
$$N\left(\frac{\alpha + \beta\sqrt{d}}{2}\right) = \frac{\alpha^2 + |\Delta|\beta^2}{4} \quad d \equiv 1\ (4).$$

Für $\Delta \leq -10$ ist $2/\pi\sqrt{|\Delta|} \leq |\Delta|/4$. Also liegen die Teiler der Primzahlen p, $p < (2/\pi)\sqrt{|\Delta|}$, schon in $\mathbb{Z}$: p ist auch in A_d prim, also träge, und somit $(\Delta/p) = -1$. □

Man kann für $d > 0$ natürlich eine analoge hinreichende Bedingung formulieren.

Nun wollen wir noch das Rätsel der Polynome

$$g(x) := x^2 + x + q, \qquad q \in \mathbb{N},$$

klären, die für ganzzahliges x, $0 \leq x \leq q-2$ nur Primzahlen als Werte liefern:

Satz 15.10. *Genau dann nimmt $g(x)$ für $0 \leq x \leq q-2$ nur Primzahlen als Werte an, wenn $d = 1-4q$ quadratfrei und der Ring A_d faktoriell ist.*

Beweis. $\Leftarrow$: Beachte zunächst $d \equiv 1(4)$ und $d < 0$. Wir kommen des Rätsels Lösung näher, wenn wir die Norm von $x + w$, $x \in \mathbb{Z}$, $\omega = \omega_d$ bestimmen:

$$\begin{aligned} N(x+\omega) = (x+\omega)(x+\omega') &= (x + \frac{1}{2} + \frac{1}{2}\sqrt{d})(x + \frac{1}{2} - \frac{1}{2}\sqrt{d}) \\ = (x+\frac{1}{2})^2 - \frac{d}{4} \quad &= x^2 + x + \frac{1}{4} - \frac{d}{4} = g(x). \end{aligned}$$

Nehmen wir zunächst an, für ein x mit $0 \leq x \leq q-2$ sei $g(x)$ keine Primzahl. Es gilt

$$g(x) = x^2 + x + q \leq (q-2)^2 + q - 2 + q = q^2 - 2q + 2 < q^2$$

für $q > 1$ – der Fall $q = 1$ ist trivial. Also besitzt $g(x)$ einen Primfaktor $p \leq q-1$:

$$N(x+\omega) \equiv 0 \ (p).$$

Dies impliziert für das Ideal $\mathfrak{p} := Ap + A(x+\omega)$, daß $\mathfrak{p}\mathfrak{p}' = Ap$. Da es kein Element der Norm p in A_d gibt (dies sieht man wie im Beweis von 15.9), ist $\mathfrak{p}$ kein Hauptideal.

$\Rightarrow$: Sei nun $g(x)$ prim für $0 \leq x \leq q-2$. Für $q = 1, 2, 3$ ist $d = 1-4q$ quadratfrei und A_d faktoriell. Sei also $q \geq 4$. Wir schreiben $d = c^2\tilde{d}$, $c > 0$. Dann ist c ungerade, $\tilde{d} \equiv 1(4)$ und

$$g\left(\frac{c-1}{2}\right) = c^2\left(\frac{1-\tilde{d}}{4}\right),$$

aber $(c-1)/2 < c \leq \sqrt{d} < q$, also $(c-1)/2 \leq q-2$. Nach Voraussetzung muß $c = 1$ sein: d ist quadratfrei.

Wegen $d \equiv 1(4)$ ist $\Delta = d$. Nach 15.9 genügt es zu zeigen: $(d/p) = -1$ für alle Primzahlen p mit

$$p < \frac{2}{\pi}\sqrt{d} < \sqrt{d} < q.$$

Wir nehmen an: $(d/p) = 0$ oder $(d/p) = 1$ für eine solche Primzahl. Dann gilt in $A_d =: A$:

$$Ap = \mathfrak{p}\mathfrak{p}', \qquad \mathfrak{p} = Ap + A(a + \omega)$$

und $N(a + \omega) = g(a) \equiv 0 \ (p)$. Sei x so gewählt, daß $0 \leq x < p$ und $x \equiv a \ (p)$. Dann ist

$$g(x) \equiv g(a) \equiv 0 \ (p).$$

Da $g(x)$ eine Primzahl ist, muß $g(x) = p$ und folglich $p \geq q$ sein, ein Widerspruch! □

Die Anwendung idealtheoretischer Methoden auf diophantische Gleichungen wollen wir am Beispiel der *Bachetschen Gleichung* studieren, das spezieller schon in Aufgabe 7.15 vorkam:

$$y^2 = x^3 + k.$$

Zu dieser diophantischen Gleichung gibt es eine ausgedehnte Literatur, aber keine Aussagen, die alle Lösungen für beliebige k beschreiben. Es ist aber bekannt, daß es stets nur endlich viele Lösungen gibt (Mordell).

Satz 15.11. *Sei $k < 0$, k quadratfrei, $k \equiv 2, 3 \mod 4$. Die Klassenzahl von $\mathbb{Q}[\sqrt{k}]$ sei nicht durch 3 teilbar. Dann besitzt die diophantische Gleichung $y^2 = x^3 + k$ genau dann eine Lösung, wenn es ein $a \in \mathbb{Z}$ mit $k = \pm 1 - 3a^2$ gibt. In diesem Fall sind $(x, y) = (a^2 - k, \pm a(a^2 + 3k))$ die einzigen Lösungen.*

Beweis. In $A_k = A$ ist

$$(y - \sqrt{k})(y + \sqrt{k}) = x^3.$$

Sei $\mathfrak{p}$ ein Primteiler vor $A(y - \sqrt{k})$. Wir nehmen an, daß $\mathfrak{p}$ auch $A(y + \sqrt{k})$ teilt. Dann gilt

$$\mathfrak{p} \mid A2\sqrt{k}, \ N(\mathfrak{p}) \mid 2k \quad \text{und} \quad \mathfrak{p} \mid A2y, \ N(\mathfrak{p}) \mid 4y^2,$$

ferner ohnehin $\mathfrak{p} \mid Ax$, $N(\mathfrak{p}) \mid x^2$. Insbesondere sind $2k$, x^2, $4y^2$ nicht teilerfremd. Dies aber ist nicht möglich:

(i) Falls $2 \mid x$, folgt $y^2 - k \equiv 0 \ (8)$. Weil $y^2 \equiv 0, 1, 4 \ (8)$, aber $k \equiv 2, 3, 6, 7 \ (8)$, kann dies nicht sein.

(ii) Falls eine Primzahl p sowohl x als auch y teilt, muß ihr Quadrat k teilen, aber k ist quadratfrei.

Zusammengefaßt: $A(y - \sqrt{k})$ und $A(y + \sqrt{k})$ sind teilerfremd. Die Eindeutigkeit der Primfaktorzerlegung impliziert:

$$A(y + \sqrt{k}) = \mathfrak{a}^3.$$

Die Ordnung der Klasse von $\mathfrak{a}$ ist durch 3 teilbar, wenn sie nicht das neutrale Element in der Klassengruppe ist. Dies ist aber unmöglich nach Voraussetzung über

die Klassenzahl. Also muß $\mathfrak{a}$ selbst ein Hauptideal sein, $\mathfrak{a} = Az$: Es gilt

$$y + \sqrt{k} = \varepsilon z^3$$

mit einer Einheit $\varepsilon \in A$.

Außer im Fall $k = -1$ gilt stets $\varepsilon = \pm 1$. Da dann $\varepsilon^3 = \varepsilon$, dürfen wir

$$y + \sqrt{k} = z^3$$

annehmen. Wegen $k \equiv 2, 3 \ (4)$ ist

$$y + \sqrt{k} = (a + b\sqrt{k})^3 = (a^3 + 3ab^2k) + (3a^2b + b^3k)\sqrt{k}.$$

Also $b(3a^2 + b^2k) = 1$, somit $b = \pm 1$, $3a^2 + k = \pm 1$, und $y = a^3 + 3ak$, $x = a^2 - k$. Ersetzen von a durch $-a$ gibt die zweite Lösung.

Es bleibt der Fall $k = -1$, $\varepsilon = i$ (Beachte $-i = i(-1)^3$):

$$y + \sqrt{-1} = \sqrt{-1}(a + b\sqrt{-1})^3 = -(3a^2b - b^3) + (a^3 - 3ab^2)\sqrt{-1}.$$

Somit $a(a^2 - 3b^2) = 1$ und $a = 1$, $b = 0$, $x = 1$, $y = 0$. Genau diese Lösung nennt der Satz. □

Ein Beispiel: $y^2 = x^3 - 74$ hat genau die Lösungen $(59, \pm 985)$.

Auch im Fall $k \equiv 5(8)$ sieht man wie oben, daß $A(y + \sqrt{k})$ und $A(y - \sqrt{k})$ teilerfremd sind. Man setzt dann (für $k \neq -3$)

$$y + \sqrt{k} = \left(\frac{a + b\sqrt{3}}{2}\right)^3, \qquad a \equiv b \ (2),$$

und hat dann die Gleichung $b(3a^2 + b^2k) = 8$ zu betrachten, also die Fälle $b = \pm 1, \pm 2, \pm 4, \pm 8$. Im Fall $k = -3$ ist außerdem noch zu beachten, daß es die Einheiten ω, ω^2 gibt.

Im Fall $k > 0$ ist die Situation wegen der schwerer zu überblickenden Einheiten erheblich komplizierter (man muß für ε die Fälle 1, η, η^2 mit einer Fundamentaleinheit η, betrachten). In gewissen Fällen kommt man dennoch weiter, vgl. etwa [**AdGo**], pp. 293, 294. Auch wenn 3 die Klassenzahl teilt, kann man manchmal durchkommen, vgl. ebendort.

Übungen.

15.12. Bestimme die Struktur der Klassengruppen von A_{-21} und A_{-31}.

15.13. Ebenso für A_{10}.

15.14. Seien p, q Primzahlen und $r > 0$ eine ungerade ganze Zahl, so daß $d = r^2 - 4q^p < 0$ und quadratfrei ist. Zeige, daß die Klassenzahl von A_d durch p teilbar ist.

15.15. Finde mit der Methode der vorangegangenen Aufgabe imaginär-quadratische Zahlkörper, deren Klassenzahl durch $2, 3, 5, 7, 11$ teilbar ist.

ABSCHNITT 16

Nochmals quadratische Reziprozität

In Abschnitt 9 haben wir schon angedeutet, wie man das quadratische Reziprozitätsgesetz mit Hilfe der quadratischen Zahlkörper beweisen kann. Die einschränkende Bedingung in Abschnitt 9, daß nämlich die beteiligten Ringe faktoriell sind, können wir nun durch den Einsatz des Zerlegungsgesetzes 14.7 und der Klassengruppe überwinden. Das dabei verwendete Argumentationsschema ist sehr transparent.

Zunächst rekapitulieren wir zwei Aussagen über die multiplikative Struktur von A_d, die wir schon bewiesen haben, und ergänzen sie um eine weitere:

Satz 16.1. *Sei $d \in \mathbb{Z}$ quadratfrei.*

(a) *Wenn $d > 0$ ist und einen Primteiler $p \equiv 3 \ (4)$ besitzt, hat die Fundamentaleinheit in A_d Norm* 1.

(b) *Wenn $d = 2$ oder $d \equiv 1 \ (4)$ eine Primzahl ist, hat die Fundamentaleinheit in A_d Norm -1.*

(c) *Seien $a, b \in A_d$. Wenn $N(a) = N(b) \neq 0$, dann existiert ein $c \in A_d$ mit*

$$\frac{a}{b} = \frac{c}{c'}.$$

Beweis. Die Teile (a) und (b) sind die Sätze 10.9 und 10.10. Für (c) setzen wir $c = aa' + ab' = bb' + ab'$, falls $a \neq -b$, und $c = \sqrt{d}$ im verbleibenden Fall. □

Man beachte, daß (b) in 10.10 wirklich ohne Benutzung des Reziprozitätsgesetzes oder seiner Ergänzungssätze hergeleitet worden ist. (Der Reinheit der Methode halber kann man diese Aussage aber auch aus Teil (c) und dem Zerlegungsgesetz gewinnen; siehe Aufgabe 16.4.)

Wir haben beim Beweis von (a) in 10.9 benutzt, daß $p \equiv 1 \ (4)$, falls -1 quadratischer Rest modulo p ist. Um auch den letzten Hauch eines Zykelschlusses aus diesem Abschnitt zu vertreiben, geben wir einen neuen Beweis des ersten Ergänzungssatzes:

$$\left(\frac{-1}{p}\right) = 1 \quad \Longleftrightarrow \quad p \equiv 1 \ (4)$$

Nach dem Zerlegungsgesetz folgt aus $(-4/p) = (-1/p) = 1$, daß p in A_{-1} zerlegt ist, $(p) = \mathfrak{p}\mathfrak{p}'$. Da A_{-1} faktoriell ist, ist $\mathfrak{p} = (\pi)$, $\pi = x + iy$, ein Hauptideal der Norm p. Da $N(\pi) > 0$, folgt $p = x^2 + y^2$ und daraus $p \equiv 1 \ (4)$.

Sei umgekehrt $p \equiv 1\ (4)$. Dann hat die Fundamentaleinheit in A_p gemäß 16.1(b) die Norm -1. Daraus resultiert eine Gleichung $x^2 - py^2 = -4$, und -1 ist quadratischer Rest modulo p.

Für den zweiten Ergänzungssatz und das quadratische Reziprozitätsgesetz brauchen wir eine wichtige Aussage über die Klassenzahl:

Satz 16.2. *Wenn die Diskriminante von A_d genau einen Primteiler besitzt, hat A_d eine ungerade Klassenzahl.*

Beweis. Für d kommen in Frage: $d = -1$, $d = \pm 2$ oder $d \equiv 1\ (4)$ und $|d|$ ist Primzahl. Wir wissen, daß A_{-1}, A_{-2} und A_2 Klassenzahl 1 haben, brauchen dies aber nicht auszunutzen.

Sei $A = A_d$. Wir nehmen an, die Klassenzahl h sei gerade. Dann gibt es eine Klasse $\mathfrak{C}$ der Ordnung 2. Da das Inverse einer Klasse durch Konjugation gewonnen wird, gilt

$$\mathfrak{C} = \mathfrak{C}^{-1} = \mathfrak{C}'.$$

Solche Klassen nennt man *ambig*. Für jedes Ideal $\mathfrak{a} \in \mathfrak{C}$ gilt

$$\mathfrak{a} \sim \mathfrak{a}'.$$

Wir können annehmen, daß $\mathfrak{a} \subset A$. Man kann nun nicht ohne weiteres schließen, daß $\mathfrak{a} = \mathfrak{a}'$ (vergleiche Aufgabe 16.7), aber dieses Ziel wollen wir erreichen. Zunächst gibt es jedenfalls $a, b \in A$ mit

$$a\mathfrak{a} = b\mathfrak{a}'$$

Dies impliziert $N(a) = \pm N(b)$. Im Fall $d < 0$ kommt nur $N(a) = N(b)$ in Frage, und im Fall $d > 0$ dürfen wir dies ebenfalls annehmen, denn andernfalls ersetzen wir b durch ηb, wobei η die Fundamentaleinheit ist: es gilt ja $N(\eta) = -1$ gemäß Satz 16.1(b). Nach Teil (c) dieses Satzes findet man nun ein $c \in A_d$ mit $a/b = c/c'$, was

$$c\mathfrak{a} = c'\mathfrak{a}'$$

impliziert. Für $\mathfrak{b} = c\mathfrak{a}$ ist dann $\mathfrak{b} = \mathfrak{b}'$, und die Primfaktorzerlegung von $\mathfrak{b}$ hat wegen ihrer Eindeutigkeit notwendig die Gestalt

$$\mathfrak{b} = (\mathfrak{p}_1\mathfrak{p}_1')^{r_1} \cdots (\mathfrak{p}_m\mathfrak{p}_m')^{r_m}\mathfrak{q}_1^{s_1} \cdots \mathfrak{q}_n^{s_n}$$

wobei $\mathfrak{q}_j = \mathfrak{q}_j'$ für $j = 1, \ldots, n$. Alle Ideale $\mathfrak{p}\mathfrak{p}'$ und $\mathfrak{q}_j^2$ werden von ganz-rationalen Zahlen erzeugt. Wenn wir durch diese Faktoren dividieren, erhalten wir ein Ideal

$$\mathfrak{c} = \mathfrak{c}',$$

in dessen Primfaktorzerlegung nur Primideale $\mathfrak{q} = \mathfrak{q}'$ und überdies der Vielfachheit 1 vorkommen. Solche Ideale heißen *ambig*. Beachte daß $\mathfrak{c}$ immer noch von der Klasse $\mathfrak{C}$ ist.

Die Primteiler von $\mathfrak{c}$ sind nach dem Zerlegungsgesetz sämtlich Primteiler der Diskriminante und daher von d. Da $|d|$ prim ist, gilt

$$\mathfrak{c} = A \qquad \text{oder} \qquad \mathfrak{c}^2 = (d).$$

Im ersten Fall ist $\mathfrak{c}$ ein Hauptideal, im zweiten aber auch, denn $(d) = (\sqrt{d})^2$. Dies ist ein Widerspruch dazu, daß $\mathfrak{C}$ die Ordnung 2 hat. □

Im Fall $d < 0$ hat Satz 16.2 eine Umkehrung (vergleiche Aufgabe 16.6). Sein Beweis zeigt, daß die Existenz einer nichttrivialen ambigen Idealklasse im Fall $d < 0$ auf eine Zerlegung der Diskriminante führt. Man kann dies zu einem Faktorisierungsverfahren mittels Klassengruppen imaginär-quadratischer Zahlkörper ausnutzen; siehe dazu [**Fors**].

Wir können jetzt auch den zweiten Ergänzungssatz beweisen:

$$\left(\frac{2}{p}\right) = \begin{cases} 1 & p \equiv \pm 1 \ (8), \\ -1 & p \equiv \pm 3 \ (8). \end{cases}$$

Sei dazu $p^* = p$ falls $p \equiv 1$ (4) und $p^* = -p$ im anderen Fall. Die Bedingung $p \equiv \pm 1$ (8) ist dann äquivalent zu $p^* \equiv 1$ (8). Dies bedeutet nach dem Zerlegungsgesetz: 2 zerfällt in A_{p^*}, $(2) = \mathfrak{p}\mathfrak{p}'$. Mit der Klassenzahl h von A_{p^*} gilt: $\mathfrak{p}^h$ ein Hauptideal, $\mathfrak{p}^h = (a)$. Im Fall $p^* < 0$ folgt sofort $N(a) = 2^h$. Im Fall $p^* = p > 0$ können wir dies aber auch annehmen: Notfalls ersetzen wir a durch ηa; die Fundamentaleinheit hat Norm -1. Dies führt auf eine Gleichung

$$4 \cdot 2^h = x^2 - p^* y^2.$$

Da h ungerade ist, ist 2 quadratischer Rest modulo p.

Wenn umgekehrt $(2/p) = 1$, ist p zerlegt in A_2. Da A_2 faktoriell ist (ungerade Klassenzahl genügt), folgt daraus mit schon bekanntem Schluß, daß $\pm p = u^2 - 2v^2$. Modulo 8 läßt das nur die Reste 1 und -1 zu. (Da die Fundamentaleinheit die Norm -1 hat, kann man $p = u^2 - 2v^2$ erreichen – das ändert aber nichts an der Schlußfolgerung.)

Nun zum quadratischen Reziprozitätsgesetz selbst:

$$\left(\frac{q}{p}\right) = \left(\frac{p}{q}\right), \quad \text{wenn } p \equiv 1 \ (4) \text{ oder } q \equiv 1 \ (4),$$
$$\left(\frac{q}{p}\right) = -\left(\frac{p}{q}\right), \quad \text{wenn } p \equiv 3 \ (4) \text{ und } q \equiv 3 \ (4).$$

Der Übersicht halber führen wir zusätzliche Variablen ein: Seien u und v verschiedene ungerade Primzahlen. Falls $u \equiv 3$ (4), setzen wir $u^* = -u$, sonst $u^* = u$. Wir behaupten:

$$\left(\frac{u^*}{v}\right) = 1 \quad \Longrightarrow \quad \left(\frac{v}{u}\right) = 1 \qquad (*)$$

Wenn nämlich $(u^*/v) = 1$, können wir v in A_{u^*} zerlegen, $(v) = \mathfrak{p}\mathfrak{p}'$. Wegen der ungeraden Klassenzahl h und weil die Fundamentaleinheit Norm -1 hat, wenn $u^* > 0$, erhalten wir mit bekanntem Schluß, daß $v^h = N(a)$ für ein $a \in A_{u^*}$. Folglich ist v quadratischer Rest modulo u.

Sei nun $p \equiv 1$ (4). Im Fall $(p/q) = 1$ folgt durch Anwenden von $(*)$ auf $u = p$, $v = q$ sofort, daß auch $(q/p) = 1$. Ist umgekehrt $(q/p) = 1$, so setzen wir $u = q$, $v = p$. Wegen $(-1/p) = 1$ ist auch $(u^*/v) = 1$, und wir erhalten aus $(*)$, daß $(p/q) = 1$.

Damit bleibt nur noch der Fall $p \equiv q \equiv 3$ (4) zu betrachten. Er wird durch den folgenden Satz erfaßt. (Im Fall $(p/q) = -1$ hilft $(*)$ allerdings noch weiter, denn $(-1/q) = -1$. Es folgt $(q/p) = 1$ wie gewünscht.)

Satz 16.3. *Seien $p \equiv q \equiv 3$ (4) Primzahlen, $p \neq q$.*

(a) *Dann sind die Primfaktoren von (p) und (q) in A_{pq} Hauptideale.*

(b) *Für genau eine Wahl des Votzeichens, nämlich $+$ im Fall $(p/q) = 1$ und $-$ im Fall $(p/q) = -1$, existieren $x, y \in \mathbb{Z}$ mit*

$$\pm 4 = px^2 - qy^2.$$

(c) *Es gilt $(q/p) = -1$, falls $(p/q) = 1$, und umgekehrt.*

Beweis. (a) Die Fundamentaleinheit η in $A = A_{pq}$ hat Norm 1 (Satz 16.1(a)). Wir wählen $c \in A_d$ mit $\eta = \eta/1 = c/c'$. Dann gilt $(c) = (c')$, und mit dem gleichen Argument wie im Beweis von Satz 16.2 erreichen wir durch Herausziehen ganz-rationaler Faktoren ein Ideal (e) mit $(e) = (e')$, in dessen Zerlegung nur Primfaktoren der Diskriminante und höchstens mit Vielfachheit 1 vorkommen können. Diese sind die Primideale $\mathfrak{p}$ mit $\mathfrak{p}^2 = (p)$ und $\mathfrak{q}^2 = (q)$.

Es gibt daher höchstens 4 Fälle:

$$(e) = A, \quad (e) = \mathfrak{p}, \quad (e) = \mathfrak{q}, \quad (e) = \mathfrak{p}\mathfrak{q}.$$

Im zweiten Fall ist $\mathfrak{p}$ ein Hauptideal und, da $\mathfrak{p}\mathfrak{q} = (\sqrt{pq})$, dann auch $\mathfrak{q}$. Ebenso folgt die Behauptung im Fall $(e) = \mathfrak{q}$. Daher genügt es, den ersten und den letzten Fall auszuschließen.

Bei $(e) = A$ ist e eine Einheit. Da nun immer noch $\eta = e/e'$ gilt und $e' = e^{-1}$ ist (alle Einheiten haben Norm 1, wenn $N(\eta) = 1$), folgt $\eta = e^2$: Die Fundamentaleinheit ist aber kein Quadrat.

By $(e) = \mathfrak{p}\mathfrak{q}$ ist $e = \varepsilon\sqrt{pq}$ mit einer Einheit ε. Mit dem gleichen Argument wie eben folgt daraus der Widerspruch $\eta = -\varepsilon^2$.

(b) Sei nun etwa $(e) = \mathfrak{p}$, $e = (u + v\sqrt{pq})/2$. Dann folgt

$$\pm 4p = u^2 - v^2 pq.$$

Offensichtlich ist u durch p teilbar, und wir können $x = u/p$, $y = v$ setzen. Wenn $(p/q) = 1$ ist, muß $+$ das Vorzeichen sein, denn $(4/q) = 1$, $(-4/q) = -1$. Im anderen Fall ist $-$ das Vorzeichen. Im Fall $(e) = \mathfrak{q}$ tauschen p und q die Rollen.

(c) folgt nun unmittelbar aus (b). □

Übungen.

16.4. Leite Teil (b) von Satz 16.1 aus dessen Teil (c) und dem Zerlegungsgesetz her. (Man kann ähnlich argumentieren wie im Beweis von Satz 16.3.)

16.5. Bestimme für die Fundamentaleinheiten η von A_{21} und A_{34} jeweils Elemente a mit $\eta = a/a'$.

16.6. Sei $d < 0$ quadratfrei, $d \equiv 1 \ (4)$. Zeige daß $|d|$ prim ist, wenn die Klassenzahl von A_d ungerade ist.

16.7. Sei $A = A_{34}$. Zeige: (a) Das Ideal $\mathfrak{p} = A3 + A(1 + \sqrt{34})$ ist prim. (b) $\mathfrak{p}^2$ ist das von $5 - \sqrt{34}$ erzeugte Hauptideal $\mathfrak{b}$, die Klasse $\mathfrak{C}$ von $\mathfrak{p}$ ist also ambig. (c) $\mathfrak{C}$ enthält kein ambiges Ideal.

Anleitung für (c): Falls $a\mathfrak{p} = a'\mathfrak{p}'$ mit $a \in A$, dann $a\mathfrak{b} = A3a'$. Was folgt daraus für die Fundamentaleinheit?

16.8. (a) Verallgemeinere das Resultat der vorangegangenen Aufgabe wie folgt: Falls $d = u^2 + v^2$ und die Fundamentaleinheit in A_d Norm 1 hat, so besitzt A_d eine anbige Idealklasse, die kein ambiges Ideal enthält.

(b) Es gilt auch die Umkehrung: Falls A_d für $d > 0$ eine ambige Klasse ohne ambiges Ideal besitzt, so hat kein Primteiler p von d Rest 3 modulo 4. (Falls mit den Bezeichnungen des Beweises von Satz 16.2 $N(a) = -N(b)$ gilt, so ist $N(ab) = c \ldots$)

Literaturverzeichnis

[AdGo] Adams, W.W. und Goldstein, L.J.: Introduction to number theory. Prentice Hall 1976

[BoSa] Borewic, S.I. und Safarevic, I.R.: Zahlentheorie. Birkhäuser 1966

[Bund] Bundschuh, P.: Einführung in die Zahlentheorie. Springer 1998

[Chan] Chandrasekharan, K.: Introduction to analytic number theory. Springer 1968

[Cohn] Cohn, H.: Advanced number theory. Dover 1980

[HaWr] Hardy, G.H. und Wright, E.M.: Zahlentheorie. Oldenburg 1958

[Hass] Hasse, H.: Vorlesungen über Zahlentheorie. Springer 1964

[IrRo] Ireland, K. und Rosen, M.: A classical introduction to modern number theory. Springer 1982

[Isch] Ischebeck, F.: Einladung zur Zahlentheorie. BI 1992

[Leut] Leutbecher, A.: Zahlentheorie. Springer 1996

[ReUl] Remmert, R. und Ullrich, P.: Elementare Zahlentheorie. Birkhäuser 1987

Klassische Werke:

[DiDe] Dirichlet, P.G. Lejeune und Dedekind, R.: Vorlesungen über Zahlentheorie. Chelsea 1968 (Nachdruck der 4. Auflage, Braunschweig)

[Gaus] Gauß, C.F.: Untersuchungen über höhere Arithmetik. Chelsea 1965 (Nachdruck der 1. Auflage, Berlin)

Zur Geschichte der Zahlentheorie:

[ScOp] Scharlau, W. und Opolka, H.: Von Fermat bis Minkowski. Springer 1980

[Weil] Weil, A.: Number theory. Ein Gang durch die Geschichte. Birkhäuser 1992

Zur algoritmischen Zahlentheorie:

[Cohe] Cohen, H.: A course in computational algebraic number theory. Springer 1993

[Fors] Forster, O.: Algorithmische Zahlentheorie. Vieweg 1996 (Dieses Buch enthält auch das Zahlentheorie-Computersystem Aribas.)

Zur Kryptographie:

[Baue] Bauer, F. L.: Entzifferte Geheimnisse. Springer 1997

[Buch] Buchmann, J.: Einführung in die Kryptographie. Springer 1999

[Kahn] Kahn, D.: The codebreakers. The story of secret writing. Scribner 1996

Interessante Internet-Seiten zur Suche nach großen Primzahlen:

[Mers] `http://www.mersenne.html, http://www_mersenne_org.html`