Kapitel 1. Teilbarkeit

Die ersten zwei Paragraphen dieses einführenden Kapitels entwickeln die Teilbarkeitstheorie im speziellen Integritätsring der ganzen Zahlen in einem Umfang, der bereits interessante Teile der "elementaren" Zahlentheorie zu begründen gestattet. Diese beiden Anfangsparagraphen beschäftigen sich mit dem multiplikativen Aufbau der ganzen Zahlen aus Primzahlen und gipfeln in zwei Beweisen für den Fundamentalsatz der Arithmetik.

Die in § 3 aufgeworfene und vollständig behandelte Frage nach der Lösbarkeit linearer diophantischer Gleichungen und nach der genauen Struktur der Lösungsgesamtheit solcher Gleichungen kann als natürliche Verallgemeinerung der Frage nach Teilbarkeit zweier ganzer Zahlen verstanden werden. Die in diesem Paragraphen besprochene Problematik wird in Kap. 4 weitergeführt und vertieft.

Die als Anwendung des Fundamentalsatzes in § 1 eingeführte Teileranzahlfunktion ebenso wie die Teilersummenfunktion sind erste Beispiele multiplikativer zahlentheoretischer Funktionen. Die wichtigsten derartigen Funktionen mit ihren wesentlichen Eigenschaften werden in § 4 aus dem Faltungsbegriff gewonnen, wie dies in moderneren Darstellungen üblich geworden ist.

Der eilige Leser kann § 4 ohne weiteres bei der ersten Lektüre übergehen, muß allerdings auf diesen Paragraphen jeweils dann zurückkommen, wenn er an späteren Stellen des Buchs auf spezielle zahlentheoretische Funktionen stößt, deren Eigenschaften er benötigt. So wird er bereits ab Kap. 2 die in 4.11 bereitgestellten Ergebnisse über die EULERsche Phifunktion immer wieder brauchen. Dagegen wird er sich mit anderen Dingen, die zweckmäßigerweise ebenfalls schon in § 4 vorbereitet sind, erst später intensiver vertraut machen müssen.

Die beiden letzten Paragraphen 5 und 6 dieses Anfangskapitels verfolgen hauptsächlich zwei Ziele:

Einerseits beschäftigen sie sich mit der Teilbarkeitstheorie beliebiger Integritätsringe. Dabei steht das Problem im Vordergrund, Bedingungen an solche Ringe zu finden, die garantieren, daß eine multiplikative Zerlegungsaussage analog zum Fundamentalsatz der Arithmetik gilt. Hierher gehören z.B. die Polynomringe über Körpern, die am Ende von § 5 studiert werden.

Andererseits leiten diese über zum zweiten Hauptziel, der Klärung der wichtigsten Eigenschaften algebraischer Zahlen und algebraischer Zahlkörper in der ersten Hälfte von § 6. Dort werden insbesondere sämtliche algebraischen Grundlagen für die Transzendenzuntersuchungen in Kap. 6 gelegt. Gegen Ende von § 6 wird die Problematik von § 5 aufgenommen, indem nun speziell die Integritätsringe der ganzen Zahlen besonders einfacher algebraischer, nämlich quadratischer Zahlkörper daraufhin untersucht werden, ob in ihnen ein Analogon zum Fundamentalsatz gilt.

Die Paragraphen 5 und 6, die der Leser zunächst überschlagen und zu denen er später bei Bedarf zurückkehren kann, enthalten nahezu alle in diesem Buch benötigten Tatsachen aus der Algebra.

§ 1. Fundamentalsatz der Arithmetik

1. Natürliche und ganze Zahlen. Mit $\mathbb{N}$ bzw. $\mathbb{Z}$ werden hier wie üblich die Mengen der natürlichen Zahlen $1, 2, 3, ...$ bzw. der ganzen Zahlen $..., -1, 0, 1, 2, ...$ bezeichnet. In der Zahlentheorie nimmt man ihre axiomatische Einführung als bereits vollzogen hin und interessiert sich für zahlreiche spezielle Eigenschaften, die diese Zahlen haben können. Für eine axiomatische Beschreibung der beiden genannten Mengen muß der Leser auf die einschlägige Lehrbuchliteratur verwiesen werden. Nur der Bequemlichkeit halber sollen hier kurz die wichtigsten Eigenschaften natürlicher bzw. ganzer Zahlen zusammengestellt werden, soweit sie zu den in diesem Buch (meist stillschweigend) benutzten unmittelbaren Konsequenzen aus der axiomatischen Beschreibung zu rechnen sind.

Die *natürlichen Zahlen* bilden eine Menge $\mathbb{N}$, aus der ein mit 1 bezeichnetes Element hervorgehoben ist und auf der eine injektive Selbstabbildung S ("Nachfolgefunktion") mit $1 \notin S(\mathbb{N})$ definiert ist, so daß gilt: *Wenn für eine Teilmenge* $M \subset \mathbb{N}$ *die Bedingungen* $1 \in M$ *und* $S(M) \subset M$ *gelten, dann ist* $M = \mathbb{N}$.

0. Das letztgenannte Axiom ist eine mengentheoretische Fassung des wohlbekannten *Prinzips der vollständigen Induktion*.

1. Sodann wird eine *Addition* $+$ und eine *Multiplikation* $\cdot$ in $\mathbb{N}$ definiert, für die man sämtliche vertrauten Rechenregeln (Assoziativ- und Kommutativgesetze sowie Distributivgesetz) nachweisen kann.

2. Des weiteren werden auf $\mathbb{N}$ Relationen $<$ bzw. $\leq$ in der üblichen Weise eingeführt: Für $m, n \in \mathbb{N}$ schreibt man $m < n$ genau dann, wenn es ein $q \in \mathbb{N}$ mit $m + q = n$ gibt; man schreibt $m \leq n$ genau dann, wenn $m < n$ oder $m = n$ zutrifft. Offenbar ist $\leq$ eine Ordnungsrelation: Die für eine Ordnung charak-

teristischen Eigenschaften (Reflexivität, Antisymmetrie, Transitivität) können nämlich für $\leq$ leicht nachgewiesen werden. Auch ergeben sich die *Monotonie* der obigen Ordnungsrelation $\leq$ bezüglich Addition und Multiplikation ebenso wie deren *Linearität*.

3. Schließlich ist jetzt das folgende, für viele Beweise der Zahlentheorie überaus nützliche *Prinzip des kleinsten Elements* einfach zu zeigen: *Jede nicht leere Teilmenge von* $\mathbb{N}$ *hat ein* (eindeutig bestimmtes) *kleinstes Element*, d.h. aus $M \subset \mathbb{N}$, $M \neq \emptyset$ folgt die Existenz (genau) eines $m \in M$ mit $m \leq n$ für alle $n \in M$.

Bemerkungen. 1) Wie üblich wird $2 := S(1)$, $3 := S(2)$ usw. geschrieben.

2) Andere geläufige Varianten des Induktionsprinzips wie z.B. die mit beliebigem Induktionsanfang oder diejenigen, welche für den Induktionsschritt nicht nur die unmittelbar vorausgehende Aussage, sondern *alle* vorangehenden ausnutzt, werden ebenso angewandt.

Aus der bezüglich der Addition $+$ kommutativen Halbgruppe $\mathbb{N}$ gewinnt man rein algebraisch (durch Bildung von Paaren natürlicher Zahlen) die additive Gruppe $\mathbb{Z}$ der *ganzen Zahlen*. Die ursprünglich nur in $\mathbb{N}$ definierte Multiplikation $\cdot$ kann so auf $\mathbb{Z}$ fortgesetzt werden, daß $\mathbb{Z}$ *bezüglich seiner Addition und Multiplikation einen Integritätsring bildet*, d.h. einen kommutativen, nullteilerfreien Ring mit Einselement.

Schließlich läßt sich die Ordnungsrelation $\leq$ von $\mathbb{N}$ auf $\mathbb{Z}$ so erweitern, daß die oben genannten Linearitäts- und Monotonieeigenschaften erhalten bleiben mit der alleinigen Maßgabe, daß aus $\ell, m, n \in \mathbb{Z}$ und $m \leq n$ die Beziehung $\ell \cdot m \leq \ell \cdot n$ (das Multiplikationszeichen $\cdot$ wird später in der Regel weggelassen) nur noch bei $0 \leq \ell$, d.h. bei $\ell \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$ folgt.

In $\mathbb{N}$ bzw. $\mathbb{Z}$ definiert man ergänzend $n \geq m$ (bzw. $n > m$) durch $m \leq n$ (bzw. $m < n$). Bequem ist auch die Einführung des Absolutbetrages durch $|n| := n$, falls $n \in \mathbb{N}_0$ bzw. $|n| := -n$, falls $-n \in \mathbb{N}_0$.

Bemerkung. 3) Sehr bald werden in diesem Buch rationale und reelle, etwas später komplexe Zahlen auftreten. Auch die grundlegende Einführung der Körper $\mathbb{Q}$, $\mathbb{R}$ bzw. $\mathbb{C}$ der rationalen, reellen bzw. komplexen Zahlen wird hier als anderweitig durchgeführt betrachtet. Genauso werden elementare Funktionen wie Wurzel- oder Logarithmusfunktionen als bekannt vorausgesetzt.

2. Teiler. Sind $m \neq 0$ und n ganze Zahlen, so heißt n durch m *teilbar*, wenn es eine (und dann auch nur eine) ganze Zahl q mit $n = mq$ gibt. Gleichbedeutend damit sind Sprechweisen wie: m ist ein *Teiler* von n, oder: n ist ein *Vielfaches*

von m, oder: *m geht in n auf.* In Zeichen wird dies durch $m|n$ ausgedrückt; $m\not|n$ bedeutet die Negation dieser Aussage, besagt also, daß n nicht durch m teilbar ist.

Es sei ausdrücklich betont, daß *Teiler* hier und im folgenden *stets als von Null verschieden vorausgesetzt* werden: Dies geschieht deshalb, weil die Gleichung $n = 0 \cdot q$ nur für $n = 0$ bestehen kann, dann allerdings für jedes ganze q.

Aus der angegebenen Definition der Teilbarkeit in $\mathbb{Z}$ folgen einige leichte Rechenregeln, die sogleich als Satz zusammengestellt seien; dabei bedeuten lateinische Buchstaben, gleichgültig ob indiziert oder nicht, stets ganze Zahlen.

Satz.

(i) *Für jedes $n \neq 0$ gilt $n|0$ und $n|n$.*

(ii) *Gilt $m|n$, so auch $-m|n$ und $m|-n$.*

(iii) *Für alle n gilt $1|n$.*

(iv) *Aus $m|n$ und $n \neq 0$ folgt $|m| \leq |n|$.*

(v) *Aus $n|1$ folgt entweder $n = 1$ oder $n = -1$.*

(vi) *Aus $m|n$ und $n|m$ folgt entweder $n = m$ oder $n = -m$.*

(vii) *Aus $\ell|m$ und $m|n$ folgt $\ell|n$.*

(viii) *Bei $\ell \neq 0$ sind $m|n$ und $\ell m|\ell n$ gleichbedeutend.*

(ix) *Gelten $m|n_1$ und $m|n_2$, so auch $m|(\ell_1 n_1 + \ell_2 n_2)$ bei beliebigen ℓ_1, ℓ_2.*

(x) *Gelten $m_1|n_1$ und $m_2|n_2$, so auch $m_1 m_2|n_1 n_2$.*

Exemplarisch sei der *Beweis* für die Regel (vii) geführt. Die dort gemachte Voraussetzung besagt, daß es ganze q_1, q_2 gibt, so daß $m = \ell q_1$ und $n = mq_2$ gelten. Daraus folgt $n = \ell(q_1 q_2)$ und dies bedeutet $\ell|n$. □

Bemerkungen. 1) Folgerungen wie in (v) bzw. (vi) werden oftmals kürzer als $n = \pm 1$ bzw. $n = \pm m$ notiert.

2) Die beiden Regeln (i) und (ii) beinhalten offenbar, daß man bei der Untersuchung der Frage, ob $m|n$ oder $m\not|n$ gilt, o.B.d.A. m und n als natürliche Zahlen voraussetzen darf.

3. Primzahlen. Regel (iv) des letzten Satzes besagt, daß jedes $n \in \mathbb{N}$ höchstens n verschiedene natürliche Teiler haben kann. Schreibt man $\tau(n)$ für die Anzahl der verschiedenen natürlichen Teiler von $n \in \mathbb{N}$, so ist stets $\tau(n) \leq n$. Nach Satz 2(iii) ist $\tau(n) \geq 1$, insbesondere $\tau(1) = 1$. Die *Teileranzahlfunktion* τ wird in 7 und später in § 4 genauer untersucht.

Kombiniert man die Regeln (i) und (iii) aus Satz 2, so erhält man $\tau(n) \geq 2$ für jedes ganze $n \geq 2$. Diejenigen n mit $\tau(n) = 2$ bekommen nun einen speziellen Namen: Eine ganze Zahl $n \geq 2$ heißt *Primzahl*, wenn 1 und n ihre einzigen positiven Teiler sind. Ist eine ganze Zahl $n \geq 2$ nicht Primzahl, so heißt sie *zusammengesetzt*.

Nach dieser Definition ist 1 keine Primzahl. Ein Grund, warum man die Definition heute stets so faßt, daß 1 nicht zu den Primzahlen rechnet, wird in Bemerkung 3 von 5 erläutert.

Die Folge der Primzahlen, der Größe nach geordnet, beginnt also mit

$$2, 3, 5, 7, 11, 13, 17, 19, 23, \ldots \quad .$$

Zur Klärung der Frage, ob die bisweilen mit $\mathbb{P}$ bezeichnete Menge aller Primzahlen unendlich ist, beweist man zunächst das

Lemma. *Der kleinste positive, von* 1 *verschiedene Teiler* $p(n)$ *jeder ganzen Zahl* $n \geq 2$ *ist Primzahl.*

Beweis. Nach Satz 2(i) hat n (≥ 2) mindestens einen von 1 verschiedenen positiven Teiler, nämlich n selbst. Die Menge aller derartigen Teiler ist also nicht leer und hat daher ein kleinstes Element (vgl. das in 1 explizit aufgeführte "Prinzip"), welches $p(n)$ genannt werde; es ist offenbar $p(n) \geq 2$. Wäre $p(n)$ nicht Primzahl, so hätte es einen von 1 und $p(n)$ verschiedenen positiven Teiler t. Nach Satz 2(iv), (vii) wäre $t < p(n)$ bzw. $t|n$ und in Verbindung mit $t \geq 2$ würde dies der Definition von $p(n)$ widersprechen. □

Die vor dem Lemma aufgeworfene Frage beantwortet der

4. Satz von EUKLID**.** *Es gibt unendlich viele Primzahlen.*

Beweis. Sind $p_0, \ldots, p_{r-1} \in \mathbb{P}$ paarweise verschieden, so definiert man die von 1 verschiedene Zahl $n \in \mathbb{N}$ durch

$$n := 1 + \prod_{\rho=0}^{r-1} p_\rho \quad . \tag{1}$$

Nach Lemma 3 ist $p(n)$ eine n teilende Primzahl. Wäre $p(n)$ gleich einem der p_ρ, so würde $p(n)|1$ gelten nach Satz 2(ix), was nicht geht. Man hat also $p_r := p(n) \in \mathbb{P} \setminus \{p_0, \ldots, p_{r-1}\}$. □

Der vorstehende Beweis liefert ein effektives Verfahren (man redet von einem *Algorithmus*) zur Gewinnung unendlicher Folgen $(p_r)_{r\in\mathbb{N}_0}$ paarweise verschiedener Primzahlen: Man startet mit beliebigem $p_0 \in \mathbb{P}$, denkt sich die r paarweise verschieden $p_0, \ldots, p_{r-1} \in \mathbb{P}$ bereits erhalten, definiert n_r durch die rechte Seite in (1) und setzt dann $p_r := p(n_r)$. Beginnt man etwa mit $p_0 := 2$, so führen die ersten Schritte dieses Verfahrens zur folgenden kleinen Tabelle:

r	0	1	2	3	4	5	6	7
n_r	–	3	7	43	1807	23479	1244335	6221671
p_r	2	3	7	43	13	53	5	6221671

Um hier zu entscheiden, ob 1807 Primzahl ist, braucht man keineswegs von allen Primzahlen $p < 1807$ festzustellen, ob sie 1807 teilen oder nicht. Es reicht, dies für die $p \le \sqrt{1807}$ zu tun, also für die Primzahlen unterhalb 43. Dies reduziert den Rechenaufwand ganz erheblich und man stützt sich dabei auf folgende

Proposition. *Eine ganze Zahl $n \ge 2$ ist genau dann Primzahl, wenn $p(n) > \sqrt{n}$ gilt.*

Beweis. Für $n \in \mathbb{P}$ ist $p(n) = n > \sqrt{n}$. Ist n zusammengesetzt, so schreibt man $n = mp(n)$; die natürliche Zahl m genügt $1 < m < n$ und man hat $p(m) \ge p(n)$, also $n \ge p(m)p(n) \ge p(n)^2$. □

Bemerkungen. 1) EUKLIDs Satz findet sich (im wesentlichen mit dem hier präsentierten Beweis) wie folgt in Buch IX, § 20 seiner *Elemente* (griechisch $\Sigma\tau o\iota\chi\varepsilon\tilde{\iota}\alpha$) formuliert: "Es gibt mehr Primzahlen als jede vorgelegte Anzahl von Primzahlen."

2) Weitere Beweise des EUKLIDschen Satzes finden sich in 4.5, 4.11 und 2.1.2. In Kap. 7 wird insbesondere die durch $\pi(x) := \#\{p \in \mathbb{P} : p \le x\}$ definierte, stückweise konstante, monoton wachsende Funktion $\pi : \mathbb{R} \to \mathbb{N}_0$ eingehend untersucht. Sie beschreibt quantitativ die Verteilung der Primzahlen in der Menge der natürlichen Zahlen.

5. Der Fundamentalsatz der Arithmetik ist das Hauptergebnis dieses ersten Paragraphen und zeigt deutlich die große Bedeutung der Primzahlen für den multiplikativen Aufbau der natürlichen Zahlen.

Fundamentalsatz der Arithmetik. *Jede von Eins verschiedene natürliche Zahl ist als Produkt endlich vieler Primzahlen darstellbar; diese Darstellung ist eindeutig, wenn man die in ihr vorkommenden Primzahlen der Größe nach ordnet.*

Ohne eine solche Ordnung ist diese Darstellung also nur bis auf die Reihenfolge der eingehenden Primzahlen eindeutig.

Existenzbeweis. Ist n Primzahl, so ist nichts weiter zu tun; insbesondere hat man damit den Induktionsanfang bei $n = 2$ erledigt. Sei nun $n \in \mathbb{N}$, $n > 2$ und es werde die Existenz einer Zerlegung für alle $m \in \{2, \ldots, n-1\}$ vorausgesetzt. O.B.d.A. darf angenommen werden, daß n zusammengesetzt ist. Nach Lemma 3 ist $p(n)$ Primzahl und es werde $n = mp(n)$ geschrieben, woraus sich $m \in \{2, \ldots, n-1\}$ ergibt. Nach Induktionsvoraussetzung hat man

$$m = \prod_{\rho=1}^{r} p_\rho$$

mit gewissen $p_\rho \in \mathbb{P}$, was mit $p_{r+1} := p(n) \in \mathbb{P}$ zu $n = \prod_{\rho=1}^{r+1} p_\rho$ führt. □

Eindeutigkeitsbeweis. Wird angenommen, die Menge der natürlichen Zahlen > 1 mit nicht eindeutiger Produktzerlegung sei nicht leer, so sei n ihr kleinstes Element. Schreibt man wieder $n = mp(n)$, so ist dies eine Produktzerlegung von n, unter deren Faktoren die Primzahl $p(n)$ vorkommt. Nach der über n gemachten Annahme hat diese Zahl eine weitere Produktzerlegung $n = \prod_{\sigma=1}^{s} q_\sigma$ mit allen $q_\sigma \in \mathbb{P}$, die nach Lemma 3 und nach Induktionsvoraussetzung sämtliche größer als $p(n)$ sein müssen. Setzt man*) $n' := n - p(n) \prod_{\sigma=2}^{s} q_\sigma$, so ist dies wegen

$$n' = (q_1 - p(n)) \prod_{\sigma=2}^{s} q_\sigma \tag{1}$$

eine natürliche Zahl, die wegen $p(n)|n$ und Satz 2(ix) durch $p(n)$ teilbar ist. Weiter ist $n' < n$ und also ist n' eindeutig in ein Produkt von Primzahlen zerlegbar, unter denen $p(n)$ vorkommt. Aus (1) sieht man dann $p(n)|(q_1 - p(n))$, also $p(n)|q_1$, was der Tatsache widerspricht, daß $q_1 \in \mathbb{P}$, $q_1 > p(n)$ gilt. □

Offenbar liefert der obige Existenzbeweis für die Produktzerlegung ein effektives Verfahren zur Gewinnung derselben: Will man ein $n \in \mathbb{N}$ mit $n \geq 2$ in seine

*) Wie üblich hat man unter *leeren Produkten* Eins zu verstehen; genauso hat man *leere Summen* stets als Null zu interpretieren.

Primfaktoren (diese Redeweise hat sich für die in der Zerlegung vorkommenden Primzahlen eingebürgert) zerlegen, so setzt man $n_0 := n$ und denkt sich die streng fallende Folge $n_0 > n_1 > \ldots > n_r$ natürlicher Zahlen schon so gewonnen, daß für $\rho = 1, \ldots, r$ gilt

$$n_\rho = \frac{n_{\rho-1}}{p(n_{\rho-1})} \tag{2}$$

Ist $n_r = 1$, so hört man auf; andernfalls setzt man $n_{r+1} := \frac{n_r}{p(n_r)}$. Insgesamt ist klar, daß das beschriebene Verfahren nach endlich vielen, etwa R Schritten sein Ende erreicht, d.h. es wird $n_0 > \ldots > n_{R-1} > n_R = 1$ und es gilt (2) für $\rho = 1, ..., R$. Letzteres zeigt $n_0 = n_R \prod_{\rho=0}^{R-1} p(n_\rho)$, also

$$n = \prod_{\rho=0}^{R-1} p(n_\rho) \tag{3}$$

und dies ist die gesuchte Darstellung von n als Primzahlprodukt. Wegen $p(n_\rho) \geq 2$ folgt aus (3) noch $R \leq \frac{\log n}{\log 2}$, eine Ungleichung, die es zu beurteilen gestattet, wie lange man bei gegebenem n schlimmstenfalls arbeiten muß, bis man die im Fundamentalsatz gesicherte Primfaktorzerlegung von n gefunden hat.

Bemerkungen. 1) Der Fundamentalsatz der Arithmetik steht nicht explizit in Euklids *Elementen*, obwohl einige der Propositionen in Buch VII bzw. IX ihm nahezu äquivalent sind. Auch in A.M. Legendres *Essai sur la Théorie des Nombres* tritt er noch nicht völlig präzisiert hervor. Seine erste klare Formulierung mit Beweis scheint von C.F. Gauss (*Disquisitiones Arithmeticae*, Art. 16) gegeben worden zu sein: "Theorema. Numerus compositus quicunque unico tantum modo in factores primos resolvi potest."*)

2) Der oben geführte Eindeutigkeitsbeweis geht auf E. Zermelo zurück, der ihn, so H. Hasse (J. Reine Angew. Math. *159*, 3-12 (1928)), mündlich an K. Hensel mitteilte. Zermelos Beweis ist hinsichtlich seiner Hilfsmittel sehr einfach, dafür in seiner Schlußweise recht kunstvoll. Ein weiterer Eindeutigkeitsbeweis, bei dem die Verhältnisse eher umgekehrt gelagert sind, wird in 2.8 geführt.

3) Die Aussage des Fundamentalsatzes wird oft ausnahmslos für alle natürlichen Zahlen formuliert; dazu hat man lediglich noch die Zahl 1 als leeres Produkt darzustellen. Würde man 1 zu den Primzahlen rechnen (vgl. 3), so würde die weitestmögliche Eindeutigkeit der im Fundamentalsatz angesprochenen Zerlegungsaussage verloren gehen: Z.B. wären $2 \cdot 3$ und $1 \cdot 2 \cdot 3$ zwei verschiedene Primfaktorzerlegungen der Zahl 6.

*) ("Satz. Jede beliebige zusammengesetzte Zahl kann nur auf eine Weise in Primfaktoren zerlegt werden.")

6. Kanonische Primfaktorzerlegung. Natürlich brauchen die in der Produktzerlegung von n gemäß Fundamentalsatz vorkommenden Primzahlen nicht verschieden zu sein; z.B. ist $72 = 2 \cdot 2 \cdot 2 \cdot 3 \cdot 3$. Für solche Fälle führt man eine kürzere Schreibweise ein: Sind $p_1, \ldots, p_k$ genau die paarweise verschiedenen, n teilenden Primzahlen und kommt p_k genau a_k-mal rechts in 5(3) vor, so schreibt man statt 5(3)

$$n = \prod_{\kappa=1}^{k} p_\kappa^{a_\kappa} \tag{1}$$

und nennt dies die *kanonische (Primfaktor-) Zerlegung* von n. Oft notiert man (1) auch in der Form

$$n = \prod_{p} p^{\nu_p(n)} \tag{2}$$

oder ähnlich, wobei das Produkt nun über *alle* $p \in \mathbb{P}$ erstreckt ist. Die Exponenten $\nu_p(n) \in \mathbb{N}_0$ in (2), oft auch als *Vielfachheit von p in n* bezeichnet, sind Null für alle $p \in \mathbb{P} \setminus \{p_1, \ldots, p_k\}$; ist p aber gleich einem der p_κ aus (1), so ist unter $\nu_p(n)$ das entsprechende a_κ aus (1) zu verstehen.

7. Teileranzahl- und Teilersummenfunktion. Hier soll zunächst folgender Hilfssatz vorausgeschickt werden.

Lemma. *Für $m, n \in \mathbb{N}$ gilt: m teilt n genau dann, wenn $\nu_p(m) \leq \nu_p(n)$ für alle $p \in \mathbb{P}$ zutrifft.*

Beweis. Es ist $m|n$ gleichwertig mit der Existenz eines $\ell \in \mathbb{N}$, für das $n = \ell m$ gilt. Aus dieser Gleichung folgt mit 6(2) und dem Fundamentalsatz $\nu_p(n) = \nu_p(\ell) + \nu_p(m)$ für alle $p \in \mathbb{P}$, wegen $\nu_p(\ell) \in \mathbb{N}_0$ insbesondere $\nu_p(n) \geq \nu_p(m)$ für alle $p \in \mathbb{P}$. Hat man jedoch diese letzte Tatsache, so sind die Differenzen $\delta_p := \nu_p(n) - \nu_p(m)$ nichtnegative ganze Zahlen für alle $p \in \mathbb{P}$, aber höchstens endlich viele δ_p sind positiv. Deswegen ist $\prod_p p^{\delta_p} \in \mathbb{N}$; bezeichnet man dieses Produkt mit ℓ, so gilt damit $\ell m = n$. □

Hat man nun eine natürliche Zahl n mit der kanonischen Primfaktorzerlegung 6(1), so erhält man nach obigem Lemma sämtliche positiven Teiler m von n in der Gestalt

$$m = \prod_{\kappa=1}^{k} p_\kappa^{\alpha_\kappa} \qquad \text{mit } \alpha_\kappa \in \{0, \ldots, a_\kappa\} \text{ für } \kappa = 1, \ldots, k. \tag{1}$$

Insbesondere hat man damit für die in 3 eingeführte Teileranzahl eines gemäß 6(1) (bzw. 6(2)) zerlegten $n \in \mathbb{N}$

$$\tau(n) = \prod_{\kappa=1}^{k}(1 + a_\kappa) \qquad (\text{bzw. } \tau(n) = \prod_{p}(1 + \nu_p(n))),$$

was übrigens auch für $n = 1$ gilt und genausogut geschrieben werden kann als

$$\tau(n) = \prod_{\kappa=1}^{k} \tau(p_\kappa^{a_\kappa}) \qquad (\text{bzw. } \tau(n) = \prod_{p} \tau(p^{\nu_p(n)})). \tag{2}$$

Bezeichnet jetzt $\sigma(n)$ die Summe aller positiven Teiler von $n \in \mathbb{N}$, und ist n wieder gemäß 6(1) zerlegt, so ist nach (1)

$$\sigma(n) = \sum_{(\alpha_1,\ldots,\alpha_k)} p_1^{\alpha_1} \cdot \ldots \cdot p_k^{\alpha_k}, \tag{3}$$

wobei rechts über alle $(\alpha_1, \ldots, \alpha_k) \in \{0, \ldots, a_1\} \times \ldots \times \{0, \ldots, a_k\}$ zu summieren ist. Aus (3) ist

$$\sigma(n) = \sum_{\alpha_1=0}^{a_1} \ldots \sum_{\alpha_k=0}^{a_k} p_1^{\alpha_1} \cdot \ldots \cdot p_k^{\alpha_k} = \prod_{\kappa=1}^{k} \sum_{\alpha_\kappa=0}^{a_\kappa} p_\kappa^{\alpha_\kappa}$$

klar und dies kann wie vorher bei τ in die äquivalente Form

$$\sigma(n) = \prod_{\kappa=1}^{k} \sigma(p_\kappa^{a_\kappa}) \qquad (\text{bzw. } \sigma(n) = \prod_{p} \sigma(p^{\nu_p(n)})) \tag{4}$$

gesetzt werden, welche ersichtlich auch für $n = 1$ zutrifft.

Bemerkung. Die soeben eingeführte *Teilersummenfunktion* σ und die *Teileranzahlfunktion* τ sind erste Beispiele sogenannter zahlentheoretischer Funktionen, welche in § 4 intensiver studiert werden sollen.

8. Vollkommene Zahlen. In der Zahlenmystik des PYTHAGORAS (um 550 v.Chr.) spielten natürliche Zahlen n, deren von n verschiedene natürliche Teiler sich zu n addieren, eine ausgezeichnete Rolle. PYTHAGORAS und seine Schule nannten derartige Zahlen vollkommen. Der christliche Theologe und Philosoph AUGUSTINUS (354–430) begründete die Erschaffung der Welt in sechs Tagen damit, daß Gott die Vollkommenheit seines Werkes auch durch die Vollkommenheit der Zahl 6 zum Ausdruck bringen wollte.

Ausgedrückt mit der in 7 eingeführten Funktion σ heißt $n \in \mathbb{N}$ also genau dann *vollkommen* (oder *perfekt*), wenn $\sigma(n) = 2n$ gilt. Über gerade vollkommene Zahlen gibt abschließende Auskunft der folgende

Satz. *Bei $n \in \mathbb{N}$ und $2|n$ sind äquivalent:*
(i) $n = 2^{k-1}(2^k - 1)$ *mit ganzem* $k \geq 2$ *und* $2^k - 1 \in \mathbb{P}$.
(ii) n *ist vollkommen.*

Beweis. Sei n wie in (i) gegeben und $M_k := 2^k - 1$ gesetzt. Wegen 7(4) und $M_k \in \mathbb{P} \setminus \{2\}$ ist

$$\sigma(n) = \sigma(2^{k-1})\sigma(M_k) = \big(\sum_{\kappa=0}^{k-1} 2^\kappa\big)(1 + M_k) = (2^k - 1)2^k = 2n,$$

also ist n vollkommen.

Um die Umkehrung einzusehen, macht man zweckmäßig den Ansatz

$$n = 2^{k-1}m \tag{1}$$

mit ungeradem $m \in \mathbb{N}$ und ganzem $k \geq 2$; wegen $2|n$ ist hier $k = 1$ unmöglich. Mittels 7(4) folgt aus (1) und der Voraussetzung (ii)

$$2^k m = 2n = \sigma(n) = \sigma(2^{k-1})\sigma(m) = (2^k - 1)\sigma(m).$$

Nun denkt man sich in der letztendlich interessierenden Gleichung

$$2^k m = (2^k - 1)\sigma(m) \tag{2}$$

für die Zahl rechts die kanonische Primfaktorzerlegung hingeschrieben: Da diese eindeutig ist und da $2 \not| (2^k - 1)$ gilt (d.h. $2^k - 1$ ungerade ist), muß die volle, in der linken Seite von (2) stehende Zweierpotenz 2^k Teiler von $\sigma(m)$ sein. Mit geeignetem $\ell \in \mathbb{N}$ hat man somit

$$\sigma(m) = 2^k \ell\,, \qquad m = (2^k - 1)\ell. \tag{3}$$

Wäre $\ell > 1$, so hätte m mindestens $1, \ell, (2^k - 1)\ell$ als verschiedene positive Teiler, was zu

$$\sigma(m) \geq 1 + \ell + (2^k - 1)\ell > 2^k \ell$$

führen würde entgegen der ersten Gleichung in (3). Dort ist also $\ell = 1$ und daher $\sigma(m) = m + 1$, weshalb m Primzahl sein muß; die zweite Gleichung von (3) zeigt schließlich $2^k - 1 \in \mathbb{P}$. □

Die einfachere Implikation (i) ⇒ (ii) des eben bewiesenen Satzes geht auf Euklid (*Elemente IX*, § 36) zurück, während (ii) ⇒ (i) erst 1747, rund zweitausend Jahre später, von L. Euler (Opera Omnia Ser. 1, V, 353–365) hinzugefügt werden konnte.

Die kleinsten geraden vollkommenen Zahlen sind die seit dem Altertum bekannten 6, 28, 496, 8128; in der Terminologie des Euklid–Eulerschen Satzes rühren diese her von $k = 2, 3, 5, 7$, die ihrerseits Primzahlen sind. Nach obigem Satz ist die Frage nach geraden vollkommenen Zahlen äquivalent mit derjenigen, für welche ganzen $k \geq 2$ die Zahl $M_k = 2^k - 1$ Primzahl ist. Eine hierfür notwendige Bedingung entnimmt man folgender

Proposition. *Für $k \in \mathbb{N}$ gilt: Ist $2^k - 1$ Primzahl, so auch k.*

Beweis. Man geht aus von folgender, bei $m \in \mathbb{N}$ gültigen Gleichung in zwei Unbestimmten

$$X^m - Y^m = (X - Y) \sum_{j=0}^{m-1} X^j Y^{m-1-j}. \tag{4}$$

Ist jetzt $k \in \mathbb{N}$ zusammengesetzt, etwa $k = \ell m$ mit $\ell, m \in \mathbb{N} \setminus \{1\}$, so ersetzt man X bzw. Y in (4) durch 2^ℓ bzw. 1 und erhält

$$M_k = M_\ell \sum_{j=0}^{m-1} 2^{j\ell},$$

also $M_\ell | M_k$, aber $M_\ell \neq 1, M_k$ wegen $\ell \neq 1, k$. Damit ist M_k als zusammengesetzt erkannt. □

Andererseits gibt es sehr viele Primzahlen k, für die $2^k - 1$ zusammengesetzt ist; derzeit sind genau 44 Primzahlen der Form $2^k - 1$ bekannt (vgl. hierzu 3.2.12). Nach dem EUKLID–EULERschen Satz kennt man heute also genau 44 gerade vollkommene Zahlen.

Man wird sich nun fragen, was man über ungerade vollkommene Zahlen weiß. In der Tat ist zur Zeit keine einzige bekannt und man *vermutet*, daß solche Zahlen nicht existieren. Das beste, was man in dieser Richtung bisher hat beweisen können, ist folgendes Resultat von P. HAGIS JR. (1980, angekündigt 1975) bzw. E.Z. CHEIN (1979): Jede ungerade vollkommene Zahl hat in ihrer kanonischen Primfaktorzerlegung mindestens *acht* verschiedene Primzahlen.

Bemerkungen. 1) Aus (4) möge der Leser $(2^{2^n} + 1) | (2^{2^n m} + 1)$ für $m, n \in \mathbb{N}_0$, $2 \nmid m$ folgern und daraus: Ist $2^k + 1$ Primzahl für ein $k \in \mathbb{N}$, so ist k eine Potenz von 2. Dies macht klar, wieso man Primzahlen der Form $2^k + 1$ sogleich in der speziellen Form $2^{2^n} + 1$ sucht, vgl. 2.1.2 und 3.2.11.

2) Der in diesem Abschnitt vermittelte Einblick in die Problematik der vollkommenen Zahlen zeigt, wie rasch man in der Zahlentheorie zu offenen Fragestellungen vorstoßen kann, um deren Lösung sich Mathematiker seit vielen Generationen bemühen. Dieser direkte, oft durch keinerlei Begriffsapparat erschwerte Zugang zu noch ungelösten Problemen macht einen der Reize aus, den die Zahlentheorie immer wieder auf mathematische Laien wie auf erfahrene Mathematiker auszuüben vermag.

9. Irrationalität. Im letzten Paragraphen von Buch X seiner *Elemente* gab EUKLID einen Beweis für die Irrationalität von $\sqrt{2}$, den man üblicherweise im

schulischen Mathematikunterricht kennenlernt und der von der Aussage des Fundamentalsatzes abhängt. Hier wird dieses Irrationalitätsresultat weitgehend verallgemeinert zu folgendem, auf GAUSS zurückgehenden

Satz. *Jede rationale Nullstelle eines Polynoms* $X^n+c_{n-1}X^{n-1}+\ldots+c_0 \in \mathbb{Z}[X]$ *ist ganz.*

Beweis. Es wird angenommen, das mit $f(X)$ bezeichnete Polynom im Satz habe eine Nullstelle $x \in \mathbb{Q} \setminus \mathbb{Z}$. Dieses x hat eine Darstellung

$$x = \frac{a}{b} \tag{1}$$

mit geeigneten $a, b \in \mathbb{Z}$, die wegen $x \notin \mathbb{Z}$ den Bedingungen $a \neq 0$, $b > 1$ genügen. Unter allen derartigen Darstellungen von x sei (1) diejenige mit kleinstem b. Die Voraussetzung $f(\frac{a}{b}) = 0$ ist äquivalent mit

$$a^n = -b \sum_{j=0}^{n-1} c_j a^j b^{n-1-j}, \tag{2}$$

wobei die Summe rechts eine ganze Zahl ist. Aufgrund des Fundamentalsatzes kann nun gesagt werden: Wegen $b > 1$ existiert eine in b aufgehende Primzahl p, die nach (2) in a^n aufgehen muß, wegen $n \geq 1$ also auch in a. Damit sind dann $a' := \frac{a}{p}$ und $b' := \frac{b}{p}$ ganz und genügen $x = \frac{a'}{b'}$ ebenso wie $a' \neq 0$, $b' > 1$; wegen $b' < b$ widerspricht dies aber der Minimalbedingung bei der Wahl von a, b in (1). □

Korollar. *Für* $m, n \in \mathbb{N}$ *ist die positive reelle Zahl* $\sqrt[n]{m}$ *entweder ganz oder irrational. Insbesondere ist* $\sqrt{m}$ *irrational, wenn* m *keine Quadratzahl ist. Noch spezieller ist* $\sqrt{p_1 \cdot \ldots \cdot p_k}$ *irrational, wenn die* $k \geq 1$ *Primzahlen* $p_1, \ldots, p_k$ *paarweise verschieden sind.*

Beweis. Für die erste Aussage wendet man den Satz an auf das Polynom $X^n - m$. Für die zweite beachtet man, daß aus der angenommenen Ganzheit von $\sqrt{m}$, also $\sqrt{m} = \ell$ mit einem $\ell \in \mathbb{N}$, die Gleichheit $m = \ell^2$ folgt, welche m als Quadratzahl ausweist. □

Bemerkung. Wie man leicht sieht, sind bei $x \in \mathbb{R}$ die beiden Aussagen "x ist irrational" und "die Zahlen 1, x sind über $\mathbb{Q}$ linear unabhängig" gleichbedeutend. Ein Resultat betreffend die lineare Unabhängigkeit mehrerer reeller Zahlen über $\mathbb{Q}$, in dessen Beweis ebenfalls wesentlich der Fundamentalsatz eingeht, ist folgendes: *Die reellen Zahlen* $\log p$, *wo* p *sämtliche Primzahlen durchläuft, sind*

über $\mathbb{Q}$ *linear unabhängig.* Daher ist nach EUKLIDs Satz 4 die Dimension von $\mathbb{R}$, aufgefaßt als Vektorraum über $\mathbb{Q}$, nicht endlich.

10. Anmerkung zum Eindeutigkeitsbeweis. Beim Nachweis des Fundamentalsatzes in 5 fällt auf, daß der Existenzbeweis für die Produktzerlegung deutlich leichter fällt als der Eindeutigkeitsbeweis und lediglich auf die multiplikative Struktur der natürlichen Zahlen sowie auf den Begriff der Primzahl zurückgreift. Daß *jeder Eindeutigkeitsbeweis im Fundamentalsatz* darüber hinaus *auch die additive Struktur der natürlichen Zahlen* irgendwie *ausnützen muß* – in 5 geschah dies durch 5(1) und davor –, wird durch folgendes, im Prinzip auf D. HILBERT zurückgehende Beispiel klar.

Man betrachtet die Teilmenge $H := \{3j + 1 : j \in \mathbb{N}_0\}$ von $\mathbb{N}$ und nennt deren Elemente vorübergehend *H-Zahlen.* Offenbar ist das Produkt zweier H-Zahlen wieder eine H-Zahl und so ist H eine Unterhalbgruppe der multiplikativen Halbgruppe $\mathbb{N}$.

Weiter bezeichnet man eine H-Zahl $n \neq 1$ als *H-Primzahl*, wenn 1 und n die einzigen *in* H gelegenen natürlichen Teiler von n sind. Die Folge der H-Primzahlen beginnt demnach mit

$$4,\ 7,\ 10,\ 13,\ 19,\ 22,\ 25,\ \ldots\ . \tag{1}$$

Genauso wie in 5 zeigt man induktiv leicht, daß jede von 1 verschiedene H-Zahl mindestens eine multiplikative Zerlegung in H-Primzahlen besitzt. Die einzige kleine Schwierigkeit dabei könnte in dieser Überlegung liegen: Gilt $m, n \in H$ und $m|n$, so ist $\frac{n}{m} \in H$. Die Situation hinsichtlich der bloßen *Existenz* einer multiplikativen Zerlegung ist hier also völlig analog zu dem in 5 behandelten klassischen Fall.

Wenn dort die *Eindeutigkeit* alleine aus der multiplikativen Struktur der natürlichen Zahlen und dem Begriff der Primzahl beweisbar wäre, müßte sich dieser Beweis auf die H-Zahlen übertragen lassen. Nun gibt es aber H-Zahlen, die *verschiedene* Zerlegungen in H-Primzahlen besitzen. Ein Beispiel dafür bietet die H-Zahl 100, die sowohl als $4 \cdot 25$ wie als $10 \cdot 10$ geschrieben werden kann, wobei 4, 10, 25 tatsächlich H-Primzahlen sind, vgl. (1). Hinsichtlich der Addition zeigen natürliche bzw. H-Zahlen ja auch ein ganz unterschiedliches Verhalten: Für $m, n \in \mathbb{N}$ ist $m + n \in \mathbb{N}$ und $m - n \in \mathbb{N}$, letzteres falls $m > n$; aber weder Summe noch Differenz zweier H-Zahlen ist wieder H-Zahl.

Bemerkung. R.D. JAMES und I. NIVEN (Proc. Amer. Math. Soc. *5*, 834–838 (1954)) haben sämtliche multiplikativen Unterhalbgruppen des Typs $H_{k,\ell} := \{kj + \ell : j \in \mathbb{N}_0\}$ von $\mathbb{N}$ bestimmt, in denen die multiplikative Zerlegung in "$H_{k,\ell}$-Primzahlen" eindeutig ist.

§ 2. Größter gemeinsamer Teiler, kleinstes gemeinsames Vielfaches

1. Größter gemeinsamer Teiler (ggT). Hier seien $n_1, \dots, n_k$ stets ganze Zahlen. Gefragt wird nach allen ganzen $d \neq 0$ mit $d|n_1, \dots, d|n_k$, mit anderen Worten, nach den gemeinsamen Teilern aller $n_1, \dots, n_k$. Ist d ein derartiger gemeinsamer Teiler, so hat $-d$ nach Satz 1.2(ii) dieselbe Eigenschaft, weshalb in Zukunft die Beschränkung auf positive gemeinsame Teiler ausreicht, zu denen übrigens die Eins nach Satz 1.2(iii) immer gehört. Weiter kann künftig vorausgesetzt werden, daß nicht alle $n_1, \dots, n_k$ Null sind; andernfalls liegt nach Satz 1.2(i) die triviale Situation vor, wo jede von Null verschiedene ganze Zahl gemeinsamer Teiler der $n_1, \dots, n_k$ ist. Nach Satz 1.2(iv) ist dann klar, daß jeder positive gemeinsame Teiler d von $n_1, \dots, n_k$ der folgenden Ungleichung genügt

$$d \leq \mathrm{Min}\{|n_i| : i = 1, \dots, k \text{ mit } n_i \neq 0\}. \tag{1}$$

Aufgrund dieser Vorbemerkungen ist ersichtlich, daß bei ganzen, nicht sämtlich verschwindenden $n_1, \dots, n_k$ die Menge aller positiven gemeinsamen Teiler nicht leer und (wegen (1)) endlich ist. Daher besitzt sie ein größtes, im weiteren als $(n_1, \dots, n_k)$ notiertes Element, das herkunftsgemäß als *größter gemeinsamer Teiler* (kurz: *ggT*; engl.: *gcd* für greatest common divisor) der $n_1, \dots, n_k$ bezeichnet wird. Man benützt traditionell für den ggT die Schreibweise mit runden Klammern, wenn Verwechslungen mit anderen Verwendungen dieser Klammern nicht zu befürchten sind, etwa mit der üblichen Notation von k-Tupeln wie z.B. in der Summationsbedingung von 1.7(3). Sollten sich beide Verwendungen im Text einmal zu nahe kommen, so wird der ggT deutlicher als $\mathrm{ggT}(n_1, \dots, n_k)$ bezeichnet.

Grundsätzlich kann man den ggT von $n_1, \dots, n_k$, nicht alle Null, effektiv berechnen, indem man für jeden (der endlich vielen) positiven Teiler der Zahl rechts in (1) prüft, ob er sämtliche n_i teilt. Von den gemeinsamen Teilern ist dann das Maximum zu nehmen. Doch ist dies Vorgehen weder elegant noch numerisch schnell, wenn die $|n_i|$ groß sind. Ein viel besseres Verfahren zur Berechnung des ggT wird später in 9 vorgestellt.

Bemerkungen. 1) Bei $n_1 \in \mathbb{Z} \setminus \{0\}$ ist $(n_1) = |n_1|$; dies ergibt Kombination von (i), (ii) und (iv) aus Satz 1.2. Daher ist der Fall $k = 1$ des ggT im weiteren zwar nicht auszuschließen, aber auch nicht sonderlich interessant.

2) Sind $n_1, n_2, \dots$ unendlich viele ganze, nicht sämtlich verschwindende Zahlen, so kann deren ggT wörtlich wie oben definiert werden.

2. Divisionsalgorithmus. In 3 sollen zwei Charakterisierungen des ggT gezeigt werden. Dabei erweist sich das Verfahren der *Division mit Rest* als überaus

hilfreich, das auch bezeichnet wird als

Divisionsalgorithmus. *Zu jedem Paar (n, m) ganzer Zahlen mit $m > 0$ existiert genau ein Paar (a, b) ganzer Zahlen, so daß gilt*

$$n = am + b \quad \text{und} \quad 0 \leq b < m. \tag{1}$$

Beweis. Es werde die Menge aller nichtnegativen ganzen Zahlen der Form $n - xm$ mit ganzem x betrachtet. Wegen $n + |n|m \in \mathbb{N}_0$ ist diese Menge nicht leer und enthält somit nach dem in 1.1 formulierten Prinzip ein kleinstes Element $b := n - am$ mit geeignetem ganzem a. Da $n-(a+1)m < 0$ ist, hat man insgesamt $0 \leq b < m$ wie behauptet. Ist (a', b') ein Paar mit denselben Eigenschaften wie (a, b), so gilt wegen (1) die Gleichung $(a - a')m + (b - b') = 0$. Wäre hier $b' \neq b$, so wäre $m \leq |b - b'|$ entgegen $0 \leq b, b' < m$. Also ist $b' = b$ und damit $a' = a$ wegen $m \neq 0$. □

Das a bzw. b aus (1) nennt man bisweilen *Quotient* bzw. *Rest* bei der Division von n durch m. Das oben bewiesene Resultat bleibt wörtlich unter der Voraussetzung $m \neq 0$ (statt $m > 0$) erhalten, wenn man $0 \leq b < |m|$ in (1) schreibt. Hat ein Paar (m, n) ganzer Zahlen mit $m \neq 0$ die spezielle Eigenschaft, daß der Divisionsrest b in (1) verschwindet, so wurde diese Situation zu Anfang von 1.2 mit der Redeweise beschrieben, n sei durch m teilbar. Der Divisionsalgorithmus ist für die Zahlentheorie ein sehr wichtiges Hilfsmittel, das auch später in diesem Buch immer wieder zum Zuge kommt (vgl. etwa in 2.1.4).

3. Zwei Charakterisierungen des ggT. Die erste Charakterisierung ist besonders von theoretischem Interesse und wird sich z.B. in 4 und 6 als wichtig erweisen, aber auch schon beim Beweis der zweiten Charakterisierung. Mit letzterer lassen sich viele Aussagen über den ggT sehr bequem zeigen (vgl. etwa 5); ebenfalls ist sie von großer Bedeutung für die Ausdehnung des ggT–Begriffs auf allgemeinere Ringe.

Satz A. *Für ganze $n_1, \ldots, n_k$, nicht alle Null, gilt*

$$(n_1, \ldots, n_k) = \mathrm{Min}\{\sum_{i=1}^{k} n_i x_i : x_1, \ldots, x_k \in \mathbb{Z}, \sum_{i=1}^{k} n_i x_i > 0\}.$$

Beweis. Wegen $\sum_{i=1}^{k} n_i^2 > 0$ ist hier die Menge $L := \{\ldots\}$ nicht leer und hat so nach dem in 1.1 formulierten Prinzip ein kleinstes Element, welches d' genannt

werden möge. Daher gibt es ganze $x_1, \ldots, x_k$, die

$$\sum_{i=1}^{k} n_i x_i = d' \tag{1}$$

genügen. Ist $d := (n_1, \ldots, n_k)$, so gilt $d|d'$ wegen (1) und der induktiv auf k Summanden ausgedehnten Regel (ix) aus Satz 1.2. Nach (iv) desselben Satzes hat man somit $d \leq d'$ und zum Beweis von Satz A braucht man jetzt noch $d' \leq d$.

Bei festem $j \in \{1, \ldots, k\}$ wendet man dazu den Divisionsalgorithmus 2 an auf das Paar (n_j, d'). Danach existieren ganze a_j, b_j mit

$$n_j = a_j d' + b_j \quad \text{und} \quad 0 \leq b_j < d', \tag{2}$$

vgl. 2(1), was wegen (1) unmittelbar zu

$$b_j = n_j - a_j d' = n_j(1 - a_j x_j) + \sum_{\substack{i=1 \\ i \neq j}}^{k} n_i(-a_j x_i) =: \sum_{i=1}^{k} n_i y_{ij} \tag{3}$$

mit ganzen y_{ij} führt. Wäre $b_j > 0$ für ein j, so wäre wegen (3) für dieses j die Summe $\sum_{i=1}^{k} n_i y_{ij}$ in L gelegen, also mindestens gleich d' entgegen $b_j < d'$, vgl. (2). Der erzielte Widerspruch zeigt $b_j = 0$ für $j = 1, \ldots, k$, also $d'|n_1, \ldots, d'|n_k$ wegen (2). Nach Definition von d ist schließlich $d' \leq d$ wie gewünscht. □

Die Beweismethode dieses Satzes liefert lediglich die *Existenz* ganzer $x_1, \ldots, x_k$, die (vgl. (1)) der Bedingung $(n_1, \ldots, n_k) = n_1 x_1 + \ldots + n_k x_k$ genügen, jedoch kein *effektives Verfahren* zur Bestimmung solcher $x_1, \ldots, x_k$, vgl. Ende von 1. Erst in 9 wird ein Verfahren präsentiert, das effektiv (und schnell) ist.

Satz B. *Für ganze $n_1, \ldots, n_k$, nicht alle Null, und ganzes $d > 0$ sind folgende Aussagen gleichwertig:*

(i) *d ist der ggT von $n_1, \ldots, n_k$.*

(ii) *$d|n_1, \ldots, d|n_k$ und aus $d' \in \mathbb{N}$, $d'|n_1, \ldots, d'|n_k$ folgt $d'|d$.*

Beweis. Sei zunächst (i) erfüllt, also $d = (n_1, \ldots, n_k)$, was jedenfalls $d|n_1, \ldots, d|n_k$ impliziert. Ist nun weiter $d' \in \mathbb{N}$ ein gemeinsamer Teiler von $n_1, \ldots, n_k$, so gilt nach Satz A, vgl. (1), auch die Teilbarkeitsbedingung $d'|d$ aus (ii). Hat umgekehrt d die Eigenschaften aus (ii), so ist nach Satz 1.2(iv) die Ungleichung $d' \leq d$ erfüllt, weshalb d der *größte* gemeinsame Teiler der $n_1, \ldots, n_k$ ist. □

Bemerkung. Der ggT ganzer $n_1, \ldots, n_k$, nicht alle Null, läßt sich nach Satz B beschreiben als *derjenige positive Teiler aller $n_1, \ldots, n_k$, der von jedem anderen derartigen Teiler geteilt wird.* Offenbar hängt diese Charakterisierung des ggT nur vom Teilbarkeitsbegriff, nicht aber von der Ordnung in $\mathbb{Z}$ (vgl. 1.1) ab. Dieser Umstand ist von Bedeutung, wenn man den Begriff des ggT in allgemeineren Ringen einführen will, in denen zwar keine Ordnungsrelation, wohl aber ein Teilbarkeitsbegriff erklärt werden kann (vgl. 5.2). In 1 wurde bewußt der ggT für $\mathbb{Z}$ in der Weise definiert, wie dieser Begriff historisch gewachsen ist und wie seine verbale Umschreibung den mathematischen Inhalt am unmittelbarsten erkennen läßt.

4. Idealtheoretische Deutung des ggT. Sind $m_1, \ldots, m_\ell$ irgendwelche ganze Zahlen, so schreibt man $m_1\mathbb{Z} + \ldots + m_\ell\mathbb{Z}$ für die Menge aller ganzen Zahlen der Form $\sum_{j=1}^{\ell} m_j x_j$, wobei der Vektor $(x_1, \ldots, x_\ell)$ ganz $\mathbb{Z}^\ell$ durchläuft. In der Sprechweise der Algebra ist $m_1\mathbb{Z} + \ldots + m_\ell\mathbb{Z}$ ein *Ideal*, genauer das von $m_1, \ldots, m_\ell$ erzeugte Ideal in $\mathbb{Z}$. In dieser Terminologie läßt sich der ggT wie folgt interpretieren (vgl. hierzu auch die Bemerkung 1 zu 5.6):

Satz. *Das von ganzen $n_1, \ldots, n_k$, nicht alle Null, erzeugte Ideal $n_1\mathbb{Z}+\ldots+n_k\mathbb{Z}$ in $\mathbb{Z}$ stimmt überein mit dem Ideal $(n_1, \ldots, n_k)\mathbb{Z}$ in $\mathbb{Z}$.*

Beweis. Man setzt $J := n_1\mathbb{Z} + \ldots + n_k\mathbb{Z}$ und $d := (n_1, \ldots, n_k)$. Ist $n \in J$, so gilt $n = \sum_{i=1}^{k} n_i x_i$ mit gewissen ganzen x_i und daher hat man $d|n$ nach Satz 1.2(ix), also $n \in d\mathbb{Z}$ und es bleibt noch die Inklusion $d\mathbb{Z} \subset J$ zu zeigen. Nach Satz 3A ist $d = \sum_{i=1}^{k} n_i x_i'$ bei geeigneter Wahl ganzer x_i'; für jedes ganze x ist also $dx \in J$, was die gewünschte Inklusion beinhaltet. □

5. Rechenregeln. Die folgenden Regeln (i) bis (v) ergeben sich unmittelbar aus der Definition des ggT, gegebenenfalls mit dessen Charakterisierung in Satz 3B; dies möge dem Leser überlassen bleiben.

Proposition. *Seien $n_1, \ldots, n_k$ ganze, nicht sämtlich verschwindende Zahlen und d ihr ggT.*

(i) *Für jede Permutation π der Indizes $1, \ldots, k$ ist $(n_{\pi(1)}, \ldots, n_{\pi(k)}) = d$.*

(ii) *Bei $k \geq 2$ und $n_k = 0$ ist $(n_1, \ldots, n_{k-1}) = d$.*

(iii) *Bei $k \geq 2$ und $n_{k-1} = n_k$ ist $(n_1, \ldots, n_{k-1}) = d$.*

(iv) *Es gilt $(n_1, \ldots, n_{k-1}, -n_k) = d$.*

(v) *Bei beliebigen ganzen* $x_1, \ldots, x_{k-1}$ *ist* $(n_1, \ldots, n_{k-1}, n_k + \sum_{i=1}^{k-1} n_i x_i) = d$.

(vi) *Für jedes ganze* $\ell \neq 0$ *ist* $(\ell n_1, \ldots, \ell n_k) = |\ell| d$.

(vii) *Es gilt* $(\frac{n_1}{d}, \ldots, \frac{n_k}{d}) = 1$.

(viii) *Ist* $k \geq 2$ *und sind* $n_1, \ldots, n_{k-1}$ *nicht alle Null, so gilt* $((n_1, \ldots, n_{k-1}), n_k) = d$.

Beweis. Zu (vi): Aus $d|n_1, \ldots, d|n_k$ folgt $\ell d|\ell n_1, \ldots, \ell d|\ell n_k$, also $\ell d|e$, wenn man $e := (\ell n_1, \ldots, \ell n_k)$ schreibt und Satz 3B ausnützt. Somit ist insbesondere $\frac{e}{\ell}$ ganz. Die Definition von e impliziert $e|\ell n_1, \ldots, e|\ell n_k$, was mit $\frac{e}{\ell}|n_1, \ldots, \frac{e}{\ell}|n_k$ äquivalent ist. Erneut Satz 3B ergibt dann $\frac{e}{\ell}|d$, also das für (vi) noch benötigte $e|\ell d$. Übrigens kann diese Regel genauso bequem mittels Satz 3A verifiziert werden.

Zu (vii): Die Zahlen $\frac{n_i}{d}$ sind für $i = 1, \ldots, k$ nach Voraussetzung ganz und nicht alle Null; f sei ihr ggT. Nach (vi) ist dann $df = (d\frac{n_1}{d}, \ldots, d\frac{n_k}{d})$; da dies gleich d ist, folgt $f = 1$.

Zu (viii): Sei $g := (n_1, \ldots, n_{k-1})$ und $h := (g, n_k)$. Wegen $h|g$, $h|n_k$ ist $h|n_i$ für $i = 1, \ldots, k$ und also $h|d$ nach Satz 3B. Nach demselben Ergebnis folgt aus $d|g$, $d|n_k$ auch $d|h$, insgesamt $h = d$ wie in (viii) behauptet. □

Die Regeln (i) bis (v) besagen, daß sich der ggT von endlich vielen ganzen Zahlen nicht ändert, wenn man ihre Reihenfolge beliebig modifiziert, wenn man Nullen bzw. gleiche Zahlen wegläßt oder hinzunimmt, wenn man sie durch ihre Absolutbeträge ersetzt oder wenn man zu einer Zahl eine beliebige Linearkombination der übrigen Zahlen mit ganzen Koeffizienten addiert.

Regel (viii) reduziert das Problem der Bestimmung des ggT von k (≥ 2) ganzen Zahlen auf dasjenige, den ggT von *zwei* Zahlen zu ermitteln, die nicht beide Null sind.

6. Teilerfremdheit. Ganze Zahlen $n_1, \ldots, n_k$, nicht alle Null, heißen *teilerfremd*, wenn ihr ggT gleich Eins ist, d.h. wenn $(n_1, \ldots, n_k) = 1$ gilt; eine solche Situation lag z.B. in Proposition 5(vii) vor. Man muß diesen Begriff jedoch sorgsam unterscheiden von folgendem: Ist $k \geq 2$, so heißen die ganzen Zahlen $n_1, \ldots, n_k$, von denen höchstens eine verschwindet, *paarweise teilerfremd*, wenn $(n_i, n_j) = 1$ für alle $i, j \in \{1, ..., k\}$ mit $i \neq j$ gilt. Offenbar besagt "teilerfremd" und "paarweise teilerfremd" für $k = 2$ dasselbe; für $k \geq 3$ ist die zweite Eigenschaft stärker als die erste. So sind z.B. die drei Zahlen 6, 10, 15 zwar teilerfremd, jedoch nicht paarweise teilerfremd.

Satz. *Seien m, m_1, m_2, n, n_1, n_2 ganz und $mm_1m_2 \neq 0$.*

(i) *Ist m Teiler von n_1n_2 und sind m, n_1 teilerfremd, so ist m Teiler von n_2.*

(ii) *Sind m_1, m_2 teilerfremd und gehen beide in n auf, so geht auch ihr Produkt m_1m_2 in n auf.*

(iii) *Sind m, n_1, n_2 teilerfremd, so gilt $(n_1, m)(n_2, m) = (n_1n_2, m)$.*

Beweis. Zu (i): Wegen $(m, n_1) = 1$ existieren ganze x, y mit $mx + n_1y = 1$, was zu $mn_2x + n_1n_2y = n_2$ führt. Wegen $m|n_1n_2$ geht hier m in der linken Seite auf, also auch in n_2.

Zu (ii): Nach Voraussetzung ist $m_2|n$, was wegen $m_1|n$ genauso gut als $m_2|m_1 \cdot \frac{n}{m_1}$ geschrieben werden darf. Diese letzte Teilbarkeitsaussage und die Voraussetzung $(m_1, m_2) = 1$ bedingen $m_2|\frac{n}{m_1}$, wenn man (i) ausnützt. Das letztere ist mit $m_1m_2|n$ äquivalent.

Zu (iii): Sei $d_i := (n_i, m)$ für $i = 1, 2$ und $d := (n_1n_2, m)$. Mit gewissen ganzen x_i, y_i ist $d_i := n_ix_i + my_i$ für $i = 1, 2$, also

$$d_1d_2 = n_1n_2(x_1x_2) + m(n_1x_1y_2 + n_2x_2y_1 + my_1y_2),$$

was $d|d_1d_2$ liefert. Da $d_i|d$ für $i = 1, 2$ unmittelbar klar ist, folgt $d_1d_2|d$ aus (ii), sobald $(d_1, d_2) = 1$ eingesehen ist. Aus $d' := (d_1, d_2)$ ergibt sich tatsächlich $d'|m$, $d'|n_1$, $d'|n_2$, also $d'|(m, n_1, n_2)$, weshalb $d' = 1$ sein muß. □

Als für die Anwendungen wichtigster Spezialfall von (iii) dieses Satzes (vgl. etwa die Beweise der Sätze 2.1.8 bzw. 2.3.4) sei noch gesondert herausgestellt das schon von EUKLID (*Elemente* VII, § 24) explizit formulierte

Korollar. *Sind $m \neq 0$, n_1, n_2 ganz und sind m, n_i teilerfremd für $i = 1, 2$, so sind auch m, n_1n_2 teilerfremd.*

Bemerkungen. 1) In (ii) des obigen Satzes bräuchte man die Teilerfremdheit von m_1, m_2 selbstverständlich nur bei $n \neq 0$ vorauszusetzen. Ein direktes Analogon zu (ii) für $k \geq 3$ teilerfremde Zahlen $m_1, \ldots, m_k$ ist nicht zu erwarten: Die teilerfremden Zahlen 6, 10, 15, nicht aber ihr Produkt, teilen die Zahl 30.

2) Sind oben in (iii) insbesondere n_1, n_2 teilerfremd, so gilt die dortige Gleichung und n_1, n_2 verschwinden nicht beide. Bei $n_1 = 0$, $n_2 = 0$ gilt jene Gleichung genau für $m = \pm 1$; trivialerweise gilt sie auch für $m = 0$, wenn man (ohne Teilerfremdheitsvoraussetzung) dann $n_1n_2 \neq 0$ verlangt.

7. Charakterisierung der Primzahlen. Eine solche kann nun leicht aus Satz 6(i) gewonnen werden, wenn man noch folgenden Hilfssatz beachtet, dessen Beweis dem Leser überlassen bleibt.

Lemma. *Für Primzahlen p und ganze n sind die Bedingungen $(p, n) = 1$ bzw. $p \not| n$ äquivalent.*

Nun zum eigentlichen Anliegen dieses Abschnitts.

Satz. *Für ganzes $m \geq 2$ sind folgende Aussagen gleichbedeutend:*

(i) *m ist Primzahl.*

(ii) *Aus $m|n_1n_2$ mit ganzen n_1, n_2 folgt $m|n_1$ oder $m|n_2$.*

Beweis. Sei zuerst $m \in \mathbb{P}$, weiter $m|n_1n_2$ mit ganzen n_1, n_2 und $m \not| n_1$. Nach obigem Lemma ist die letzte Bedingung mit $(m, n_1) = 1$ äquivalent, weshalb $m|n_2$ nach Satz 6(i) gelten muß. Ist umgekehrt $\tau(m) \geq 3$ mit dem τ aus 1.3, also m zusammengesetzt, so ist $m = n_1n_2$ mit natürlichen, von 1 und m verschiedenen n_1, n_2, weshalb weder $m|n_1$ noch $m|n_2$ gelten kann. Somit trifft die Aussage (ii) hier nicht zu. □

Bemerkungen. 1) Die "interessante" Implikation (i) ⇒ (ii) dieses Satzes findet sich bei Euklid (*Elemente* VII, § 30) formuliert.

2) Die in 1.10 kurz diskutierten H-Primzahlen haben die Eigenschaft (ii) des obigen Satzes nicht: Zwar wird das Produkt der beiden H-Zahlen 4 und 25 von der H-Primzahl 10 geteilt, jedoch keine der beiden H-Zahlen selbst.

8. Nochmals: Eindeutigkeit im Fundamentalsatz. Wie bei dem in 1.5 geführten Eindeutigkeitsbeweis nach Zermelo sei $n > 1$ die kleinste natürliche Zahl mit nicht eindeutiger Produktzerlegung in Primzahlen. Mit gewissen Primzahlen p_ρ, q_σ hat man also

$$(1) \qquad n = p_1 \cdot \ldots \cdot p_r \quad \text{und} \quad n = q_1 \cdot \ldots \cdot q_s.$$

Nun ist $p_r|n$ und somit $p_r|q_1 \cdot \ldots \cdot q_s$ nach (1). Wegen (i) ⇒ (ii) von Satz 7 ist $p_r|q_1 \cdot \ldots \cdot q_{s-1}$ oder $p_r|q_s$. Induktiv führt dies zu $p_r|q_\sigma$ für mindestens ein $\sigma \in \{1, \ldots, s\}$ und o.B.d.A. sei etwa $p_r|q_s$, was sofort $p_r = q_s$ bedeutet. Die natürliche Zahl $\frac{n}{p_r}$, die kleiner als n ist, hätte dann aber nach (1) die beiden verschiedenen Primfaktorzerlegungen $p_1 \cdot \ldots \cdot p_{s-1}$ und $q_1 \cdot \ldots \cdot q_{s-1}$, was der Wahl von n widerspricht.

Daß auch dieser Eindeutigkeitsbeweis für den Fundamentalsatz der Arithmetik die additive Struktur von $\mathbb{N}$ ausnützt (vgl. 1.10), wird klar, wenn man überlegt, daß die verwendete Implikation (i) ⇒ (ii) in Satz 7 via Satz 3A letztlich auf den Divisionsalgorithmus 2 zurückgreift.

9. Euklidischer Algorithmus und ggT. Wie bereits in 1 und 3 angekündigt, soll nun ein effektives Verfahren zur Berechnung des ggT *zweier* ganzer Zahlen angegeben werden, die nicht beide Null sind. Wie am Ende von 5 (implizit) festgestellt, ist man damit in der Lage, den ggT von $k\,(\geq 2)$ ganzen Zahlen effektiv zu berechnen.

In leichter Abwandlung der bisherigen Bezeichungsweise seien r_0, r_1 die beiden ganzen Zahlen, deren ggT zu bestimmen ist. Nach (i) und (iv) der Proposition 5 darf o.B.d.A. $r_1 > 0$ vorausgesetzt werden.

Auf r_0, r_1 wird nun der Divisionsalgorithmus 2 angewandt und es ergibt sich nach 2(1)

$$r_0 = a_0 r_1 + r_2 \quad \text{und} \quad 0 \leq r_2 < r_1$$

mit ganzen a_0, r_2. Ist $r_2 = 0$, so stoppt man das Verfahren; ist $r_2 > 0$, so wendet man den Divisionsalgorithmus erneut an, jetzt auf r_1, r_2: Mit ganzen a_1, r_3 ist

$$r_1 = a_1 r_2 + r_3 \quad \text{und} \quad 0 \leq r_3 < r_2.$$

Auf diese Weise kann man fortfahren. Da die Folge $r_1, r_2, \ldots$ natürlicher Zahlen, die bei den wiederholten Anwendungen des Divisionsalgorithmus als Reste auftreten, streng monoton fällt, muß nach endlich vielen, etwa j Schritten der Rest 0 erscheinen. Dann endet das Verfahren und man hat insgesamt folgende Gleichungen erhalten:

$$(1) \qquad \begin{cases} r_0 &= a_0 r_1 + r_2 \\ r_1 &= a_1 r_2 + r_3 \\ &\vdots \\ r_{j-1} &= a_{j-1} r_j + r_{j+1} \\ r_j &= a_j r_{j+1}. \end{cases}$$

Dabei sind alle a_i, r_i ganz und die r_i genügen überdies

$$(2) \qquad r_1 > r_2 > \ldots > r_j > r_{j+1} > 0.$$

Da mit (1)

$$(3) \qquad \frac{r_i}{r_{i+1}} = a_i + \frac{r_{i+2}}{r_{i+1}} \ (i = 0, \ldots, j-1, \text{ falls } j \geq 1), \ \frac{r_j}{r_{j+1}} = a_j$$

äquivalent ist, ist a_i für $i = 0, \ldots, j$ die größte ganze, $\frac{r_i}{r_{i+1}}$ nicht übertreffende Zahl; dazu hat man (2) zu beachten. Nach der zuletzt gegebenen Interpretation der a_i kann somit auf die Ungleichungen $a_1 \geq 1, \ldots, a_{j-1} \geq 1$, $a_j \geq 2$ geschlossen werden, falls $j \geq 1$ ist.

Satz. *Wendet man auf die ganzen Zahlen r_0, r_1 mit $r_1 > 0$ in der bei (1) beschriebenen Weise sukzessive den Divisionsalgorithmus an, so ist $(r_0, r_1) = r_{j+1}$.*

Beweis. Definiert man $d := (r_0, r_1)$, so folgt $d|r_2$ aus der ersten, dann $d|r_3$ aus der zweiten usw., schließlich $d|r_{j+1}$ aus der vorletzten Gleichung in (1). Andererseits folgt $r_{j+1}|r_j$ aus der letzten, $r_{j+1}|r_{j-1}$ aus der vorletzten usw., endlich $r_{j+1}|r_1$ aus der zweiten und $r_{j+1}|r_0$ aus der ersten Gleichung (1). Nach Satz 3B gilt dann $r_{j+1}|d$, was zusammen mit $d|r_{j+1}$ die Behauptung beweist. □

Bemerkungen. 1) Die wiederholte Anwendung des Divisionsalgorithmus zur Gewinnung der Gleichungen (1) bezeichnet man als *euklidischen Algorithmus*. In der Tat findet sich das Verfahren völlig allgemein beschrieben bei EUKLID (*Elemente* VII, § 2).

2) Ein numerisches Beispiel zum euklidischen Algorithmus wird in 3.4 vorgeführt anläßlich einer geringfügig weitergehenden Aufgabenstellung als sie hier mit der bloßen Ermittlung des ggT zweier ganzer Zahlen vorliegt.

10. Regelmäßiger Kettenbruch rationaler Zahlen. Sofort an dieser Stelle soll die Gelegenheit ergriffen werden, um aus dem System 9(1) für eine beliebige rationale Zahl $\frac{r_0}{r_1}$ mit ganzen r_0, r_1 und $r_1 > 0$ eine spezielle Darstellung zu gewinnen, die sich bereits hier aufdrängt und an die später in 5.3.1 angeknüpft wird. Dazu werden für eine Folge $X_0, X_1, \ldots$ von Unbestimmten die Symbole $[X_0; X_1, \ldots, X_i]$, $i = 0, 1, \ldots$ rekursiv definiert durch die Festsetzungen

$$[X_0] := X_0 \tag{1}$$

bzw.

$$[X_0; X_1, \ldots, X_i] := \left[X_0; X_1, \ldots, X_{i-2}, X_{i-1} + \frac{1}{X_i}\right] \tag{2}$$

für $i \geq 1$. Sind nun r_0, r_1 wie oben und die $a_0, \ldots, a_j$ dem System 9(1) entnommen, so wird

$$\frac{r_0}{r_1} = \left[a_0; a_1, \ldots, a_{i-1}, \frac{r_i}{r_{i+1}}\right] \quad \text{für } i = 0, \ldots, j \tag{3}$$

behauptet. Nach (1) ist dies für $i = 0$ richtig. Ist $j > 0$ und (3) für ein ganzes i mit $0 \leq i < j$ bewiesen, so liefern (2), (3) und die $(i+1)$–te Gleichung 9(3)

$$\frac{r_0}{r_1} = \left[a_0; a_1, \ldots, a_{i-1}, a_i + \frac{r_{i+2}}{r_{i+1}}\right] = \left[a_0; a_1, \ldots, a_{i-1}, a_i, \frac{r_{i+1}}{r_{i+2}}\right]$$

und dies ist (3) für $i+1$ statt i. Wendet man (3) für $i=j$ an und beachtet die letzte Gleichung in 9(3), so erhält man für die rationale Zahl $\frac{r_0}{r_1}$ die Darstellung

$$\frac{r_0}{r_1} = [a_0; a_1, \ldots, a_j]\,. \tag{4}$$

Man sagt, man habe in (4) die rationale Zahl $\frac{r_0}{r_1}$ in einen *endlichen regelmäßigen* (oder *regulären*) *Kettenbruch* entwickelt. Die im Kettenbruch rechts in (4) auftretenden a_i heißen *Elemente* (oder auch *Teilnenner*) des Kettenbruchs. Die Feststellungen über die a_i im Anschluß an 9(3) zeigen: Bei $j \geq 1$ sind $a_1, \ldots, a_j$ natürliche Zahlen und überdies ist $a_j \geq 2$; die ganze Zahl a_0 kann auch Null oder negativ ausfallen. Um an dieses unterschiedliche Verhalten von a_0 und den a_i mit $i \geq 1$ zu erinnern, trennt man a_0 von den übrigen Elementen durch einen Strichpunkt ab.

Bemerkung. Vorwiegend in der älteren Literatur wurde der Kettenbruch rechts in (4) folgendermaßen expliziter notiert:

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots \cfrac{1}{a_{j-1} + \cfrac{1}{a_j}}}}}$$

Dies erklärt zwar die Bezeichnung "Kettenbruch"sehr gut, hat aber gegenüber der von O.Perron [19], S. 27 eingeführten einzeiligen Notation $[a_0; a_1, \ldots, a_j]$ den evidenten Nachteil viel größeren Platzverbrauchs.

11. Kleinstes gemeinsames Vielfaches (kgV). Hier wird noch ein Begriff besprochen, der aus dem Schulunterricht im Zusammenhang mit der Auffindung des Hauptnenners mehrerer, in gekürzter Form vorliegender rationaler Zahlen geläufig ist.

Es seien $n_1, \ldots, n_k$ stets ganze Zahlen, von denen *keine* Null ist. Man interessiert sich für die gemeinsamen Vielfachen aller n_i, also für alle ganzen m mit $n_1|m, \ldots, n_k|m$. Trivialerweise sind 0 und $\prod_{i=1}^{k} n_i$ solche gemeinsame Vielfache. Ist m gemeinsames Vielfaches, so auch $-m$, weshalb man sich weiterhin auf positive gemeinsame Vielfache beschränken kann.

Aufgrund dieser Vorbemerkung ist klar, daß die Menge der positiven gemeinsamen Vielfachen von $n_1, \ldots, n_k$ nicht leer ist; $\prod_{i=1}^{k} |n_i|$ gehört ja dazu. Diese Menge besitzt daher ein kleinstes Element, das herkunftsgemäß als *kleinstes gemeinsames Vielfaches* (kurz *kgV*; engl. *lcm* für least common multiple) der $n_1, \ldots, n_k$ bezeichnet wird. Ohne Verwechslungen mit anderen Verwendungen eckiger Klammern (wie z.B. in 10) zu riskieren, kann hier der Gepflogenheit gefolgt werden, das kgV von $n_1, \ldots, n_k$ als $[n_1, \ldots, n_k]$ zu notieren; gelegentlich wird dafür deutlicher $\mathrm{kgV}(n_1, \ldots, n_k)$ geschrieben.

Für ganzes $n_1 \neq 0$ ist $[n_1] = |n_1|$ sofort zu sehen, weshalb der Fall $k = 1$ hier wie in 1 beim ggT uninteressant ist.

Das kgV kann in analoger Weise charakterisiert werden wie dies durch Satz 3B für den ggT erledigt wurde:

Satz. *Für ganze, nicht verschwindende $n_1, \ldots, n_k$ und natürliches m sind folgende Aussagen gleichwertig:*

(i) *m ist das kgV von $n_1, \ldots, n_k$.*

(ii) *$n_1|m, \ldots, n_k|m$ und aus $m' \in \mathbb{N}$, $n_1|m', \ldots, n_k|m'$ folgt $m|m'$.*

Beweis. Sei erst (i) erfüllt, also $m = [n_1, \ldots, n_k]$, was $n_1|m, \ldots, n_k|m$ impliziert. Sei weiterhin $m' \in \mathbb{N}$ ein gemeinsames Vielfaches von $n_1, \ldots, n_k$. Wendet man auf m', m den Divisionsalgorithmus 2 an, so hat man mit ganzen a, b die Gleichung $m' = am + b$ sowie $0 \leq b < m$. Wegen $n_i|m$ und $n_i|m'$ ist $n_i|b$ für $i = 1, \ldots, k$, weshalb b nach Definition von m Null sein muß. Das zeigt $m|m'$. Die Umkehrung (ii) $\Rightarrow$ (i) ist trivial einzusehen, da aus $m|m'$ sofort $m \leq m'$ folgt. □

Das *kgV* ganzer $n_1, \ldots, n_k$, alle nicht Null, läßt sich somit beschreiben als *dasjenige positive Vielfache aller $n_1, \ldots, n_k$, das jedes andere derartige Vielfache teilt.* Der restliche Teil der Bemerkung in 3 zur (zweiten) Charakterisierung des ggT gilt hier wörtlich auch für das kgV.

Einige einfache Rechenregeln für das kgV können vom Leser selbst, eventuell gestützt auf die im Satz gegebene Charakterisierung, bewiesen werden; sie seien hier zusammengefaßt als

Proposition. *Seien $n_1, \ldots, n_k$ ganze, nichtverschwindende Zahlen und m ihr kgV.*

(i') *Für jede Permutation π der Indizes $1, \ldots, k$ ist $[n_{\pi(1)}, \ldots, n_{\pi(k)}] = m$.*

(ii') *Bei $k \geq 2$ und $n_k = 1$ ist $[n_1, \ldots, n_{k-1}] = m$.*

(iii') *Bei $k \geq 2$ und $n_{k-1} = n_k$ ist $[n_1, \ldots, n_{k-1}] = m$.*

(iv') *Es gilt* $[n_1, \ldots, n_{k-1}, -n_k] = m$.
(vi') *Für jedes ganze* $l \neq 0$ *ist* $[\ell n_1, \ldots, \ell n_k] = |\ell| m$.
($viii'$) *Bei* $k \geq 2$ *ist* $[[n_1, \ldots, n_{k-1}], n_k] = m$.

Hier entspricht die Numerierung offenbar derjenigen für die analogen Regeln über den ggT in 5.

Sind nun n_i für $i = 1, \ldots, k$ natürliche Zahlen mit der kanonischen Primfaktorzerlegung $n_i = \prod_p p^{\nu_p(n_i)}$, vgl. 1.6(2), so bestätigt der Leser leicht die Richtigkeit von

$$[n_1, \ldots, n_k] = \prod_p p^{\mathrm{Max}(\nu_p(n_1), \ldots, \nu_p(n_k))}.$$

Eine analoge Formel gilt übrigens für den ggT $(n_1, \ldots, n_k)$ mit Min statt Max in den Exponenten rechts.

Schließlich sei dem Leser der Beweis der folgenden idealtheoretischen Deutung des kgV überlassen:

Für ganze, nichtverschwindende $n_1, \ldots, n_k$ *stimmt der Durchschnitt der Ideale* $n_1\mathbb{Z}, \ldots, n_k\mathbb{Z}$ *in* $\mathbb{Z}$ *überein mit dem Ideal* $[n_1, \ldots, n_k]\mathbb{Z}$ *in* $\mathbb{Z}$.

12. Zusammenhang zwischen ggT und kgV. Der Leser hat bemerkt, daß die Erörterungen über das kgV in 11 weitgehend parallel zu denjenigen über den ggT verliefen. Es ist daher an der Zeit, den Zusammenhang zwischen beiden Begriffen aufzudecken. Hierzu dient folgender

Satz A. *Sind* $n_1, \ldots, n_k$, $n'_1, \ldots, n'_k$ *und* n *ganze, von Null verschiedene Zahlen mit* $n_i n'_i = n$ *für* $i = 1, \ldots, k$, *so gilt*

$$[n_1, \ldots, n_k](n'_1, \ldots, n'_k) = |n|. \tag{1}$$

Insbesondere hat man für beliebige ganze, von Null verschiedene Zahlen n_1,n_2

$$[n_1, n_2](n_1, n_2) = |n_1 n_2|. \tag{2}$$

Beweis. Mit $d := (n'_1, \ldots, n'_k)$ folgt $n_i \frac{n'_i}{d} = \frac{n}{d}$ für $i = 1, \ldots, k$ nach Voraussetzung, weshalb die ganze Zahl $\frac{n}{d}$ gemeinsames Vielfaches von $n_1, \ldots, n_k$ ist; nach Satz 11 geht also $m := [n_1, \ldots, n_k]$ in $\frac{n}{d}$ auf und für (1) ist lediglich noch $n | md$ einzusehen. Dazu definiert man ganze $m_1, \ldots, m_k$ durch $m_i := \frac{m}{n_i}$, was mit $nm_i = n'_i m$ für $i = 1, \ldots, k$ gleichwertig ist. Nach Proposition 5(vi) ist $|n|(m_1, \ldots, m_k) = m(n'_1, \ldots, n'_k) = md$, was tatsächlich $n | md$ beinhaltet. Damit ist (1) gezeigt.

Um (2) nachzuweisen, hat man (1) anzuwenden mit $n'_1 := n_2$, $n'_2 := n_1$ und $n := n_1 n_2$. □

Formel (2) führt das kgV *zweier* ganzer, von Null verschiedener Zahlen vollständig auf den ggT dieser Zahlen zurück. In Verbindung mit Proposition 11(viii′) ist damit die Bestimmung von $[n_1, \ldots, n_k]$ ganz allgemein auf die Berechnung größter gemeinsamer Teiler von *zwei* ganzen Zahlen reduziert.

Einen ähnlich einfachen Zusammenhang zwischen ggT und kgV von $n_1, \ldots, n_k$ wie im Falle $k = 2$ in Gestalt von Formel (2) hat man für $k \geq 3$ nicht. Der Leser kann aber für von Null verschiedene ganze n_1, n_2, n_3 durch geeignete Spezialisierung aus (1) gewinnen

$$(n_1, n_2, n_3)[n_1n_2, n_2n_3, n_3n_1] = |n_1n_2n_3| = [n_1, n_2, n_3](n_1n_2, n_2n_3, n_3n_1).$$

Anläßlich der Definition des kgV in 11 wurde implizit festgestellt, daß stets $[n_1, \ldots, n_k] \leq |n_1 \cdot \ldots \cdot n_k|$ gilt. Abschließend soll charakterisiert werden, wann genau in dieser Ungleichung das Gleichheitszeichen eintritt.

Satz B. *Für $k \geq 2$ ganze, von Null verschiedene Zahlen $n_1, \ldots, n_k$ sind folgende Aussagen äquivalent:*

(i) $n_1, \ldots, n_k$ *sind paarweise teilerfremd.*

(ii) *Es gilt* $[n_1, \ldots, n_k] = |n_1 \cdot \ldots \cdot n_k|$.

Beweis. Man setze $n := n_1 \cdot \ldots \cdot n_k$ und $n_i' := \frac{n}{n_i}$ für $i = 1, \ldots, k$. Formel (1) von Satz A zeigt die Äquivalenz von (ii) mit $(n_1', \ldots, n_k') = 1$, was sich als zu (i) gleichwertig erweisen muß. Diese letzte Äquivalenz ist aber deswegen gegeben, weil aufgrund der obigen Definition $n_i' = n_1 \cdot \ldots \cdot n_{i-1} \cdot n_{i+1} \cdot \ldots \cdot n_k$ für jede Primzahl p gilt: p geht in allen $n_1', \ldots, n_k'$ auf genau dann, wenn es in mindestens zweien der $n_1, \ldots, n_k$ aufgeht. □

§ 3. Lineare diophantische Gleichungen

1. Warum "diophantisch"? Im letzten Paragraphen wurde verschiedentlich die Tragweite des Satzes 2.3A deutlich, der hier aus einem etwas anderen Blickwinkel betrachtet werden soll. Sind nämlich $a_1, \ldots, a_k$ ganze teilerfremde (nicht sämtlich verschwindende) Zahlen, so beinhaltet der angesprochene Satz die Existenz ganzer $x_1, \ldots, x_k$ mit

$$a_1x_1 + \ldots + a_kx_k = 1. \tag{1}$$

Oder etwas anders gewandt: Die Gleichung

$$a_1X_1 + \ldots + a_kX_k = 1 \tag{2}$$

in den Unbestimmten $X_1, \ldots, X_k$ mit ganzen teilerfremden Koeffizienten $a_1, \ldots$ $\ldots, a_k$ ist in ganzen Zahlen lösbar, d.h. es gibt *mindestens ein* der Gleichung (1) genügendes $(x_1, \ldots, x_k) \in \mathbb{Z}^k$.

Wenn man die *Lösbarkeit* einer unbestimmten Gleichung wie (2) erst einmal gesichert hat, wird man sich als nächstes fragen, *wieviele Lösungen* die betrachtete Gleichung hat, d.h. wieviele (1) genügende $(x_1, \ldots, x_k) \in \mathbb{Z}^k$ es gibt. Und schließlich wird man sich noch weitergehend für die *Struktur der Menge aller Lösungen* interessieren.

Fragestellungen der hier angedeuteten Art werden, historisch belegt, seit über zweieinhalb Jahrtausenden behandelt. So hat bereits PYTHAGORAS bemerkt, die Gleichung $3^2 + 4^2 = 5^2$ impliziere die geometrische Tatsache, daß jedes ebene Dreieck mit dem Seitenlängenverhältnis $3 : 4 : 5$ rechtwinklig ist. Deshalb suchte er nach anderen Quadratzahlen, die wie 5^2 Summe zweier Quadratzahlen sind, d.h. er suchte weitere Lösungen in natürlichen Zahlen der Gleichung

$$X_1^2 + X_2^2 = X_3^2. \tag{3}$$

Tatsächlich schreibt man PYTHAGORAS die Entdeckung zu, daß jedes Tripel $(m, \frac{1}{2}(m^2 - 1), \frac{1}{2}(m^2 + 1))$ mit ungeradem ganzem $m \geq 3$ eine Lösung von (3) in natürlichen Zahlen ist. Die Lösungen (8, 15, 17) und (12, 35, 37) von (3) belegen aber, daß in der Lösungsschar des PYTHAGORAS keineswegs sämtliche Lösungen von (3) in natürlichen Zahlen enthalten sind. Wie die Gesamtheit aller derartigen Lösungen von (3) aussieht, hat EUKLID (*Elemente X*, §§ 28, 29) beschrieben; der Leser findet dies Resultat in 4.2.1–2.

PYTHAGORAS hat mit seiner Entdeckung der unendlich vielen Lösungen von (3) eine Entwicklung eingeleitet, die erst einige Jahrhunderte später durch DIOPHANT (um 250 n. Chr. ?) einen vorläufigen Höhepunkt und Abschluß erreichte. DIOPHANT lebte im ägyptischen Alexandria, jahrhundertelang dem wissenschaftlichen Zentrum der antiken Welt. Seine Lebensdaten sind recht unsicher und schwanken zwischen 100 v. Chr. und 350 n. Chr. Diese Grenzen ergeben sich indirekt einerseits aus Erwähnungen älterer Mathematiker in DIOPHANTs *Arithmetika* (*Αριθμητικά*), andererseits durch Zitierungen dieses Werks in der späteren Literatur.

Die *Arithmetika* gilt als erste große, ausschließlich zahlentheoretischen Problemen gewidmete Abhandlung, deren Einfluß auf die Entwicklung der Zahlentheorie kaum zu überschätzen ist. Dabei haben von den überlieferten Teilen dieses Werks diejenigen Untersuchungen DIOPHANTs bis heute die stärksten Impulse gegeben, die sich mit unbestimmten Gleichungen des Typs

$$P(X_1, \ldots, X_k) = 0, \qquad P \in \mathbb{Z}[X_1, \ldots, X_k] \tag{4}$$

wie (2) oder (3) befaßten. Aus zahlreichen Beispielen solcher Gleichungen bis zum Gesamtgrad vier von P, die in der *Arithmetika* explizit vorgeführt wur-

den, ließen sich sehr allgemeine Methoden zur Gewinnung der Lösungen in ganzen (oder wie bei DIOPHANT meistens in positiven) rationalen Zahlen herauspräparieren. Zwei der DIOPHANTschen Methoden werden später in 4.2.3–4 vorgestellt; 4.2.5–8 vermitteln einen Eindruck vom Nachwirken DIOPHANTs bis in die Neuzeit.

DIOPHANT zu Ehren bezeichnet man jede unbestimmte Gleichung des Typs (4), für die man arithmetisch charakterisierte Lösungen $(x_1, \ldots, x_k)$ (also $(x_1, \ldots, x_k)$ aus $\mathbb{Z}^k$ oder aus $\mathbb{Q}^k$ oder ähnlich) sucht, als *diophantische Gleichung*, genauer als *polynomiale* diophantische Gleichung. Den Zusatz "polynomial" fügt man bei Gleichungen der Form (4) an, seit auch unbestimmte Gleichungen verstärkt behandelt werden, die nicht von der Form (4) sind. Weiter heißen die speziellen polynomialen Gleichungen des Typs (2) *lineare* diophantische Gleichungen, natürlich auch bei Ersetzung der Eins rechts in (2) durch irgendeine feste ganze Zahl.

Bemerkung. Von den dreizehn Büchern der *Arithmetika* schienen bis vor etwa dreißig Jahren lediglich sechs erhalten, sämtliche wie das Original in griechischer Sprache. Selbstverständlich wurden Übersetzungen und Kommentierungen seit DIOPHANT auch in zahlreichen anderen Sprachen herausgegeben. Als deutsche Übersetzung sei genannt

CZWALINA, A.: *Arithmetik des Diophantos aus Alexandria*, Vandenhoeck–Ruprecht, Göttingen, 1952.

Als ausgezeichnete Kommentierung von DIOPHANTs Werk und der Weiterentwicklung seiner Methoden kann dem Leser empfohlen werden

BASMAKOVA, I.G.: *Diophant und diophantische Gleichungen*, Birkhäuser, Basel-Stuttgart, 1974.

Vor rund dreißig Jahren sind vier weitere Bücher der *Arithmetika* aufgetaucht und zwar in arabischer Sprache. In ihrem Titel wiesen sie sich als Bücher IV bis VII*) aus:

SESIANO, J.: *Books IV to VII of Diophantus' Arithmetica: In the Arabic Translation Attributed to Qusṭâ Ibn Lûqâ*, Springer, New York, 1982.

2. Lösbarkeitsbedingung. Die Ergebnisse der vorangegangenen Paragraphen erlauben in diesem eine vollständige Behandlung der 1(2) verallgemeinernden linearen diophantischen Gleichung

$$a_1X_1 + \ldots + a_kX_k = c \tag{1}$$

*) Im vorliegenden Buch ist die *ältere* Numerierung von I bis VI bei Zitaten der zuerst bekannten DIOPHANT–Bücher beibehalten.

mit ganzen $a_1, \ldots, a_k, c$. Sind hier alle a_i Null, so ist (1) für $c \neq 0$ unlösbar; ist aber auch noch $c = 0$, so ist jedes $(x_1, \ldots, x_k) \in \mathbb{Z}^k$ Lösung von (1). Seien in (1) also künftig nicht alle a_i Null; dann liefert der folgende Satz eine notwendige und hinreichende Bedingung für die Lösbarkeit von (1).

Satz. *Für ganze $a_1, \ldots, a_k, c$, nicht alle a_i gleich Null, ist die diophantische Gleichung (1) genau dann lösbar, wenn der ggT der $a_1, \ldots, a_k$ in c aufgeht.*

Beweis. Ist (1) lösbar, so gibt es ganze $x_1, \ldots, x_k$ mit

$$a_1 x_1 + \ldots + a_k x_k = c. \tag{2}$$

Nach Satz 1.2(ix) wird dann c von $d := (a_1, \ldots, a_k)$ geteilt. Umgekehrt folgt aus Satz 2.3A die Existenz ganzer $y_1, \ldots, y_k$ mit

$$a_1 y_1 + \ldots + a_k y_k = d. \tag{3}$$

Geht man nun von der Voraussetzung $d|c$ aus, so gilt $c = qd$ mit ganzem q; man setzt $x_i := qy_i$ für $i = 1, \ldots, k$ und sieht nach Multiplikation von (3) mit q, daß (2) für den Vektor $(x_1, \ldots, x_k)$ zutrifft, dieser also (1) löst. □

Zur weiteren Behandlung der Gleichung (1) kann vorausgesetzt werden, daß die Lösbarkeitsbedingung $(a_1, \ldots, a_k)|c$ des obigen Satzes erfüllt ist und daß o.B.d.A. $a_1 \cdot \ldots \cdot a_k \neq 0$ gilt. Im Fall $k = 1$ ist die ganze Zahl $\frac{c}{a_1}$ offenbar die einzige Lösung von (1) und so kann man ab sofort noch $k \geq 2$ verlangen. Die weitere Diskussion von (1) vollzieht sich nun so, daß in 3 und 4 der Fall $k = 2$ komplett erledigt wird. In 5 wird dann bewiesen, daß sich der Fall von mehr als zwei Unbestimmten in (1) auf den schon behandelten mit genau zwei Unbestimmten zurückführen läßt.

3. Der Fall zweier Unbestimmten wird weiter reduziert durch folgenden

Satz. *Seien a, b, c ganze Zahlen mit $ab \neq 0$, der ggT d von a, b teile c und für die ganzen Zahlen x_0, y_0 gelte*

$$ax_0 + by_0 = d. \tag{1}$$

Dann hat die lineare diophantische Gleichung

$$aX + bY = c \tag{2}$$

genau die ganzzahligen Lösungen

$$\left(\frac{cx_0+bt}{d},\ \frac{cy_0-at}{d}\right) \tag{3}$$

mit ganzem t.

Bemerkung. Unter den gemachten, völlig natürlichen Voraussetzungen über a, b, c sind damit alle Lösungen von (2) bekannt genau dann, wenn man ganze x_0, y_0 gefunden hat, die (1) genügen. Das letztgenannte Problem wird in 4 vollständig gelöst.

Beweis. Mühelos prüft man zunächst, daß jedes in (3) genannte Paar (wegen $d|a, b, c$) ganzer Zahlen mit Rücksicht auf (1) die Gleichung (2) löst.

Ist nun umgekehrt $(x, y) \in \mathbb{Z}^2$ irgendeine Lösung von (2), so erhält man

$$\frac{a}{d}(x-x_1) = -\frac{b}{d}(y-y_1), \tag{4}$$

wenn man die ganzen Zahlen x_1 bzw. y_1 durch $\frac{c}{d}x_0$ bzw. $\frac{c}{d}y_0$ definiert und wieder (1) berücksichtigt. Nach Proposition 2.5(vii) sind $\frac{a}{d}$, $\frac{b}{d}$ teilerfremd; wegen Satz 2.6(i) muß $\frac{b}{d}$ in $x - x_1$ aufgehen, d.h. es gibt ein ganzes t mit $x - x_1 = \frac{b}{d}t$ und nach (4) ist dann $y - y_1 = -\frac{a}{d}t$. Nach Definition von x_1, y_1 sind die zuletzt erhaltenen Gleichungen äquivalent mit $x = \frac{cx_0+bt}{d}$, $y = \frac{cy_0-at}{d}$. □

4. Spezielle Lösung, numerisches Beispiel. Das in 3 offen gebliebene Problem des effektiven Auffindens ganzer x_0, y_0, die 3(1) erfüllen, wird ganz explizit erledigt durch die

Proposition. *Seien a, b von Null verschiedene ganze Zahlen und d ihr ggT. Wendet man den euklidischen Algorithmus 2.9(1) an mit den Startwerten $r_0 := |a|$, $r_1 := |b|$, entnimmt man j und die $a_0, \dots, a_{j-1}$ aus 2.9(1) und definiert damit rekursiv*

$$p_0 := 0, \quad q_0 := 1, \quad p_{i+1} := q_i, \quad q_{i+1} := p_i - a_{j-1-i} \cdot q_i \quad (0 \le i < j), \tag{1}$$

so gilt

$$|a|p_j + |b|q_j = d, \tag{2}$$

d.h. $x_0 := p_j \operatorname{sgn} a$, $y_0 := q_j \operatorname{sgn} b$ genügen 3(1).

Beweis. Für die $r_0, \dots, r_{j+1}$ aus 2.9(1) sind die Gleichungen

$$r_{j-i}p_i + r_{j+1-i}q_i = r_{j+1} \qquad (i = 0, \dots, j) \tag{3}$$

wegen (1) induktiv sofort klar. Da $r_{j+1} = d$ nach Satz 2.9 mit Rücksicht auf Proposition 2.5(iv) gilt, ist (3) im Spezialfall $i = j$ mit (2) identisch. □

Korollar. *Sind a, b, d wie in Satz 3, so können ganze, der Gleichung 3(1) genügende x_0, y_0 mittels euklidischem Algorithmus bestimmt werden.*

Beispiel. Es soll die Gleichung

$$9973X - 2137Y = 1 \tag{4}$$

untersucht werden, die nach Satz 2 genau dann lösbar ist, wenn die Koeffizienten 9973 und 2137 links in (4) teilerfremd sind. Um ihre Teilerfremdheit zu prüfen, wendet man den euklidischen Algorithmus 2.9(1) an, der hier wie folgt abläuft:

$$\begin{array}{rcrcr} 9973 & = & 4 \cdot 2137 & + & 1425 \\ 2137 & = & 1 \cdot 1425 & + & 712 \\ 1425 & = & 2 \cdot 712 & + & 1 \\ 712 & = & 712 \cdot 1 & & \end{array} \tag{5}$$

Die Koeffizienten links in (4) sind damit als teilerfremd erkannt (es sind sogar beides Primzahlen) und so hat (4) unendlich viele Lösungen, die man 3(3) entnehmen kann, sobald man eine spezielle Lösung (x_0, y_0) von (4) kennt. Dazu hat man laut obiger Proposition die $j, a_0, \ldots, a_{j-1}$ aus 2.9(1) im Spezialfall (5) festzustellen: Hier ist $j = 3$, $a_0 = 4$, $a_1 = 1$, $a_2 = 2$ (und $a_3 = 712$), was gemäß (1) zu $p_1 = 1$, $p_2 = -2$, $p_3 = 3$ bzw. $q_1 = -2$, $q_2 = 3$, $q_3 = -14$ führt. Also ist das Paar $(x_0, y_0) = (3, 14)$ eine spezielle Lösung von (4), was nach Satz 3 die allgemeine Lösung $(3 + 2137t,\ 14 + 9973t)$, $t \in \mathbb{Z}$, hat.

Bemerkungen. 1) Gemäß 2.10 hat sich hier ganz nebenbei der regelmäßige Kettenbruch der rationalen Zahlen $\frac{9973}{2137}$ zu $[4; 1, 2, 712]$ ergeben oder ausgeschrieben

$$\frac{9973}{2137} = 4 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{712}}}$$

2) Selbstverständlich gibt es Situationen, wo man zur Lösung einer Gleichung des Typs 3(2) nicht erst den euklidischen Algorithmus benützen wird, sondern wo genaues Hinsehen schon weiterhilft. So sieht man z.B. der Gleichung $7X + 10Y = 1$ unmittelbar *eine* Lösung $(3, -2)$ an, weshalb ihre allgemeine Lösung mit $(3 + 10t, -2 - 7t)$, $t \in \mathbb{Z}$, bereits hingeschrieben werden kann.

3) Merkwürdigerweise findet sich die vollständige Behandlung der Gleichung 3(2) *schriftlich* erst bei den indischen Astronomen ARYABHATA und BRAHMAGUPTA (zwischen 500 und 600), deren Methode auf dem euklidischen Algorithmus basiert. Dies verwundert umso mehr, als z.B. die großen griechischen Mathematiker der Antike von EUKLID bis DIOPHANT viel schwierigere Gleichungen höheren Grades komplett lösen konnten, vgl. etwa 4.2.1 und 4.3.1.

5. Reduktion des allgemeinen Falls. Bei ganzen $a_1, \ldots, a_k, c$ wird hier die Gleichung 2(1), also

$$a_1X_1 + \ldots + a_kX_k = c \tag{1}$$

im Fall $k \geq 3$ diskutiert unter den am Ende von 2 genannten Voraussetzungen $a_1 \cdot \ldots \cdot a_k \neq 0$, $\mathrm{ggT}(a_1, \ldots, a_k)|c$. Dazu wird mit $a := \mathrm{ggT}(a_{k-1}, a_k)$ neben (1) das folgende System von zwei linearen diophantischen Gleichungen in den $k+1$ Unbestimmten $X_1, \ldots, X_k, Y$ betrachtet:

$$\begin{aligned} a_1X_1 + \ldots + a_{k-2}X_{k-2} + aY &= c, \\ a_{k-1}X_{k-1} + a_kX_k - aY &= 0. \end{aligned} \tag{2}$$

Sei nun $\mathbf{x} := (x_1, \ldots, x_k) \in \mathbb{Z}^k$ eine Lösung von (1) und es werde y definiert durch $y := \frac{1}{a}(a_{k-1}x_{k-1} + a_kx_k)$. Nach Definition von a ist auch y ganz und ersichtlich löst der Vektor $(\mathbf{x}, y) := (x_1, \ldots, x_k, y) \in \mathbb{Z}^{k+1}$ das System (2). Klar ist auch, daß die durch $\mathbf{x} \mapsto (\mathbf{x}, y)$ definierte Abbildung der Lösungsmenge von (1) in die von (2) injektiv und surjektiv, also bijektiv ist. Damit kann man sicher sein, *alle* Lösungen $\mathbf{x} \in \mathbb{Z}^k$ von (1) zu erhalten, wenn man alle Lösungen $(\mathbf{x}, y) \in \mathbb{Z}^{k+1}$ von (2) kennt.

Dabei ist zu beachten, daß wiederholte Anwendung von Proposition 2.5(viii) zu $\mathrm{ggT}(a_1, \ldots, a_{k-2}, a) = \mathrm{ggT}(a_1, \ldots, a_k)$ führt, weshalb die oben vorausgesetzte Lösbarkeitsbedingung für (1) mit $\mathrm{ggT}(a_1, \ldots, a_{k-2}, a)|c$, also mit der nach Satz 2 notwendigen und hinreichenden Lösbarkeitsbedingung für die erste Gleichung in (2) äquivalent ist.

Die Lösungen $(\mathbf{x}, y)$ von (2) bekommt man jetzt so: Die Induktionsvoraussetzung garantiert, daß man die erste Gleichung in (2) vollständig lösen kann. Von jeder solchen Lösung $(x_1, \ldots, x_{k-2}, y)$ nimmt man die letzte Komponente y und behandelt damit via 3 und 4 die Gleichung

$$a_{k-1}X_{k-1} + a_kX_k = ay. \tag{3}$$

6. Struktur der Lösungsgesamtheit. Die in 5 vorgenommene Reduktion gestattet nun eine vollständige Beschreibung der Lösungen von 5(1).

Satz. *Seien $a_1, \ldots, a_k$ ganze, von Null verschiedene Zahlen, deren ggT die ganze Zahl c teilt. Dann hat die Lösungsmenge der diophantischen Gleichung 5(1) die Form $\mathbf{p} + \mathbb{Z}\mathbf{q}_1 + \ldots + \mathbb{Z}\mathbf{q}_{k-1}$, wobei sämtliche $\mathbf{q}_i \in \mathbb{Z}^k$ nur von $a_1, \ldots, a_k$ abhängen, während $\mathbf{p} \in \mathbb{Z}^k$ außerdem von c abhängt und (genau) bei $c = 0$*

als Nullvektor wählbar ist. Überdies ist der Rang der aus den Zeilenvektoren $\mathbf{q}_1, \ldots, \mathbf{q}_{k-1}$ *gebildeten Matrix maximal, also gleich* $k-1$.

Beweis. (Induktion über k). Für $k = 1$ ist die Aussage des Satzes trivial und für $k = 2$ entnimmt man sie aus Satz 3. Sei nun $k \geq 3$ und gezeigt, daß die Lösungsmenge linearer diophantischer Gleichungen in $k-1$ Unbestimmten, insbesondere also der ersten Gleichung in 5(2) die Form $\mathbf{p}^* + \mathbb{Z}\mathbf{q}_1^* + \ldots + \mathbb{Z}\mathbf{q}_{k-2}^*$ hat, wobei die $\mathbf{p}^*, \mathbf{q}_1^*, \ldots, \mathbf{q}_{k-2}^* \in \mathbb{Z}^{k-1}$ die übrigen im Satz behaupteten Eigenschaften bereits aufweisen mögen. Mit $\mathbf{p}^* = (p_1^*, \ldots, p_{k-1}^*)$ und $\mathbf{q}_i^* = (q_{i1}^*, \ldots, q_{i,k-1}^*)$ für $i = 1, \ldots, k-2$ gilt für die allgemeine Lösung $(x_1, \ldots, x_{k-2}, y)$ der ersten Gleichung in 5(2) somit

$$(1) \qquad x_j = p_j^* + \sum_{i=1}^{k-2} q_{ij}^* t_i \quad (j = 1, \ldots, k-2), \quad y = p_{k-1}^* + \sum_{i=1}^{k-2} q_{i,k-1}^* t_i.$$

Nach Induktionsvoraussetzung hängen alle hier vorkommenden q_{ij}^* nur von $a_1, \ldots, a_{k-2}$ und $a = \mathrm{ggT}(a_{k-1}, a_k)$, also nur von $a_1, \ldots, a_k$ ab, bei $c = 0$ kann $p_j^* = 0$ für $j = 1, \ldots, k-1$ gewählt werden und es ist Rang $(q_{ij}^*)_{1 \leq i \leq k-2, 1 \leq j \leq k-1} = k-2$.

Für jedes feste y aus (1) hat man jetzt die Gleichung 5(3) anzusehen. Nach Satz 3 ist die allgemeine Lösung (x_{k-1}, x_k) von 5(3) von der Form

$$(2) \qquad x_{k-1} = x_0 y + \frac{a_k}{a} t_{k-1}, \quad x_k = y_0 y - \frac{a_{k-1}}{a} t_{k-1},$$

wobei t_{k-1} ganz $\mathbb{Z}$ durchläuft und die x_0, y_0 alleine von a_{k-1}, a_k abhängen, der Gleichung $a_{k-1}x_0 + a_k y_0 = a$ genügen und daher nicht beide verschwinden. Substituiert man in (2) noch für y aus (1), so findet man nach der Überlegung aus 5 die allgemeine Lösung $\mathbf{x} = (x_1, \ldots, x_k)$ von 5(1) in der Form $\mathbf{p} + \mathbf{q}_1 t_1 + \ldots + \mathbf{q}_{k-1} t_{k-1}$, $(t_1, \ldots, t_{k-1}) \in \mathbb{Z}^{k-1}$, mit

$$(3) \qquad \begin{cases} \mathbf{q}_i := (q_{i1}^*, \ldots, q_{i,k-2}^*, x_0 q_{i,k-1}^*, y_0 q_{i,k-1}^*) & \text{für } i = 1, \ldots, k-2, \\ \mathbf{q}_{k-1} := (0, \ldots, 0, \frac{a_k}{a}, -\frac{a_{k-1}}{a}), \quad \mathbf{p} := (p_1^*, \ldots, p_{k-2}^*, x_0 p_{k-1}^*, y_0 p_{k-1}^*). \end{cases}$$

Hieraus sind die Behauptungen des Satzes über die $\mathbf{p}, \mathbf{q}_1, \ldots, \mathbf{q}_{k-1}$ aufgrund der oben erwähnten Induktionsvoraussetzung und der Eigenschaften von x_0, y_0 evident bis auf die Rangaussage. Für gewisse ganze $\tau_1, \ldots, \tau_{k-1}$ sei $\sum_{i=1}^{k-1} \tau_i \mathbf{q}_i = \mathbf{0}$, d.h. $\sum_{i=1}^{k-2} \tau_i q_{ij}^* = 0$ für $j = 1, \ldots, k-2$ sowie

$$x_0 \sum_{i=1}^{k-2} \tau_i q_{i,k-1}^* + \tau_{k-1} \frac{a_k}{a} = 0, \quad y_0 \sum_{i=1}^{k-2} \tau_i q_{i,k-1}^* - \tau_{k-1} \frac{a_{k-1}}{a} = 0,$$

was äquivalent ist mit

$$\sum_{i=1}^{k-2} \tau_i q^*_{i,k-1} = 0, \quad \tau_{k-1} = 0.$$

Aus $\sum_{i=1}^{k-2} \tau_i \mathbf{q}^*_i = \mathbf{0}$ folgt $\tau_1 = 0, \ldots, \tau_{k-2} = 0$ nach Induktionsvoraussetzung, womit auch die Rangaussage bewiesen ist.

Die Aussage des "genau" im Satz ist trivial: Ist $\mathbf{p}$ als Nullvektor wählbar, so muß dieser offenbar 5(1) lösen, d.h. es muß $c = 0$ sein. □

Schließlich sei noch als Kurzfassung des Satzes formuliert

Korollar. *Unter den Voraussetzungen des Satzes bildet die Lösungsmenge von 5(1) einen $(k-1)$-dimensionalen freien $\mathbb{Z}$-Modul, der genau für $c = 0$ sogar ein $\mathbb{Z}$-Modul ist.*

§ 4. Zahlentheoretische Funktionen

1. Einige Definitionen. Jede Abbildung $f : \mathbb{N} \to \mathbb{C}$ bezeichnet man als *zahlentheoretische Funktion.* In der Sprache der Analysis ist dies nichts anderes als eine komplexwertige Folge, also ein Element aus $\mathbb{C}^{\mathbb{N}}$; nur schreibt man in der Zahlentheorie traditionell $f(n)$ statt f_n für die Folgenglieder. Außerdem wird hier $\mathbb{C}^{\mathbb{N}}$ mit Z abgekürzt.

Als Beispiele zahlentheoretischer Funktionen, die bisher schon in natürlicher Weise im Rahmen der Teilbarkeitstheorie auftraten, seien genannt: Die Teileranzahlfunktion τ (in 1.3 und 1.7), die Teilersummenfunktion σ (in 1.7) und für jede feste Primzahl p die Vielfachheit ν_p (in 1.6). In diesem Paragraphen werden folgende weiteren zahlentheoretischen Funktionen immer wieder vorkommen:

a) Die durch $\mathbf{0}(n) := 0$ für alle natürlichen n definierte Funktion $\mathbf{0}$;

b) die durch $\varepsilon(1) := 1$ und $\varepsilon(n) := 0$ für alle ganzen $n \geq 2$ definierte Funktion ε;

c) die mit ι abgekürzte Identität auf $\mathbb{N}$ und schließlich für jedes reelle α die durch $\iota_\alpha(n) := n^\alpha$ für alle $n \in \mathbb{N}$ definierten Funktionen ι_α; es ist insbesondere $\iota_1 = \iota$ und $\iota_0(n) = 1$ für alle $n \in \mathbb{N}$.

2. Multiplikative und additive Funktionen. Aufgrund der großen Allgemeinheit des Begriffs einer zahlentheoretischen Funktion ist plausibel, daß man im Rahmen der Zahlentheorie nicht an all diesen Funktionen gleichermaßen interessiert ist. Eine für die Zahlentheorie besonders wichtige Teilmenge von Z, hier mit M abgekürzt, besteht aus den multiplikativen Funktionen.

Dabei heißt $f \in Z$ *multiplikativ*, wenn

$$(1) \qquad f(n_1 n_2) = f(n_1) f(n_2) \quad \text{für alle teilerfremden } n_1, n_2 \in \mathbb{N}$$

gilt; trifft (1) für alle $n_1, n_2 \in \mathbb{N}$ ohne die Einschränkung "teilerfremd" zu, so heißt f *streng multiplikativ*.

Beispiele streng multiplikativer Funktionen sind $\mathbf{0}$, ε und sämtliche ι_α mit $\alpha \in \mathbb{R}$; daß weder τ noch σ streng multiplikativ sind, ist Spezialfall einer Feststellung in 10. Dagegen sind τ und σ beide multiplikativ, wie man 1.7(2) bzw. 1.7(4) in Verbindung mit (ii) der nächsten Proposition entnimmt. Die Multiplikativität von τ, σ wird sich nochmals in 10 ergeben.

Proposition.

(i) *Für jedes $f \in M \setminus \{\mathbf{0}\}$ gilt $f(1) = 1$.*

(ii) *Für $f \in Z$ gilt die Äquivalenz*

$$f \in M \iff f(n) = \prod_p f\left(p^{\nu_p(n)}\right) \qquad \text{für alle } n \in \mathbb{N}.$$

Beweis. Zu (i): Wegen $f(1) = f(1 \cdot 1) = f(1)^2$ gemäß (1) ist entweder $f(1) = 1$ oder $f(1) = 0$. Die zweite Alternative würde $f(n) = f(n \cdot 1) = f(n) f(1) = 0$ (erneut nach (1)) für alle $n \in \mathbb{N}$ implizieren, also $f = \mathbf{0}$.

Zu (ii): Ist $f \in M$, so folgt die behauptete Zerlegung von $f(n)$, indem man (1) endlich oft auf die Produktzerlegung $\prod_p p^{\nu_p(n)}$ von n anwendet und $f(1) = 1$ bei $f \neq \mathbf{0}$ beachtet, welch letzteres o.B.d.A. vorausgesetzt werden darf. Die Implikation "$\Leftarrow$" ist trivial. □

Teil (ii) der Proposition deckt einen Grund für die Bedeutung der *multiplikativen zahlentheoretischen Funktionen* auf: Sie *sind durch ihre Werte an den sämtlichen Primzahlpotenzen bereits vollständig festgelegt.*

Bemerkung. Gelegentlich sind auch sogenannte additive $f \in Z$ in der Zahlentheorie von Bedeutung. Dabei heißt $f \in Z$ *additiv*, wenn $f(n_1 n_2) = f(n_1) + f(n_2)$ für alle teilerfremden $n_1, n_2 \in \mathbb{N}$ gilt; trifft dies ohne die Einschränkung "teilerfremd" zu, so heißt f *streng additiv*.

Hier seien noch einige Kleinigkeiten über additive Funktionen zusammengestellt, die sich der Leser selbst überlegen möge: Für jede Primzahl p ist ν_p streng additiv. Die durch $\omega(n) := \#\{p \in \mathbb{P} : p|n\}$ definierte Funktion ω ist additiv, jedoch nicht streng additiv. Ist $f : \mathbb{N} \to \mathbb{R}_+$ (streng) multiplikativ, so ist offenbar die Zusammensetzung $\log \circ f$ (streng) additiv. Für additives $f \in Z$ gilt stets $f(1) = 0$.

3. Produktdarstellung unendlicher Reihen. Dieser und der folgende Abschnitt stellen Hilfsmittel bereit, die vor allem in Kap. 7 zur genaueren Untersuchung der Primzahlverteilung unabdingbar sein werden. Sie werden aber auch schon in 5 und 12 mit Erfolg angewandt.

Satz. *Ist $f \in Z$ multiplikativ und $\sum_{n=1}^{\infty} f(n)$ absolut konvergent, so gilt*

$$\sum_{n=1}^{\infty} f(n) = \prod_{p} \sum_{\nu=0}^{\infty} f(p^\nu). \tag{1}$$

Oft nützlich ist noch folgendes Ergebnis, bei dem die Voraussetzungen leicht abgewandelt sind, dessen Beweis aber bis zu einem gewissen Punkt demjenigen des Satzes folgt.

Proposition. *Ist $f \in Z$ reellwertig, nichtnegativ und multiplikativ und konvergiert $\sum_{\nu \geq 0} f(p^\nu)$ für jede Primzahl p, so gilt für alle reellen x*

$$\sum_{n \leq x} f(n) \leq \prod_{p \leq x} \sum_{\nu=0}^{\infty} f(p^\nu). \tag{2}$$

Beweise. O.B.d.A. sei $f \neq \mathbf{0}$. Für reelles $x < 2$ ist (2) trivial wegen $f(1) = 1$ und der Konvention über leere Summen bzw. Produkte. Sei ab jetzt $x \geq 2$ und seien $p_1, \ldots, p_k$ genau die verschiedenen, x nicht übersteigenden Primzahlen. Dann ist

$$\prod_{p \leq x} \sum_{\nu \geq 0} f(p^\nu) = \prod_{\kappa=1}^{k} \sum_{a_\kappa=0}^{\infty} f\left(p_\kappa^{a_\kappa}\right) = \sum_{(a_1,\ldots,a_k) \in \mathbb{N}_0^k} f\left(p_1^{a_1} \cdot \ldots \cdot p_k^{a_k}\right), \tag{3}$$

wobei zuletzt die absolute Konvergenz der Reihen $\sum f(p^\nu)$ und die Multiplikativität von f ausgenützt wurde. Nach dem Fundamentalsatz der Arithmetik

in 1.5 kommen rechts in (3) unter den $p_1^{a_1} \cdot \ldots \cdot p_k^{a_k}$ alle natürlichen $n \leq x$ vor. Daraus folgt bereits (2); man braucht hierzu nur die Nichtnegativität aller f-Werte auszunützen.

Um den Satz einzusehen, startet man erneut mit (3) und bemerkt, daß dort die Summe rechts die Form

$$\sum_{n=1}^{\infty} f(n) - \sum{}' f(n) \tag{4}$$

hat. Dabei bedeutet $\sum'$, daß die Summation genau über die $n \in \mathbb{N}$ zu erstrecken ist, die von mindestens einer Primzahl, die größer als x ist, geteilt werden; erst recht sind also diese n größer als x. Gibt man sich nun ein $\varepsilon \in \mathbb{R}_+$ beliebig vor, so gilt bei geeignetem $x_0(\varepsilon)$ nach Voraussetzung die Ungleichung $\sum' |f(n)| < \varepsilon$ für alle $x > x_0(\varepsilon)$. Kombination von (3) und (4) liefert (1). □

Bemerkung. Die Gleichung (1) ist nichts anderes als ein analytisches Äquivalent zum Fundamentalsatz der Arithmetik. Offenbar machen beide Beweise *keinen* Gebrauch von EUKLIDs Satz 1.4; man beachte, daß die Summe $\sum'$ in (4) leer wäre, wenn es oberhalb x keine Primzahl mehr gäbe.

4. Riemannsche Zetafunktion. Bei $n \in \mathbb{N}$ und $s \in \mathbb{C}$ definiert man wie üblich die komplexe Zahl n^{-s} durch $\exp(-s \log n)$, wobei log den reellen Logarithmus und (im weiteren) Re s den Realteil von s bedeutet.

Satz. *Ist g eine beschränkte multiplikative zahlentheoretische Funktion, so definiert die Reihe*

$$\sum_{n=1}^{\infty} g(n) n^{-s} \tag{1}$$

in Re $s > 1$ *eine holomorphe Funktion G mit der Produktentwicklung*

$$G(s) = \prod_p \sum_{\nu=0}^{\infty} g(p^\nu) p^{-\nu s} \qquad \text{in Re } s > 1. \tag{2}$$

Beweis. Die durch $f(n) := g(n) n^{-s}$ für $n \in \mathbb{N}$ definierte Funktion $f \in Z$ ist wegen $g \in M$ multiplikativ. Wegen der Beschränktheit von g konvergiert die Reihe (1) in Re $s > 1$ absolut (man beachte $|n^{-s}| = n^{-\mathrm{Re}\ s}$) und kompakt gleichmäßig, definiert daher dort eine holomorphe Funktion G. Formel (2) folgt direkt aus 3(1). □

Von besonderer Bedeutung ist der Spezialfall $g = \iota_0$ des vorstehenden Satzes. Hier setzt man

$$\zeta(s) := \sum_{n=1}^{\infty} n^{-s} \qquad \text{für Re } s > 1 \tag{3}$$

und hat nach (2) für dieselben $s \in \mathbb{C}$ die Produktformel

$$\zeta(s) = \prod_{p} (1 - p^{-s})^{-1}. \tag{4}$$

Dabei ist lediglich $\sum_{\nu \geq 0} p^{-\nu s}$ für jede Primzahl p als geometrische Reihe aufsummiert worden.

Die hier eingeführte, in der Halbebene Re $s > 1$ holomorphe und (wegen (4)) nullstellenfreie Funktion ζ heißt RIEMANN*sche Zetafunktion*. Sie spielt beim Studium der Primzahlverteilung eine überragende Rolle (vgl. 7.3.1); daß sie mit den Primzahlen engstens zu tun hat, ist aus ihrer Produktdarstellung (4) evident.

5. Zweimal Euklids Satz über die Existenz unendlich vieler Primzahlen.

Eine mehr amüsante Beweisvariante. Für jedes ganze $s \geq 2$ ist jeder Faktor rechts in 4(4) rational. Unter der Annahme, es gäbe nur endlich viele Primzahlen, wäre dann das Produkt rechts in 4(4) rational, also auch $\zeta(s)$. Nun ist aber die Irrationalität (sogar Transzendenz) von $\zeta(2t)$ für alle $t \in \mathbb{N}$ wohlbekannt: Z.B. ist $\zeta(2) = \sum n^{-2} = \frac{1}{6}\pi^2$, allgemeiner $\zeta(2t) = r_t \pi^{2t}$ mit gewissen $r_t \in \mathbb{Q}_+$ und so folgt EUKLIDs Satz aus der Irrationalität von π^{2t} für alle $t \in \mathbb{N}$, vgl. 6.3.2. □

Für die zweite Variante wird ein Teil des folgenden Hilfssatzes über die Partialsummen der harmonischen Reihe benötigt.

Lemma. *Für alle natürlichen n gelten die Ungleichungen*

$$\log(n+1) < \sum_{m=1}^{n} \frac{1}{m} \leq 1 + \log n.$$

Beweis. Durch Vergleich der Summe mit $\int_1^{n+1} \frac{dt}{t}$ bzw. $\int_1^n \frac{dt}{t}$. □

Eine seriöse Beweisvariante. Ohne ein so unangemessen starkes Hilfsmittel wie z.B. die Irrationalität von π^2 für EUKLIDs Satz anzuwenden, wird nun Proposition 3 ausgenützt: Man nimmt dort $f := \iota_{-1}$ und erhält unter Berücksichtigung der linken Hälfte des obigen Lemmas

$$\log x < \log([x]+1) < \sum_{n\le x} \frac{1}{n} \le \prod_{p\le x} (1-\frac{1}{p})^{-1} \tag{1}$$

für alle reellen $x > 0$. Da hier die linke Seite beliebig groß werden kann, muß dies auch für das Produkt rechts gelten, d.h. dieses kann nicht von einer Stelle an konstant sein. □

Bemerkungen. 1) In der ersten Variante hätte man auch $s = 3$ nehmen können; die Irrationalität von $\zeta(3)$ wurde von R. APERY (Astérisque *61*, 11–13 (1979)) bewiesen. Bei $\zeta(5), \zeta(7), \zeta(9), \ldots$ ist das Problem der Irrationalität noch offen; allerdings ist in den letzten Jahren, angestoßen durch T. RIVOAL (C. R. Acad. Sci. Paris Sér. I Math. *331*, 267–279 (2000)), neue "Bewegung" in diese Fragestellung gekommen.

2) Bei der ersten Variante ergibt sich die rein *qualitative* Aussage der Unendlichkeit der Menge $\mathbb{P}$ durch Betrachtung der Funktion ζ an *einer* einzigen geeignet gewählten Stelle. Es ist plausibel, daß Informationen über $\zeta(s)$ auf einer *reichhaltigeren* Menge von Punkten s der komplexen Ebene zu genaueren *quantitativen* Aussagen über das Anwachsen der in 1.4 definierten Anzahlfunktion $\pi(x) := \#\{p \in \mathbb{P} : p \le x\}$ bei $x \to \infty$ führen wird. Dies wird sich in § 3 von Kap. 7 tatsächlich bestätigen.

3) *Eine* derartige quantitative Aussage über $\pi(x)$ läßt sich übrigens aus der zweiten Variante gewinnen: Beachtet man, daß $p \ge 2$ für alle $p \in \mathbb{P}$ gilt, so ist das Produkt rechts in (1) höchstens gleich $2^{\pi(x)}$, so daß (1) unmittelbar zu $\pi(x) > \frac{\log\log x}{\log 2}$ für alle reellen $x > 1$ führt. Verglichen etwa mit der in 7.2.3 angegebenen Abschätzung (4) ist dies allerdings ein sehr schwaches Ergebnis.

4) Beachtet man die vom Leser zu verifizierende, für alle reellen $t \in [0, \frac{1}{2}]$ gültige Ungleichung $(1-t)^{-1} \le 4^t$, so folgt aus (1) noch $\frac{\log\log x}{\log 4} < \sum_{p\le x} \frac{1}{p}$ für alle $x > 1$. Insbesondere divergieren also $\prod_p (1-p^{-1})^{-1}$ und $\sum_p \frac{1}{p}$, was bereits EULER bekannt war, vgl. 14.

6. Faltung. In den Abschnitten 6 bis 11 werden die zahlentheoretischen Funktionen von einem relativ abstrakten, strukturellen Gesichtspunkt aus betrachtet. Dabei werden die Teilerfunktionen τ und σ in allgemeinere Zusammenhänge eingeordnet, bei deren Studium man einigen weiteren klassischen zahlentheoretischen Funktionen begegnen wird.

Für $f, g \in Z$ und $\kappa \in \mathbb{C}$ definiert man neue zahlentheoretische Funktionen $f+g$ bzw. $\kappa \cdot f$ punktweise durch die Festsetzungen

$$(f+g)(n) := f(n) + g(n) \quad \text{bzw.} \quad (\kappa \cdot f)(n) := \kappa f(n)$$

für alle $n \in \mathbb{N}$. Klar ist, daß $\langle Z, + \rangle$ eine abelsche Gruppe bildet, die durch die oben definierte skalare Multiplikation $\cdot$ zu einem $\mathbb{C}$-Vektorraum wird.

Eine weitere Verknüpfung in Z, die in der Zahlentheorie von größter Bedeutung ist, wird nun folgendermaßen definiert: Bei $f, g \in Z$ sei

$$(f * g)(n) := \sum_{d|n} f(\frac{n}{d}) g(d) \qquad \text{für alle } n \in \mathbb{N}; \tag{1}$$

dabei bedeutet die Bedingung $d|n$ stets, daß über *alle* positiven Teiler d von n zu summieren ist, also 1 und n eingeschlossen. Die durch (1) eingeführte zahlentheoretische Funktion $f * g$ heißt die *Faltung von f mit g*.

Man beachte sogleich, daß sich die Summe rechts in (1) in die symmetrische Form

$$\sum_{\substack{(c,d) \in \mathbb{N}^2 \\ cd=n}} f(c) g(d) \tag{2}$$

setzen läßt. Damit gestalten sich manche Rechnungen mit der Faltung zweier Funktionen bequemer; z.B. ist (iii) der folgenden Proposition dann evident.

Proposition. *Für beliebige $f, g, h \in Z$ gilt*
(i) $(f * g) * h = f * (g * h)$,
(ii) $f * \varepsilon = f$,
(iii) $f * g = g * f$.

Beweis. Für (i) arbeitet man bei $n \in \mathbb{N}$ die Summe

$$\Sigma(n) := \sum_{\substack{(b,c,d) \in \mathbb{N}^3 \\ bcd=n}} f(b) g(c) h(d)$$

unter Beachtung von (1) und (2) zunächst folgendermaßen um

$$\begin{aligned} \Sigma(n) &= \sum_{\substack{(a,d) \in \mathbb{N}^2 \\ ad=n}} h(d) \sum_{\substack{(b,c) \in \mathbb{N}^2 \\ bc=a}} f(b) g(c) = \sum_{\substack{(a,d) \in \mathbb{N}^2 \\ ad=n}} (f * g)(a) h(d) \\ &= ((f * g) * h)(n). \end{aligned}$$

Offensichtlich kann man genauso gut $\Sigma(n) = (f * (g * h))(n)$ erhalten, was (i) beweist.

Zu (ii): Berechnet man $(f * \varepsilon)(n)$ gemäß (1), so bleibt in der dortigen Summe wegen $\varepsilon(d) = 0$ für $d > 1$ alleine der Summand $f(n)\varepsilon(1) = f(n)$ zurück. □

Regel (i) macht Klammersetzung bei Faltung von beliebig (aber endlich) vielen zahlentheoretischen Funktionen überflüssig. Regel (iii) erlaubt es, von der Faltung von f *und* g zu sprechen. Insgesamt beinhaltet die Proposition, daß $\langle Z, * \rangle$ *eine kommutative Halbgruppe* bildet; die Rolle des neutralen Elements spielt ε.

Bezüglich der *beiden* oben eingeführten Verknüpfungen $+$ und $*$ gilt nun der

Satz. $\langle Z, +, * \rangle$ *ist ein Integritätsring.*

Beweis. Nach den anfänglichen Feststellungen über $\langle Z, + \rangle$ und wegen der Proposition ist nur noch Distributivgesetz und Nullteilerfreiheit zu prüfen, wobei das erstere dem Leser überlassen sei. Für die Nullteilerfreiheit ist $f * g \neq \mathbf{0}$ bei $f, g \in Z \setminus \{\mathbf{0}\}$ zu zeigen.

Seien c_0 bzw. d_0 aus $\mathbb{N}$ jeweils kleinstmöglich gewählt mit $f(c_0) \neq 0$, $g(d_0) \neq 0$; dann ist $(f * g)(c_0 d_0) = f(c_0)g(d_0) \neq 0$, also $f * g \neq \mathbf{0}$. Für $n = c_0 d_0$ reduziert sich die Summe (1) nämlich auf $f(c_0)g(d_0)$, weil $g(d) = 0$ für $d < d_0$ und $f\left(\frac{c_0 d_0}{d}\right) = 0$ für $d > d_0$ gilt (beachte hier $\frac{c_0 d_0}{d} < c_0$). □

7. Inverse bezüglich Faltung. Während sich $\langle Z, * \rangle$ in 6 als kommutative Halbgruppe herausstellte, klärt der nächste Satz, genau wann ein $f \in Z$ bezüglich $*$ eine Inverse hat.

Satz. *Für $f \in Z$ sind folgende Aussagen äquivalent:*
(i) $f(1) \neq 0$.
(ii) Es gibt ein $g \in Z$ mit $f * g = \varepsilon$.

Bemerkung. Für g wie in (ii) gilt $g(1) = \frac{1}{f(1)}$, also $g(1) \neq 0$.

Beweis. Sei zunächst $f(1) \neq 0$; dann wird *ein* g wie in (ii) punktweise folgendermaßen rekursiv konstruiert. Man setzt erst $g(1) := \frac{1}{f(1)}$ und hat damit $(f * g)(1) = 1\,(= \varepsilon(1))$. Ist dann $n > 1$ und $g(d)$ schon für $d = 1, \ldots, n-1$ definiert, so setzt man

$$g(n) := -\frac{1}{f(1)} \sum_{\substack{d \mid n \\ d \neq n}} f(\frac{n}{d}) g(d);$$

nach 6(1) ist dies ja mit $(f * g)(n) = 0\,(= \varepsilon(n))$ äquivalent.

Ist andererseits (ii) vorausgesetzt, so ist nach 6(1) insbesondere $f(1)g(1) = 1$, also $f(1) \neq 0$. □

Genügt $f \in Z$ der Bedingung $f(1) \neq 0$ und haben $g, h \in Z$ die Eigenschaft $f * g = \varepsilon$, $f * h = \varepsilon$, so ist $g = g * \varepsilon = (g * f) * h = \varepsilon * h = h$ nach Proposition 6. Somit existiert zu jedem $f \in Z$ mit $f(1) \neq 0$ bezüglich $*$ genau eine *Inverse* in Z, die üblicherweise mit $\check{f}$ bezeichnet wird und für die $\check{f}(1) = \frac{1}{f(1)} \neq 0$ gilt.

Der obige Satz legt die Einführung folgender Abkürzung nahe:

$$Z_1 := \{f \in Z : f(1) \neq 0\}.$$

Offenbar ist Z_1 bezüglich $*$ abgeschlossen; man hat ja nur $(f*g)(1) = f(1)g(1) \neq 0$ für $f, g \in Z_1$ zu beachten. Diese Beobachtung in Verbindung mit Proposition 6 und obigem Satz führt unmittelbar zum

Korollar. *$\langle Z_1, * \rangle$ ist eine abelsche Gruppe.*

8. Die Gruppe der multiplikativen Funktionen. Vorausgeschickt sei hier das einfache

Lemma. *Für teilerfremde $n_1, n_2 \in \mathbb{N}$ gelte $d | n_1 n_2$ mit einem $d \in \mathbb{N}$. Dann gibt es ein eindeutig bestimmtes Paar (d_1, d_2) natürlicher Zahlen mit $d_1 d_2 = d$, $d_1 | n_1$, $d_2 | n_2$, wobei außerdem d_1, d_2 zueinander teilerfremd sind ebenso wie $\frac{n_1}{d_1}$, $\frac{n_2}{d_2}$.*

Beweis. Nach Lemma 1.7 ist $\nu_p(d) \leq \nu_p(n_1 n_2)$ für alle $p \in \mathbb{P}$, insbesondere $\nu_p(d) = 0$ für $p \nmid n_1 n_2$. Bei $p | n_1 n_2$ ist nach Satz 2.7 genau ein $\nu_p(n_i)$ Null und man hat entweder $\nu_p(d) \leq \nu_p(n_1)$, $\nu_p(n_2) = 0$ oder $\nu_p(d) \leq \nu_p(n_2)$, $\nu_p(n_1) = 0$, also

$$d = \prod_{p | n_1 n_2} p^{\nu_p(d)} = \prod_{p | n_1} p^{\nu_p(d)} \cdot \prod_{p | n_2} p^{\nu_p(d)} =: d_1 \cdot d_2.$$

Aus der Definition ist $d_i | n_i$ für $i = 1, 2$ klar. Haben d_1', d_2' dieselben Eigenschaften wie d_1, d_2 im Lemma, so folgt aus $d_1 d_2 = d_1' d_2'$ wegen der Teilerfremdheit von d_1,d_2' und wegen Satz 2.6(i) die Bedingung $d_1 | d_1'$, also $d_1' = \delta d_1$, $d_2 = \delta d_2'$ mit ganzem $\delta > 0$. Wegen der Teilerfremdheit von n_1, n_2 muß $\delta = 1$ gelten. □

Proposition. *Die Menge M der multiplikativen zahlentheoretischen Funktionen ist bezüglich Faltung abgeschlossen, ebenso die Menge $M \setminus \{\mathbf{0}\}$.*

Beweis. Seien $f, g \in M$ und $n_1, n_2 \in \mathbb{N}$ zueinander teilerfremd. Dann ist nach 6(1), dem vorangestellten Lemma und der vorausgesetzten Multiplikativität von f, g

$$(f*g)(n_1n_2) = \sum_{d\mid n_1n_2} f\left(\frac{n_1n_2}{d}\right) g(d) = \sum_{\substack{(d_1,d_2)\in\mathbb{N}^2\\ d_1\mid n_1,\, d_2\mid n_2}} f\left(\frac{n_1n_2}{d_1d_2}\right) g(d_1d_2)$$

$$= \sum_{d_1\mid n_1}\sum_{d_2\mid n_2} f\left(\frac{n_1}{d_1}\right) g(d_1) f\left(\frac{n_2}{d_2}\right) g(d_2) = \prod_{j=1}^{2}\sum_{d_j\mid n_j} f\left(\frac{n_j}{d_j}\right) g(d_j)$$

$$= \prod_{j=1}^{2} (f*g)(n_j),$$

was $f*g \in M$ beweist. Sind insbesondere $f, g \in M\setminus\{\mathbf{0}\}$, so ist $f(1) = 1$, $g(1) = 1$ nach Proposition 2(i) und somit $(f*g)(1) = f(1)g(1) = 1$, also hat man $f*g \in M\setminus\{\mathbf{0}\}$. □

Wie soeben schon festgestellt, gilt $f(1) = 1$ für $f \in M\setminus\{\mathbf{0}\}$, insbesondere also $f \in Z_1$. Nun wird über die Teilmenge $M\setminus\{\mathbf{0}\}$ von Z_1 behauptet der

Satz. *$\langle M\setminus\{\mathbf{0}\}, *\rangle$ ist eine Untergruppe der Gruppe $\langle Z_1, *\rangle$.*

Beweis. Wegen Korollar 7 und obiger Proposition bleibt nur noch zu zeigen, daß $M\setminus\{\mathbf{0}\}$ zu jedem Element f auch die Inverse $\check{f}$ enthält. Klar ist zunächst $f \in Z_1$ und also hat f nach 7 *in* Z_1 die Inverse $\check{f}$, die als in M enthalten erkannt werden muß. Dazu definiert man $g \in Z$ vermöge

$$g(n) := \prod_{p\mid n} \check{f}\left(p^{\nu_p(n)}\right). \tag{1}$$

Offenbar ist $g(1) = 1$ und $g \in M$, also $g \in M\setminus\{\mathbf{0}\}$ und somit $f*g \in M\setminus\{\mathbf{0}\}$ nach obiger Proposition. Wenn $f*g = \varepsilon$ nachgewiesen ist, folgt $\check{f} = g$ aus $f*\check{f} = \varepsilon$ und somit $\check{f} \in M\setminus\{\mathbf{0}\}$ wie gewünscht. Für $f*g = \varepsilon$ reicht nach der Anmerkung am Ende von 2 der Nachweis, daß die beiden multiplikativen Funktionen $f*g$ und ε auf den Primzahlpotenzen übereinstimmen. Tatsächlich ist für alle $j \in \mathbb{N}_0$, $p \in \mathbb{P}$ nach (1)

$$(f*g)(p^j) = \sum_{i=0}^{j} f(p^{j-i})g(p^i) = \sum_{i=0}^{j} f(p^{j-i})\check{f}(p^i) = \varepsilon(p^j). \qquad \square$$

Ist $f \in Z$ und ι_0 gemäß 1 definiert, so heißt $Sf := \iota_0 * f$ die *summatorische Funktion* von f. Wegen $\iota_0(d) = 1$ für alle $d \in \mathbb{N}$ ist also nach 6(1)

$$(Sf)(n) = \sum_{d\mid n} f(d) \qquad \text{für alle } n \in \mathbb{N},$$

was die Bezeichnungsweise verständlich macht. Damit ergeben die bisherigen Schlußweisen noch folgendes

Korollar. *Für $f \in Z$ gilt die Äquivalenz*

$$f \in M \iff Sf \in M.$$

Beweis. Ist $f \in M$, so folgt aus $\iota_0 \in M$ (vgl. 2) und obiger Proposition $Sf = \iota_0 * f \in M$. Wegen $\iota_0 \in M \setminus \{\mathbf{0}\}$ und dem Satz ist $\check{\iota}_0 \in M \setminus \{\mathbf{0}\}$; setzt man jetzt $Sf \in M$ voraus, so ist $f = \check{\iota}_0 * (\iota_0 * f) = \check{\iota}_0 * Sf \in M$, wieder nach der Proposition. □

Bemerkung. Bei $f \in Z$ gilt offenbar auch $f \in M \setminus \{\mathbf{0}\} \Leftrightarrow Sf \in M \setminus \{\mathbf{0}\}$.

9. Möbiussche Müfunktion. Die im Beweis von Korollar 8 aufgetretene Funktion $\check{\iota}_0$ ist in der Zahlentheorie sehr wichtig; man bezeichnet sie üblicherweise mit μ und nennt sie die MÖBIUS*sche Müfunktion.* Ihre wichtigsten Eigenschaften seien zusammengestellt als

Satz. *Über die* MÖBIUS*sche Funktion μ hat man folgende Aussagen:*

(i) *μ ist multiplikativ.*

(ii) *Für jede Prinzahl p und jedes ganze $j \geq 2$ ist $\mu(p) = -1$, $\mu(p^j) = 0$.*

(iii) *Es ist* $\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{für } n = 1, \\ 0 & \text{für alle ganzen } n > 1. \end{cases}$

(iv) *Ist f eine zahlentheoretische Funktion und F ihre summatorische Funktion, so gilt für alle natürlichen n*

$$f(n) = \sum_{d|n} \mu(\frac{n}{d}) F(d).$$

Beweis. Im Beweis von Korollar 8 wurde $\mu := \check{\iota}_0 \in M \setminus \{\mathbf{0}\}$ erledigt, weshalb (i) gilt. Nach Proposition 2(i) ist insbesondere $\mu(1) = 1$.

Zu (ii): Wegen $\varepsilon = \mu * \iota_0$ und den Definitionen von ε bzw. ι_0 ist mit 6(1)

$$\sum_{i=0}^{j} \mu(p^i) = \varepsilon(p^j) = 0 \tag{1}$$

für alle $j \in \mathbb{N}$, $p \in \mathbb{P}$. Wegen $\mu(1) = 1$ ist $\mu(p) = -1$, wenn man (1) mit $j = 1$ anwendet. Damit folgt $\sum_{i=2}^{j} \mu(p^i) = 0$ für $j = 2, 3, \ldots$ aus (1), was zu $\mu(p^j) = 0$ für die eben genannten j führt. (iii) ist eine ausführliche Version der Gleichung $\iota_0 * \mu = \varepsilon$.

Zu (iv): Da F für Sf steht, ist $F = \iota_0 * f$ nach Definition von Sf, also $\mu * F = (\mu * \iota_0) * f = \varepsilon * f = f$. Die Formel in (iv) besagt dasselbe wie $\mu * F = f$. □

Bemerkungen. 1) Kombination von (i) und (ii) des Satzes zeigt $\mu(\mathbb{N}) = \{-1, 0, 1\}$.

2) Eine ganze Zahl n heißt *quadratfrei*, wenn $p^2 \not| n$ für alle Primzahlen p gilt. Danach ist 1 quadratfrei, 0 jedoch nicht. Der Leser möge sich für $n \in \mathbb{N}$ die folgenden Äquivalenzen überlegen:

$$n \text{ ist quadratfrei} \quad \Leftrightarrow \quad (\mu(n))^2 = 1 \quad \Leftrightarrow \quad \mu(n) \neq 0.$$

3) Die Formel in (iv) heißt MÖBIUS*sche Umkehrformel*. In ihr ist einer der Gründe für die zahlentheoretische Bedeutung der μ-Funktion zu sehen: Durch die Umkehrformel gelingt die Rückgewinnung der ursprünglichen Funktion aus ihrer summatorischen Funktion.

4) Nach Bemerkung 1 gilt mit $M(x) := \sum_{n \leq x} \mu(n)$ trivialerweise die Ungleichung $|M(x)| < x$ für alle reellen $x > 1$. In einem Brief vom 11. Juli 1885 an C. HERMITE kündigte T.J. STIELTJES an, er habe einen Beweis dafür, daß $|M(x)|x^{-1/2}$ bei $x \to \infty$ beschränkt bleibt; in Klammern fügte er an, man könne wohl 1 als Schranke nehmen. Da er keinen Beweis für seine Behauptungen veröffentlichte, diese aber für die Untersuchungen der analytischen Eigenschaften der RIEMANNschen Zetafunktion aus 4 weitreichende Konsequenzen gehabt hätten, beschäftigten sich Ende des vorigen Jahrhunderts viele Mathematiker mit der Funktion $M(x)$. Insbesondere veröffentlichte F. MERTENS (Sitz.–Ber. Akad. Wiss. Wien IIa *106*, 761–830 (1897)) eine Tabelle der Werte $\mu(n)$, $M(n)$ für $n = 1, \ldots, 10000$, aufgrund deren er schloß, die Ungleichung

$$|M(x)| < x^{1/2} \quad \text{für } x > 1 \tag{2}$$

sei "sehr wahrscheinlich". Diese als MERTENS*sche Vermutung* in die Literatur eingegangene Behauptung wurde von A.M. ODLYZKO und H.J.J. TE RIELE (J. Reine Angew. Math. *357*, 138–160 (1985)) widerlegt. Ihr indirekter Beweis liefert allerdings kein x_0, für das (2) falsch ist; sie erwarten solche x_0 nicht unterhalb 10^{20}. J. PINTZ (Astérisque *147/148*, 325–333 (1987)) hat die Existenz solcher x_0 unterhalb $\exp(3,21 \cdot 10^{64})$ bewiesen.

10. Weitere spezielle multiplikative Funktionen. Ist ι_α bei reellem α wie in 1 definiert, so setzt man nun $\sigma_\alpha := \iota_0 * \iota_\alpha$ oder ausführlicher

$$\sigma_\alpha(n) := \sum_{d|n} d^\alpha \qquad \text{für } n \in \mathbb{N}. \tag{1}$$

Wegen $\iota_0, \iota_\alpha \in M \setminus \{\mathbf{0}\}$ gilt $\sigma_\alpha \in M \setminus \{\mathbf{0}\}$ nach Proposition 8. Wegen $(\sigma_\alpha(2))^2 = (1+2^\alpha)^2 = 1+2^{\alpha+1}+4^\alpha \neq 1+2^\alpha+4^\alpha = \sigma_\alpha(4)$ ist kein σ_α *streng* multiplikativ.

Wie die folgende Formel lehrt, gehen σ_α und $\sigma_{-\alpha}$ in einfacher Weise auseinander hervor:

$$\sigma_{-\alpha}(n) = \sum_{d|n} d^{-\alpha} = n^{-\alpha} \sum_{d|n} \left(\frac{n}{d}\right)^\alpha = n^{-\alpha}(\iota_\alpha * \iota_0)(n) = n^{-\alpha}\sigma_\alpha(n).$$

Insbesondere ist $\sigma_0(n) = \#\{d \in \mathbb{N} : d|n\} = \tau(n)$ für alle $n \in \mathbb{N}$, wenn man (1) mit $\alpha = 0$ und die Definition von τ in 1.3 anwendet. Weiter ist $\sigma_1(n) = \sum_{d|n} d = \sigma(n)$ nach (1) und der Definition von σ in 1.7. *Die Funktionen σ_α verallgemeinern also die früheren τ, σ.*

Weiter definiert man bei beliebigem reellem α die zahlentheoretische Funktion ψ_α implizit durch die Forderung $S\psi_\alpha = \iota_\alpha$, d.h.

$$\iota_0 * \psi_\alpha = \iota_\alpha. \tag{2}$$

Nach Definition der MÖBIUS–Funktion ist dies mit

$$\psi_\alpha = \mu * \iota_\alpha \tag{3}$$

äquivalent; insbesondere ist $\psi_0 = \varepsilon$. Wegen $\mu, \iota_\alpha \in M \setminus \{\mathbf{0}\}$ ist auch $\psi_\alpha \in M \setminus \{\mathbf{0}\}$ für jedes α. *Streng* multiplikativ ist ψ_α genau dann, wenn $\alpha = 0$ ist. Ausführlicher als in (2) bzw. (3) hat man

$$\sum_{d|n} \psi_\alpha(d) = n^\alpha \quad \text{bzw.} \quad \psi_\alpha(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) d^\alpha \tag{4}$$

für alle $n \in \mathbb{N}$. Die Bedeutung der speziellen ψ_α mit $\alpha \in \mathbb{N}$ wird in 11 vollständig aufgeklärt.

Bemerkung. Multiplikative zahlentheoretische Funktionen treten bei ganz unterschiedlichen Fragestellungen in natürlicher Weise auf. So z.B. ψ_1 $(= \varphi_1 =: \varphi$, vgl. 11) in 2.3.4ff ebenso wie in Kap. 2, § 5 und Kap. 3, § 1, weiterhin ρ_f in 2.4.2ff, δ, Δ und σ_u in 4.1.8.

11. Eulers Phifunktion und Verallgemeinerungen. Bei festem $\alpha \in \mathbb{N}$ wird für alle natürlichen n gesetzt

$$(1) \qquad \varphi_\alpha(n) := \#\{(\ell_1, \dots, \ell_\alpha) \in \{1, \dots, n\}^\alpha : \ \mathrm{ggT}(\ell_1, \dots, \ell_\alpha, n) = 1\}.$$

Insbesondere ist $\varphi_1(n)$ die Anzahl der natürlichen, n nicht übersteigenden Zahlen, die zu n teilerfremd sind; üblicherweise schreibt man kürzer $\varphi := \varphi_1$ und bezeichnet dies als EULERsche *Phifunktion*. Die allgemeinen φ_α aus (1) scheinen erstmals von C. JORDAN eingeführt worden zu sein.

Über den Zusammenhang der φ_α mit den ψ_α aus 10(3) gibt Auskunft folgende

Proposition. *Es ist $\varphi_\alpha = \psi_\alpha$ für alle natürlichen α.*

Beweis. Bei $n \in \mathbb{N}$ ist trivialerweise stets $\mathrm{ggT}(\ell_1, \dots, \ell_\alpha, n) | n$. Es werde nun ein beliebiges $d \in \mathbb{N}$ mit $d|n$ festgehalten. Hat der Vektor

$$(2) \qquad (\ell_1, \dots, \ell_\alpha) \in \{1, \dots, n\}^\alpha \text{ die Eigenschaft } \mathrm{ggT}(\ell_1, \dots, \ell_\alpha, n) = d$$

und definiert man ganze ℓ'_i durch $\ell'_i := \frac{1}{d}\ell_i$ für $i = 1, \dots, \alpha$, so hat der Vektor

$$(3) \qquad (\ell'_1, \dots, \ell'_\alpha) \in \{1, \dots, \frac{n}{d}\}^\alpha \text{ die Eigenschaft } \mathrm{ggT}(\ell'_1, \dots, \ell'_\alpha, \frac{n}{d}) = 1,$$

vgl. Proposition 2.5(vii). Gilt umgekehrt (3) für einen Vektor $(\ell'_1, \dots, \ell'_\alpha)$ und setzt man $\ell_i := d\ell'_i$ für $i = 1, \dots, \alpha$, so genügt der Vektor $(\ell_1, \dots, \ell_\alpha)$ den Bedingungen (2). Die Anzahl der α–Tupel, die bei (3) gezählt werden, ist $\varphi_\alpha\left(\frac{n}{d}\right)$ und daher ist $\sum_{d|n} \varphi_\alpha\left(\frac{n}{d}\right)$ die Anzahl aller $(\ell_1, \dots, \ell_\alpha) \in \{1, \dots, n\}^\alpha$, also n^α. Man hat also $\varphi_\alpha * \iota_0 = \iota_\alpha$ gefunden, woraus mit 10(2) die Behauptung folgt. □

Die wichtigsten Eigenschaften der φ_α seien zusammengestellt als

Satz. *Für alle* JORDANschen *Verallgemeinerungen φ_α, $\alpha = 1, 2, \dots$, der* EULERschen *Funktion $\varphi = \varphi_1$ hat man folgende Aussagen:*
(i) *φ_α ist multiplikativ.*
(ii) *Für jede Primzahl p und jedes natürliche j ist $\varphi_\alpha(p^j) = p^{(j-1)\alpha}(p^\alpha - 1)$.*
(iii) *Es ist $\sum_{d|n} \varphi_\alpha(d) = n^\alpha$ für alle natürlichen n.*
(iv) *Es ist $\sum_{d|n} \mu\left(\frac{n}{d}\right) d^\alpha = \varphi_\alpha(n)$ für alle natürlichen n.*

Beweis. Durch Kombination der Proposition und der Resultate in 10 folgen sämtliche vier Aussagen. Für (ii), (iii) und (iv) hat man insbesondere 10(4) auszunützen, wobei man (ii) so einsieht:

$$\varphi_\alpha(p^j) = \psi_\alpha(p^j) = \sum_{i=0}^{j} \mu(p^{j-i})p^{i\alpha} = p^{j\alpha} - p^{(j-1)\alpha}. \qquad \square$$

Einige weitere Eigenschaften der φ_α sind bisweilen von Nutzen und hier zusammengefaßt.

Korollar. *Über die φ_α des obigen Satzes hat man:*

(i) *Für alle natürlichen n gilt $\varphi_\alpha(n) = n^\alpha \prod_{p|n}(1 - p^{-\alpha})$.*

(ii) *Für alle natürlichen n gilt $1 \leq \varphi_\alpha(n) \leq n^\alpha$.*

(iii) *Es gilt die Äquivalenz $\varphi_\alpha(n) = n^\alpha - 1 \Leftrightarrow n \in \mathbb{P}$.*

(iv) *Es gelten die Äquivalenzen*

$$2 \not| \varphi_\alpha(n) \Leftrightarrow n \in \{1, 2\} \Leftrightarrow \varphi_\alpha(n) \text{ ist } 1 \text{ oder } 2^\alpha - 1.$$

Beweis. Kombination von (i) und (ii) des Satzes liefert (i), was dann auch (ii) sowie die Implikation $\Leftarrow$ von (iii) nach sich zieht. Sei umgekehrt $\varphi_\alpha(n) = n^\alpha - 1$; wegen $\varphi_\alpha(1) = 1$ ist sicher $n > 1$. Wäre n zusammengesetzt und etwa p eine in n aufgehende Primzahl, so würden mindestens die beiden verschiedenen α–Tupel $(p, \ldots, p)$ und $(n, \ldots, n)$ bei (1) nicht gezählt, so daß $\varphi_\alpha(n) \leq n^\alpha - 2$ wäre.

Zu (iv): Sei $2 \not| \varphi_\alpha(n)$. Ist $n = 2^j$, so $j \in \{0, 1\}$ wegen (ii) des Satzes; würde n von einer ungeraden Primzahl p geteilt, so wäre $(p^\alpha - 1) | \varphi_\alpha(n)$ wegen (i) und (ii) des Satzes, also $\varphi_\alpha(n)$ gerade. Insgesamt ist $n \in \{1, 2\}$. Der Rest ist trivial einzusehen. □

Ein weiterer *Beweis des* EUKLID*schen Satzes* 1.4. Es werde $\mathbb{P}$ endlich, etwa $\mathbb{P} = \{p_1, \ldots, p_r\}$ angenommen. Bildet man $n := p_1 \cdot \ldots \cdot p_r$, so gilt $n > 2$ wegen $2, 3 \in \mathbb{P}$. Andererseits ist für jedes $m \in \{2, \ldots, n\}$ das in 1.3 eingeführte $p(m)$ eine Primzahl, d.h. aus $\{p_1, \ldots, p_r\}$, und so sind m, n nicht teilerfremd. Deswegen ist $\varphi(n) = 1$, was nach (iv) des Korollars zu $n \leq 2$ äquivalent ist. Der erhaltene Widerspruch beweist EUKLIDs Satz aufs neue. □

12. Eine Aussage "im Mittel". Um den nächsten Satz ebenso wie asymptotische Aussagen an späteren Stellen bequem formulieren zu können, wird nun eine abkürzende Schreibweise eingeführt.

Sind f, g reellwertige Funktionen einer Variablen, die für alle genügend großen reellen Argumentwerte definiert sind und ist überdies g positiv, so schreibt man

(i) $f(x) = O(g(x))$ bei $x \to \infty$, falls $\frac{f(x)}{g(x)}$ bei diesem Grenzübergang beschränkt bleibt;

(ii) $f(x) = o(g(x))$ bei $x \to \infty$, falls $\lim\limits_{x \to \infty} \frac{f(x)}{g(x)}$ existiert und Null ist;

(iii) $f(x) \sim g(x)$ bei $x \to \infty$, falls $\lim\limits_{x \to \infty} \frac{f(x)}{g(x)}$ existiert und Eins ist.

Man liest (i) bzw. (ii) als "$f(x)$ ist *groß–oh* bzw. *klein–oh* von $g(x)$", (iii) als "$f(x)$ ist *asymptotisch gleich* $g(x)$". Klar ist, daß sowohl (ii) als auch (iii) einzeln (i) implizieren. Weiter ist evident, daß $\sim$ auf der Menge der für alle großen

Argumente definierten, positivwertigen Funktionen eine Äquivalenzrelation definiert; daher sagt man auch "$f(x)$ und $g(x)$ sind asymptotisch gleich". Die Notationen (i) und (ii) gehen auf P. Bachmann und E. Landau zurück, (iii) scheint systematisch zuerst von G.H. Hardy und J.E. Littlewood benutzt worden zu sein.

Satz. *Es gilt* $\sum_{n\le x}\varphi(n) = 3\pi^{-2}x^2 + O(x\log x)$ *bei* $x\to\infty$.

Beweis. Wegen 10(4) und Proposition 11, jeweils mit $\alpha = 1$ angewandt, gilt für $N := [x]$

$$\Phi(N) := \sum_{n=1}^{N}\varphi(n) = \sum_{n=1}^{N}\sum_{\substack{(c,d)\in\mathbb{N}^2\\ cd=n}}\mu(c)d = \sum_{\substack{(c,d)\in\mathbb{N}^2\\ cd\le N}}\mu(c)d$$

$$= \sum_{c=1}^{N}\mu(c)\sum_{d=1}^{[\frac{N}{c}]}d = \frac{1}{2}\sum_{c=1}^{N}\mu(c)\left[\frac{N}{c}\right]\left(\left[\frac{N}{c}\right]+1\right) \tag{1}$$

$$= \frac{1}{2}\sum_{c=1}^{N}\mu(c)\left(\frac{N}{c}-\vartheta\right)\left(\frac{N}{c}+1-\vartheta\right) =: \frac{1}{2}N^2\sum_{c=1}^{N}\frac{\mu(c)}{c^2} + R(N).$$

Dabei sind die rationalen Zahlen $\vartheta := \frac{N}{c} - [\frac{N}{c}]$ zwar von N und c abhängig, aber stets aus dem Intervall $[0,1[$. Wegen $|\mu(c)|\le 1$, $|1-2\vartheta|\le 1$, $|\vartheta(1-\vartheta)|\le\frac{1}{4}$ für alle $c, N\in\mathbb{N}$ mit $c\le N$ hat man für das "Restglied" $R(N)$

$$2|R(N)| = |N\sum_{c=1}^{N}\frac{\mu(c)}{c}(1-2\vartheta) - \sum_{c=1}^{N}\mu(c)\vartheta(1-\vartheta)| \le N\sum_{c=1}^{N}\frac{1}{c} + \frac{1}{4}N.$$

Damit führt die rechte Hälfte von Lemma 5 zu

$$|R(N)| \le \frac{1}{2}N\log N + \frac{5}{8}N. \tag{2}$$

Aus der Gleichungskette (1) folgt

$$\Phi(N) - \frac{1}{2}N^2\sum_{c=1}^{\infty}\frac{\mu(c)}{c^2} = R(N) - \frac{1}{2}N^2\sum_{c>N}\frac{\mu(c)}{c^2}. \tag{3}$$

Um die unendliche Reihe links in (3) auszuwerten, wendet man Satz 4 an mit $g := \mu$ (vgl. auch Bemerkung 1 in 9) und erhält für ihren Wert

$$\prod_p(1-p^{-2}) = \frac{1}{\zeta(2)} = 6\pi^{-2},$$

wobei man noch Satz 9(ii) ebenso wie 4(4) beachtet hat. Für die Summe rechts in (3) gilt

$$|\sum_{c>N} \frac{\mu(c)}{c^2}| < \sum_{c>N} \frac{1}{c^2} < \sum_{c>N} \left(\frac{1}{c-1} - \frac{1}{c}\right) = \frac{1}{N}. \tag{4}$$

Nach Definition von N ist $\Phi(N)$ genau die Summe im Satz und aus (3) folgt mittels (2) und (4)

$$\begin{aligned} |\sum_{n\le x} \varphi(n) - 3\pi^{-2}x^2| &\le 3\pi^{-2}|x^2 - N^2| + \frac{1}{2}\,N\log N + \frac{9}{8}\,N \\ &\le \left(6\pi^{-2} + \frac{9}{8}\right) x + \frac{1}{2}\,x\log x, \end{aligned}$$

wenn man $N = [x]$, also $0 \le x - N < 1$ beachtet. Die letzte Ungleichungskette gibt nach der Bachmann–Landauschen Konvention (i) den Satz. □

Bemerkung. Der Satz besagt offenbar, daß die Werte von φ, über einen "langen" Anfangsabschnitt $\{1, 2, \dots, N\}$ der natürlichen Zahlen gemittelt, in der Größenordnung $3\pi^{-2}N$ liegen. Mit der hier gegebenen Genauigkeit $O(x\log x)$ des Restglieds wurde er zuerst bewiesen von Mertens (J. Reine Angew. Math. *77*, 289– 291 (1874)).

Selbstverständlich wurden auch andere zahlentheoretische Funktionen in analoger Weise auf ihre Größenordnung "im Mittel" untersucht, worauf hier jedoch nicht eingegangen werden kann.

13. Wahrscheinlichkeit für Teilerfremdheit. Als Folgerung aus Satz 12 sei noch abgeleitet das

Korollar. *Die Wahrscheinlichkeit dafür, daß zwei zufällig gewählte natürliche Zahlen zueinander teilerfremd sind, beträgt* $6\pi^{-2} = 0,6079\dots$.

Beweis. Für das in 12(1) eingeführte Φ gilt nach Definition von φ in 11(1)

$$\begin{aligned} \Phi(N) = \sum_{n=1}^{N} \sum_{\substack{d=1 \\ (d,n)=1}}^{n} 1 &= \#\{(d,n) \in \mathbb{N}^2 :\ d \le n \le N,\ \mathrm{ggT}(d,n) = 1\} \\ &= \frac{1}{2}(1 + \#\{(d,n) \in \mathbb{N}^2 :\ d, n \le N,\ \mathrm{ggT}(d,n) = 1\}). \end{aligned}$$

Mit Satz 12 folgt hieraus

$$\frac{\#\{(d,n)\in\mathbb{N}^2 : d,n\le N,\ \mathrm{ggT}(d,n)=1\}}{\#\{(d,n)\in\mathbb{N}^2 : d,n\le N\}} \longrightarrow 6\pi^{-2}$$

bei $N\to\infty$; im Nenner links steht ja genau N^2. □

Bemerkung. Die Aussage des Korollars wurde gefunden von E. CESARO (Mathesis *1*, 184 (1881)) und 1883 von J.J. SYLVESTER (Collected Papers III, 672–676; IV, 84–87). Jedoch scheint sie schon 1849, wenn auch auf etwas anderem Wege, von P.G.L. DIRICHLET (Werke II, 51-66) entdeckt worden zu sein.

14. Historische Anmerkungen. B. RIEMANN schlug 1859 in seiner berühmt gewordenen Arbeit *Ueber die Anzahl der Primzahlen unter einer gegebenen Größe* (Werke, 136–144) vor, das genaue Verhalten der in 1.4 eingeführten Funktion $\pi(x)$ für große x durch Untersuchung der analytischen Eigenschaften der komplexen Funktion ζ in 4(3) zu studieren. Dieser Vorschlag erwies sich in der Folgezeit als überaus fruchtbar und führte 1896 zum Beweis des *Primzahlsatzes* $\pi(x)\sim\frac{x}{\log x}$, in dem Kap. 7 gipfeln wird. RIEMANN zu Ehren trägt die Funktion 4(3) seinen Namen.

Doch die Geschichte der Zetafunktion begann rund 125 Jahre vor RIEMANN. Sowohl P. MENGOLI (*Novae quadraturae arithmeticae*, Bologna, 1650) als auch J. WALLIS (*Arithmetica infinitorum*, Oxford, 1655) hatten das Problem gestellt, den Wert der Reihe $\sum_{n\ge1} n^{-2}$ (also $\zeta(2)$) zu berechnen. G.W. LEIBNIZ ebenso wie die älteren BERNOULLI–Brüder konnten ab 1670 nur Näherungswerte angeben, die später von D. BERNOULLI und C. GOLDBACH (1728), J. STIRLING (1730) und EULER (1731) sukzessive verbessert wurden. 1734 gelang dann EULER der Nachweis von $\zeta(2)=\frac{\pi^2}{6}$, allgemeiner von $\zeta(2t)=(-1)^{t-1}\frac{2^{2t-1}}{(2t)!}B_{2t}\pi^{2t}$ für alle $t\in\mathbb{N}$ mit den in 4.2.8 einzuführenden (rationalen) BERNOULLI–Zahlen B_k (vgl. 5).

Auf EULER (1737) geht auch die Entdeckung des Produkts 4(4) für $\zeta(s)$ zurück; daher nennt man heute Produktentwicklungen des Typs 4(2) für Reihen der Form 4(1) EULER–Produkte. Allerdings beschränkte sich EULER auf reelle s. Interessant ist, daß er 4(4) noch für $s=1$ anwandte und daraus auf die Divergenz von $\prod_p(1-p^{-1})^{-1}$ und $\sum_p p^{-1}$ schloß, vgl. Bemerkung 4 in 5.

Ohne die Schreibweise $\varphi(n)$ zu benutzen, führte EULER (Opera Omnia Ser. 1, II, 531–555) im Zusammenhang mit seiner Verallgemeinerung 2.3.4 des Satzes 2.3.3 von P. FERMAT die Anzahl der zu $n\in\mathbb{N}$ teilerfremden natürlichen Zahlen ein, vgl. 11(1) für $\alpha=1$. Später notierte er diese Anzahl als πn; das Symbol φ

geht auf GAUSS (*Disquisitiones Arithmeticae*, Art. 38) zurück, der φn schrieb. EULER selbst hatte die Eigenschaften (i) und (ii) in Satz 11 bzw. (i) in Korollar 11 für die φ–Funktion entdeckt. Die Aussage (iii) von Satz 11 wurde für $\alpha = 1$ von GAUSS (a.a.O., Art. 39) bewiesen, während (iv) desselben Satzes 1856 von E. BETTI gefunden wurde, nachdem 1831 die Funktion μ aus 9 von A.F. MÖBIUS (Werke IV, 591–613) definiert und systematisch untersucht worden war.

Die in 6 bis 11 gegebene Einführung der zahlentheoretischen Funktionen scheint auf Originalarbeiten von E.T. BELL ab 1915 zurückzugehen, die er in seinem Buch *Algebraic Arithmetic* (AMS Coll. Publ. VII, New York, 1927) unter sehr allgemeinen, vereinheitlichenden Gesichtspunkten dargestellt hat.

§ 5. Teilbarkeit in Integritätsringen

Für die in 1 bis 3 zu gebenden Definitionen genügt es, R lediglich als *kommutativen Ring mit* vom Nullelement 0 verschiedenem *Einselement* 1 vorauszusetzen. In 4 bis 7 sei R überdies nullteilerfrei, also *Integritätsring*.

1. Teiler, Einheiten, Assoziiertheit. Sind $m, n \in R$, $m \neq 0$, so heißt (in Verallgemeinerung von 1.2) n durch m *teilbar*, wenn es ein $q \in R$ mit $n = mq$ gibt. (Bei nullteilerfreiem R kann es höchstens ein derartiges q geben.) Ist n durch m teilbar, so notiert man dies als $m|n$ (die Negation als $m \not| n$) und sagt, m sei ein *Teiler* von n; wie in 1.2 sind auch hier Teiler generell von Null verschieden.

Weiter heißt jeder Teiler $\varepsilon \in R\backslash\{0\}$ von 1 eine *Einheit*; die Menge aller Einheiten von R wird mit $E(R)$ bezeichnet und $1 \in E(R)$ ist klar. Man hat nun folgende

Proposition.

(i) *Zu jedem $\varepsilon \in E(R)$ existiert genau ein $\varepsilon' \in E(R)$ mit $\varepsilon\varepsilon' = 1$.*

(ii) *Es ist $\langle E(R), \cdot \rangle$ eine abelsche Gruppe.*

Die Gruppe in (ii) heißt *Einheitengruppe* von R.

Beweis. Zu (i): Wegen $\varepsilon|1$ existiert ein $\varepsilon' \in R$ mit $\varepsilon\varepsilon' = 1$; ersichtlich ist $\varepsilon' \neq 0$ und $\varepsilon'|1$, also $\varepsilon' \in E(R)$. Hat ε'' dieselben Eigenschaften wie ε', so folgt aus $\varepsilon\varepsilon'' = \varepsilon\varepsilon'$ sofort $\varepsilon'' = \varepsilon'$.

Zu (ii) reicht es nach (i), die Abgeschlossenheit von $E(R)$ bezüglich der Multiplikation zu zeigen: Bei $\varepsilon_1, \varepsilon_2 \in E(R)$ existieren $\varepsilon_1', \varepsilon_2' \in E(R)$ mit $\varepsilon_1\varepsilon_1' = 1 = \varepsilon_2\varepsilon_2'$, weswegen auch $(\varepsilon_1\varepsilon_2)(\varepsilon_1'\varepsilon_2') = 1$ gilt, was $\varepsilon_1\varepsilon_2 \neq 0$ und $(\varepsilon_1\varepsilon_2)|1$, also $\varepsilon_1\varepsilon_2 \in E(R)$ beinhaltet. □

Bemerkung. Aus Satz 1.2 gewinne der Leser $E(\mathbb{Z}) = \{-1, 1\}$. Eine weitere Einheitengruppe ist in Bemerkung 3 zu 2.1.8 zu berechnen.

Sind $m, n \in R$, so heißt m *assoziiert zu* n, wenn es ein $\varepsilon \in E(R)$ mit $\varepsilon m = n$ gibt. Ist m assoziiert zu n, so schreibt man $m \sim n$; die Negation hiervon wird als $m \not\sim n$ notiert. Man erkennt mit der Proposition unmittelbar, daß die Relation $\sim$ auf R eine Äquivalenzrelation definiert; daher wird man bei $m \sim n$ einfacher sagen, m und n seien (zueinander) assoziiert.

Klar ist, daß eine der Äquivalenzklassen von R bezüglich $\sim$ alleine aus dem Nullelement 0 von R besteht und daß eine weitere mit $E(R)$ übereinstimmt. Hat man irgend zwei Elemente m, n aus einer von $\{0\}$ verschiedenen Klasse, so teilen sich diese gegenseitig, d.h.

$$m \sim n \implies m|n,\ n|m. \tag{1}$$

Bei nullteilerfreiem R gilt in (1) auch die umgekehrte Implikation. Speziell sind $\{-n, n\}$, $n = 1, 2, \ldots$, genau die von $\{0\}$ verschiedenen Äquivalenzklassen von $\mathbb{Z}$.

Offenbar ist Teilbarkeit eine Eigenschaft, die sich jeweils auf ganze Äquivalenzklassen bezüglich $\sim$ bezieht, d.h. $m|n$ für $m, n \in R$, $m \neq 0$ ist gleichwertig mit $m'|n'$ für alle $m', n' \in R$, $m' \neq 0$, $m' \sim m$, $n' \sim n$. Um den nichttrivialen Teil dieser Behauptung zu beweisen, stützt man sich auf (i) in der Proposition. In $\mathbb{Z}$ bedeutet dies: Zur Untersuchung der Teilbarkeit von ganzen Zahlen darf man sich auf deren Absolutbeträge beschränken; vgl. Bemerkung 2 zu 1.2.

2. Die Begriffe ggT und kgV. Seien $n_1, \ldots, n_k \in R$ nicht alle Null; $d \in R \setminus \{0\}$ heißt *ein ggT* von $n_1, \ldots, n_k$ genau dann, wenn gilt: $d|n_1, \ldots, d|n_k$ und aus $d' \in R \setminus \{0\}$, $d'|n_1, \ldots, d'|n_k$ folgt $d'|d$.

Es mögen jetzt $n_1, \ldots, n_k \in R$, nicht alle Null, einen ggT $d \in R \setminus \{0\}$ besitzen. Nach Proposition 1(i) ist dann jedes $d^* \in R \setminus \{0\}$ mit $d^* \sim d$ ein ggT von $n_1, \ldots, n_k$. Ist andererseits $d^* \in R \setminus \{0\}$ ein weiterer ggT von $n_1, \ldots, n_k$, so gelten nach Definition eines ggT

$$d^*|d \quad \text{und} \quad d|d^*. \tag{1}$$

a) Ist insbesondere $d \in E(R)$, so folgt $d^* \in E(R)$ aus (1) und man kann sagen: Haben $n_1, \ldots, n_k \in R$, nicht alle Null, überhaupt einen ggT und ist dieser eine Einheit, so ist die Menge aller ggT von $n_1, \ldots, n_k$ gleich $E(R)$; genau in diesem Falle nennt man $n_1, \ldots, n_k$ zueinander *teilerfremd*.

b) Wie nach 1(1) festgestellt, folgt bei nullteilerfreiem R aus (1) die Assoziiertheit von d, d^*. Ist also R ein Integritätsring und sind $n_1, \ldots, n_k \in R$ nicht alle Null, so ist ihr ggT, falls er überhaupt existiert, bis auf Assoziiertheit eindeutig bestimmt. Dann wird die Äquivalenzklasse unter $\sim$ aller ggT von $n_1, \ldots, n_k$ mit $(n_1, \ldots, n_k)$ bezeichnet. Dies ist auch im Falle a) sinnvoll, wo man selbstverständlich $(n_1, \ldots, n_k) = E(R)$ hat.

Der Begriff eines ggT ist hier in einer Weise eingeführt worden, wie dies durch den Charakterisierungssatz 2.3B in Verbindung mit der dortigen Bemerkung nahegelegt war. Analog läßt man sich nun von Satz 2.11 leiten, um in R ein kgV zu definieren.

Seien $n_1, \ldots, n_k \in R \setminus \{0\}$; $m \in R \setminus \{0\}$ heißt *ein kgV* von $n_1, \ldots, n_k$ genau dann, wenn gilt: $n_1|m, \ldots, n_k|m$ und aus $m' \in R \setminus \{0\}$, $n_1|m', \ldots, n_k|m'$ folgt $m|m'$.

Überträgt der Leser die obigen Betrachtungen zum ggT–Begriff bis hin zu (1) und die in b) daran anschließenden auf den kgV–Begriff, so wird er feststellen: Ist R ein Integritätsring und sind $n_1, \ldots, n_k \in R \setminus \{0\}$, so ist ihr kgV, falls es überhaupt existiert, bis auf Assoziiertheit eindeutig bestimmt; dann wird die Äquivalenzklasse unter $\sim$ aller kgV von $n_1, \ldots, n_k$ mit $[n_1, \ldots, n_k]$ bezeichnet.

Bemerkung. In 6.11 wird sich zeigen, daß nicht in jedem Integritätsring zu vorgegebenen Elementen ein ggT oder ein kgV existiert.

3. Unzerlegbare Elemente, Primelemente. Bei $m, n \in R$, $m \neq 0$ heißt m *echter Teiler* von n, falls m Teiler von n ist, der weder Einheit noch zu n assoziiert ist. Speziell ist $m \in \mathbb{Z} \setminus \{0\}$ echter Teiler von $n \in \mathbb{Z}$ genau dann, wenn $m|n$, $m \neq \pm 1$, $m \neq \pm n$ gilt.

Sei nun $n \in R$ weder Null noch Einheit; n heißt *unzerlegbar* (oder *irreduzibel*), wenn es keine echten Teiler hat, und andernfalls *zerlegbar* (oder *reduzibel*).

Danach sind die unzerlegbaren Elemente von $\mathbb{Z}$ genau diejenigen Elemente aus $\mathbb{Z} \setminus \{-1, 0, 1\}$ der Form p oder $-p$ mit $p \in \mathbb{P}$.

Sei erneut $n \in R$ weder Null noch Einheit; n heißt *Primelement* genau dann, wenn aus $n|n_1 n_2$, $n_1, n_2 \in R$ stets $n|n_1$ oder $n|n_2$ folgt.

Nach dem Charakterisierungssatz 2.7 sind die Primelemente von $\mathbb{Z}$ nichts anderes als die unzerlegbaren Elemente von $\mathbb{Z}$. Die Tatsache, daß die beiden zuletzt eingeführten Begriffe im Ring $\mathbb{Z}$ zusammenfallen, beruht auf einer speziellen Eigenschaft desselben, vgl. Satz 5A, Satz 6 und die Bemerkung 1 in 6.

Stets jedoch hat man die nachfolgende Implikation.

Satz. *In Integritätsringen ist jedes Primelement unzerlegbar.*

Beweis. Ist R ein Integritätsring und $n \in R$ ein Primelement, so ist $n \neq 0$ und $n \notin E(R)$. Sei $n_1 \in R \setminus \{0\}$ ein beliebiger Teiler von n; mit geeignetem $n_2 \in R \setminus \{0\}$ ist also

$$n = n_1 n_2. \tag{1}$$

Da n Primelement ist, folgt $n|n_1$ oder $n|n_2$ aus (1). Bei $n|n_1$ folgt aus $n_1|n$ die Assoziiertheit von n_1 und n (vgl. nach 1(1)), bei $n|n_2$ folgt aus $n_2|n$ analog $n_2 \sim n$ und somit $n_1 \in E(R)$ wegen (1). Jedenfalls ist n_1 Einheit oder zu n assoziiert und so hat n keine echten Teiler, ist also unzerlegbar. □

In Satz 5A wird eine große, $\mathbb{Z}$ umfassende Klasse von Integritätsringen angegeben, in denen die Umkehrung der Satzaussage ebenfalls gilt. Andererseits findet sich in 6.11 ein Ring, in dem nicht jedes unzerlegbare Element auch Primelement ist.

Für die Untersuchungen in 4 wird noch ein Hilfssatz bereitgestellt, dessen Beweis sich am zweiten Beweis für die Eindeutigkeitsaussage des Fundamentalsatzes der Arithmetik orientiert, vgl. 2.8.

Lemma. *Gilt für $r \geq 1$ Primelemente $p_1, \ldots, p_r$, für $s \geq 1$ unzerlegbare Elemente $q_1, \ldots, q_s$ und für eine Einheit ε eines Integritätsrings die Gleichung*

$$p_1 \cdot \ldots \cdot p_r = \varepsilon q_1 \cdot \ldots \cdot q_s, \tag{2}$$

so ist $r = s$ und es gibt eine Permutation π der Zahlen $1, \ldots, r$ mit $q_{\pi(\rho)} \sim p_\rho$ für $\rho = 1, \ldots, r$.

Bemerkung. Unter den Voraussetzungen des Lemmas erweisen sich also auch die q's als Primelemente.

Beweis durch Induktion über $\mathrm{Min}(r, s)$. Sei erst $\mathrm{Min}(r, s) = 1$. Ist $r = 1$, so geht die linke Seite p_1 von (2) in einem der $q_1, \ldots, q_s$ auf, da p_1 als Primelement die Einheit ε nicht teilen kann; nach Kürzen in Gleichung (2) durch p_1 würde bei $s \geq 2$ rechts ein q_σ zurückbleiben, welches dann das Einselement des Integritätsrings R teilen müßte entgegen $q_\sigma \notin E(R)$. So ist $s = 1$ und (2) beinhaltet $p_1 \sim q_1$. Bei $s = 1$ besagt (2) soviel wie $q_1 = \varepsilon' p_1 \cdot \ldots \cdot p_r$ mit $\varepsilon' := \varepsilon^{-1} \in E(R)$; bei $r \geq 2$

hätte q_1 z.B. den echten Teiler p_1, was der Unzerlegbarkeit von q_1 widerspricht. So ist $r = 1$ und wieder $p_1 \sim q_1$, was den Induktionsbeginn erledigt.

Sei nun $\mathrm{Min}(r, s) \geq 2$ und die Behauptung für $\mathrm{Min}(r, s) - 1 = \mathrm{Min}(r - 1, s - 1)$ bewiesen. Aus (2) folgt $p_r | q_t$ für ein $t \in \{1, \ldots, s\}$, da p_r Primelement ist. Andererseits hat q_t als unzerlegbares Element keine echten Teiler und so muß $q_t \sim p_r$ sein, d.h. $q_t = \varepsilon_1 p_r$ mit einem $\varepsilon_1 \in E(R)$. Danach ist (2) äquivalent zu

$$p_1 \cdot \ldots \cdot p_{r-1} = \varepsilon' q'_1 \cdot \ldots \cdot q'_{s-1} \tag{3}$$

mit $\varepsilon' := \varepsilon\varepsilon_1 \in E(R)$ und unzerlegbaren $q'_\sigma := q_\sigma$, $(\sigma = 1, \ldots, t-1)$, $q'_\sigma := q_{\sigma+1}$, $(\sigma = t, \ldots, s-1)$. Auf (3) ist nun die Induktionsvoraussetzung anwendbar. □

4. Faktorielle Ringe. Ein Integritätsring heißt *faktorieller Ring*, wenn sich jedes seiner von Null und den Einheiten verschiedenen Elemente als Produkt endlich vieler Primelemente darstellen läßt.

In einem faktoriellen Ring ist die geforderte Produktdarstellung der Elemente nach Lemma 3 in Verbindung mit Satz 3 automatisch eindeutig bis auf die Reihenfolge der Faktoren und bis auf Assoziiertheit. In dem durch Lemma 3 völlig präzisierten Sinne wird im folgenden gesagt, eine Produktdarstellung sei *im wesentlichen eindeutig*. Damit kann behauptet werden der

Satz. *Für Integritätsringe R sind folgende Aussagen äquivalent:*

(i) *R ist faktorieller Ring.*

(ii) *Jedes von Null und den Einheiten verschiedene Element von R läßt sich im wesentlichen eindeutig als Produkt endlich vieler irreduzibler Elemente darstellen.*

Beweis. (i) $\Rightarrow$ (ii) folgt unmittelbar aus der Definition eines faktoriellen Rings sowie aus Lemma 3 und Satz 3. Für (ii) $\Rightarrow$ (i) braucht lediglich eingesehen zu werden, daß unter der Voraussetzung (ii) jedes unzerlegbare Element von R bereits Primelement ist. Sei $n \in R$ unzerlegbar und es gelte $n | n_1 n_2$ mit $n_1, n_2 \in R$, o.B.d.A. beide von Null und Einheiten verschieden. Wegen (ii) lassen sich beide n_i im wesentlichen eindeutig als Produkte endlich vieler unzerlegbarer Elemente $q_{1,i}, \ldots, q_{s(i),i}$ darstellen und das Produkt dieser $s(1) + s(2)$ Elemente q ist eine Darstellung von $n_1 n_2$, die ihrerseits im wesentlichen eindeutig ist. Weil n ein unzerlegbarer Teiler von $n_1 n_2$ ist, muß es zu einem der obigen q's assoziiert sein und somit in einem der beiden n_i aufgehen, womit n als Primelement erkannt ist. □

Bemerkungen. 1) Oft bezeichnet man einen faktoriellen Ring auch als ZPE–*Ring* (Ring mit *Z*erlegung in *P*rimelemente *e*indeutig). Man beachte, daß faktorielle Ringe direkt dadurch *definiert* sind, daß in ihnen ein Analogon zum Fundamentalsatz 1.5 verlangt wird.

2) In faktoriellen Ringen haben endlich viele Elemente (unter den üblichen Voraussetzungen über ihr Nichtverschwinden) stets einen ggT und ein kgV.

3) E.D. Cashwell und C.J. Everett (Pacific J. Math. *9*, 975–985 (1959)) haben bewiesen, daß der in Satz 4.6 untersuchte Integritätsring der zahlentheoretischen Funktionen faktoriell ist.

5. Hauptidealringe. Bekanntlich heißt eine nichtleere Teilmenge J eines kommutativen Rings R ein *Ideal in* R, falls gilt:

(i) $\langle J, + \rangle$ ist Untergruppe von $\langle R, + \rangle$,

(ii) $JR \subset J$, wobei $JR := \{n \cdot x : n \in J,\ x \in R\}$.

$\{0\}$ bzw. R sind offenbar Ideale in R, das *Null–* bzw. *Einheitsideal* in R. Sind $n_1, \ldots, n_k \in R$ fest vorgegeben, so ist auch die Teilmenge

$$n_1 R + \ldots + n_k R := \left\{ \sum_{i=1}^{k} n_i x_i :\ (x_1, \ldots, x_k) \in R^k \right\}$$

von R ein Ideal in R, *das von* $n_1, \ldots, n_k$ *erzeugte Ideal*, vgl. 2.4. Ein Ideal in R, das von einem einzigen Element von R erzeugt wird, heißt *Hauptideal*. Ein Integritätsring, in dem jedes Ideal schon Hauptideal ist, heißt *Hauptidealring*. Bevor nun die wichtigsten Ergebnisse über Teilbarkeit in Hauptidealringen vorgestellt werden, wird noch ein einfacher Hilfssatz bereitgestellt.

Lemma. *Für Elemente* a, b *eines Integritätsrings* R *gilt*

(i) $a \sim b \ \Leftrightarrow\ aR = bR$;

bei $b \neq 0$ *gelten außerdem*

(ii) b *ist Teiler von* $a \ \Leftrightarrow\ aR \subset bR$,

(iii) b *ist echter Teiler von* $a \ \Leftrightarrow\ aR \underset{\neq}{\subset} bR \underset{\neq}{\subset} R$.

Beweis. Zu (ii): $b|a$ besagt $a = bq$ mit geeignetem $q \in R$; ist nun $x \in aR$, so auch $x \in bR$, also $aR \subset bR$. Gilt nun diese Mengeninklusion, so ist $a \in bR$, d.h. $a = bq$ mit einem $q \in R$, also $b|a$.

Zu (i) darf o.B.d.A. $a \neq 0$, $b \neq 0$ vorausgesetzt werden (sonst gilt die Äquivalenz trivialerweise). Im Anschluß an 1(1) wurde $a \sim b \ \Leftrightarrow\ a|b, b|a$ festgestellt und

nach (ii) sind letztere Teilbarkeitsbedingungen mit $aR \supset bR$, $aR \subset bR$, also $aR = bR$ äquivalent.

Kombination von (i) und (ii) liefert (iii) sofort. □

Satz A. *In Hauptidealringen ist jedes unzerlegbare Element auch Primelement.*

Beweis. Sei R Hauptidealring und sei $n \in R$ ein unzerlegbarer Teiler von $n_1 n_2$ mit $n_1, n_2 \in R$. Man betrachte das Ideal

$$J := nR + n_1 R \tag{1}$$

in R. Weil R Hauptidealring ist, wird J von einem $d \in R$ erzeugt, was $J = dR$ bedeutet; wegen $0 \neq n \in J$ ist $d \neq 0$. Aus (1) folgt $dR \supset nR$, also $d|n$ nach (ii) im Lemma. Wegen seiner Unzerlegbarkeit hat n keinen echten Teiler, was zu $d \sim n$ oder $d \sim 1$ ($\Leftrightarrow d \in E(R)$) führt. Im ersten Fall ist $nR = dR \supset n_1 R$ wegen (1), also $n|n_1$ nach dem Lemma. Im zweiten Fall ist $R = nR + n_1 R$ und daher hat man $1 = nx + n_1 y$ bei geeigneter Wahl von $x, y \in R$, also $n_2 = n(n_2 x) + (n_1 n_2)y$, was wegen $n|n_1 n_2$ zu $n|n_2$ führt. Somit ist n als Primelement erkannt. □

Satz B. *In einem Hauptidealring läßt sich jedes von Null und den Einheiten verschiedene Element als Produkt endlich vieler unzerlegbarer Elemente darstellen.*

Beweis. Ist R der betrachtete Hauptidealring, so sei M die Menge aller $m \in R \setminus \{0\}$, $m \notin E(R)$, so daß für alle $k \in \mathbb{N}$ und für alle k–Tupel $(r_1, \ldots, r_k)$ unzerlegbarer Elemente von R die Ungleichung $r_1 \cdot \ldots \cdot r_k \neq m$ gilt. Die Behauptung des Satzes ist offenbar mit $M = \emptyset$ gleichbedeutend.

Nimmt man jetzt $M \neq \emptyset$ an, so gelingt die rekursive Konstruktion einer geeigneten Folge $(m_i) \in M^{\mathbb{N}_0}$: Man setzt $m_0 \in M$ beliebig fest. Ist dann $i \geq 0$ und $m_i \in M$ bereits fixiert, so muß m_i einen echten Teiler (etwa m_i') haben, weil es sonst schon selbst unzerlegbar wäre und daher M nicht angehören könnte. Mit geeignetem $m_i'' \in R$ ist also $m_i = m_i' m_i''$ und daher $m_i'' \neq 0$, $m_i'' \notin E(R)$, $m_i'' \not\sim m_i$, weshalb auch m_i'' echter Teiler von m_i sein muß. Sicher ist nun $m_i' \in M$ oder $m_i'' \in M$, weil andernfalls $m_i' m_i'' = m \notin M$ wäre. Man wählt m_{i+1} beliebig in $\{m_i', m_i''\} \cap M$.

Nach (iii) im Lemma und nach Konstruktion der Folge (m_i) gelten die Mengeninklusionen

$$\{0\} \neq m_0 R \underset{\neq}{\subset} m_1 R \underset{\neq}{\subset} \ldots \underset{\neq}{\subset} m_i R \underset{\neq}{\subset} \ldots, \tag{2}$$

wobei noch $m_0 \neq 0$ beachtet ist. Definiert man schließlich

$$J := \bigcup_{i \in \mathbb{N}_0} m_i R, \tag{3}$$

so ist $\emptyset \neq J \subset R$ klar und es wird behauptet, daß J ein Ideal in R ist.

Dazu hat man (i), (ii) der Definition nachzuweisen: Sind $m, n \in J$, so gibt es $i, j \in \mathbb{N}_0$ mit $m \in m_i R$, $n \in m_j R$ wegen (3); setzt man o.B.d.A. $i \leq j$ voraus, so ist auch $m \in m_j R$ nach (2), also $m + n \in m_j R$ und somit $m + n \in J$ nach (3), was (i) beweist. Ist $n \in J$, $x \in R$, so folgt aus $n \in m_j R$ sofort $nx \in m_j R$, also $nx \in J$ mit (3), was auch (ii) beweist.

Da R Hauptidealring ist, ist das Ideal J aus (3) ein Hauptideal, weshalb man $J = mR$ bei geeignetem $m \in J \setminus \{0\}$ hat. Wegen (3) gibt es ein $i_0 \in \mathbb{N}_0$ mit $m \in m_{i_0} R$, also mit $m_{i_0} | m$ oder äquivalent $(J =) mR \subset m_{i_0} R$ nach (ii) im Lemma. Andererseits beinhaltet (3) die Inklusion $m_i R \subset J$ für alle $i \in \mathbb{N}_0$ und damit wegen (2)

$$J \subset m_{i_0} R \underset{\neq}{\subset} m_{i_0+1} R \subset J.$$

Der hier erzielte Widerspruch zeigt die Unhaltbarkeit der Annahme $M \neq \emptyset$. □

Kombination der Sätze A und B ergibt unter Berücksichtigung der Definition eines faktoriellen Rings in 4 unmittelbar den

Satz C. *Jeder Hauptidealring ist faktorieller Ring.*

6. Euklidische Ringe. Wie kann man einem Integritätsring ansehen, ob er Hauptidealring ist? Um eine einigermaßen befriedigende Antwort auf diese Frage geben zu können, sei folgende Definition vorausgeschickt.

Ein Integritätsring R heißt *euklidischer Ring*, wenn in R ein "Divisionsalgorithmus" und eine Abbildung $G : R \setminus \{0\} \to \mathbb{N}$ mit folgenden Eigenschaften existiert: Zu jedem Paar $(n, m) \in R^2$ mit $m \neq 0$ gibt es ein Paar $(a, b) \in R^2$, so daß gilt

$$n = am + b \text{ mit entweder } b = 0 \text{ oder } G(b) < G(m). \tag{1}$$

Die Abbildung G heißt *Gradfunktion* (oder *euklidische Normfunktion*).

Über euklidische Ringe hat man folgenden wichtigen Satz, der eine Antwort auf die eingangs gestellte Frage gibt.

Satz. *Jeder euklidische Ring ist Hauptidealring (und somit faktorieller Ring).*

Beweis. Sei R euklidischer Ring und J ein beliebiges Ideal in R, welches o.B.d.A. nicht das Nullideal $\{0\}$ ist. Offenbar ist $\{G(x) : x \in J, x \neq 0\}$ eine nichtleere Teilmenge von $\mathbb{N}$, die somit ein kleinstes Element besitzt. Es werde $m \in J$, $m \neq 0$ mit minimalem G–Wert fixiert. Ist $n \in J$ beliebig, so existieren Elemente $a, b \in R$, die den Bedingungen (1) genügen; dabei ist sogar $b \in J$ wegen $m, n \in J$, wenn man noch beide definierende Eigenschaften eines Ideals ausnützt. Bei $b \neq 0$ wäre b wegen der letzten Bedingung in (1) ein Element von $J \setminus \{0\}$ mit *kleinerem* G–Wert als m, was nicht geht. So ist $b = 0$, also $n = am$ oder $n \in mR$, was $J \subset mR$ bedeutet. Da $mR \subset J$ aus $m \in J$ folgt, ist $J = mR$ bewiesen und somit J als Hauptideal erkannt. □

Bemerkungen. 1) Wie in 2.2 nachgewiesen, ist $\mathbb{Z}$ ein euklidischer Ring; als Gradfunktion ist die durch $G(x) := |x|$ für $x \in \mathbb{Z} \setminus \{0\}$ festgelegte geeignet. Nach dem hier gezeigten Satz ist $\mathbb{Z}$ auch Hauptidealring, weswegen Satz 2.4 so formuliert werden kann: Das von $n_1, \ldots, n_k \in \mathbb{Z}$, nicht alle Null, erzeugte Ideal in $\mathbb{Z}$ ist identisch mit dem von $\mathrm{ggT}(n_1, \ldots, n_k)$ erzeugten Hauptideal.

2) In euklidischen Ringen kann *ein* ggT zweier Elemente, die nicht beide Null sind, analog wie in 2.9 durch mehrfache Anwendung des in (1) geforderten Divisionsalgorithmus ermittelt werden.

7. Polynome. Hier sei R Integritätsring und k eine natürliche Zahl. Eine formale k-fache Summe

$$f := \sum_{\mathbf{i} \in \mathbb{N}_0^k} a(i_1, \ldots, i_k) X_1^{i_1} \cdot \ldots \cdot X_k^{i_k} \tag{1}$$

mit allen $a(\mathbf{i}) := a(i_1, \ldots, i_k) \in R$, von denen *höchstens endlich viele* von Null verschieden sind, heißt *Polynom* in den k Unbestimmten $X_1, \ldots, X_k$ über R, die $a(\mathbf{i})$ heißen die *Koeffizienten* von f.

Die Menge aller Polynome (1) über R wird als $R[X_1, \ldots, X_k]$ notiert. Ist f wie in (1) und

$$g := \sum_{\mathbf{i} \in \mathbb{N}_0^k} b(i_1, \ldots, i_k) X_1^{i_1} \cdot \ldots \cdot X_k^{i_k} \in R[X_1, \ldots, X_k], \tag{2}$$

so heißt f *gleich* g (in Zeichen $f = g$), falls $a(\mathbf{i}) = b(\mathbf{i})$ für alle $\mathbf{i} \in \mathbb{N}_0^k$ gilt. Ist nicht f gleich g, so schreibt man $f \neq g$.

Das Polynom f aus (1) heißt *nichtkonstant*, wenn $a(\mathbf{i}) \neq 0$ für mindestens ein $\mathbf{i} \in \mathbb{N}_0^k \setminus \{\mathbf{0}\}$, andernfalls *konstant*; sind insbesondere alle $a(\mathbf{i})$ Null, so heißt f das *Nullpolynom*, welches als 0 notiert wird. Bei $f \neq 0$ heißt

$$\partial(f) := \mathrm{Max}\{i_1 + \ldots + i_k : \mathbf{i} \in \mathbb{N}_0^k,\ a(\mathbf{i}) \neq 0\} \tag{3}$$

der *Gesamtgrad von* f und

$$\partial_\kappa(f) := \mathrm{Max}\{i_\kappa : \mathbf{i} \in \mathbb{N}_0^k,\ a(\mathbf{i}) \neq 0\} \tag{4}$$

der *Grad von* f *in* X_κ, $\kappa = 1, \ldots, k$.

Hat man insbesondere ein Polynom $f \neq 0$ in *einer* Unbestimmten (die man X ohne Index schreibt), so fallen (3) und (4) zusammen und man spricht nur vom *Grad von* f und schreibt

$$f = \sum_{i=0}^{\partial(f)} a_i X^i.$$

Dabei heißt $a_{\partial(f)} \neq 0$ der *Leitkoeffizient* (oder *höchste Koeffizient*) von f; ist dieser insbesondere gleich 1, so heißt f *normiert*.

Zu $f, g \in R[X_1, \ldots, X_k]$ wie in (1), (2) werden nun zwei neue formale k–fache Summen definiert gemäß

$$f + g := \sum_{\mathbf{i} \in \mathbb{N}_0^k} c(i_1, \ldots, i_k) X_1^{i_1} \cdot \ldots \cdot X_k^{i_k} \text{ mit } c(\mathbf{i}) := a(\mathbf{i}) + b(\mathbf{i}) \tag{5}$$

bzw.

$$f \cdot g := \sum_{\mathbf{i} \in \mathbb{N}_0^k} d(i_1, \ldots, i_k) X_1^{i_1} \cdot \ldots \cdot X_k^{i_k} \text{ mit } d(\mathbf{i}) := \sum_{\substack{\mathbf{j}, \mathbf{l} \in \mathbb{N}_0^k \\ \mathbf{j}+\mathbf{l}=\mathbf{i}}} a(\mathbf{j}) b(\mathbf{l}), \tag{6}$$

jeweils für alle $\mathbf{i} \in \mathbb{N}_0^k$. Nach Voraussetzung gibt es ein $I \in \mathbb{N}_0$, so daß die $a(\mathbf{i})$, $b(\mathbf{i})$ für alle $\mathbf{i} \in \mathbb{N}_0^k$ mit $i_1 + \ldots + i_k > I$ Null sind. Für dieselben $\mathbf{i}$ ist dann $c(\mathbf{i})$ Null und hieraus folgt insbesondere

$$\partial(f + g) \leq \mathrm{Max}(\partial(f), \partial(g)) \tag{7}$$

für alle $f, g \in R[X_1, \ldots, X_k] \setminus \{0\}$ mit $f + g \neq 0$. Weiter ist auch $d(\mathbf{i})$ Null für die $\mathbf{i}$ mit $i_1 + \ldots + i_k > 2I$; denn für diese gilt in jedem Summanden $a(\mathbf{j})b(\mathbf{l})$ ganz rechts in (6) eine der Bedingungen $j_1 + \ldots + j_k > I$ oder $l_1 + \ldots + l_k > I$. Daher sind $f + g$, $f \cdot g \in R[X_1, \ldots, X_k]$; sie heißen *Summe* bzw. *Produkt* von f und g. Man rechnet nun elementar nach, daß $R[X_1, \ldots, X_k]$ bezüglich der hier definierten Addition (die Inverse zu f wird $-f$ geschrieben) und Multiplikation einen kommutativen Ring mit Einselement (der Eins von R) bildet; dieser heißt der *Polynomring* in k Unbestimmten über R.

Bisher hat man die Nullteilerfreiheit von R noch nicht ausgenützt; dies wird nun nötig, wenn man noch die Nullteilerfreiheit des Polynomrings $R[X_1, \ldots, X_k]$ haben möchte für folgenden

Satz. *Für jedes $k \in \mathbb{N}$ ist der Polynomring in k Unbestimmten über einem Integritätsring selbst wieder ein Integritätsring.*

Beweis. Man definiert zunächst in $\mathbb{N}_0^k$ eine Relation $\preceq$, indem man für $\mathbf{i}, \mathbf{i}' \in \mathbb{N}_0^k$ schreibt $\mathbf{i} \preceq \mathbf{i}'$ (und sagt, $\mathbf{i}$ kommt nicht nach $\mathbf{i}'$), falls aus $i_1 = i_1', \ldots, i_{\kappa-1} = i_{\kappa-1}'$, $i_\kappa \neq i_\kappa'$ für ein $\kappa \in \{1, \ldots, k\}$ folgt $i_\kappa < i_\kappa'$. Weiter schreibt man $\mathbf{i} \prec \mathbf{i}'$ ($\mathbf{i}$ kommt vor $\mathbf{i}'$), falls $\mathbf{i} \preceq \mathbf{i}'$, $\mathbf{i} \neq \mathbf{i}'$ gilt. Man sieht leicht, daß $\preceq$ bzw. $\prec$ eine Ordnung bzw. strikte Ordnung sind, deren Linearität und deren Monotonie bezüglich der Addition (in $\mathbb{N}_0^k$) sofort klar ist.

Sei nun f bzw. g aus $R[X_1, \ldots, X_k] \setminus \{0\}$ und sei $\mathbf{J} = (J_1, \ldots, J_k)$ bzw. $\mathbf{L} = (L_1, \ldots, L_k)$ das bezüglich der eingeführten Ordnung letzte $\mathbf{j}$ bzw. $\mathbf{l}$ aus $\mathbb{N}_0^k$ mit $a(\mathbf{j}) \neq 0$ bzw. $b(\mathbf{l}) \neq 0$. Nach (6) ist

$$d(\mathbf{J}+\mathbf{L}) = \sum_{\substack{\mathbf{j},\mathbf{l} \in \mathbb{N}_0^k \\ \mathbf{j}+\mathbf{l}=\mathbf{J}+\mathbf{L}}} a(\mathbf{j})b(\mathbf{l}) = a(\mathbf{J})b(\mathbf{L}) \neq 0, \tag{8}$$

also $f \cdot g \neq 0$. Dabei ist folgendes beachtet: In der Summe in (8) brauchen nur solche $\mathbf{j}$, $\mathbf{l}$ berücksichtigt zu werden, für die $\mathbf{j} \preceq \mathbf{J}$ und $\mathbf{l} \preceq \mathbf{L}$ gilt; ist jedoch $\mathbf{j} \prec \mathbf{J}$, $\mathbf{l} \preceq \mathbf{L}$ oder $\mathbf{j} \preceq \mathbf{J}$, $\mathbf{l} \prec \mathbf{L}$, so ist $\mathbf{j}+\mathbf{l} \prec \mathbf{J}+\mathbf{l} \preceq \mathbf{J}+\mathbf{L}$, also $\mathbf{j}+\mathbf{l} \prec \mathbf{J}+\mathbf{L}$ im ersten Fall und die letzte Beziehung folgt auch im zweiten. Daher war in (8) alleine der Summand mit $\mathbf{j} = \mathbf{J}$, $\mathbf{l} = \mathbf{L}$ zu berücksichtigen. □

Aus diesem Beweis ergibt sich als Korollar der

Grad–Satz. *Ist R Integritätsring, so gilt für $f, g \in R[X] \setminus \{0\}$ erstens $f \cdot g \in R[X] \setminus \{0\}$ und zweitens $\partial(f \cdot g) = \partial(f) + \partial(g)$.*

Man beachte, daß die Ordnung $\preceq$ in $\mathbb{N}_0^k$ (die sogenannte *lexikographische Ordnung*) für $k = 1$ nichts anderes ist als die in 1.1 eingeführte Ordnung $\leq$ in $\mathbb{N}_0$.

Bemerkung. Der Leser weise $E(R[X]) = E(R)$ für Integritätsringe R nach ebenso wie $E(K) = K^\times$ für Körper K.

8. Polynome über Körpern. Für Körper K ist $K[X]$ nach Satz 7 Integritätsring und in diesem hat man folgenden

Divisionsalgorithmus. *Ist K Körper, so gibt es zu jedem Paar (f, g) mit $f, g \in K[X]$, $g \neq 0$ genau ein Paar (q, r) mit $q, r \in K[X]$, so daß $f = qg + r$ und entweder $r = 0$ oder $\partial(r) < \partial(g)$ gilt.*

Beweis. Ist entweder $f = 0$ oder $\partial(f) < \partial(g)$, so leisten offenbar $q := 0$, $r := f$ und nur diese das Gewünschte. Sei jetzt $(n :=)\ \partial(f) \geq \partial(g)\ (=: m)$ und $f = \sum_{i=0}^{n} a_i X^i$, $g = \sum_{j=0}^{m} b_j X^j$ mit $a_n b_m \neq 0$.

Wenn dann ein Paar (q, r) den Bedingungen des Satzes genügt, ist $q \neq 0$ und $\partial(q) = n - m$, und mit $q = \sum_{\ell=0}^{n-m} c_\ell X^\ell$ ist

$$r = f - qg = \sum_{i=0}^{n} \Big(a_i - \sum_{j=\mathrm{Max}(0,i+m-n)}^{\mathrm{Min}(m,i)} b_j c_{i-j} \Big) X^i \tag{1}$$

sowie

$$a_i = \sum_{j=\mathrm{Max}(0,i+m-n)}^{m} b_j c_{i-j} \qquad \text{für } i = m, \ldots, n. \tag{2}$$

Betrachtet man andererseits (2) als lineares inhomogenes System von $n-m+1$ Gleichungen für die $n-m+1$ Unbekannten $c_0, \ldots, c_{n-m}$, so ist dieses eindeutig lösbar, da die Elemente in der Hauptdiagonalen seiner quadratischen Koeffizientenmatrix ersichtlich alle gleich b_m $(\neq 0)$ sind. Bildet man dann q mit den gefundenen c_ℓ wie oben und definiert damit $r := f - qg$, so folgt mit (2) aus der rechten Seite von (1), daß entweder $r = 0$ oder $\partial(r) < m = \partial(g)$ gilt. □

Für den späteren Gebrauch wird hieraus noch abgeleitet das

Abspaltungslemma. *Ist K Körper, $f \in K[X]$ und ist $c \in K$ Nullstelle von f, so gilt $f(X) = (X - c)q(X)$ mit eindeutig bestimmtem $q \in K[X]$; ist hier $f \neq 0$, so auch $q \neq 0$ und $\partial(f) = 1 + \partial(q)$.*

Beweis. Nach dem Divisionsalgorithmus ist $f(X) = (X - c)q(X) + r(X)$ mit konstantem Polynom r, wobei $r(X) = r(c) = f(c) = 0$ ist. □

Nun kann man behaupten den

Satz. *Für Integritätsringe R sind folgende Aussagen äquivalent:*

(i) *R ist Körper.*

(ii) *$R[X]$ ist euklidischer Ring.*

(iii) *$R[X]$ ist Hauptidealring.*

Beweis. Ist (i) erfüllt, so definiert man $G(f) := 2^{\partial(f)}$ für $f \in R[X] \setminus \{0\}$ und hat damit in G eine Gradfunktion, für die 6(1) gilt; damit trifft (ii) zu. Die

Implikation (ii) ⇒ (iii) entnimmt man Satz 6, während der Nachweis von (iii) ⇒ (i) in diesem Buch nicht geführt werden soll. □

9. Polynomringe über faktoriellen Ringen. Da $\mathbb{Z}$ kein Körper ist, beinhaltet die hier nicht bewiesene Implikation (iii) ⇒ (i) von Satz 8 den folgenden

Satz A. *Der Polynomring $\mathbb{Z}[X]$ ist kein Hauptidealring.*

Hierfür soll ein von Satz 8 unabhängiger direkter *Beweis* gegeben werden. Man betrachtet das von den Polynomen 2 und X erzeugte Ideal J in $\mathbb{Z}[X]$. Wäre J Hauptideal, so müßte es von einem $d \in J \setminus \{0\}$ erzeugt werden, was $d|2$ (nach dem Grad–Satz 7 also die Konstanz von d), $d|X$ und die Existenz von $f, g \in \mathbb{Z}[X]$ mit

$$d = 2f(X) + Xg(X) \tag{1}$$

nach sich zöge. Wegen $d|X$ bliebe nur $d \in \{-1, 1\}$, aber andererseits folgt $d = 2f(0)$ aus (1), was $2|d$ bedeutet. □

Schließlich soll ohne Beweis noch folgendes Ergebnis mitgeteilt werden, das im Falle $R = \mathbb{Z}$ auf Gauss (Konsequenz des Satzes in *Disquisitiones Arithmeticae*, Art. 42) zurückgeht und das man allgemein in der Algebra zeigt.

Satz B. *Für Integritätsringe R gilt: R ist faktorieller Ring genau dann, wenn $R[X]$ faktorieller Ring ist.*

Bemerkungen. 1) Da $\mathbb{Z}$ faktorieller Ring ist, trifft dies auch für $\mathbb{Z}[X]$ zu. Insbesondere ist $\mathbb{Z}[X]$ nach den Sätzen A, B ein Beispiel für einen faktoriellen Ring, der *kein* Hauptidealring ist; Satz 5C ist also nicht umkehrbar.

2) Da $\mathbb{Z}$ faktoriell ist, ist $\mathbb{Z}[X_1]$ faktoriell, also $\mathbb{Z}[X_1, X_2]$ usw. Allgemein ist der *ganzzahlige* Polynomring $\mathbb{Z}[X_1, \ldots, X_k]$ faktorieller Ring für alle natürlichen k.

§ 6. Algebraische Zahlkörper, insbesondere quadratische

1. Algebraische Zahlen, Minimalpolynom. Sei $K|L$ irgendeine Körpererweiterung. Man nennt $\alpha \in K$ *algebraisch über* L, wenn es ein $f \in L[X] \setminus \{0\}$ gibt mit $f(\alpha) = 0$; andernfalls heißt α *transzendent über* L. Ist speziell $L = \mathbb{Q}$ und $\alpha \in \mathbb{C}$, so läßt man den Zusatz "über $\mathbb{Q}$" meist weg: Man sagt in diesem Fall also,

α sei *algebraisch* (bzw. *transzendent*), wenn es ein (bzw. kein) $f \in \mathbb{Q}[X] \setminus \{0\}$ gibt mit $f(\alpha) = 0$; im ersten Fall ist $\partial(f) \in \mathbb{N}$ klar. Selbstverständlich kann f hier bei Bedarf als ganzzahliges Polynom vorausgesetzt werden.

Sei jetzt $\alpha \in \mathbb{C}$ algebraisch und $\delta \in \mathbb{N}$ minimal gewählt, so daß es ein $f \in \mathbb{Q}[X] \setminus \{0\}$ mit $f(\alpha) = 0$, $\partial(f) = \delta$ gibt. Hat g dieselben Eigenschaften wie f, so gilt nach dem Divisionsalgorithmus 5.8: Es ist $f = qg + r$ mit $q, r \in \mathbb{Q}[X]$ und entweder $r = 0$ oder $\partial(r) < \partial(g) = \delta$. Wegen $f(\alpha) = 0 = g(\alpha)$ ist $r(\alpha) = 0$ und nach Definition von δ ist $r = 0$, also $f = qg$, wobei $q \in \mathbb{Q}^\times$ nach dem Grad–Satz 5.7 gelten muß. *Es gibt* also *genau ein normiertes Polynom f_α kleinsten positiven Grades mit rationalen Koeffizienten, welches α annulliert*; dies f_α heißt das *Minimalpolynom von α*. Ebenso existiert genau ein α annullierendes Polynom P_α desselben Grades wie f_α mit teilerfremden ganzen Koeffizienten und positivem Leitkoeffizienten; dieses P_α werde hier das *ganzzahlige Minimalpolynom* von α genannt. Offenbar ist P_α gleich f_α mal dem Leitkoeffizienten von P_α. Der gemeinsame Grad von f_α und P_α heißt der *Grad von α*, in Zeichen: $\partial(\alpha)$.

Satz. *Das Minimalpolynom einer algebraischen Zahl ist (in $\mathbb{Q}[X]$) irreduzibel.*

Bemerkung. Das Minimalpolynom f_α ist Element des nach Satz 5.8 euklidischen Rings $\mathbb{Q}[X]$ und dort ist klar, was unter irreduziblen (oder unzerlegbaren) und Primelementen dieses Rings zu verstehen ist; nach Satz 5.5A besagen beide Begriffe hier dasselbe. Der Zusatz "in $\mathbb{Q}[X]$" ist der Deutlichkeit halber beigefügt; in umfassenderen Polynomringen kann f_α sehr wohl zerlegbar sein.

Beweis. Wegen $\partial(f_\alpha) \in \mathbb{N}$ und der Bemerkung zu 5.7 ist f_α keine Einheit von $\mathbb{Q}[X]$. Wäre f_α zerlegbar, so gäbe es echte Teiler $g, h \in \mathbb{Q}[X] \setminus \{0\}$ von f_α mit $f_\alpha = g \cdot h$ und $0 < \partial(g), \partial(h) < \partial(f_\alpha)$; wegen $g(\alpha)h(\alpha) = f_\alpha(\alpha) = 0$ wäre $g(\alpha) = 0$ oder $h(\alpha) = 0$, was nicht sein kann. □

2. Konjugierte. Vorab wird bereitgestellt folgendes

Lemma.

(i) *Ist α Nullstelle eines $g \in \mathbb{Q}[X] \setminus \{0\}$, so wird g vom Minimalpolynom f_α von α geteilt.*

(ii) *Ist α Nullstelle eines irreduziblen $g \in \mathbb{Q}[X]$, so sind f_α und g zueinander assoziiert.*

(iii) Irreduzible Polynome aus $\mathbb{Q}[X]$ mit einer gemeinsamen Nullstelle sind stets zueinander assoziiert.

Beweis. (i) Nach Voraussetzung ist α algebraisch und nach dem Divisionsalgorithmus 5.8 existieren $q, r \in \mathbb{Q}[X]$, so daß gilt $g = qf_\alpha + r$ mit $r = 0$ oder $\partial(r) < \partial(f_\alpha)$. Wegen $0 = g(\alpha) = r(\alpha)$ muß $r = 0$ sein, also $g = qf_\alpha$.

(ii) Ist g irreduzibel, so ist $g \neq 0$ und dann $q \neq 0$ wegen $g = qf_\alpha$, weshalb $q \in \mathbb{Q}^\times$ gelten muß, d.h. $g \sim f_\alpha$ nach der Bemerkung zu 5.7.

(iii) Sind $g_1, g_2 \in \mathbb{Q}[X]$ irreduzibel und ist $\alpha \in \mathbb{C}$ eine gemeinsame Nullstelle, so gilt $g_i \sim f_\alpha$ für $i = 1, 2$ nach (ii); die Transitivität der Relation $\sim$ liefert die Behauptung. □

Ist jetzt α eine algebraische Zahl und $\delta := \partial(\alpha)$ $(= \partial(f_\alpha))$, so hat f_α in $\mathbb{C}$ genau δ Nullstellen $\alpha_1 := \alpha, \alpha_2, \ldots, \alpha_\delta$; dies ist Inhalt des Fundamentalsatzes der Algebra. Als erste Folgerung aus dem Lemma wird benötigt

Korollar. *Sei α algebraisch, f_α sein Minimalpolynom und $\alpha_2, \ldots, \alpha_\delta$ (falls $\delta := \partial(\alpha) \geq 2$) dessen weitere Nullstellen. Dann ist f_α das Minimalpolynom jedes $\alpha_2, \ldots, \alpha_\delta$ und somit $\partial(\alpha_j) = \partial(\alpha)$ für $j = 2, \ldots, \delta$.*

Beweis. Sei $g_j \in \mathbb{Q}[X]$ das Minimalpolynom von α_j für $j = 2, \ldots, \delta$; nach Satz 1 sind g_j und f_α irreduzibel. Da beide α_j als Nullstelle haben, ist $g_j = q_j f_\alpha$ mit einem $q_j \in \mathbb{Q}^\times$ nach (iii) im Lemma. Weil aber beide Polynome normiert sind, ist $q_j = 1$, also $g_j = f_\alpha$ für $j = 2, \ldots, \delta$. □

Über die Nullstellen des Minimalpolynoms einer algebraischen Zahl gibt nun Auskunft der folgende

Satz. *Die Nullstellen des Minimalpolynoms einer algebraischen Zahl sind sämtliche einfach.*

Beweis. Sind $\alpha_1 := \alpha, \alpha_2, \ldots, \alpha_\delta$ wie im Korollar und käme etwa β mehrfach unter den α's vor, so wäre $\delta \geq 2$ und β Nullstelle von f_α ebenso wie von dessen Ableitung $f'_\alpha \in \mathbb{Q}[X] \setminus \{0\}$. Nach dem Korollar ist $f_\beta = f_\alpha$ und nach (i) des Lemmas wird f'_β von f_β geteilt, was aus Gradgründen nicht geht. □

Die paarweise verschiedenen Nullstellen $\alpha_1 := \alpha, \alpha_2, \ldots, \alpha_\delta$ einer algebraischen Zahl α des Grades $\delta := \partial(\alpha)$ heißen die *Konjugierten* von α (bezüglich $\mathbb{Q}$).

3. Algebraische Zahlkörper. Die Menge aller komplexen algebraischen Zahlen wird mit $\overline{\mathbb{Q}}$ bezeichnet. Mit Hilfe des Satzes über symmetrische Funktionen beweist man in der Algebra leicht, daß aus $\alpha, \beta \in \overline{\mathbb{Q}}$ folgt $\alpha + \beta$, $\alpha \cdot \beta \in \overline{\mathbb{Q}}$. Trivialerweise gilt bei $\beta \in \overline{\mathbb{Q}}$ auch $-\beta, \frac{1}{\beta} \in \overline{\mathbb{Q}}$, letzteres bei $\beta \neq 0$, so daß sich insgesamt $\overline{\mathbb{Q}}$ als Körper mit $\mathbb{Q} \subset \overline{\mathbb{Q}} \subset \mathbb{C}$ erweist. $\overline{\mathbb{Q}}$ heißt der *algebraische Abschluß von* $\mathbb{Q}$ *in* $\mathbb{C}$.

Gewisse Zwischenkörper von $\mathbb{Q}$ und $\overline{\mathbb{Q}}$ werden vor allem in Kap. 6 benötigt und zwar diejenigen Körper K, die über $\mathbb{Q}$, als Vektorräume aufgefaßt, endliche Dimension haben. Genau diese Körper K bezeichnet man als *algebraische Zahlkörper*; ihre Dimension über $\mathbb{Q}$ schreibt man als $[K : \mathbb{Q}]$ und bezeichnet sie als *Grad* von K.

Man beachte aber, daß $\overline{\mathbb{Q}}$ über $\mathbb{Q}$ *nicht* von endlicher Dimension ist.

4. Normen. Sei α algebraisch und $f_\alpha \in \mathbb{Q}[X]$ sein Minimalpolynom. Sind $\alpha_1 := \alpha, \alpha_2, \ldots, \alpha_\delta$ mit $\delta := \partial(\alpha)$ die Konjugierten von α, so folgt aus $f_\alpha(X) = (X - \alpha_1) \cdot \ldots \cdot (X - \alpha_\delta)$ sofort

$$\prod_{j=1}^{\delta} \alpha_j = (-1)^\delta f_\alpha(0). \tag{1}$$

Das Produkt links in (1) heißt die *Norm von* α (in Zeichen: $N(\alpha)$), über die folgendes notiert werden kann.

Proposition. *Die Norm einer algebraischen Zahl ist eine rationale Zahl, welche genau dann Null ist, wenn die algebraische Zahl Null ist.*

Beweis. Ist α die algebraische Zahl, so ist aus (1) die Rationalität von $N(\alpha)$ klar. Bei $N(\alpha) = 0$ ist $f_\alpha(0) = 0$ und so wird $f_\alpha(X)$ von X geteilt, was $f_\alpha(X) = X$ wegen der Unzerlegbarkeit und der Normiertheit von f_α nach sich zieht; daher ist $\alpha = 0$. □

Es wird nun ein weiterer Normbegriff entwickelt, der mit dem obigen eng zusammenhängt, der aber eine zusätzliche Eigenschaft hat, die für manche Zwecke sehr nützlich ist.

Sei K ein algebraischer Zahlkörper mit $\kappa := [K : \mathbb{Q}]$. Nach dem aus der Algebra bekannten Satz vom primitiven Element existiert ein $\vartheta \in K$ mit $\partial(\vartheta) = \kappa$, so daß sich jedes $\alpha \in K$ eindeutig darstellen läßt als $\alpha = a_0 + a_1\vartheta + \ldots a_{\kappa-1}\vartheta^{\kappa-1}$ mit

$$A := a_0 + a_1 X + \ldots + a_{\kappa-1} X^{\kappa-1} \in \mathbb{Q}[X]. \tag{2}$$

Sind $\vartheta_1 := \vartheta, \vartheta_2, \ldots, \vartheta_\kappa$ die verschiedenen Konjugierten von ϑ bezüglich $\mathbb{Q}$, so sei σ_ι für $\iota = 1, \ldots, \kappa$ der $\mathbb{Q}$–Isomorphismus von $K = \mathbb{Q}(\vartheta)$ auf den konjugierten Körper $\mathbb{Q}(\vartheta_\iota)$ mit der Zusatzeigenschaft $\sigma_\iota(\vartheta) = \vartheta_\iota$.

Sei nun $\alpha \in K$ und $f_\alpha, \delta, \alpha_1, \ldots, \alpha_\delta$ wie zu Anfang dieses Abschnitts. Wegen $\alpha = A(\vartheta)$ mit A wie in (2) kommen alle komplexen Zahlen $A(\vartheta_\iota) = \sigma_\iota(A(\vartheta)) = \sigma_\iota(\alpha)$, $\iota = 1, \ldots, \kappa$, unter den $\alpha_1, \ldots, \alpha_\delta$ vor. Nach dem Satz über symmetrische Funktionen ist

$$\prod_{\iota=1}^{\kappa}(X - A(\vartheta_\iota)) \in \mathbb{Q}[X]. \tag{3}$$

Sei $h_1(X) \cdot \ldots \cdot h_t(X)$ eine Zerlegung des Polynoms in (3) mit irreduziblen $h_\tau \in \mathbb{Q}[X]$. Dann hat jedes h_τ eine Nullstelle α_j und nach Lemma 2(ii) in Verbindung mit Korollar 2 ist h_τ zu f_α assoziiert. Wegen der Normiertheit von f_α ist somit

$$\prod_{\iota=1}^{\kappa}(X - A(\vartheta_\iota)) = f_\alpha(X)^t, \tag{4}$$

was zu $\kappa = t \cdot \partial(\alpha)$ $(= t \cdot \delta)$ führt. Nun beachtet man, daß sich aus (1) und (4)

$$\prod_{\iota=1}^{\kappa} A(\vartheta_\iota) = \Big(\prod_{j=1}^{\partial(\alpha)} \alpha_j\Big)^{\kappa/\partial(\alpha)} = (-1)^\kappa f_\alpha(0)^{\kappa/\partial(\alpha)} \tag{5}$$

ergibt. Aus (5) ist ersichtlich, daß das Produkt links bei festem algebraischem Zahlkörper K alleine vom Element $\alpha \in K$ abhängt, nicht jedoch von dem gewählten erzeugenden Element ϑ der Erweiterung $K|\mathbb{Q}$.

Die Zahl in (5) heißt *Norm von α bezüglich des Körpers K*, geschrieben als $N_{K|\mathbb{Q}}(\alpha)$. Kombination von (1) und (5) liefert unmittelbar den folgenden Zusammenhang zwischen beiden Normen

$$N_{K|\mathbb{Q}}(\alpha) = N(\alpha)^{[K:\mathbb{Q}]/\partial(\alpha)}, \tag{6}$$

was insbesondere $N_{\mathbb{Q}(\alpha)|\mathbb{Q}}(\alpha) = N(\alpha)$ beinhaltet. Nun hat man den

Satz. *Für algebraische Zahlkörper K und $\alpha, \beta \in K$ gilt:*
(i) $N_{K|\mathbb{Q}}(\alpha)$ *ist rational;* $N_{K|\mathbb{Q}}(\alpha) = 0 \;\Leftrightarrow\; \alpha = 0$.
(ii) $N_{K|\mathbb{Q}}(\alpha \cdot \beta) = N_{K|\mathbb{Q}}(\alpha) \cdot N_{K|\mathbb{Q}}(\beta)$.

Beweis. Während sich (i) sofort aus der Proposition in Verbindung mit (6) ergibt, sieht man (ii) so: Seien

$$B := b_0 + \ldots + b_{\kappa-1}X^{\kappa-1} \quad \text{bzw.} \quad C := c_0 + \ldots + c_{\kappa-1}X^{\kappa-1}$$

aus $\mathbb{Q}[X]$, so daß $\beta = B(\vartheta)$ bzw. $\alpha\beta = C(\vartheta)$ die zu $\alpha = A(\vartheta)$ analogen Darstellungen seien. Nach Definition von $N_{K|\mathbb{Q}}$ ist dann

$$N_{K|\mathbb{Q}}(\alpha\beta) = \prod_{\iota=1}^{\kappa} C(\vartheta_\iota) = \prod_{\iota=1}^{\kappa}(A(\vartheta_\iota)\cdot B(\vartheta_\iota)) = N_{K|\mathbb{Q}}(\alpha)\cdot N_{K|\mathbb{Q}}(\beta). \qquad \square$$

Bemerkung. Die in (ii) zum Ausdruck kommende Multiplikativität von $N_{K|\mathbb{Q}}$ ist für manche Anwendungen (vgl. etwa 9) eine vorteilhafte Eigenschaft, die die zuerst eingeführte Norm N nicht besitzt.

5. Ganzheit. Eine algebraische Zahl α heißt *ganz*, wenn $f_\alpha \in \mathbb{Z}[X]$ gilt. Der Deutlichkeit halber sagt man hier auch, α sei *ganzalgebraisch*; in diesem Zusammenhang nennt man die Elemente von $\mathbb{Z}$ *ganzrational*. In dieser neuen Terminologie kann Satz 1.9 so ausgesprochen werden: Jede ganzalgebraische Zahl, die rational ist, ist ganzrational. Offenbar fallen genau dann, wenn α ganzalgebraisch ist, Minimalpolynom f_α und ganzzahliges Minimalpolynom P_α zusammen, vgl. 1.

Erneut mit Hilfe des Satzes über symmetrische Funktionen beweist man, daß mit α und β auch $\alpha+\beta$ und $\alpha\cdot\beta$ ganzalgebraisch sind. Danach ist für jeden Zwischenkörper K von $\mathbb{Q}$ und $\overline{\mathbb{Q}}$, gleichgültig ob von endlichem Grad über $\mathbb{Q}$ oder nicht, klar, daß

$$O_K := \{\alpha \in K :\ f_\alpha \in \mathbb{Z}[X]\}$$

Integritätsring ist. Dieser heißt *Ganzheitsring* von K.

Satz.

(i) *Für ganzalgebraische α ist $N(\alpha)$ ganzrational; dabei ist $N(\alpha)=0$ genau für $\alpha=0$.*

Ist K ein algebraischer Zahlkörper, so gilt für $\alpha \in O_K$:

(ii) *Es ist $N_{K|\mathbb{Q}}(\alpha)$ ganzrational und überdies gleich Null genau für $\alpha=0$.*

(iii) *Genau dann ist α Einheit in O_K, wenn $N_{K|\mathbb{Q}}(\alpha)$ und $N(\alpha)$ gleich 1 oder -1 sind.*

Beweis. Zu (i) und (ii): Für ganzes α sind $f_\alpha(0)$ und damit nach 4(1) auch $N(\alpha)$ ganzrational; ist außerdem $\alpha \in O_K$, so ist $N_{K|\mathbb{Q}}(\alpha)$ ganzrational nach 4(6). Die Zusatzaussagen über das Verschwinden von $N(\alpha)$ bzw. $N_{K|\mathbb{Q}}(\alpha)$ entnimmt man Proposition 4 bzw. Satz 4(i).

Zu (iii): Bei $\alpha \in E(O_K)$ existiert $\alpha' \in O_K$ mit $\alpha\alpha' = 1$, woraus man mit Satz 4(ii) erhält $N_{K|\mathbb{Q}}(\alpha)N_{K|\mathbb{Q}}(\alpha') = N_{K|\mathbb{Q}}(1) = 1$. Unter Beachtung von (ii) heißt

dies $N_{K|\mathbb{Q}}(\alpha) \in \{-1, 1\}$ und wegen 4(6) auch $N(\alpha) \in \{-1, 1\}$, wobei man noch berücksichtigt hat, daß $N(\alpha)$ nach (i) ganzrational ist. Ist umgekehrt $N(\alpha) = \pm 1$, so ist dies nach 4(1) mit $\pm\alpha_2 \cdot \ldots \cdot \alpha_\delta = \frac{1}{\alpha}$ gleichbedeutend, wenn wie früher $\delta = \partial(\alpha)$ ist und $\alpha_2, \ldots, \alpha_\delta$ die von α_1 verschiedenen Konjugierten von $\alpha_1 := \alpha$ bezeichnen. Nach Korollar 2 haben $\alpha_2, \ldots, \alpha_\delta$ das Minimalpolynom f_α und sind somit genauso wie ihr Produkt ganzalgebraisch und so ist $\frac{1}{\alpha}$ ganzalgebraisch und in K. Also ist $\frac{1}{\alpha} \in O_K$ und α als Einheit in O_K erkannt. □

Bemerkung. Jeder Transzendenzbeweis verwendet in dieser oder jener Form folgende Konsequenz aus Teil (i) des Satzes: *Die Norm einer nichtverschwindenden ganzen algebraischen Zahl ist absolut nicht kleiner als Eins.* Man vergleiche etwa den Beweis von Lemma 6.4.2.

6. Quadratische Zahlkörper. Dies sind genau die algebraischen Zahlkörper K mit $[K : \mathbb{Q}] = 2$. Ist überdies $K \subset \mathbb{R}$, so heißt K *reell–quadratisch*, andernfalls *imaginär–quadratisch.* Ist α aus einem reell–quadratischen (bzw. imaginär–quadratischen) Zahlkörper, jedoch nicht rational, so nennt man α eine *reell–* (bzw. *imaginär–*) *quadratische Irrationalität* (oder *Irrationalzahl*).

Wie sehen nun die Elemente eines quadratischen Zahlkörpers K aus?

Sei $\alpha \in K \setminus \mathbb{Q}$ (also $\partial(\alpha) = 2$) und $P_\alpha := rX^2 + sX + t$ sein ganzzahliges Minimalpolynom; insbesondere ist $r > 0$. Wegen $r\alpha^2 + s\alpha + t = 0$ ist $\alpha = \big(-s \pm \sqrt{s^2 - 4rt}\big)/(2r)$, wo $s^2 - 4rt$ wegen $\alpha \notin \mathbb{Q}$ keine Quadratzahl ist. Man kann nun in eindeutiger Weise eine quadratfreie (vgl. Bemerkung 2 zu 4.9) Zahl $d \in \mathbb{Z} \setminus \{1\}$ und eine Zahl $u \in \mathbb{Z} \setminus \{0\}$ finden, so daß

$$s^2 - 4rt = du^2 \tag{1}$$

gilt, womit dann

$$\alpha = \frac{1}{2r}(-s \pm u\sqrt{d}) \tag{2}$$

mit $r, s, u \in \mathbb{Z}$, $ru \neq 0$ entsteht. Dabei ist α reell– bzw. imaginär–quadratisch genau dann, wenn d positiv bzw. negativ ist.

Ist α wie oben gewählt, so bilden 1, α eine Basis von K über $\mathbb{Q}$. Nach den Überlegungen von soeben stellen auch 1, $\sqrt{d}$ eine derartige Basis dar; d ist ein Charakteristikum von K und man schreibt $K = \mathbb{Q}(\sqrt{d})$.

7. Deren Ganzheitsring. Wie sieht nun der Ganzheitsring $O_d := O_{\mathbb{Q}(\sqrt{d})}$ des Körpers $\mathbb{Q}(\sqrt{d})$ bei quadratfreiem $d \in \mathbb{Z} \setminus \{1\}$ aus? Die vollständige Antwort gibt folgender

Satz. *Ist $d \neq 1$ ganzrational und quadratfrei, so sind die Zahlen $\frac{x+y\sqrt{d}}{z}$ mit teilerfremden ganzrationalen $z > 0, x, y$ des quadratischen Zahlkörpers $\mathbb{Q}(\sqrt{d})$ genau dann ganz, wenn entweder $z = 1$ oder $z = 2$, $2 \nmid xy$, $4|(d-1)$ gilt.*

Beweis. Ist $\alpha \in \mathbb{Q}(\sqrt{d}) \setminus \mathbb{Q}$ und P_α sein ganzzahliges Minimalpolynom wie in 6, so ist die Ganzheit von α definitionsgemäß mit $r = 1$ äquivalent, nach 6(2) also mit $\alpha = \frac{1}{2}(-s \pm u\sqrt{d})$, wobei überdies

$$4|(s^2 - du^2) \tag{1}$$

wegen 6(1) gelten muß. Für $4|(d-2)$ und $4|(d-3)$ erweist sich (1) mit $2|s$, $2|u$ äquivalent, weshalb α dann von der Form $x + y\sqrt{d}$ mit ganzrationalen x, y ist. Ist $4|(d-1)$, so ist (1) mit $4|(s^2 - u^2)$ und dies mit $2|(s-u)$ äquivalent; sind dann s, u beide gerade, so hat α dieselbe Form wie soeben. Oder aber es sind s, u beide ungerade; dann ist α von der Form $\frac{1}{2}(x + y\sqrt{d})$ mit $2 \nmid xy$.

Ist $\alpha \in \mathbb{Q} \cap O_d$, so ist $\alpha \in \mathbb{Z}$ nach Satz 1.9 (vgl. auch Anfang von 5); dann ist α von der im Satz angegebenen Form, man kann dort $x := \alpha$, $y := 0$, $z := 1$ wählen. □

Bemerkung. Bei $z = 2$ und ungeraden x, y ist $\frac{1}{2}(x+y\sqrt{d}) = \frac{1}{2}(x-y)+\frac{1}{2}y(1+\sqrt{d})$, wobei $\frac{1}{2}(x-y) \in \mathbb{Z}$ ist. Demnach ist

$$O_d = \begin{cases} \mathbb{Z} + \mathbb{Z}\frac{1+\sqrt{d}}{2}, & \text{falls } 4|(d-1) \\ \mathbb{Z} + \mathbb{Z}\sqrt{d}, & \text{falls } 4|(d-2) \text{ oder } 4|(d-3). \end{cases} \tag{2}$$

Wegen der Quadratfreiheit von d kann der fehlende Fall $4|d$ selbstverständlich nicht auftreten. Man kann übrigens aus (2) leicht ablesen, daß O_d ein Integritätsring ist; hier kommt man ohne den Satz über symmetrische Funktionen aus (vgl. vor Satz 5).

8. Einheiten quadratischer Zahlringe. Hierüber gibt abschließende Auskunft folgender

Satz A. *Ist $d \neq 1$ ganzrational und quadratfrei, so ist ε Einheit in O_d genau dann, wenn mit ganzrationalen x, y gilt: $\varepsilon = \frac{1}{2}(x+y\sqrt{d})$, $2|(x-y)$, $x^2 - dy^2 = 4$ oder $x^2 - dy^2 = -4$, falls $4|(d-1)$ bzw. $\varepsilon = x + y\sqrt{d}$, $x^2 - dy^2 = 1$ oder $x^2 - dy^2 = -1$, falls $4 \nmid (d-1)$.*

Beweis. Nach Satz 5(iii) gilt $\varepsilon \in E(O_d)$ genau dann, wenn $N_{\mathbb{Q}(\sqrt{d})|\mathbb{Q}}(\varepsilon) = \pm 1$ ist. Nach Definition im Anschluß an 4(5) ist diese Norm gleich $\frac{1}{2}(x + y\sqrt{d}) \cdot \frac{1}{2}(x - y\sqrt{d}) = \frac{1}{4}(x^2 - dy^2)$ im ersten bzw. gleich $(x+y\sqrt{d}) \cdot (x-y\sqrt{d}) = x^2 - dy^2$ im zweiten Fall, woraus alle Behauptungen folgen. □

Mit dieser Charakterisierung der Einheiten eines quadratischen Zahlrings O_d kann nun leicht bewiesen werden

Satz B. *Ist $d \in \mathbb{Z}$ negativ und quadratfrei, so hat der Ganzheitsring des imaginär–quadratischen Zahlkörpers $\mathbb{Q}(\sqrt{d})$ stets genau die beiden Einheiten 1 und -1; zu diesen treten im Fall $d = -1$ noch i und $-i$, im Fall $d = -3$ noch die vier Zahlen $\frac{1}{2}(\pm 1 \pm \sqrt{-3})$ als Einheiten hinzu.*

Beweis. Daß bei $d < 0$ der Ausdruck $x^2 - dy^2$ aus Satz A weder -4 noch -1 sein kann, ist klar. Bei $d < -3$ und quadratfrei ist $x^2 - dy^2 = 4$ bzw. $x^2 - dy^2 = 1$ offenbar gleichbedeutend damit, daß (x, y) gleich $(\pm 2, 0)$ bzw. $(\pm 1, 0)$ ist, was den ersten Teil der Behauptung liefert. Bei $d = -3$ hat man $x^2 + 3y^2 = 4$ zu beachten und dies ist äquivalent mit $(x, y) = (\pm 2, 0), (\pm 1, \pm 1)$, was zu den angegebenen sechs Einheiten führt. In den Fällen $d = -1, -2$ schließlich hat man $x^2 - dy^2 = 1$ zu betrachten, was bei $d = -2$ mit $(x, y) = (\pm 1, 0)$ und bei $d = -1$ mit $(x, y) = (\pm 1, 0), (0, \pm 1)$ äquivalent ist. □

Das Problem der Bestimmung aller Einheiten im Ganzheitsring reell–quadratischer Zahlkörper kann erst in 4.3.6 behandelt werden. Dort wird sich zeigen, daß im reell–quadratischen Fall stets unendlich viele Einheiten existieren, ganz im Gegensatz zum imaginär–quadratischen Fall.

9. Euklidische quadratische Zahlringe. Die imaginär–quadratischen Zahlkörper $\mathbb{Q}(\sqrt{d})$, deren Ganzheitsringe euklidisch im Sinne von 5.6 sind, sind sämtliche bekannt. Man entnimmt sie folgendem

Satz. *Ist d eine der Zahlen -11, -7, -3, -2, -1, 2, 3, 5, 13, so ist der Ganzheitsring von $\mathbb{Q}(\sqrt{d})$ euklidisch.*

Beweis. Offenbar muß man die Bedingung 5.6(1) für einen euklidischen Ring sichern. Sei $d \in \mathbb{Z} \setminus \{1\}$ quadratfrei und $\eta := \frac{1}{2}(1 + \sqrt{d})$ bei $4 | (d-1)$ bzw. $\eta := \sqrt{d}$ bei $4 \nmid (d-1)$ gesetzt. Sind $\nu, \mu \in O_d$, $\mu \neq 0$, so gilt mit gewissen $a, b \in \mathbb{Q}$ die Gleichung $\frac{\nu}{\mu} = a + b\eta$. Mit eindeutig bestimmten $B \in \mathbb{Z}$ und $t \in]-\frac{1}{2}, \frac{1}{2}] \cap \mathbb{Q}$ hat man $b = B + t$. Analog schreibt man $a = A + s$ mit $A \in \mathbb{Z}$, $s \in \mathbb{Q}$, wobei man jetzt allerdings nach $4 | (d-1)$ bzw. $4 \nmid (d-1)$ unterscheidet: Im ersten Fall sorgt man für $s \in]-\frac{1}{2}(1+t), \frac{1}{2}(1-t)]$, im zweiten für $s \in]-\frac{1}{2}, \frac{1}{2}]$. In jedem Falle setzt man $\alpha := A + B\eta \in O_d$ und erhält $\frac{\nu}{\mu} = \alpha + (s + t\eta)$ und hier ist das folgende β aus O_d

$$\beta := \nu - \alpha\mu = \mu(s + t\eta). \tag{1}$$

Als Abbildung $G : O_d \setminus \{0\} \to \mathbb{N}$ im Sinne von 5.6 wählt man die durch $\gamma \mapsto |N_{\mathbb{Q}(\sqrt{d})|\mathbb{Q}}(\gamma)|$ gegebene, man beachte hierzu Satz 5(ii). Nach Satz 4(ii) folgt aus (1)

$$G(\beta) = G(\mu)|N_{\mathbb{Q}(\sqrt{d})|\mathbb{Q}}(s + t\eta)| =: G(\mu)|\Delta|,$$

wo man nur noch nachzuweisen hat, daß $|\Delta| < 1$ für die genannten d gilt. Wegen

$$\Delta = N_{\mathbb{Q}(\sqrt{d})|\mathbb{Q}}(s + t\eta) = (s + t\eta) \cdot \begin{cases} (s + t - t\eta) & \text{für } 4|(d-1) \\ (s - t\eta) & \text{für } 4\not|(d-1) \end{cases} \tag{2}$$

kann gesagt werden:

Ist $4|(d-1)$ und $-12 < d < 16$, so gilt wegen (2) und $\eta = \frac{1}{2}(1 + \sqrt{d})$ nach leichter Rechnung

$$-1 = -\frac{16}{4} \cdot \frac{1}{4} < (s + \frac{t}{2})^2 - \frac{d}{4}t^2 = \Delta < \frac{1}{4} + \frac{12}{4} \cdot \frac{1}{4} = 1.$$

Die quadratfreien $d \neq 1$, die den beiden Bedingungen nach (2) genügen, sind genau diese: -11, -7, -3, 5, 13.

Ist $4\not|(d-1)$ und $-3 < d < 4$, so gilt mit (2) und $\eta = \sqrt{d}$

$$-1 = -4 \cdot \frac{1}{4} < s^2 - dt^2 = \Delta < \frac{1}{4} + 3 \cdot \frac{1}{4} = 1;$$

die quadratfreien d, die jetzt beiden genannten Bedingungen genügen, sind genau diese: -2, -1, 2, 3. □

Bemerkung. Über die neun im Satz genannten d hinaus ist der Ganzheitsring von $\mathbb{Q}(\sqrt{d})$ noch genau dann euklidisch *bezüglich der Norm* (vgl. obigen Beweis), wenn d eine der folgenden zwölf Zahlen ist

$$6,\ 7,\ 11,\ 17,\ 19,\ 21,\ 29,\ 33,\ 37,\ 41,\ 57,\ 73. \tag{3}$$

Dieses Ergebnis wurde schrittweise von mehreren Mathematikern gesichert. Man vergleiche etwa H. CHATLAND und H. DAVENPORT (Canad. J. Math. *2*, 289 - 296 (1950)), beachte aber E.S. BARNES und H.P.F. SWINNERTON–DYER (Acta Math. *87*, 259–323 (1952)), die nachwiesen, daß der Ganzheitsring von $\mathbb{Q}(\sqrt{97})$ jedenfalls nicht bezüglich der Norm euklidisch ist, wie dies früher behauptet worden war. Über reell-quadratische Zahlkörper, deren Ganzheitsringe euklidisch bezüglich einer *beliebigen* Gradfunktion im Sinne von 5.6 sind, liegen bisher keine ähnlich abschließende Resultate vor.

Nach Satz 5.6 ist der Ganzheitsring O_d von $\mathbb{Q}(\sqrt{d})$ für die einundzwanzig d–Werte im Satz bzw. in (3) erst recht Hauptidealring. Die Frage, für welche

negativen quadratfreien d der Ring O_d Hauptidealring ist, geht auf GAUSS (*Disquisitiones Arithmeticae*, Art. 303) zurück, der folgende neun angab

$$-1,\ -2,\ -3,\ -7,\ -11,\ -19,\ -43,\ -67,\ -163. \tag{4}$$

A.a.O. schrieb GAUSS*) "nullum dubium esse videtur, quin series adscriptae revera abruptae sint" und fuhr später fort "demonstrationes autem *rigorosae* harum observationum perdifficiles esse videntur." Tatsächlich lieferte erst H.M. STARK (Michigan Math. J. *14*, 1–27 (1967)) einen vollständigen Beweis dafür, daß der Ganzheitsring imaginär–quadratischer Zahlkörper $\mathbb{Q}(\sqrt{d})$ genau dann Hauptidealring ist, wenn d eine der neun in (4) angegebenen Zahlen ist.

10. Primzahlen als Summe zweier Quadrate. Wenn wie üblich $i := \sqrt{-1}$ gesetzt ist, besagt Satz 9 insbesondere, daß der Ganzheitsring $\mathbb{Z}+\mathbb{Z}i$ von $\mathbb{Q}(i)$ euklidisch ist. Diese Tatsache soll nun angewandt werden zum Nachweis folgender

Proposition. *Ist eine Primzahl als Summe zweier Quadratzahlen darstellbar, so ist diese Darstellung bis auf die Reihenfolge der Summanden eindeutig.*

Beweis. Sei p eine Primzahl und $x^2+y^2 = p = x_1^2+y_1^2$ mit ganzrationalen x, y, x_1, y_1; die letzte Gleichung ist mit

$$(x+yi)(x-yi) = p = (x_1+y_1i)(x_1-y_1i) \tag{1}$$

äquivalent. Alle vier Klammerausdrücke müssen hier Primelemente in $\mathbb{Z}+\mathbb{Z}i$ sein. Sicher ist nämlich keiner Null oder eine Einheit, letzteres wegen Satz 8B: Gilt etwa $x+yi = \alpha\cdot\beta$ mit $\alpha,\beta\in\mathbb{Z}+\mathbb{Z}i$, so folgt aus $\alpha\cdot\beta\cdot(x-yi)=p$ wegen Satz 4(ii)

$$N_{\mathbb{Q}(i)|\mathbb{Q}}(\alpha)\cdot N_{\mathbb{Q}(i)|\mathbb{Q}}(\beta)\cdot(x^2+y^2) = p^2,$$

also $N_{\dots}(\alpha)\cdot N_{\dots}(\beta) = p$, weshalb hier genau einer der Faktoren 1 ist; nach Satz 5(iii) ist genau eines der α,β Einheit.

Da die Primelementzerlegung in $\mathbb{Z}+\mathbb{Z}i$ im wesentlichen eindeutig ist, folgt aus (1) entweder $x+yi = \varepsilon(x_1+y_1i)$, $x_1-y_1i = \varepsilon(x-yi)$ oder $x+yi = \varepsilon_1(x_1-y_1i)$, $x_1+y_1i = \varepsilon_1(x-yi)$ mit Einheiten $\varepsilon,\varepsilon_1$, die nach Satz 8B gleich einer der Zahlen 1, -1, i, $-i$ sein müssen. Ist z.B. $\varepsilon = -i$, so heißt dies $x+yi = -i(x_1+y_1i) = y_1-ix_1$, also $(x,y)=(y_1,-x_1)$; völlig analog können die restlichen Fälle diskutiert werden. □

*) ("Es scheint außer Zweifel, daß die angegebenen Folgen tatsächlich abbrechen." "Aber *strenge* Beweise dieser Feststellungen dürften äußerst schwierig sein.")

Bemerkungen. 1) Die hier angeschnittene Problematik wird in 3.3.4 und 4.1.1–2 vertieft. Insbesondere wird in 4.1.1 vollständig geklärt, welche Primzahlen tatsächlich als Summe zweier Quadratzahlen darstellbar sind.

2) Der hier betrachtete Ring $\mathbb{Z}+\mathbb{Z}i$, bisweilen auch $\mathbb{Z}[i]$ notiert, heißt *Gauss-scher Zahlring*. In der Tat hat ihn zuerst Gauss 1807 untersucht und zur Lösung arithmetischer Probleme benutzt. Seine diesbezüglichen Ergebnisse hat er jedoch erst 1831 publiziert (Werke II, S. 93ff., S. 169ff.). Mit dem Schritt, zur Behandlung von Aufgaben im Ganzrationalen umfassendere Ringe wie den obigen heranzuziehen, hat Gauss auch als Wegbereiter der algebraischen Zahlentheorie zu gelten.

11. Dedekinds Beispiel. R. Dedekind (Gesammelte mathematische Werke III, 278–281) hat gezeigt, daß der Ganzheitsring $\mathbb{Z}+\mathbb{Z}\sqrt{-5}$ des imaginär-quadratischen Zahlkörpers $\mathbb{Q}(\sqrt{-5})$ nicht faktoriell ist. Dieser Ring wird kurz mit $\mathbb{D}$ bezeichnet; nach Satz 8B hat er nur die beiden Einheiten 1 und -1.

Proposition.

(i) *Ist $\beta\in\mathbb{D}$ (echter) Teiler von $\alpha\in\mathbb{D}$, so ist $N_{\mathbb{Q}(\sqrt{-5})|\mathbb{Q}}(\beta)$ (echter) Teiler von $N_{\mathbb{Q}(\sqrt{-5})|\mathbb{Q}}(\alpha)$ in $\mathbb{Z}$.*

(ii) *Haben $\alpha,\beta\in\mathbb{D}$ dieselbe Norm bezüglich $\mathbb{Q}(\sqrt{-5})$ und ist $\alpha\neq\pm\beta$, so gilt $\alpha\nmid\beta$, $\beta\nmid\alpha$.*

(iii) *Die Norm bezüglich $\mathbb{Q}(\sqrt{-5})$ jedes Elements von $\mathbb{D}$ läßt bei Division durch 5 einen der Reste 0, 1 oder 4.*

Beweis. (i) Ist $\alpha=\beta\gamma$ mit einem $\gamma\in\mathbb{D}$, so folgt die Behauptung aus Satz 4(ii); man muß nur noch beachten, daß nach Satz 5(iii) Einheiten von $\mathbb{D}$ dadurch charakterisiert sind, daß ihre Norm bezüglich $\mathbb{Q}(\sqrt{-5})$ gleich 1 ist.

(ii) Unter den gemachten Voraussetzungen ist zunächst $\alpha\beta\neq 0$. Haben nun α, β gleiche Normen und gilt etwa $\beta|\alpha$, also $\alpha=\beta\gamma$ mit $\gamma\in\mathbb{D}$, so muß γ bezüglich $\mathbb{Q}(\sqrt{-5})$ Norm 1 haben, also Einheit sein; dann gilt $\alpha=\beta$ oder $\alpha=-\beta$.

(iii) Hat $\alpha\in\mathbb{D}$ die Form $x+y\sqrt{-5}$ mit $x,y\in\mathbb{Z}$, so ist $N_{\mathbb{Q}(\sqrt{-5})|\mathbb{Q}}(\alpha)=x^2+5y^2$; da x^2 bei Division durch 5 nur die Reste 0, 1 oder 4 lassen kann, ist die Behauptung klar. □

Nun wird behauptet, daß *die Elemente* 3, $2+\sqrt{-5}$, $2-\sqrt{-5}$ *von* $\mathbb{D}$ *unzerlegbar* sind. Jeder echte Teiler eines dieser drei Elemente (deren Normen bezüglich $\mathbb{Q}(\sqrt{-5})$ sämtliche 9 sind), müßte nach (i) die Norm 3 haben, was (iii) widerspricht. Nach (ii) ist außerdem klar, daß keines der drei Elemente in einem der

beiden anderen aufgehen kann. Wegen

$$3 \cdot 3 = 9 = (2 + \sqrt{-5}) \cdot (2 - \sqrt{-5})$$

hat also $9 \in \mathbb{D}$ zwei wesentlich verschiedene Produktzerlegungen in unzerlegbare Elemente. Nach Satz 5.4 ist $\mathbb{D}$ *kein faktorieller Ring.* Wegen $(2 + \sqrt{-5}) | 3 \cdot 3$, aber $(2 + \sqrt{-5}) \nmid 3$ ist $2 + \sqrt{-5}$ *kein Primelement* (*obwohl unzerlegbar*, vgl. Satz 5.3).

Weiter haben die Zahlen 9 *und* $3(2 + \sqrt{-5})$ *weder einen ggT noch ein kgV.*

Man setzt $\sigma := 3$, $\tau := 2 + \sqrt{-5}$. Wäre $\delta \in \mathbb{D}$ ein ggT von 3σ und 3τ (die beide Norm 81 haben), so müßte nach (i) die Norm von δ in 81 aufgehen. Da andererseits σ und τ (beide von der Norm 9) gemeinsame Teiler von 3σ und 3τ sind und somit auch δ teilen, muß die Norm von δ mit Rücksicht auf (iii) gleich 9 oder 81 sein. Im ersten Fall sind die Normen von δ, σ, τ gleich; wegen $\sigma | \delta$, $\tau | \delta$ und (ii) ist $\sigma = \pm\delta$, $\tau = \pm\delta$, also $\sigma = \pm\tau$, was offenbar falsch ist. Im zweiten Fall sind die Normen von δ, 3σ, 3τ gleich; wegen $\delta | 3\sigma$, $\delta | 3\tau$ und (ii) ist $\delta = \pm 3\sigma$, $\delta = \pm 3\tau$ und man erhält denselben Widerspruch wie im ersten Fall.

Zum gleichen Widerspruch führt auch die Annahme, 3σ und 3τ hätten ein $\mu \in \mathbb{D}$ als kgV; hier hat man lediglich davon Gebrauch zu machen, daß 9σ und 9τ gemeinsame Vielfache von 3σ, 3τ sind. Die Ausführung der Einzelheiten kann dem Leser als Übung überlassen bleiben.