D. Wolke

Manuskript zur Vorlesung
Elementare Zahlentheorie

SS 2005

Einleitung. Die Vorlesung gibt eine Einführung in die elementare Zahlentheorie. Das Wort „elementar“ bedeutet dabei erstens, daß die Fragestellungen sich fast ausschließlich auf Eigenschaften der natürlichen und ganzen Zahlen beziehen. Zweitens sollen außer Grundkenntnissen in Analysis und Algebra keine weiteren Hilfsmittel verwandt werden.

Die Zahlentheorie ist neben der Geometrie der älteste Teil der Mathematik. Aus Babylonien, dem alten Ägypten, und China sind erste theoretische Quellen überliefert (z.B. die Darstellung einer rationalen Zahl $a/q \in (0,1]$ als Summe $\frac{1}{n_1} + \cdots + \frac{1}{n_k}$ mit $1 < n_1 < \cdots < n_k$, „ägyptische Brüche“).

Die alten Griechen untersuchten Probleme, die teilweise noch heute aktuell sind, z.B. „diophantische Gleichungen“, d.h. die Suche nach ganzzahligen Lösungen von Gleichungen wie $x^2 + y^2 = z^2$ („pythagoräische Tripel“), oder das höchst rätselhafte Verhalten der Folge der Primzahlen.

Da einerseits die Bausteine, die Elemente von $\mathbb{Z}$, begrifflich leicht zugänglich sind, andererseits so viele höchst schwierige, zum Teil noch ungelöste Probleme bestehen, gehörte die Zahlentheorie stets zu den bevorzugten Arbeitsgebieten der Mathematiker. Einige der bekanntesten Namen, wie Euler, Lagrange, Gauss, werden im Folgenden mehrfach auftreten. Durch die Entwicklung schneller Rechner sind zahlentheoretische Methoden in den letzten Jahrzenten für Anwendungen, z.B. die Kryptografie, sehr wichtig geworden.

Mit dem Ausbau der Mathematik, vor allem seit Beginn des 19. Jahrhunderts, erweiterte sich die Zahlentheorie in Bezug auf Fragestellungen und Methoden erheblich. Die Untersuchung von algebraischen, transzendenten und p–adischen Zahlen, von Folgen ganzer Zahlen, unendlichen Reihen mit zahlentheoretisch interessanten Koeffizienten und vielem anderen gehört heute zu den Zweigen des uralten, aber immer noch rasch wachsenden Baumes der Zahlentheorie. Dementsprechend werden Hilfsmittel aus nahezu allen Teilen der Mathematik verwandt, vor allem aus Algebra (algebraische Zahlentheorie) und komplexer Analysis (analytische Zahlentheorie).

In Freiburg werden regelmäßig Fortsetzungsveranstaltungen angeboten, insbesondere über transzendente Zahlen, algebraische und analytische Zahlentheorie. Hierfür ist der

elementare Teil die verbindende Grundlage.

Literatur. Es gibt zahllose Einführungen in die Zahlentheorie. Bei den folgenden Büchern handelt es sich um bewährte „Klassiker“.

G.H. Hardy, E.M. Wright: An Indroduction to the Theory of Numbers. Oxford, Clarendon

Loo–Keng Hua: Introduction to Number Theory. Springer Verlag.

Bezeichnungen
$\alpha, \beta, t \in \mathbb{R};\ a, b, c, r, g, h, z \in \mathbb{Z};\ d, k, \ell, m, n, \in \mathbb{N};\ p, q$ Primzahlen.

1. Kapitel. Teilbarkeit.

Während die Umkehrung der Addition, die Subtraktion, in $\mathbb{Z}$ unbeschränkt ausführbar ist, läßt sich die Division nicht immer durchführen. $\mathbb{N}, \mathbb{Z}\backslash\{0\}$ und $\mathbb{Z}$ sind bezüglich der Multiplikation nur Halbgruppen. Der wesentliche Begriff hierzu ist der der **Teilbarkeit**.

1.1. Def. (1) a **teilt** b (oder: a ist **Teiler** von b, b wird von a **geteilt**, b ist **Vielfaches** von a) $\underset{\text{Df}}{\Leftrightarrow} \exists\ c : b = ac.$
Kurz: $a|b$, andernfalls $a \nmid b$

Bsp. $1|5,\ 5|5,\ 2 \nmid 5,\ 10|0,\ -2|6,\ 0|0,\ 0 \nmid a \quad \forall\, a \neq 0.$
(2) a heißt **echter Teiler** von $b \underset{\text{Df}}{\Leftrightarrow} a|b \wedge |a| < |b|.$

Folgerungen.

(1) $a|b \ \Rightarrow\ \forall\, c : a|bc,$
(2) $a|b \ \wedge\ b|c \Rightarrow a|c,$
(3) $a|b \ \wedge\ a|c \Rightarrow \forall\ x, y \in \mathbb{Z} :\ a|xb + yc,$
(4) $a|b \ \wedge\ b|a \Rightarrow |a| = |b|,$
(5) $a|b \ \wedge\ b \neq 0 \Rightarrow |a| \leq |b|,$
(6) $a|b \ \Rightarrow\ \forall\, c :\ ca|cb.$

Hinweis zu (4). Die Voraussetzungen ergeben $b = c_1 a,\ a = c_2 b$, also $b = c_1 c_2 b$. Im Fall $b = 0$ folgt $a = 0$. Im Fall $b \neq 0$ folgt $c_1 c_2 = 1$, also $c_1, c_2 \in \{\pm 1\}$.

1.2. Def. $[\alpha] \underset{\text{Df}}{=} \max\{a,\ a \leq \alpha\}$ ($[\alpha]$ = **größte ganze** Zahl $\leq \alpha$.
Kurz: Größtes Ganzes oder **Gauss–Klammer)**.
Bsp. $[n] = n,\ [\pi] = 3,\ [-\pi] = -4.$

1.3. Divisionsalgorithmus.

$$\forall\, a \in \mathbb{Z}\ \forall\, n \in \mathbb{N}\ \exists\, r \in \mathbb{N}_0 : 0 \leq r < n \wedge a = \left[\frac{a}{n}\right] n + r.$$

In der Darstellung $a = bn + r\ (b \in \mathbb{Z}, r \in \mathbb{N}_0, 0 \leq r < n)$ sind b und r eindeutig festgelegt.

Beweis. Nach Definition ist $\left[\frac{a}{n}\right] \leq \frac{a}{n} < \left[\frac{a}{n}\right] + 1$, also $0 \leq a - \left[\frac{a}{n}\right]n < n$.

Dies ist die Ungleichung für r. Es existiert somit eine Darstellung $a = bn+r$ mit $0 \leq r < n$. Sei $a = b'n + r',\ 0 \leq r' < n$ eine weitere. Dann ergibt sich $0 = (b - b')n + (r - r')$, wobei $-n < r - r' < n$. Dies ist nur möglich mit $b = b'$ und $r = r'$. □

Das zu a und n eindeutig bestimmte r heißt der **kleinste nichtnegative Rest** von a bei Division durch n. Es kann r auch durch die Forderung $|r| \leq \frac{n}{2}$ (**absolut kleinster Rest**) festgelegt werden. Dann ist es nicht immer eindeutig festgelegt ($30 = 7 \cdot 4 + 2 = 8 \cdot 4 - 2$).

1.4. Def.

(1) d heißt **gemeinsamer Teiler** von a und $b \underset{\text{Df}}{\Leftrightarrow} d|a \wedge d|b$.

(2) $a^2 + b^2 > 0$. d heißt **größter gemeinsamer Teiler** von a und b (ggT), wenn $d = \max\{c,\ c|a \wedge c|b\}$.
Geschrieben: $d = (a, b)$.

(3) Für $a_1 \neq 0$ sei $(a_1) = |a_1|$. Für $a_1^2 + \cdots + a_n^2 > 0$ wird

$$(a_1, \ldots, a_n, a_{n+1}) = \big((a_1, \ldots, a_n), a_{n+1}\big) \text{ gesetzt.}$$
$$(0, \ldots, 0, a_{n+1}) \underset{\text{Df}}{=} |a_{n+1}| \text{ für } a_{n+1} \neq 0$$

(ggT der Zahlen $a_1, \ldots, a_{n+1}$).

(4) a und b heißen **teilerfremd** oder **relativ prim**, wenn $(a, b) = 1$. $a_1, \ldots, a_n$ $(n \geq 2)$ heißen teilerfremd, wenn $(a_1, \ldots, a_n) = 1$. $a_1, \ldots, a_n$ heißen **paarweise teilerfremd**, wenn

$$(a_j, a_k) = 1 \quad \text{für alle} \quad 1 \leq j < k \leq n.$$

Aus der paarweisen Teilerfremdheit folgt die Teilerfremdheit, das Umgekehrte braucht nicht zu gelten.
Beispiel. (6, 10, 15) = 1, aber (6,10) = 2, (6,15) = 3, (10,15) = 5.

1.5. Satz. $a_1^2 + \cdots + a_n^2 > 0$.

Beh. $(a_1, \ldots, a_n) = \min\{d \,|\, \exists\, z_1, \ldots, z_n \in \mathbb{Z} : d = z_1a_1 + \cdots + z_na_n\}$.

Folgerungen.

(1) $d = (a, b) \Leftrightarrow d|a \wedge d|b \wedge \forall\, c : (c|a \wedge c|b \Rightarrow c|d)$

(2) $(ca, cb) = |c|(a, b)$, falls $c \neq 0$.

(3) Bei $(a_1, \ldots, a_n)$ kommt es nicht auf die Reihenfolge der a_ν an.

(1) kann auch so ausgedrückt werden: Sei $T(a)$ die Menge der Teiler von a ($T(0) = \mathbb{Z},\ |T(a)| < \infty$ für $a \neq 0$). Dann gilt für $a^2 + b^2 > 0$

$$T(a) \cap T(b) = T\big((a, b)\big).$$

Der Beweis zu 1.5. beruht auf dem „Euklidischen Algorithmus“, dem ersten, nicht auf der Hand liegenden algorithmischen Verfahren der Mathematik.

1.6. Euklidischer Algorithmus (Eukleides von Alexandria, um 300 vor der Zeitrechnung).
Zu gegebenen $n_1, n_2 \in \mathbb{N}$ führe man das folgende Schema von Divisionen mit Rest aus

$$\begin{aligned} n_1 &= a_1 n_2 + n_3, && 0 < n_3 < n_2 \\ n_2 &= a_2 n_3 + n_4, && 0 < n_4 < n_3 \\ &\vdots \\ n_{j-2} &= a_{j-2}\, n_{j-1} + n_j, && 0 < n_j < n_{j-1} \\ n_{j-1} &= a_{j-1} n_j. \end{aligned}$$

Beh. $n_j = (n_1, n_2)$.

Beweis. 1. Sei d ein gemeinsamer Teiler von n_1 und n_2. Aus der ersten Zeile des Schemas folgt $d|n_3$, aus der zweiten $d|n_4$, also

(1) $\qquad d|n_1 \wedge d|n_2 \Rightarrow d|n_j.$

2. Umgekehrt sieht man

$$n_j|n_{j-1},\ n_j|n_{j-2}, \ldots,\ n_j|n_2,\ n_j|n_1$$

(2) n_j ist gemeinsamer Teiler von n_1 und n_2.
Aus (1) folgt $(n_1, n_2) \mid n_j$, also $(n_1, n_2) \leq n_j$. Mit (2) ergibt sich die Behauptung. □

Zusatzbemerkungen.

1. Für die Praxis ist es wichtig, bei Paaren großer Zahlen rasch festzustellen, ob sie teilerfremd sind. Der Euklidische Algorithmus ist hierfür gut geeignet.
Am Beweis sieht man, daß statt mit den kleinsten positiven Resten auch mit absolut kleinsten Resten gerechnet werden kann, d.h. in jedem Schritt erfolgt mindestens Halbierung, der Algorithmus stoppt nach $\leq C \ln \min(n_1, n_2)$ Divisionen.

2. Auch bei den kleinsten positiven Resten stoppt er ähnlich schnell. Es gilt, daß nach spätestens zwei Divisionen Halbierung erfolgt. Sei $n_2 < n_1$, dann ist $n_3 \leq \frac{1}{2} n_1$. Denn ist bereits $n_2 \leq \frac{1}{2} n_1$, dann ist es klar. Im Fall $n_2 > \frac{1}{2} n_1$ lautet die erste Zeile

$$n_1 = n_2 + n_3 \quad \text{mit} \quad n_3 < \frac{1}{2} n_1.$$

Ebenso bei den weiteren Divisionen.

Beweis zu Satz 1.5.
1. Aus dem Euklidischen Algorithmus entnimmt man

(∗) $$\exists\, z_1, z_2 : (n_1, n_2) = z_1 n_1 + z_2 n_2.$$

Denn nach der ersten Zeile läßt sich n_3 als ganzzahlige Linearkombination von n_1 und n_2 schreiben, nach der zweiten n_4, usw.

2. Der Beweis zur Behauptung des Satzes wird induktiv geführt. Für $n = 1$ ist nichts zu zeigen. Sei $n \geq 1$ und $a_1^2 + \cdots + a_{n+1}^2 > 0$. Die Fälle $a_1 = \ldots = a_n = 0$ oder $a_{n+1} = 0$ sind leicht einzusehen.

Sei also

$$d_n \underset{\text{Df}}{=} (a_1, \ldots, a_n) > 0, \quad a_{n+1} \neq 0, \quad d_{n+1} \underset{\text{Df}}{=} (d_n, a_{n+1}) > 0.$$

Aus $(*)$ entnimmt man

$$\exists\, z', z_{n+1} \in \mathbb{Z} : \; d_{n+1} = z' d_n + z_{n+1}\, a_{n+1}.$$

Die Induktionsvoraussetzung für $a_1, \ldots, a_n$ liefert

$(**)$ $$\exists\, z_1, \ldots, z_{n+1} \in \mathbb{Z} : \; d_{n+1} = z_1 a_1 + \cdots + z_{n+1}\, a_{n+1}.$$

Ist k das im Satz genannte Minimum, dann folgt $0 < k \leq d_{n+1}$. Da umgekehrt $d_{n+1}|a_1, \ldots, d_{n+1}|a_{n+1}$, folgt $d_{n+1}|k$. Dann bleibt nur $d_{n+1} = k$.

3. Zu Folgerung (1). Die Richtung von links nach rechts ergibt sich unmittelbar aus dem Satz. Ist umgekehrt d eine Zahl, die die Eigenschaft der rechten Seite hat, dann folgt mit $c = (a, b) : c|d$, also $(a, b) \leq d$. Bleibt wegen $d|a \wedge d|b$ nur $d = (a, b)$.

4. Folgerungen (2) und (3) sieht man unmittelbar. □

1.7. Hilfssatz.
(1) Aus $(a, c) = (b, c) = 1$ folgt $(ab, c) = 1$.
(2) Aus $c|ab$ und $(c, b) = 1$ folgt $c|a$.

Beweis zu (1). Mit 1.5., Folgerung (1) sieht man $d \underset{\text{Df}}{=} (ab, c)|ab \wedge d|c$, also $d|ab \wedge d|ac,\ d\,|\,(ab, ac)$. Nach Folgerung (2) ist $(ab, ac) = |a|(b, c) = |a|$, und somit $d|a \wedge d|c$, d.h. $d|(a, c) = 1,\ d = 1$.

Zu (2). Es gilt $c|ab,\ c|ac$, also $c|(ab, ac) = |a|(b, c) = |a|$, d.h. $c|a$. □

1.8. Def. $a_1, \ldots, a_n \neq 0$. k heißt **kleinstes gemeinsames Vielfaches** von $a_1, \ldots, a_n$ (kgV) $\underset{\text{Df}}{\Leftrightarrow}\ k = \min\{m, \forall\, j \leq n : a_j|m\}$.
Kurz: $k = [a_1, \ldots, a_n]$.

Hinweis. Im Fall $n = 1$ darf die eckige Klammer nicht mit der Gauß–Klammer verwechselt werden. Für $a_1 \in \mathbb{N}$ ist kgV $(-a_1) = a_1$, aber Gauß $(-a_1) = -a_1$.

1.9. Satz. $a_1, \ldots, a_n \neq 0$.
(1) b ist gemeinsames Vielfaches von $a_1, \ldots, a_n$ (d.h. $a_1|b, \ldots, a_n|b$) $\Leftrightarrow$ b ist Vielfaches von $[a_1, \ldots, a_n]$.

(2) $[a_1, a_2] \cdot (a_1, a_2) = |a_1 a_2|$.

Satz 1.5 und 1.9 entsprechen einander:

a) Die Menge der gemeinsamen Teiler von $a_1, \ldots, a_n$ ist gleich der Menge der Teiler von $(a_1, \ldots, a_n)$.

b) Die Menge der gemeinsamen Vielfachen von $a_1, \ldots, a_n$ ist gleich der Menge der Vielfachen von $[a_1, \ldots, a_n]$.

Beweis **1.** Zu (2). Sei m Vielfaches von a_1 und a_2. Dann ist $m = a_1 b$, es ist zugleich Vielfaches von a_2, also $\dfrac{m}{a_2} = \dfrac{a_1 b}{a_2} \in \mathbb{Z}$. Sei $d = (a_1, a_2), a_1 = da_1', a_2 = da_2'$ mit $(a_1', a_2') = 1$ nach 1.5., Folgerung (2).
Dies ergibt $\dfrac{m}{a_2} = \dfrac{a_1' b}{a_2'}$, also $a_2' | a_1' b$. Wegen $(a_2', a_1') = 1$ und 1.7. (2) folgt
$a_2' | b,\ b = a_2' b'$. Es bleibt

$$m = a_1 b = a_1 a_2' b' = \frac{a_1 a_2}{d}\, b'.$$

Konsequenz: Jedes gemeinsame Vielfache $m \in \mathbb{N}$ von a_1 und a_2 wird von $\dfrac{a_1 a_2}{(a_1, a_2)}$ geteilt.
Der kleinstmögliche Wert für m ist $\dfrac{|a_1 a_2|}{(a_1, a_2)}$. Dies ist das
gesuchte kgV.

2. Die letzten Überlegungen beinhalten Aussage (1) für $n = 2$.

3. Die Erweiterung von (1) auf mehr als zwei Zahlen a_ν erfolgt induktiv. Man erhält ähnlich wie beim ggT die Rekursion

$$[a_1, \ldots, a_{n+1}] = \big[[a_1, \ldots, a_n], a_{n+1}\big].$$

Formel (2) gilt i.a. nicht für $n \geq 3$.
Richtig ist dagegen:
$a_1, \ldots, a_n$ sind paarweise teilerfremd $\Leftrightarrow [a_1, \ldots, a_n] = |a_1 \cdot \ldots \cdot a_n|$. □

Die Primzahlen als multiplikative Bausteine der ganzen Zahlen bilden eine wichtige Teilmenge von $\mathbb{N}$ und haben von Beginn an die Aufmerksamkeit der Mathematiker auf sich gezogen.

1.10. Def.

(1) $p \in \mathbb{N}$ heißt **Primzahl**, wenn $p > 1$ ist und nur die natürlichen Teiler $d = 1$ und $d = p$ besitzt.

(2) n heißt **zusammengesetzt**, wenn $n > 1$ und keine Primzahl ist.

Folgerung. $p > 1$ ist Primzahl $\Leftrightarrow (\forall\, a, b \in \mathbb{N} :\ p | ab \Rightarrow p | a \vee p | b)$.
Zum Beweis sei

1. p Primzahl und $p | ab$. Falls $d = (p, a) > 1$, muß $d = p$ sein, also $p | a$. Im Fall $(p, a) = 1$ folgt $p | b$ nach 1.7.(2).

2. Die Umkehrung ist simpel. Ist n zusammengesetzt, d.h. $n = n_1 n_2$ mit $1 < n_1, n_2 < n$, dann gilt $n | n_1 n_2$, aber weder $n | n_1$ noch $n | n_2$. □

1.11. Satz (Euklid). Es existieren unendlich viele Primzahlen.

Erster Beweis (Euklid). Jedes $n > 1$ besitzt mindestens einen Primteiler (= Teiler, der Primzahl ist), beispielsweise den kleinsten Teiler d von n mit $1 < d \leq n$. Seien $2 \leq p_1 < \cdots < p_k$ k verschiedene Primzahlen. Dann ist jeder Primteiler von $n = p_1 \cdot \ldots \cdot p_k + 1 > 1$ von $p_1, \ldots, p_k$ verschieden. Denn aus $q|n,\ q = p_\nu$ $(1 \leq \nu \leq k)$ folgte $q|1$, was nicht sein kann. Auf die Weise können unendlich viele Primzahlen gewonnen werden.

Zweiter Beweis. Angenommen, $p_1, \ldots, p_k$ seien alle Primzahlen, dann ist das Produkt

$$\prod_{\nu=1}^{k} \Big(1 - \frac{1}{p_\nu}\Big)^{-1} = \prod_{\nu=1}^{k} \Big(1 + \frac{1}{p_\nu} + \frac{1}{p_\nu^2} + \ldots\Big)$$

konvergent. Mit dem Satz 1.13. über die Eindeutigkeit der Primfaktorzerlegung (zu dessen Beweis die Unendlichkeit der Primzahlmenge nicht benutzt wird) sieht man, daß das Produkt mit $\displaystyle\sum_n \frac{1}{n}$ übereinstimmt. Die Divergenz hiervon ergibt einen Widerspruch.

Dritter Beweis. Im Zusammenhang mit der Frage nach der Konstruierbarkeit des regulären n–Ecks mit Zirkel und Lineal taucht das Problem auf, welche der Zahlen $2^m + 1$ prim sind. Wegen

$$(1) \qquad m = k\ell,\ \ell \text{ ungerade } \Rightarrow 2^{k\ell} + 1 = (2^k + 1)(2^{k(\ell-1)} - 2^{k(\ell-2)} + \cdots + 1)$$

kann dies nur der Fall sein, wenn m selbst Zweierpotenz ist. Zu Ehren ihres ersten Untersuchers Pierre de Fermat (1601–1665) nennt man die Zahlen

$$F_n = 2^{2^n} + 1 \qquad (n \in \mathbb{N}_0)$$

Fermat–Zahlen. Diese Zahlen sind paarweise teilerfremd.

$$(2) \qquad (F_n, F_m) = 1 \quad \text{für} \quad n \neq m$$

(denn $F_{n+k} - 2 = F_n\big(d^{2^k-1} - d^{2^k-2} + \cdots - 1), d = 2^{2^n}\big)$.

Jede unendliche Folge paarweise teilerfremder Zahlen liefert unendlich viele Primteiler. □

Die Tatsache, daß $F_1, \ldots, F_4$ prim sind, führte Fermat zu der Vermutung, daß dies für alle F_n zutrifft. Euler widerlegte es durch das Beispiel

$$F_5 = 641 \cdot 6\,700\,417.$$

Ebenso sind F_6, F_7, F_8 Produkte aus zwei Primzahlen. Bis heute kennt man kein weiteres primes F_n. Die Frage, ob es unter den F_n weitere, oder gar unendlich viele Primzahlen gibt, dürfte für lange Zeit noch unangreifbar sein.

Ein einfaches Verfahren zur Aufstellung von Primzahllisten ist das

1.12. Sieb des Eratosthenes (276?–194? vor ZR)
Sei $N \in \mathbb{N},\ N \geq 2$.
1) Man schreibe die Zahlen $2, \ldots, N$ hin.
2_1) Man streiche die echten Vielfachen von 2.
2_2) Man gehe zur nächsten nicht gestrichenen Zahl und streiche hiervon alle echten Vielfachen, usw.
3) Man höre auf, wenn die nächste ungestrichene Zahl $> N^{1/2}$ ist.

Beh. Die nicht gestrichenen Zahlen sind die Primzahlen $\leq N$.

Beweis. Es geht keine Primzahl verloren, denn es werden nur echte Vielfache von Zahlen ≥ 2 gestrichen.

Jedes zusammengesetzte $n \leq N$ wird gestrichen, denn es hat einen Primteiler $p \leq \sqrt{N}$. Dies p wird nicht gestrichen, n als echtes Vielfaches von p fällt weg. □

Zur Kenntnis der multiplikativen Struktur von $\mathbb{Z}$ ist der folgende Satz grundlegend. Obwohl er intuitiv wesentlich früher benutzt wurde, ist er erst von Gauß in exakter Form angegeben worden.

1.13. Satz von der eindeutigen Primfaktorzerlegung.
Jedes $n > 1$ besitzt genau eine Darstellung („kanonische Zerlegung“)

$$n = p_1^{k_1} \cdot \ldots \cdot p_\ell^{k_\ell} \qquad (2 \leq p_1 < \cdots < p_\ell)$$

(bzw. $n = \prod\limits_p p^{a_p}, a_p \in \mathbb{N}_0$. Dabei ist das Produkt über alle Primzahlen erstreckt; $a_p \neq 0$ nur für endliche viele p. Für $n = 1$ ist $\forall\, p : a_p = 0$).

Beweis. 1. Existenz. Falls n nicht prim ist, zerfällt es in zwei Faktoren $1 < n_1, n_2 < n$. Diese sind Produkte von Primzahlen (Induktion!), also auch n.

2. Eindeutigkeit. Es gebe Zahlen mit zwei Darstellungen, $n(> 1)$ sei unter diesen die kleinste.

$$(*) \qquad n = p_1^{a_1} \cdot \ldots \cdot p_k^{a_k} = q_1^{b_1} \cdot \ldots \cdot q_\ell^{b_\ell} (p_1 < \cdots < p_k;\ q_1 < \cdots < q_\ell;\ a_\nu, b_\nu \in \mathbb{N}).$$

$p_1 \neq q_1, \ldots, q_\ell$, da sonst durch p_1 dividiert werden könnte, und man ein kleineres n' mit zwei Darstellungen erhielte.
$(p_1, q_1) = \cdots = (p_1, q_\ell) = 1$ (denn $(p_1, q_j) > 1$ bedingte $p_1 = q_j$). Aus $(*)$ sieht man

$$p_1 \,|\, q_1 \cdot q_1^{b_1 - 1}\, q_2^{b_2} \ldots q_\ell^{b_\ell} = q_1 n'.$$

1.7 (2) und $(p_1, q_1) = 1$ bewirken $p_1 | n'$. Fortsetzung dieses Verfahrens führt schließlich zu $p_1 | q_\ell$, was wegen $(p_1, q_\ell) = 1$ ausgeschlossen ist. □

Für das Weitere ist eine einfache Feststellung wichtig. Für

$$n = \prod_p p^{a_p}, \quad d = \prod_p p^{b_p} \qquad \text{gilt}$$

$(*)$ $\qquad d|n \Leftrightarrow \forall\, p : b_p \leq a_p.$

Die Richtung $\Leftarrow$ ist klar. Sei $d|n$ und $b_p > a_p$ für mindestens ein p, zum Beispiel für $p = q$. Aus $n = dm$ folgt $n\, q^{-a_p} = d\, q^{b_q - a_q} m$. Links steht ein $n' \in \mathbb{N}$, in dessen Primfaktorzerlegung q nicht vorkommt, während es rechts mit einem Exponenten $\geq b_q - a_q > 0$ auftritt. Dies widerspricht 1.13. □

Bezeichnet $d(n)$ die Anzahl der natürlichen Teiler von n, so gilt

$$d\big(p_1^{a_1} \ldots p_k^{a_k}\big) = (a_1 + 1) \cdot \ldots \cdot (a_k + 1).$$

Mit der Aussage $(*)$ sieht man unmittelbar

1.14. Satz. Sei $n_\nu = \prod_p p^{a_{p,\nu}}$ $(\nu = 1, \ldots, k)$,

$$A_p = \min(a_{p,1}, \ldots, a_{p,k}), \quad B_p = \max(a_{p,1}, \ldots, a_{p,k}).$$

Dann gilt

$$(n_1, \ldots, n_k) = \prod_p p^{A_p}, \quad [n_1, \ldots, n_k] = \prod_p p^{B_p}.$$

Der Satz von der eindeutigen Primfaktorzerlegung besagt, daß jedes Element $\neq 0$ des Ringes $\mathbb{Z}$ eindeutig als Produkt von unzerlegbaren Elementen p und einer Einheit $e \in \{1, -1\}$ geschrieben werden kann. Die nächsteinfachen Bereiche sind die Ringe

$$\mathbb{Z}[\sqrt{a}] = \big\{b_1 + b_2\sqrt{a}; b_1, b_2 \in \mathbb{Z}\big\}$$

für $a \in \mathbb{Z}\backslash\{0\}$, a keine Quadratzahl. Hier können in naheliegender Weise Einheiten und Primzahlen definiert werden. Wie das Beispiel

$$6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5}) \quad \text{in} \quad \mathbb{Z}[\sqrt{-5}]$$

zeigt, kann die ZPE–Eigenschaft verloren gehen. Diese Probleme bilden den Ausgangspunkt zur algebraischen Zahlentheorie.

2. Kapitel. Kongruenzen, Restsysteme.

Bei Division durch eine feste Zahl m bilden die kleinsten nichtnegativen Reste eine m–periodische Folge. Für zahlreiche Fragen reicht es aus, das Verhalten innerhalb einer Periode zu studieren.

2.1. Def. $m \in \mathbb{N}$. a und b heißen **kongruent modulo** m, wenn $m|b - a$, bzw. wenn $b = a + gm$ ist. Geschrieben $a \equiv b \bmod m$ oder $a \equiv b(m)$. m heißt der **Modul der Kongruenz**.

Folgerungen.

(1) $a \equiv b(m) \Leftrightarrow a$ und b lassen bei Division durch m denselben Rest.
(2) $a \equiv b(m), \quad b \equiv c(m) \Rightarrow a \equiv c(m)$.
(3) $a_1 \equiv b_1(m), \quad a_2 \equiv b_2(m) \Rightarrow a_1 + a_2 \equiv b_1 + b_2(m)$, ebenso mit $\cdot$ statt $+$.
(4) f Polynom mit ganzen Koeffizienten, $a \equiv b(m) \Rightarrow f(a) \equiv f(b)(m)$.

(5) $na \equiv nb(m) \Rightarrow a \equiv b\big(\frac{m}{(n,m)}\big)$. Insbesondere $a \equiv b(m)$, falls $(n,m) = 1$.
(6) $a \equiv b(m_j)$ $(j = 1, \ldots, k) \Leftrightarrow a \equiv b\big([m_1, \ldots, m_k]\big)$.
(7) $a \equiv b(m) \Rightarrow (a,m) = (b,m)$.

Die Eigenschaften (1), (2), (3) sind unmittelbar einzusehen, (4) entsteht durch mehrfache Anwendung von (3).

Zu (5). Mit der Definition sieht man $\dfrac{m}{(m,n)} \Big| \dfrac{n}{(m,n)}(b-a)$. Wegen $\Big(\dfrac{m}{(m,n)}, \dfrac{n}{(m,n)}\Big) = 1$ und 1.7.(2) gilt $\dfrac{m}{(m,n)} \big| b-a$, also die Behauptung.
(6) ist klar. (7) ergibt sich aus

$$a \equiv b(m) \Rightarrow (d|m \wedge d|a \Leftrightarrow d|m \wedge d|b).$$

□

Die Relation „$\equiv$ mod m“ ist auf $\mathbb{Z}$ offenbar eine Äquivalenzrelation, zerlegt $\mathbb{Z}$ also in m paarweise disjunkte Äquivalenzklassen.

2.2. Def. Die Äquivalenzklassen der Relation „$\equiv$ mod m“ auf $\mathbb{Z}$ heißen **Restklassen mod m**.

Folgerungen.
(1) Die Restklassen mod m haben die Gestalt

$$x + m\mathbb{Z} \underset{\text{Df}}{=} \{x + ma,\ a \in \mathbb{Z}\} \quad (x \in \mathbb{Z}),$$
$$x_1 + m\mathbb{Z} = x_2 + m\mathbb{Z} \Leftrightarrow x_1 \equiv x_2(m).$$

(2) Durch

$$(x_1 + m\mathbb{Z}) + (x_2 + m\mathbb{Z}) \underset{\text{Df}}{=} (x_1 + x_2) + m\mathbb{Z}$$

wird auf $\mathbb{Z}_m$, der Menge der Restklassen mod m, eine Verknüpfung definiert, die $\mathbb{Z}_m$ zu einer Gruppe macht. $(\mathbb{Z}_m, +)$ heißt die **additive Restklassengruppe mod m**.

(3) $(\mathbb{Z}_m, +)$ ist zyklisch. $a + m\mathbb{Z}$ ist erzeugendes Element genau dann, wenn $(a,m) = 1$.

(1) ist klar. Die Definition der Addition in (2) ist unabhängig von den Repräsentanten x_1 und x_2. Denn ist

$$x_j' \in x_j + m\mathbb{Z}, \quad \text{d.h.} \quad x_j' \equiv x_j(m),$$

dann folgt $x_1' + x_2' \equiv x_1 + x_2(m)$, also

$$(x_1' + m\mathbb{Z}) + (x_2' + m\mathbb{Z}) = (x_1 + x_2) + m\mathbb{Z}.$$

Assoziativität und Kommutativität der Verknüpfung gelten wie in $\mathbb{Z}$. $0 + m\mathbb{Z}$ ist das neutrale Element, zu $x + m\mathbb{Z}$ ist $(m - x) + m\mathbb{Z}$ das Inverse. $1 + m\mathbb{Z}$ ist Erzeugendes der Gruppe. Die letzte Bemerkung wird im Anschluß an den nächsten Begriff begründet.

2.3. Def. $\{x_1, \ldots, x_m\}$ heißt **vollständiges Restsystem mod m**, wenn die x_j paarweise $\not\equiv$ mod m sind, bzw. jede Restklasse mod m genau ein x_j enthält.

Folgerung. $\{x_1, \ldots, x_m\}$ vollständiges Restsystem mod m. $a \in \mathbb{Z}$, $(b, m) = 1 \Rightarrow$ $\{x_1 + a, \ldots, x_m + a\}$, $\{x_1 b, \ldots, x_m b\}$ vollständiges Restsystem mod m. Falls $(b, m) > 1$, ist $\{x_1 b, \ldots, x_m b\}$ kein vollständiges Restsystem mod m.

Mit x und y sind auch $x + a$ und $y + a$ mod m inkongruent. Dasselbe gilt nach 2.1.(5) für xb und yb. Sei $m > 1$, $d = (m, b) > 1$. Dann ist $1 \leq \frac{m}{d} < m$ und $\frac{m}{d} \not\equiv 0(m)$. Gilt oBdA $x_1 \equiv 0(m)$, $x_2 \equiv \frac{m}{d}(m)$, dann folgt

$$x_1 b \equiv 0(m),\ x_2 b \equiv \frac{m}{d}\, b \equiv m \equiv 0(m),$$

d.h. $\{x_1 b, \ldots, x_m b\}$ bildet kein vollständiges Restsystem.

Dies beinhaltet 2.2.(3), denn $x + m\mathbb{Z}$ erzeugt $(\mathbb{Z}_m, +)$ genau dann, wenn $\{x \cdot 0,\ x \cdot 1, \ldots, x \cdot (m-1)\}$ ein vollständiges Restsystem mod m ist. □

$(\mathbb{Z}_m, \cdot)$ ist assoziativ und kommutativ, $1 + m\mathbb{Z}$ wirkt als neutrales Element, aber nicht für alle $x + m\mathbb{Z}$ existiert ein multiplikatives Inverses, z.B. im Fall $m \neq 1$ für $0 + m\mathbb{Z}$.

2.4. Satz und Def.

(1) $a + m\mathbb{Z}$ heißt **reduzierte** oder **prime Restklasse mod m**, wenn $(a, m) = 1$ (nach 2.1.(7) besteht die gesamte Restklasse aus zu m teilerfremden Zahlen, wenn dies für nur ein Element zutrifft).

(2) Zu $a + m\mathbb{Z}$ existiert ein multiplikatives Inverses (d.h. ein a^* mit

$$(a + m\mathbb{Z}) \cdot (a^* + m\mathbb{Z}) = 1 + m\mathbb{Z})$$

genau dann, wenn $(a, m) = 1$.

Die maximale Teilmenge der Restklassen mod m, auf der die Multiplikation zur Gruppen–Eigenschaft führt, ist somit die Menge der reduzierten Restklassen.

Beweis zu (2). Sei $(a, m) = 1$. Nach Satz 1.5. existieren a^* und $z \in \mathbb{Z}$, so daß $aa^* + mz = 1$. Dies bedeutet $aa^* \equiv 1(m)$ oder $(a + m\mathbb{Z}) \cdot (a^* + m\mathbb{Z}) = 1 + m\mathbb{Z}$.
Ist umgekehrt $d = (a, m) > 1$ und $a^* + m\mathbb{Z}$ Inverses zu $a + m\mathbb{Z}$, dann gilt $aa^* \equiv 1(m)$, $aa^* + 1 = gm$. Aus $d|a$ und $d|m$ ergibt sich $d|1$, was nicht sein kann. □

2.5. Def.

(1) Die Anzahl der reduzierten Restklassen mod m wird als $\varphi(m) =$ **Euler–Funktion** bezeichnet. (Leonhard Euler, 1707–1783).

$$\text{Oder:} \qquad \varphi(m) = \#\{1 \leq a \leq m,\ (a, m) = 1\}.$$

(2) Die Menge der $\varphi(m)$ reduzierten Restklassen mod m wird mit $\mathbb{Z}_m^*$ abgekürzt. Die abelsche Gruppe $(\mathbb{Z}_m^*, \cdot)$ (nach 2.4.(2)) heißt **multiplikative Restklassengruppe mod m**.

(3) Jedes Vertretersystem $\{x_1, \ldots, x_{\varphi(m)}\}$ der $\varphi(m)$ reduzierten Restklassen mod m heißt **reduziertes** oder **primes Restsystem mod m**.(„prim“ besagt hier nicht, daß die x_j Primzahlen sein sollen).

Folgerung. Ist $\{x_1, \ldots, x_{\varphi(m)}\}$ ein reduziertes Restsystem mod m und $(a, m) = 1$, dann ist auch $\{ax_1, \ldots, ax_{\varphi(m)}\}$ eins.
Der Beweis verläuft wie der zu 2.3.

Für $m > 1$ sind die $m = p$ offenbar die einzigen Moduln, zu denen alle $a + m\mathbb{Z}$ mit $a \not\equiv 0$ mod m ein multiplikatives Inverses besitzen.

$$\mathbb{Z}_p(+, \cdot) \qquad \text{ist ein Körper.}$$

Für kein anderes m hat $\mathbb{Z}_m(+, \cdot)$ diese Eigenschaft. In der Algebra wird gezeigt, daß es exakt zu den $m = p^k$ $(k \in \mathbb{N})$ einen Körper mit m Elementen gibt. Dieser ist bis auf Isomorphie eindeutig bestimmt und wird als $GF(p^k)$ (= Galois–Feld; Evariste Galois, 1811–1832) bezeichnet. Für $k > 1$ ist die Konstruktion der $GF(p^k)$ um einiges verwickelter als für $k = 1$.

2.6. Satz.
(1) $(m, n) = 1$; $\{x_1, \ldots, x_{\varphi(m)}\}, \{y_1, \ldots, y_{\varphi(n)}\}$ seien reduzierte Restsysteme mod m bzw. n.
Beh. $\{x_j n + y_k m,\ 1 \le j \le \varphi(m),\ 1 \le k \le \varphi(n)\}$
ist ein reduziertes Restsystem mod mn.

(2) $\quad (m, n) = 1 \Rightarrow \varphi(mn) = \varphi(m)\,\varphi(n)$
(d.h. die zahlentheoretische Funktion $\varphi : \mathbb{N} \to \mathbb{N}$ ist „multiplikativ“).

(3) $\quad \varphi(p^k) = (p - 1)\,p^{k-1}$,

$$\varphi(p_1^{a_1} \cdot \cdots \cdot p_\ell^{a_\ell}) = \prod_{j=1}^{\ell} (p_j - 1)\,p_j^{a_j - 1}\ (p_1 < \cdots < p_\ell,\ a_j \ge 1).$$

Beweis zu (1). Die angegebenen $\varphi(n) \cdot \varphi(m)$ Zahlen $\quad z_{jk} = x_j n + y_k m \quad$ sind zu mn teilerfremd. Denn hat mn mit z_{jk} einen Primteiler p gemeinsam, dann oBdA $p | n$, also $p | y_k m = z_{jk} - x_j n$. Wegen $(y_k, n) = 1$ und 1.7.(2) folgt $p | m$, also $p | (m, n) = 1$, was nicht sein kann.

Die z_{jk} sind mod mn paarweise inkongruent. Denn aus

$$x_j n + y_k m \equiv x_{j'} n + y_{k'} m (mn)$$

folgt

$$m \,\big|\, (x_j - x_{j'})n + (y_k - y_{k'})m, \quad m \,\big|\, (x_j - x_{J'})n.$$

Mit $(m, n) = 1$ und 1.7.(2) ergibt sich $m | x_j - x_{j'},\ x_j \equiv x_{j'}(m),\ j = j'$. Ebenso $k = k'$.

Drittens ist jedes z mit $(z, mn) = 1$ zu einem der z_{jk} mod mn, kongruent. Denn nach Satz 1.5. existieren x' und $y' \in \mathbb{Z}$ mit $z = x'n + y'm$. Hier muß $(x', m) = (y', n) = 1$ sein, da ansonsten $(z, mn) > 1$ wäre. Es gibt also $j \le \varphi(m)$ und $k \le \varphi(n)$, so daß

$$x' \equiv x_j(m),\ y' \equiv y_k(n), \quad \text{also} \quad z \equiv x_j n + y_j m (mn).$$

(1) beinhaltet (2).

Zu (3).

$$\begin{aligned}\varphi(p^k) &= \#\{1 \le a \le p^k,\ p \nmid a\} \\ &= \#\{1 \le a \le p^k\} - \#\{1 \le a \le p^k,\ a \equiv 0(p)\} \\ &= p^k - p^{k-1}.\end{aligned}$$

Die letzte Formel entsteht durch mehrfaches Anwenden von (2) und dem Vorigen. □

Die folgende Kongruenz gehört zu den wichtigsten Aussagen der elementaren Zahlentheorie. Zum Abschluß des Kapitels soll auf eine Anwendung in der Kryptografie (RSA–Verfahren) eingegangen werden.

2.7. Eulersche Kongruenz.
Für $(a, m) = 1$ gilt

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

Im Spezialfall $m = p$ und $p \nmid a$ ist insbesondere $a^{p-1} \equiv 1(p)$ bzw. für alle a gilt $a^p \equiv a(p)$ (**Fermat–Kongruenz**).

Beweis. Sind $\{x_1, \ldots, x_{\varphi(m)}\}$ und $\{y_1, \ldots, y_{\varphi(m)}\}$ reduzierte Restsysteme, dann folgt nach eventueller Umbenennung

$$x_j \equiv y_j(m) \qquad \big(1 \le j \le \varphi(m)\big).$$

Mehrfache Anwendung von 2.1.(3) ergibt

$$(*) \qquad P \underset{\text{Df}}{=} \prod_{j=1}^{\varphi(m)} x_j \equiv \prod_{j=1}^{\varphi(m)} y_j \bmod m,\ (P, m) = 1.$$

Nach der Folgerung zu 2.5.(3) darf als $\{y_j\}$ das System $\{ax_j\}$ genommen werden. Dann wird aus $(*)$

$$P \equiv a^{\varphi(m)}\ P \bmod m.$$

Da $(P, m) = 1$, kann nach 2.1.(5) P gekürzt werden. □

Die Eulersche Kongruenz ist ein Spezialfall des gruppentheoretischen Satzes: Ist G eine Gruppe mit n Elementen und dem neutralen Element e. Dann gilt für jedes $g \in G$:

$$g^n = e.$$

Der Beweis beruht auf der gleichen Idee wie der eben ausgeführte.

Während die Struktur der Gruppe $(\mathbb{Z}_m, +)$ auf der Hand liegt, ist $(\mathbb{Z}_m^*, \cdot)$ wesentlich mühsamer zu analysieren. Wie der nächste Satz - der erste in dieser Vorlesung mit einigem Tiefgang - zeigen wird, ist die Zyklizität der Gruppe eher die Ausnahme.

2.8. Def. $(a, m) = 1$. a heißt **Primitivwurzel mod m**, wenn $\{a, a^2, \ldots, a^{\varphi(m)}\}$ ein reduziertes Restsystem mod m bildet. Mit anderen Worten: a ist erzeugendes Element

der Gruppe $(\mathbb{Z}_m^*, \cdot)$.

$(\mathbb{Z}_m^*, \cdot)$ ist genau dann zyklisch, wenn Primitivwurzeln mod m existieren.

2.9. Satz von Euler.
Zu genau den folgenden Moduln m existieren Primitivwurzeln:

$$m = 1, 2, 4, p^k, 2p^k \qquad (p > 2,\ k \in \mathbb{N}).$$

Beweis.
1. Die Überlegungen zu dem Ordnungs–Begriff sind auch unabhängig von diesem Satz vonInteresse.

1.1 Sei $m \in \mathbb{N}$ und $(a, m) = 1$. Nach Fermat–Euler gibt es $d \in \{1, \ldots, \varphi(m)\}$ mit $a^d \equiv 1(m)$. Das kleinste solche d heißt die **Ordnung** von a **mod** $\boldsymbol{m}$ $\big(d = \mathrm{ord}_m(a)\big)$. Dies entspricht dem Begriff der Ordnung des Elements $a + m\mathbb{Z}$ in der Gruppe $(\mathbb{Z}_m^*, \cdot)$. a ist demnach Primitivwurzel mod m genau dann, wenn $\mathrm{ord}_m(a) = \varphi(m)$. Man beachte auch, daß $\mathrm{ord}_m(a)$ nur für $(a, m) = 1$ definiert ist.

1.2. Sei $d = \mathrm{ord}_m(a)$. Dann sind die Zahlen $a^0 = 1,\ a, \ldots, a^{d-1}$ mod m paarweise inkongruent.

Denn aus $a^\ell \equiv a^k \bmod m$ mit $0 \le \ell < k < d$ folgt mit 2.1.(5) $a^{k-\ell} \equiv 1(m)$, im Widerspruch zur Minimalität von d.

1.3. Sei $a^\ell \equiv a^k \bmod m$ $(\ell, k \ge 0)$. Dann ist $\ell \equiv k \bmod d$ $\big(d = \mathrm{ord}_m(a)\big)$. Umgekehrt folgt $a^\ell \equiv a^k(m)$ aus $\ell \equiv k(d)$.

ℓ und k werden mit Rest durch d dividiert.

$$\ell = b_1 d + r_1, \quad k = b_2 d + r_2, \quad 0 \le r_j < d,$$
$$a^\ell = (a^d)^{b_1} a^{r_1} \equiv a^{r_1}(m), \text{ ebenso } a^k \equiv a^{r_2}(m).$$

Nach 1.2. kann $a^\ell \equiv a^k(m)$ nur für $r_1 = r_2$, d.h. $\ell \equiv k(d)$ gelten. Die Umkehrung sieht man unmittelbar.

1.4. $\mathrm{ord}_m(a)$ ist Teiler von $\varphi(m)$.

Denn es gilt $a^{\varphi(m)} \equiv 1 = a^0(m)$, also nach 1.3. $\varphi(m) \equiv 0(d)$.

1.5. Aus $\mathrm{ord}_m(a) = d_1 d_2$ folgt $\mathrm{ord}_m(a^{d_1}) = d_2$.

Sei $d' = \mathrm{ord}_m(a^{d_1})$. Dann gilt einerseits

$$1 \equiv (a^{d_1})^{d'} = a^{d_1 d'} \bmod m,$$

also nach 1.3. $d_1 d_2 \mid d_1 d', d_2 | d'$. Andererseits ist nach Voraussetzung $(a^{d_1})^{d_2} = a^{d_1 d_2} \equiv 1(m)$, also $d' | d_2$. Bleibt nur $d_2 = d'$.

1.6. Aus $\operatorname{ord}_m(a_\nu) = d_\nu$ $(\nu = 1, 2)$ und $(d_1, d_2) = 1$ folgt $\operatorname{ord}_m(a_1 a_2) = d_1 d_2$.

Sei $d' = \operatorname{ord}_m(a_1 a_2)$, also $(a_1 a_2)^{d'} \equiv 1 \bmod m$. Potenzieren mit d_1 ergibt

$$a_1^{d_1 d'} a_2^{d_1 d'} \equiv 1 \mod m$$

Wegen $a_1^{d_1 d'} \equiv 1 \bmod m$ wird daraus $a_2^{d_1 d'} \equiv 1 \bmod m$. Anwendung von 1.3. ergibt $d_2 | d_1 d'$, und mit 1.7.(2) $d_2 | d'$. Ebenso sieht man $d_1 | d'$. Mit $(d_1, d_2) = 1$ folgt

(1) $\qquad d_1 d_2 | d'$.

Umgekehrt sieht man

$$(a_1 a_2)^{d_1 d_2} = (a_1^{d_1})^{d_2} (a_2^{d_2})^{d_1} \equiv 1 \mod m$$

und daraus mit 1.3. $d' | d_1 d_2$.

2. Zum Beweis wird Satz 3.5. (Lagrange) benötigt, dessen Herleitung hier vorgezogen wird:
Sei $f(x) = a_0 + a_1 + \cdots + a_n x^n \in \mathbb{Z}[x]$ ein Polynom mit ganzen Koeffizienten, p eine Primzahl mit $p \nmid a_n$. Dann existieren höchstens n Zahlen $x \in \{0, 1, \ldots, p-1\}$ (bzw. n Elemente eines vollständigen RS mod p) mit

$(*)$ $\qquad f(x) \equiv 0 (\bmod p)$.

Beweis. Die Aussage ist klar, wenn man bedenkt, daß „$\equiv \bmod p$ in $\mathbb{Z}$“ dasselbe bedeutet wie „$=$ im Körper $\mathbb{Z}_p$“. Die Bedingung $p \nmid a_n$ besagt, daß f als Polynom über $\mathbb{Z}_p$ den Grad n hat. Nach einem allgemeinen Satz der Algebra - dessen Beweis im Folgenden im Prinzip gegeben wird - besitzt f über $\mathbb{Z}_p$ höchstens n Nullstellen, was der Behauptung entspricht.

Seien – im Fall daß es überhaupt Lösungen gibt – $x_1, \ldots, x_r \in \{0, \ldots, p-1\}$ die Lösungen zu $(*)$. Polynomdivision ergibt

$$f(x) = (x - x_1)\, g(x) + b_1$$

mit einem Polynom g vom Grad $n-1$, dessen Leitkoeffizient $\not\equiv 0 \bmod p$ ist. $x = x_1$ zeigt $b_1 \equiv 0 \bmod p$. Für $2 \le \nu \le r$ folgt

$$0 \equiv f(x_\nu) \equiv (x_\nu - x_1)\, g(x_\nu) \bmod p,$$

also $g(x_\nu) \equiv 0 \bmod p$. Die Behauptung ergibt sich hiermit leicht induktiv. Für $n = 1$ lautet $(*)$

$$a_1 x + a_0 \equiv 0 \bmod p.$$

Wegen $(a_1, p) = 1$ existiert ein a_1^* mit $a_1 a_1^* \equiv 1 \bmod p$, also

$$x + a_0 a_1^* \equiv 0 \bmod p,$$

d.h. im Fall $n = 1$ gibt es mod p genau ein x.
Hinweis: Es darf hieraus nicht der Schluß gezogen werden, daß die Lösungszahl stets $= n$ ist. Es kann z.B. sein, daß für $n \geq 2$ gar keine Lösungen existieren.

3. Beweis des Satzes für $m = p > 2$.
3.1. Seien $d_1, \ldots, d_k$ alle auftretenden Ordnungen mod p,

$$d = [d_1, \ldots, d_k].$$

Es wird sich herausstellen, daß $d = \varphi(p) = p - 1$ ist und selbst als Ordnung angenommen wird, was der Behauptung entspricht.

3.2. Da alle d_j nach 1.4. $p-1$ teilen, gilt

$$d | p - 1, \qquad d \leq p - 1.$$

3.3. Für jedes $a \in \mathbb{Z}_p^*$ ist $a^d \equiv 1 \bmod p$, da $\mathrm{ord}_p(a)$ die Zahl d teilt. Die Kongruenz

$$x^d - 1 \equiv 0 \bmod p$$

hat $p-1$ Lösungen, nämlich alle x mit $p \nmid x$. Nach dem Satz von Lagrange muß $d \geq p-1$ sein, mit 3.2. also $d = p - 1$. Wegen $p > 2$ ist auch $d > 1$.

3.4. Sei $d = q_1^{b_1} \ldots q_l^{b_l}$ die kanonische Zerlegung von d. Es gibt nach Definition von d ein $c_1 \in \mathbb{Z}_p^*$ mit

$$\mathrm{ord}_p(c_1) = q_1^{b_1} \cdot d_1', \quad q_1 \nmid d_1'.$$

Nach 1.5. existiert $a_1 (= c_1^{d_1'})$ mit $\mathrm{ord}_p(a_1) = q_1^{b_1}$. Analog $a_2, \ldots, a_k$. Da die $q_j^{b_j}$ paarweise teilerfremd sind, ergibt 1.6.

$$\mathrm{ord}_p(a_1 \ldots a_k) = q_1^{b_1} \ldots q_k^{b_k} = d.$$

Mit dem Vorigen hat man $\mathrm{ord}_p(a_1 \ldots a_k) = p - 1$, wie behauptet.

4. Die Existenz von Primitivwurzeln mod p^k $(k > 1)$ bzw. $2p^k$.

4.1. Es sei g eine Primitivwurzel mod $p (> 2)$. Es werden die Zahlen

$$c_\nu = (g + p\nu)^{p-1} \qquad (\nu = 0, \ldots, p-1)$$

betrachtet.

$$\begin{aligned} c_0 &= g^{p-1} = 1 + b_0 p \\ c_\nu &= (g + p\nu)^{p-1} = g^{p-1} + (p-1)\, g^{p-2} p\nu + p^2\, y_\nu \\ &= 1 + p\big(b_0 - g^{p-2}\,\nu + p(y_\nu + g^{p-2}\nu)\big) = 1 + b_\nu\, p. \end{aligned}$$

Wegen $(g, p) = 1$ durchlaufen mit ν die b_ν ein volles Restsystem mod p. Insbesondere gibt es ein $\nu \in \{0, \ldots, p-1\}$, so daß $p \nmid b_\nu$. Dieses ν werde im Folgenden benutzt.

4.2. Sei $\operatorname{ord}_{p^k}(g+p\nu)=d$. Dann gilt $(g+p\nu)^d\equiv 1 \bmod p$, und, da mit g auch $g+p\nu$ Primitivwurzel mod p ist, $p-1|d$.

4.3 Nach 1.4. gilt $d|\varphi(p^k)=p^{k-1}(p-1)$, also mit 4.2.

$$d=p^{\ell-1}(p-1), \quad \text{wobei} \quad 1\le \ell \le k.$$

4.4 Aus $(g+p\nu)^{p-1}=1+b_\nu p$ folgt schrittweise - und hier wird $p>2$ benutzt -

$$(g+p\nu)^{p(p-1)}=1+p^2\,b_{\nu,1},$$
$$(g+p\nu)^{p^2(p-1)}=1+p^3\,b_{\nu,2}, \text{ usw.}$$

mit zu p teilerfremden $b_{\nu,1}, b_{\nu,2}, \ldots$ Aus der Kongruenz

$$(g+p\nu)^d=(g+p\nu)^{p^{\ell-1}(p-1)}\equiv 1 \bmod p^k$$

wird daher

$$1+p^\ell\, b_{\nu,\ell-1}\equiv 1 \bmod p^k.$$

Also ist $\ell=k$ und $g+p\nu$ Primitivwurzel mod p^k.

4.5. Ist für ungerades p g Primitivwurzel mod p^k, so ist die gerade unter den Zahlen g und $g+p^k$ eine zu $2p^k$.

Denn nach Satz 2.5. ist $\varphi(2p^k)=\varphi(p^k)$. Falls ein ungerades x die Kongruenz $x^d\equiv 1$ mod p^k erfüllt, dann auch mod $2p^k$, und umgekehrt. Für ungerade x ist also

$$\operatorname{ord}_{p^k}(x)=\operatorname{ord}_{2p^k}(x).$$

5. $m=2^k$ besitzt Primitivwurzeln nur für $k=1$ und $k=2$. $g=1$ ist eine mod 2, $g=3$ eine mod 4.

Für $k\ge 3$ ist $\varphi(2^k)=2^{k-1}$. Ein ungerades a hat jedoch mod 2^k höchstens die Ordnung 2^{k-2}. Denn man sieht induktiv

$$a^{2^1}=1+8\,b_1,$$
$$a^{2^2}=1+16\,b_2,$$
$$a^{2^{k-2}}=1+2^k\,b_{k-2}\equiv 1 \bmod 2^k.$$

Es sei bemerkt, daß für $k\ge 3$ die 2^{k-1} Zahlen

$$\pm 5^0, \pm 5^1, \ldots, \pm 5^{2^{k-2}-1}$$

ein reduziertes Restsystem mod 2^k bilden. Man benutzt

$$5^{2^{\ell-3}}\equiv 1+2^{\ell-1} \bmod 2^\ell \quad (\ell\ge 3).$$

6. Sei $1<m=p_1^{k_1}\cdot\ldots\cdot p_\ell^{k_\ell}$ in kanonischer Zerlegung gegeben und $(a,m)=1$. Für jedes $\nu\le\ell$ gilt

$$a^{\varphi(p_\nu^{k_\nu})}\equiv 1 \bmod p_\nu^{k_\nu}.$$

Ist $d = \big[\varphi(p_1^{k_1}), \ldots, \varphi(p_\ell^{k_\nu})\big]$, so folgt $a^d \equiv 1 \bmod p_\nu^{k_\nu}\ \forall\, \nu \leq \ell$, also $a^d \equiv 1 \bmod m$. Wegen $d|\varphi(m)$ existieren Primitivwurzeln mod m nur, wenn

$$\varphi(p_1^{k_1}) \cdot \ldots \cdot \varphi(p_\ell^{k_\ell}) = \varphi(m) = d = \big[\varphi(p_1^{k_1}), \ldots, \varphi(p_\ell^{k_\ell})\big].$$

Falls m mindestens zwei ungerade Primteiler besitzt, ist $d \leq \frac{\varphi(m)}{2}$. Im Fall $m = 2^{k_1}\, p_2^{k_2}, k_1 \geq 2, p_2 > 2$ gilt dies ebenfalls.

Also bleiben wegen 5. nur die im Satz genannten m.

7. Falls zu m eine Primitivwurzel g existiert, bilden die Zahlen g^ν, $1 \leq \nu \leq \varphi(m)$ ein reduziertes Restsystem mod m. Man überzeugt sich leicht, daß g^ν genau dann Primitivwurzel ist, wenn $\big(\nu, \varphi(m)\big) = 1$ gilt. Somit gibt es zu den genannten m genau $\varphi\big(\varphi(m)\big)$ mod m verschiedene Primitivwurzeln.

8. Die Struktur der abelschen Gruppe $\mathbb{Z}_m^*$ kann vollständig beschrieben werden. Ist $1 < m = p_1^{k_1} \cdot \ldots \cdot p_\ell^{k_\ell}$, dann zeigt Satz 2.6.

$$\mathbb{Z}_m^* = \mathbb{Z}_{p_1^{k_1}}^* \times \ldots \times \mathbb{Z}_{p_\ell^{k_\ell}}^*$$

($\times$ bedeutet das direkte Produkt). Für $p_\nu > 2$ ist $\mathbb{Z}_{p_\nu^{k_\nu}}^*$ isomorph zur zyklischen Gruppe $\big(\mathbb{Z}_{\varphi(p_\nu^{k_\nu})}, +\big)$. Für $p = 2$ und $k \geq 3$ ist nach der Bemerkung in 5.

$$\mathbb{Z}_{2^k}^* \simeq \mathbb{Z}_2 \times \mathbb{Z}_{2^{k-2}}$$

(links mit der Multiplikation, rechts mit der Addition). □

Im folgenden Satz wird die Darstellbarkeit natürlicher Zahlen in Ziffernsystemen behandelt, im Anschluß daran die wohlbekannten Teilbarkeitsregeln im Zehnersystem.

2.10. Satz. Sei $g \geq 2$. Dann existieren zu jedem $n \in \mathbb{N}$ eindeutig ein $k \geq 0$ ($k + 1 =$ Stellenzahl) und $a_0, \ldots, a_k \in \{0, \ldots, g-1\}$ mit $a_k \geq 1$ (Ziffern), so daß

$$n = \sum_{\nu=0}^{k} a_\nu\, g^\nu.$$

Beweis. Zu jedem $n \in \mathbb{N}$ existiert genau ein $k \in \mathbb{N}_0$ mit

$$g^k \leq n < g^{k+1}.$$

Der Existenzbeweis wird durch Induktion nach k geführt. Für $k = 0$ ist nichts zu zeigen. Sei $g^{k+1} \leq n < g^{k+2}$ für $k \geq 0$. Man setze

$$n_1 = n - [n/g^{k+1}]g^{k+1}.$$

Mit der Definition der Gauß–Klammer folgt $0 \leq n_1 < g^{k+1}$, d.h. auf n_1 ist die Induktionsvoraussetzung anwendbar. Wegen $1 \leq \frac{n}{g^{k+1}} < g$ ist $1 \leq \big[\frac{n}{g^{k+1}}\big] < g$, d.h. $\big[\frac{n}{g^{k+1}}\big]$ ist als Ziffer $a_{k+1} \neq 0$ verwendbar.
Zur Eindeutigkeit.

Seien $\sum_{\nu=0}^{k} a_\nu\, g^\nu$ und $\sum_{\nu=0}^{r} b_\nu\, g^\nu$ zwei verschiedene Darstellungen von n, sei $\ell \geq 0$ der größte Index ν mit $a_\nu \neq b_\nu$. Gelte oBdA $b_\nu > a_\nu$. Dann folgt

$$(b_\ell - a_\ell)\, g^\ell = \sum_{\nu=0}^{\ell-1} (a_\nu - b_\nu)\, g^\nu.$$

Im Fall $\ell = 0$ ist dies offensichtlich widersprüchlich. Für $\ell \geq 1$ ist die linke Seite $\geq g^\ell$, während die rechte Seite im Betrag

$$\leq (g-1) \sum_{\nu=0}^{\ell-1} g^\nu = (g-1)\frac{g^\ell - 1}{g-1} < g^\ell$$

ist, was nicht zusammenpaßt. □

Die wohlbekannten Teilbarkeitsregeln erweisen sich als einfache Anwendung der Kongruenzrechnung.

2.11. Satz. Sei $n = \sum_{\nu=0}^{k} a_\nu\, 10^\nu$ mit $0 \leq a_\nu < 10$. Dann gelten die folgenden Teilbarkeitsregeln.

(1) $2|n \Leftrightarrow 2|a_0$,
(2) $4|n \Leftrightarrow 4|a_0 + 10\, a_1$,
(3) $8|n \Leftrightarrow 8|a_0 + 10\, a_1 + 100\, a_2$,
(4) $5|n \Leftrightarrow 5|a_0$,
(5) $3|n \Leftrightarrow 3|a_0 + \cdots + a_k$,
(6) $9|n \Leftrightarrow 9|a_0 + \cdots + a_k$,
(7) $11|n \Leftrightarrow 11|a_0 - a_1 + a_2 - \cdots + (-1)^k\, a_k$.

Das Beweisprinzip soll an Hand von (7) illustriert werden. Wegen $10 \equiv -1 \bmod 11$ ist $10^{2\nu} \equiv 1 \bmod 11$ und $10^{2\nu+1} \equiv -1 \bmod 11$.
Also gilt

$$\begin{aligned} 11|n &\Leftrightarrow a_0 + 10\, a_1 + \cdots + 10^k a_k \equiv 0(11) \\ &\Leftrightarrow a_0 - a_1 + a_2 - + \cdots + (-1)^k\, a_k \equiv 0(11). \end{aligned}$$

□

Zum Abschluß des Kapitels ein paar Bemerkungen zum **RSA–Verfahren** (= R.L. Rivest, A. Shamir, L. Adleman, 1978), einem der heute gebräuchlichsten Verfahren der Kryptografie (= Ver– und Entschlüsseln von Nachrichten). Es beruht auf der Erfahrungstatsache, daß es numerisch wesentlich einfacher ist, von einer natürlichen Zahl festzustellen ob sie Primzahl ist (Primzahltests, bei 500 stelligen Dezimalzahlen heute in wenigen Minuten machbar), als sie in Primfaktoren zu zerlegen (Faktorisierungsverfahren, bei mehr als 200 Stellen i.a. praktisch unmöglich).

Seien p_1, p_2 zwei große Primzahlen mit etwa gleich vielen Stellen, $N = p_1 p_2$. Die Zerlegung von N und damit auch $\varphi(N)$ sei nur dem Versender S (z.B. Bank) bekannt. N kann

veröffentlicht werden, da Nicht–Eingeweihte nicht in der Lage sind, es zu faktorisieren. Eine Nachricht k (= Klartext), mit $1 < k < \min(p_1, p_2)$, im Dezimalsystem dargestellt, soll dem Empfänger R in verschlüsselter Form $v(k)$ zugeschickt werden. Nur R soll in der Lage sein, aus $v(k)$ auf k zurückzuschließen. R wird ein Zahl t mit $\big(t, \varphi(N)\big) = 1$ zugeordnet. R und nur R erfährt den Schlüssel s, eine Zahl mit $ts \equiv 1 \bmod \varphi(N)$. Die Verschlüsselung des Textes k geschieht durch

$$k \to v(k) = k^t \mod N,$$

R entschlüsselt gemäß

$$v(k) \to \big(v(k)\big)^s \equiv k^{st} = k^{1+g\varphi(N)} \equiv k \bmod N$$

(nach Euler, da $(k, N) = 1$).

Alle angebenen Rechnungen, Auffinden von Primzahlen, Angeben großer t mit $\big(t, \varphi(N)\big) = 1$, Potenzieren, sind mit vertretbarem Rechenaufwand ausführbar. Wegen der - bislang nicht bewiesenen - höheren Komplexität des Faktorisierens kann ein von R verschiedener Empfänger mit $v(k)$ nichts anfangen, selbst wenn er über N und t verfügt. Nur die Kenntnis von s, bzw. von t und $\varphi(N)$ könnte ihm beim Entschlüsseln helfen.

3. Kapitel. Kongruenzen in einer Unbekannten.

3.1. Def. Sei $f(x) = a_0 + a_1 x + \cdots + a_n x^n$ ein Polynom mit ganzen Koeffizienten, $m \in \mathbb{N}$. Die Anzahl $\rho(m) = \rho(m, f)$ der $x \in \{0, \ldots, m-1\}$ (bzw. irgendeinem vollständigen Restsystem mod m) mit

$$f(x) \equiv 0(m)$$

heißt die **Lösungszahl der Kongruenz**.

3.2. Satz. Die lineare Kongruenz

$$ax + b \equiv 0(m)$$

ist lösbar genau dann, wenn $d = (a, m)$ Teiler von b ist. Im Fall der Lösbarkeit gilt $\rho(m) = d$.

Beweis. 1. Im Fall der Lösbarkeit existieren x und $g \in \mathbb{Z}$ mit $ax + b = gm$. Dann gilt $d|b$.

2. Es werde $d|b$ vorausgesetzt. Nach Definition ist die Ausgangskongruenz

$$(*) \qquad ax + b \equiv 0(m)$$

äquivalent zu

$$(**) \qquad \frac{a}{d}\,x + \frac{b}{d} \equiv 0\Big(\frac{m}{d}\Big)$$

mit $\Big(\frac{a}{d}, \frac{m}{d}\Big) = 1$. Wie im vorigen Kapitel sieht man mit dem Euklidischen Algorithmus (oder mit Hilfe der Eulerschen Kongruenz), daß ein $\tilde{a}$ mit $\frac{a}{d} \cdot \tilde{a} \equiv 1\Big(\frac{m}{d}\Big)$ existiert.

Dadurch wird $(**)$ äquivalent zu

$$(***) \qquad x \equiv -\frac{b}{d}\,\tilde{a}\Big(\frac{m}{d}\Big),$$

d.h. $(**)$ hat genau eine Lösung. Die d Zahlen

$$x_0\Big(\underset{\text{Df}}{=} -\frac{b}{d}\,\tilde{a}\Big), \quad x_0+\frac{m}{d},\ldots,x_0+(d-1)\frac{m}{d}$$

sind modulo m verschieden. Jedes x, das $(***)$ erfüllt, ist modulo m zu einer dieser Zahlen kongruent. Die einzige Lösung x_0 mod m/d induziert somit d Lösungen mod m. □

3.3. Chinesischer Restsatz.
Seien $m_1,\ldots,m_k$ paarweise teilerfremd.
Beh. Zu jedem k–Tupel $(x_1,\ldots,x_k)$ gibt es modulo $m_1\cdot\ldots\cdot m_k$ genau ein x_0, so daß

$$x\equiv x_j(m_j) \qquad \forall\, j\le k \qquad \Leftrightarrow\ x\equiv x_0\,(m_1\cdot\ldots\cdot m_k).$$

Andere Formulierung: Der Durchschnitt der Restklassen $x_j+m_j\mathbb{Z}$ $(j=1,\ldots,k)$ ist gleich einer Restklasse $x_0+m_1\cdot\ldots\cdot m_k\mathbb{Z}$.

Bemerkungen.
1. Die Namensgebung geht zurück auf chinesische Quellen um 1200. Das Prinzip des Satzes, Ersetzung eines Systems von Kongruenzen durch eine einzige, tritt unabhängig davon in zahlreichen früheren Schriften auf.

2. Die Bedingung der paarweisen Teilerfremdheit ist nötig. Man kann sich leicht überlegen, daß andernfalls die Aussage modifiziert werden muß: Gar keine Lösung mod $m_1\cdot\ldots\cdot m_k$ oder mehr als eine.

Beweis.
1. Das gesuchte x_0 kann explizit angegeben werden. Sei $m=m_1\cdot\ldots\cdot m_k$, $M_j=\dfrac{m}{m_j}$.
Dann bewirkt die Voraussetzung $(M_j,m_j)=1$ und es existieren M_j^* mit

$$M_j\,M_j^*\equiv 1\,(m_j) \qquad (j=1,\ldots,k).$$

Es werde - bei gegebenem k-Tupel $(x_1,\ldots,x_k)$ -

$$x_0=M_1\,M_1^*\,x_1+\cdots+M_k\,M_k^*\,x_k$$

gesetzt.

2. Es gelte $x\equiv x_0(m)$. Wegen $m_1|M_2,\ldots,m_1|M_k$ folgt

$$x\equiv M_1\,M_1^*\,x_1+0\equiv 1\cdot x_1=x_1(m_1).$$

Ebenso $x\equiv x_j(m_j),\ 2\le j\le k$.

3. x erfülle

$$x\equiv x_j(m_j) \quad \text{für alle} \quad j\le k.$$

Da $x_j \equiv M_j M_j^* x_j (m_j)$ und für $\ell \neq j$ $M_\ell M_\ell^* x_\ell \equiv 0(m_j)$, folgt

$$x \equiv x_0(m_j) \quad \text{für alle} \quad j \leq k.$$

Wegen der paarweisen Teilerfremdheit ergibt sich daraus $x \equiv x_0(m)$.

4. Es ist klar, daß die Lösungen x_0 und $\tilde{x}_0$ mod m zu Tupeln $(x_1, \ldots, x_k)$ und $(\tilde{x}_1, \ldots, \tilde{x}_k)$ mod m verschieden sind, wenn die Tupel sich in mindestens einer Komponente mod m_j unterscheiden. □

3.4. Satz. Aus $(m_1, m_2) = 1$ folgt $\rho(m_1 m_2) = \rho(m_1)\rho(m_2)$ (d.h. die Lösungsanzahl $\rho(m, f)$ ist „multiplikativ“ im Modul).

Der Beweis ist eine einfache aussagenlogische Anwendung des chinesischen Restsatzes. Seien $x_1, \ldots, x_k$ $\big(k = \rho(m_1)\big)$ Vertreter der Lösungsrestklassen mod m_1, ebenso $y_1, \ldots, y_\ell$ $\big(\ell = \rho(m_2)\big)$ mod m_2. Wegen $(m_1, m_2) = 1$ ist

$$(*) \qquad f(x) \equiv 0(m_1 m_2)$$

äquivalent zu

$$f(x) \equiv 0(m_1) \quad \wedge \quad f(x) \equiv 0(m_2)$$

und dies gleichbedeutend mit

$$\big(x \equiv x_1(m_1) \vee \cdots \vee x \equiv x_k(m_1)\big) \wedge \big(x \equiv y_1(m_2) \vee \cdots \vee x \equiv y_\ell(m_2)\big)$$

Dies wiederum läßt sich als Disjunktion von $k \cdot \ell$ Zweiersystemen

$$x \equiv x_j(m_1) \wedge x \equiv y_r(m_2)$$

schreiben. Jedes entspricht nach 3.3. einer Kongruenz

$$x \equiv x_{jr}(m_1 m_2).$$

Verschiedenen Paaren (x_j, y_r) entsprechen mod $m_1 m_2$ verschiedene x_{jr}. $(*)$ hat demnach $\rho(m_1)\rho(m_2)$ Lösungen. □

Der chinesische Restsatz zeigt, wie aus den Lösungen $\bmod m_1$ und $\bmod m_2$ die zu $m_1 m_2$ konstruiert werden können.
Es reicht demnach aus, Kongruenzen $f(x) \equiv 0(p^k)$ zu betrachten und durch mehrfache Anwendung von 3.4. auf die Lösungen zu beliebigen m aufzusteigen. Im weiteren wird gezeigt, daß es im Prinzip ausreicht, Primzahlmoduln zu behandeln. Befriedigende Aussagen, was Lösbarkeit, Lösungsanzahl und die Gestalt der Lösungen angeht, sind nur im Fall linearer oder quadratischer Kongruenzen möglich.

3.5. Satz von Lagrange.
Sei $f(x) = a_0 + a_1 x + \cdots + a_n x^n, \quad p \nmid a_n$. Dann gilt

$$\rho(p, f) \leq n.$$

Der Beweis wurde im vorigen Kapitel gegeben. Daß die Lösungsanzahl stark schwanken kann, zeigen folgende Beispiele. Für die Kongruenz

$$f(x) = x^3 + 2x - 7 \equiv 0 \bmod p$$

berechnet man

$$\rho(2) = 1, \quad \rho(3) = 0, \quad \rho(5) = 2, \quad \rho(7) = 1.$$

Die Kongruenz

$$f(x) = x^p - x \equiv 0 \bmod p$$

hat nach Fermat p Lösungen.

Eine Anwendung, die allerdings für praktische Primzahltests ungeeignet ist, ist der

3.6. Satz von Wilson (M.B. Wilson, 1741–1793).
Sei $n > 1$. Dann sind äquivalent

a) n ist Primzahl,
b) $(n-1)! \equiv -1 \bmod n$.

Beweis.
1. Ist n zusammengesetzt, etwa $n = pn'$ mit $n' > 1$, dann teilt p die Zahlen n und $(n-1)!$. Die Kongruenz b) kann also nicht bestehen.

2. Sei p prim und oBdA > 2. Die Kongruenz

$$(x-1)(x-2)\dots\big(x-(p-1)\big) - (x^{p-1}-1) \equiv 0(p)$$

wird nach Fermat–Euler von den Zahlen $x = 1, \dots, p-1$ gelöst. Die linke Seite, als Polynom geschrieben, hat den Grad $p-2$. Ist einer der Koeffizienten $\not\equiv 0 \bmod p$, dann ergibt sich ein Widerspruch zum Satz von Lagrange. Insbesondere hat man

$$a_0 = \prod_{n=1}^{p-1}(-n) + 1 = (p-1)! + 1 \equiv 0(p). \qquad \square$$

Im folgenden Satz wird beschrieben, wie man von den Lösungen der Kongruenz $f(x) \equiv 0 \bmod p$ zu denen $\bmod p^2, \bmod p^3, \dots$ aufsteigen kann.

Es ist klar, daß für $k \geq 1$ aus $f(x) \equiv 0(p^{k+1})$ die schwächere Bedingung $f(x) \equiv 0(p^k)$ folgt. Eine Restklasse $x_0 + p^k\mathbb{Z} \bmod p^k$ zerfällt in p Restklassen

$$(*) \qquad x_0 + bp^k + p^{k+1}\mathbb{Z} \quad (0 \leq b < p) \bmod p^{k+1}.$$

Sei also $x_0 + p^k\mathbb{Z}$ eine Lösungsrestklasse mod p^k. Dann prüft man, welche der Restklassen $(*)$ Lösungen mod p^{k+1} sind. So kann man von allen Lösungen mod p^k auf die mod p^{k+1} schließen.

3.7. Satz. Sei $k \geq 1, \quad f(x_0) \equiv 0(p^k)$.
Sei $g = g(x_0, k)$ die Anzahl der $b \in \{0, \dots, p-1\}$, für die $x_0 + bp^k$ Lösung mod p^{k+1} ist. Dann gilt

1. $g = 1$, falls $f'(x_0) \not\equiv 0(p)$,
2. $g = p$, falls $f'(x_0) \equiv 0(p)$ und $f(x_0) \equiv 0(p^{k+1})$,
3. $g = 0$, falls $f'(x_0) \equiv 0(p)$ und $f(x_0) \not\equiv 0(p^{k+1})$.

Beweis. Durch Taylor–Entwicklung sieht man

$$\begin{aligned} f(x_0 + bp^k) &= f(x_0) + bp^k\, f'(x_0) + cp^{k+1} \\ &\equiv f(x_0) + bp^k\, f'(x_0) \bmod p^{k+1}. \end{aligned}$$

Hierbei wurde benutzt, daß die Faktoren $f^{(\nu)}(x_0)/\nu!$ ganzzahlig sind. $x_0 + bp^k$ ist somit Lösung mod p^{k+1} genau dann, wenn

$$f(x_0)p^{-k} + bf'(x_0) \equiv 0 \bmod p \tag{1}$$

gilt.

1. Fall: $f'(x_0) \not\equiv 0(p)$, also $\big(p, f'(x_0)\big) = 1$. Nach Satz 3.2. hat (1) genau eine Lösung b mod p. Das ist 1.

2. Fall: $f'(x_0) \equiv 0(p), \quad f(x_0) \equiv 0(p^{k+1})$. Dann erfüllen alle b die Kongruenz (1).

3. Fall: $f'(x_0) \equiv 0(p), \quad f(x_0) \not\equiv 0(p^{k+1})$. Hier ist $\big(p, f'(x_0)\big) = p$, aber $p \nmid f(x_0)p^{-k}$, (1) wird von keinem b gelöst. □

Beispiel. $f(x) = x^4 + 7x + 4$. Durch Einsetzen sieht man, daß $x_0 = 1$ die einzige Lösung mod 3 ist. Wegen $f'(x_0) = 4 + 7 \equiv 2(3)$ tritt der erste Fall ein. (1) wird zu

$$\frac{12}{3} + 11b \equiv 0(3), \quad \text{d.h.} \quad 4 + 2b \equiv 0(3), \quad \text{d.h.} \quad b \equiv 1(3).$$

Somit ist $x_0 + 1 \cdot 3 \equiv 4 \bmod 9$ die einzige Lösung mod 9.
Die zweite Hälfte des Kapitels ist der von Euler und Gauß entwickelten Theorie der quadratischen Kongruenzen gewidmet. Diese stellt einen Höhepunkt der elementaren Zahlentheorie dar.

Es sei zuerst bemerkt, daß es ausreicht, Kongruenzen der Gestalt

$$x^2 \equiv a \bmod p, \quad p > 2, \quad p \nmid a$$

zu betrachten. Der Fall $p = 2$ kann offenbar ausgenommen werden, ebenso $a \equiv 0(p)$.
Ist

$$f(x) = a_2x^2 + a_1x + a_0, \quad p > 2, \quad (a_2, p) = 1,$$

so sind die Kongruenzen

$$f(x) \equiv 0 \quad \text{und} \quad (2a_2x + a_1)^2 \equiv a_1^2 - 4a_0a_2 \bmod p$$

äquivalent. Man löse also zuerst

$$y^2 \equiv a_1^2 - 4a_0a_2 \bmod p$$

und sodann für jede Lösung y_0 hiervon

$$2a_2x + a_1 \equiv y_0 \bmod p.$$

Der Übergang von p zu p^2, usw. wird am Ende diskutiert.

3.8. Def. $p > 2, \quad (a,p) = 1.$

(1) a heißt **quadratischer Rest mod p** (qR), falls die Kongruenz $x^2 \equiv a(p)$ lösbar ist. Andernfalls heißt a **quadratischer Nicht–Rest mod p** (qNR).

(2) $\left(\frac{a}{p}\right) \underset{\text{Df}}{=} \begin{cases} 1, & \text{falls} \quad a\ qR \mod p, \\ -1, & \text{falls} \quad a\ qNR \mod p. \end{cases}$

(**Legendre–Symbol** „a über p“, Adrien–Marie L., 1752–1833).

Man bedenke, daß $\left(\frac{a}{p}\right)$ nur für a mit $p \nmid a$ definiert ist.

Folgerungen.

(1) $\qquad a \equiv b(p) \Rightarrow \left(\frac{a}{p}\right) = \left(\frac{b}{p}\right).$

(2) $\qquad \left(\frac{a^2}{p}\right) = 1.$

(3) Unter den Zahlen $1, \ldots, p-1$ sind $\frac{p-1}{2}$ qR und $\frac{p-1}{2}$ qNR mod p. Anders ausgedrückt: $\sum\limits_{j=1}^{p-1} \left(\frac{j}{p}\right) = 0.$

(4) Im Fall $\left(\frac{a}{p}\right) = 1$ hat die Kongruenz $x^2 \equiv a(p)$ genau zwei Lösungen.

Zu (3). Die Zahlen

$$(*) \qquad 1^2, 2^2, \ldots, \left(\frac{p-1}{2}\right)^2$$

sind mod p paarweise inkongruent. Denn $k^2 \equiv \ell^2 \mod p \ \left(1 \le k, \ell \le \frac{p-1}{2}\right)$ impliziert $(k-\ell)(k+\ell) \equiv 0(p)$. Wegen $2 \le k+\ell \le p-1, \quad p \nmid k+\ell$ folgt $\quad p \mid k-\ell, \quad$ also $k = \ell, \quad$ da $\quad |k-\ell| < p.$

Jede Zahl $x^2, \quad p \nmid x$ ist zu einer der Zahlen $(*)$ mod p kongruent. Denn sei

$$x^2 = (y+cp)^2 \equiv y^2 \mod p \quad \text{mit} \quad 1 \le y \le p-1.$$

Im Fall $\frac{p-1}{2} + 1 \le y \le p-1$ ist aber $1 \le p - y \le \frac{p-1}{2}$ und $y^2 \equiv (p-y)^2 \mod p$.
Ein quadratischer Rest a mod p ist daher zu einem der $\frac{p-1}{2}$ qR aus $(*)$ kongruent, d.h. es existieren je $\frac{p-1}{2}$ qR und qNR mod p.

Zu (4). Die Kongruenzen $x^2 \equiv a \mod p$, wobei a alle $\frac{p-1}{2}$ qR durchläuft, haben zusammen $p-1$ Lösungen. Jede einzelne hat nach dem Satz von Lagrange höchstens zwei Lösungen, also jede genau zwei. $\square$

3.9. Satz. $p > 2, \quad (ab, p) = 1.$

(1) $\qquad \left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}}(p) \quad$ (**Euler–Kriterium**).

(2) $\qquad \left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right) \quad$ (**Multiplikationssatz**).

Beweis zu (1). Aus der Eulerschen Kongruenz ergibt sich

$(*) \qquad \left(a^{\frac{p-1}{2}} - 1\right)\left(a^{\frac{p-1}{2}} + 1\right) \equiv 0 \bmod p.$

Nur einer der zwei Faktoren links, deren Differenz $= 2$ ist, wird von p geteilt. Ist $\left(\frac{a}{p}\right) = 1$, also $\quad x^2 \equiv a \bmod p \quad$ lösbar, dann gilt nach Euler

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

Der erste Faktor links in $(*)$ wird somit für die $\frac{p-1}{2}$ qR mod p von p geteilt. Nach dem Satz von Lagrange gibt es keine weiteren Lösungen. Also gilt für alle qNR a mod p

$$a^{\frac{p-1}{2}} + 1 \equiv 0(p), \quad \text{d.h.} \quad a^{\frac{p-1}{2}} \equiv \left(\frac{a}{p}\right) \bmod p.$$

Zu (2). (1) bewirkt

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

Wegen $\quad \left|\left(\frac{ab}{p}\right) - \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)\right| \leq 2 \quad$ und $\quad p > 2 \quad$ folgt hieraus die Gleichheit. $\square$

3.10. Gaußsches Lemma. $p > 2,\ (a, p) = 1$. Seien $c_1, \ldots, c_{\frac{p-1}{2}}$ die kleinsten positiven Reste der Zahlen $a, 2a, \ldots, \frac{p-1}{2}\, a$ bei Division durch p. Seien μ der c_n größer als $p/2$. Dann gilt

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

Beweis. 1. Wie angegeben sei

$$na = g_n p + c_n, \quad 0 < c_n \leq p - 1.$$

Seien

$$b_1, \ldots, b_\mu \quad \text{die} \quad c_n \quad \text{mit} \quad \frac{p+1}{2} \leq c_n \leq p - 1,$$

$$d_1, \ldots, d_\nu \quad \text{die} \quad c_n \quad \text{mit} \quad 1 \leq c_n \leq \frac{p-1}{2},$$

$\nu + \mu = \frac{p-1}{2}$. Die c_n und somit die b_j und d_j sind paarweise inkongruent mod p.

2. Für $\quad 1 \leq j \leq \mu, \quad 1 \leq k \leq \nu \quad$ gilt

$$p - b_j \not\equiv d_k \bmod p.$$

Denn aus der Richtigkeit einer solchen Kongruenz folgte $b_j + d_k \equiv 0 \bmod p$, also $(n_1 + n_2)\, a \equiv 0 \bmod p$ mit einem Paar (n_1, n_2) mit $1 \le n_1, n_2 \le \frac{p-1}{2}$, $n_1 \neq n_2$, was wegen $p \nmid a$ nicht sein kann. Die Aussage kann auch so formuliert werden, daß die Mengen

$$\{1, \ldots, \frac{p-1}{2}\} \quad \text{und} \quad \{d_1, \ldots, d_\nu,\ p - b_1, \ldots, p - b_\mu\}$$

identisch sind.

3. Nach 1. gilt mit 3.9.(1)

$$P \underset{\text{Df}}{=} \prod_{j=1}^{\mu} b_j \prod_{k=1}^{\nu} d_k \equiv a^{\frac{p-1}{2}} \Big(\frac{p-1}{2}\Big)! \equiv \Big(\frac{a}{p}\Big)\Big(\frac{p-1}{2}\Big)! \bmod p.$$

also, mit 2.,

$$(-1)^\mu \Big(\frac{p-1}{2}\Big)! \equiv \Big(\frac{a}{p}\Big)\Big(\frac{p-1}{2}\Big)! \bmod p.$$

Wegen $\Big(p, \Big(\frac{p-1}{2}\Big)!\Big) = 1$ darf dividiert werden.

$$(-1)^\mu \equiv \Big(\frac{a}{p}\Big) \bmod p.$$

Da beide Seiten im Betrag ≤ 1 sind, und $p > 2$ ist, folgt die Behauptung. □

3.11. Quadratisches Reziprozitätsgesetz (Gauß, 1801). Für zwei verschiedene, ungerade Primzahlen p und q gilt

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

Ergänzungsgesetze.

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

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

Die Aussagen können auch so formuliert werden

QRG: $$\Big(\frac{p}{q}\Big) = \begin{cases} \Big(\frac{q}{p}\Big), & \text{falls} \quad p \equiv 1 \bmod 4 \text{ oder } q \equiv 1 \bmod 4, \\ -\Big(\frac{q}{p}\Big), & \text{falls} \quad p, q \equiv 3 \bmod 4. \end{cases}$$

1. EG. $$\Big(\frac{-1}{p}\Big) = \begin{cases} 1, & \text{falls} \quad p \equiv 1\,(4), \\ -1, & \text{falls} \quad p \equiv 3\,(4). \end{cases}$$

2. EG. $$\Big(\frac{2}{p}\Big) = \begin{cases} 1, & \text{falls} \quad p \equiv 1 \quad \text{oder} \quad 7\,(8), \\ -1, & \text{falls} \quad p \equiv 3 \quad \text{oder} \quad 5\,(8). \end{cases}$$

Mit Hilfe der früheren Ergebnisse und des Reziprozitätsgesetzes kann im Prinzip jedes Legendre–Symbol $\left(\frac{a}{p}\right)$ relativ rasch berechnet werden. Man verwendet Verschiebung des Zählers mod p, multiplikative Zerlegung des Zählers und Invertierung nach dem qRG. Die rechnerisch aufwändige Faktorisierung, kann, wie im Anschluß gezeigt wird, mit Hilfe des „Jacobi–Symbols“ umgangen werden.

Beispiel. Ist die Kongruenz $x^2 + 77 \equiv 0 \bmod 43$ lösbar? Da 43 prim ist, berechnet man

$$\begin{aligned}\left(\frac{-77}{43}\right) &= \left(\frac{-1}{43}\right)\left(\frac{7}{43}\right)\left(\frac{11}{43}\right)\\ &= (-1)\Big(-\left(\frac{43}{7}\right)\Big)\Big(-\left(\frac{43}{11}\right)\Big)\\ &= -\Big(\frac{1}{7}\Big)\left(\frac{10}{11}\right) = -\Big(\frac{2}{11}\Big)\left(\frac{5}{11}\right)\\ &= -(-1)\Big(\frac{11}{5}\Big) = \Big(\frac{1}{5}\Big) = 1.\end{aligned}$$

Die Kongruenz ist somit lösbar. Für die Bestimmung der Lösungen x_0 und $43 - x_0$ gibt es kein ähnlich einfaches Verfahren.

Beweis des Reziprozitätsgesetzes.
1. Wie beim Gauß–Lemma sei

$$nq = \left[\frac{nq}{p}\right] p + c_n, \quad 0 < c_n < p, \quad 1 \le n \le \frac{p-1}{2}.$$

$$b_1, \ldots, b_\mu \quad \text{seien die} \quad c_n \in \Big[\frac{p+1}{2}, p-1\Big],$$

$$d_1, \ldots, d_\nu \quad \text{die} \quad \in \Big[1, \frac{p-1}{2}\Big].$$

Es werde

$$S_1 = \sum_{1 \le n \le \frac{p-1}{2}} \left[\frac{qn}{p}\right]$$

gesetzt. Summation über n ergibt

$$q\,\frac{(p-1)(p+1)}{8} = \sum_{1 \le n \le \frac{p-1}{2}} nq = pS_1 + \sum_{1 \le n \le \frac{p-1}{2}} c_n.$$

Wie in 2. im Beweis zu 3.10. sieht man

$$\begin{aligned}\sum_{1\le n\le \frac{p-1}{2}} c_n &= \sum_{j=1}^{\mu} b_j + \sum_{k=1}^{\nu} d_k \\ &= 2\sum_{j=1}^{\mu} b_j + \sum_{j=1}^{\mu}(p-b_j) + \sum_{k=1}^{\nu} d_k - \mu p \\ &= 2\sum_{j} b_j + \sum_{1\le n\le \frac{p-1}{2}} n - \mu p \\ &= \frac{(p-1)(p+1)}{8} + 2\sum_{j} b_j - \mu p.\end{aligned}$$

Zusammenfassung ergibt

$$\mu p = p\,S_1 + (1-q)\frac{(p-1)(p+1)}{8} + 2\sum_{j} b_j \equiv p\,S_1 \bmod 2,$$

also, wegen $p > 2$,

$$\mu \equiv S_1 \bmod 2,$$

und mit dem Gaußschen Lemma,

$$\left(\frac{q}{p}\right) = (-1)^{S_1}.$$

2. In ähnlicher Weise sieht man

$$\left(\frac{p}{q}\right) = (-1)^{S_2}, \quad \text{wobei} \quad S_2 = \sum_{1\le n\le \frac{q-1}{2}} \left[\frac{kp}{q}\right].$$

3. Es wird sich

$$S_1 + S_2 = \frac{p-1}{2}\cdot\frac{q-1}{2} \tag{3.1}$$

herausstellen, was nach 1. und 2. die Behauptung des Satzes nach sich zieht.
Zum Beweis sei oBdA $p > q$. Bezeichne R das abgeschlossene Rechteck in der (n,k)–Ebene mit

$$1 \le n \le \frac{p-1}{2} \quad \text{und} \quad 1 \le k \le \frac{q-1}{2}.$$

G (= Gitterpunkte) sei die Anzahl der $(n,k) \in \mathbb{N}^2$ in R. Dann gilt offensichtlich

$$G = \frac{p-1}{2}\cdot\frac{q-1}{2}. \tag{3.2}$$

Die Hauptdiagonale D in R läßt sich charakterisieren durch

$$k = \frac{q-1}{p-1}\,n \tag{D}$$

Die Gerade

$$D') \qquad k = \frac{q}{p}\, n$$

hat wegen $p > q$ größere Steigung als D. Man rechnet leicht nach, daß der Schnittpunkt $\left(\frac{p-1}{2}, s\right)$ von D' mit der Geraden $n = \frac{p-1}{2}$ die Ungleichung $\frac{q-1}{2} < s < \frac{q-1}{2} + 1$ erfüllt.

Auf D' liegen keine „Gitterpunkte“ (n, k) in R. Denn $k = \frac{q}{p}n,\ n \le \frac{p-1}{2},\ k \le \frac{q-1}{2}$ bewirkt $p|qn$, was hierfür nicht sein kann.

G kann berechnet werden durch Aufsummieren von G_1, der Anzahl der (n, k) unterhalb D', und G_2, der Anzahl der (n, k) oberhalb D'.

$$\begin{aligned} G_1 + G_2 &= \sum_{1 \le n \le \frac{p-1}{2}} \sum_{\substack{1 \le k \le \frac{q-1}{2} \\ k \le \frac{q}{p} n}} 1 \;+ \sum_{1 \le k \le \frac{q-1}{2}} \sum_{\substack{1 \le n \le \frac{p-1}{2} \\ n \le \frac{p}{q} k}} 1 \\ &= \sum_{1 \le n \le \frac{p-1}{2}} \sum_{1 \le k \le \frac{q}{p} n} 1 \;+ \sum_{1 \le k \le \frac{q-1}{2}} \sum_{1 \le n \le k \frac{p}{q}} 1 \\ &= \sum_{1 \le n \le \frac{p-1}{2}} \left[\frac{q}{p} n\right] + \sum_{1 \le k \le \frac{q-1}{2}} \left[k \frac{p}{q}\right] \\ &= S_1 + S_2. \end{aligned}$$

Mit (3.2) ergibt dies (3.1), womit der Beweis geführt ist. □

Für das spezielle Polynom $f(x) = x^2 - a$ ist das Lösungsverhalten mod p ($p > 2$) und mod p^k ($k \in \mathbb{N}$) identisch, während für das allgemeine Polynom zweiten Grades von Fall zu Fall nach Satz 3.7 untersucht werden muß. Der Vollständigkeit wegen wird schließlich die Kongruenz $x^2 \equiv a(2^k)$ untersucht.

3.12. Satz. $k \in \mathbb{N}$

(1) Sei $p > 2,\ (a, p) = 1$. Dann hat die Kongruenz

$$x^2 \equiv a(p^k)$$

genau $1 + \left(\frac{a}{p}\right)$ Lösungen.

(2) $2 \nmid a$. Die Lösungszahl der Kongruenz

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

ist

$= 1,$ falls $k = 1,$

$$= \begin{cases} 2 & \text{für} \quad a \equiv 1\,(4), \\ & \\ 0 & \text{für} \quad a \equiv 3\,(4) \end{cases} \text{, falls} \quad k = 2,$$

$$= \begin{cases} 4 & \text{für} \quad a \equiv 1\,(8), \\ & \\ 0 & \text{für} \quad a \not\equiv 1\,(8). \end{cases} \text{, falls} \quad k > 2,$$

Beweis zu (1). Für $k = 1$ ist dies Folgerung (4) zu 3.8. Für höhere k wendet man 3.7 auf $f(x) = x^2 - a$ an. Wegen $p \nmid 2a$ gilt $f'(x)\,2x_0 \neq 0(p)$ für jede Lösung x_0 von $x_0^2 - a \equiv 0(p^k)$. Es tritt also stets der erste Fall in 3.7 ein, und man erhält die Behauptung.

Zu (2). $k = 1$ und 2 rechnet man ohne Schwierigkeiten nach. Ist

$$(*) \qquad x^2 \equiv a(2^k)$$

lösbar, dann muß wegen $2 \nmid a$ x ungerade sein, $x = 2b + 1,$ also

$$a \equiv (2b+1)^2 = 4b\,(b+1) = 8\frac{b(b+1)}{2} + 1 \equiv 1 \bmod 8.$$

Für $a \equiv 1(8)$ hat $(*)$ bei $k = 3$ die vier Lösungen $x = 1, 3, 5, 7.$ Dies diene als Induktionsanfang. Für $k \geq 4$ sei x Lösung zu $(*) \bmod 2^{k-1}$. Es werde

$$x^* x \equiv 1\,(2^k), \quad b = x^* \frac{a - x^2}{2^{k-1}}$$

gesetzt. Damit gilt

$$(x + 2^{k-2}b)^2 \equiv x^2 + 2^{k-1} x\, b \equiv x^2 + a - x^2 \equiv a\,(2^k).$$

Es gibt also Lösungen $\bmod 2^k$.

Seien x_1, x_2 zwei Lösungen mod 2^k,

$$0 \equiv x_1^2 - x_2^2 = (x_1 - x_2)\,(x_1 + x_2) \bmod 2^k.$$

Da x_1 und x_2 beide ungerade sind, kann durch 4 dividiert werden

$$\frac{x_1 - x_2}{2} \cdot \frac{x_1 + x_2}{2} \equiv 0(2^{k-2}).$$

$\frac{x_1 - x_2}{2}$ und $\frac{x_1 + x_2}{2}$ können nicht zugleich gerade oder ungerade sein, da sonst ihre Summe, nämlich $x_1\ ,\ \equiv 0 \bmod 2$ wäre. Sei also im ersten Fall $\frac{x_1 - x_2}{2} \equiv 0(2^{k-2})$, das heißt $x_2 \equiv x_1 \bmod 2^{k-1}$. Dies induziert mod 2^k die zwei Werte x_1 und $x_1 + 2^{k-1}$. Ähnlich erhält man im Fall
$\frac{x_1 + x_2}{2} \equiv 0(2^{k-2})$ die zwei Werte $-x_1$ und $-x_1 + 2^{k-1}$. Man überzeugt sich, daß diese vier Zahlen mod 2^k verschieden sind. Andere Lösungen kann es nicht geben. □

Zur raschen Berechnung des Legendre–Symbols $\left(\frac{a}{p}\right)$ wurde 1846 durch Carl Gustav Jacobi (1804–1851) das später nach ihm benannte Symbol eingeführt.

3.13. Def. Sei m ungerade,

$$m = \prod_{j=1}^{k} p_i^{b_j} \qquad (2 < p_1 < \cdots < p_k), \quad (a, m) = 1.$$

Dann wird das **Jacobi–Symbol** definiert durch

$$\left(\frac{a}{m}\right) \underset{\text{Df}}{=} \prod_{j=1}^{k} \left(\frac{a}{p_j}\right)^{b_j}.$$

Bemerkung. Für $m = p > 2$ stimmen Legendre– und Jacobi–Symbol offenbar überein. Für zusammengesetztes m bedeutet $\left(\frac{a}{m}\right) = 1$ nicht notwendig, daß die Kongruenz $x^2 \equiv a \bmod m$ lösbar ist (Bsp. $\left(\frac{2}{9}\right) = 1$, aber $x^2 \equiv 2\,(9)$ unlösbar).

Die Rechengesetze für das Legendre–Symbol übertragen sich auf das Jacobi–Symbol.

3.14. Satz m, n ungerade. Dann gilt

(1) $$\left(\frac{a}{m}\right) = \left(\frac{b}{m}\right), \quad \text{falls} \quad (a, m) = 1 \quad \text{und} \quad a \equiv b \bmod m,$$

(2) $$\left(\frac{ab}{m}\right) = \left(\frac{a}{m}\right)\left(\frac{b}{m}\right), \quad \text{falls} \quad (ab, m) = 1,$$

(3) $$\left(\frac{a^2}{m}\right) = 1, \quad \text{für} \quad (a, m) = 1,$$

(4) $$\left(\frac{a}{mn}\right) = \left(\frac{a}{m}\right)\left(\frac{a}{n}\right), \quad \text{falls} \quad (a, mn) = 1,$$

(5) $$\left(\frac{-1}{m}\right) = (-1)^{\frac{m-1}{2}},$$

(6) $$\left(\frac{2}{m}\right) = (-1)^{\frac{m^2-1}{8}},$$

(7) $$\left(\frac{m}{n}\right)\left(\frac{n}{m}\right) = (-1)^{\frac{m-1}{2}\cdot\frac{n-1}{2}}, \quad \text{falls} \quad (m, n) = 1.$$

Die Aussagen können alle ohne Mühe auf die entsprechenden Beziehungen für das Legendre–Symbol zurückgeführt werden. Es werde am Beispiel (7) ausgeführt.
Sei $m = \prod_{i=1}^{r} p_i$, $n = \prod_{j=1}^{s} q_j$ mit nicht notwendig verschiedenen ungeraden Primzahlen p_i und q_j, aber $p_i \neq q_j$ für alle i, j. Dann ist nach Definition und den Gesetzen für das

Legendre–Symbol

$$\begin{aligned}\Big(\frac{m}{n}\Big) &= \prod_{j=1}^{s}\prod_{i=1}^{r}\Big(\frac{p_i}{q_j}\Big)\\ &= \prod_{j=1}^{s}\prod_{i=1}^{r}\Big(\frac{q_j}{p_i}\Big)\,(-1)^{(p_i-1)/2\cdot(q_j-1)/2} = \Big(\frac{n}{m}\Big)(-1)^{\alpha} \quad \text{mit}\\ \alpha &= \sum_{j=1}^{s}\sum_{i=1}^{r}\frac{p_i-1}{2}\cdot\frac{q_j-1}{2} = \sum_{i=1}^{r}\frac{p_i-1}{2}\cdot\sum_{j=1}^{s}\frac{q_j-1}{2}.\end{aligned}$$

Man überprüft leicht durch Induktion nach r bzw. s

$$\sum_{i=1}^{r}\frac{p_i-1}{2} \equiv \frac{m-1}{2} \bmod 2, \quad \sum_{j=1}^{s}\frac{q_j-1}{2} \equiv \frac{n-1}{2} \bmod 2,$$

womit die Behauptung folgt. □

Satz 3.14 erlaubt es, $\big(\frac{a}{p}\big)$ ohne Faktorisieren des Zählers zu berechnen. Denn es sind folgende Operationen ausreichend.

a) Reduzieren des Zählers, so daß der Betrag des Zählers kleiner wird als die Hälfte des Nenners.
b) Herausziehen von Zweierpotenzen im Zähler.
c) Berechnen von $\big(\frac{-1}{m}\big)$ und $\big(\frac{2}{m}\big)$.
d) Anwenden des Reziprozitätsgesetzes.

Beispiel: 443 ist eine Primzahl.

$$\begin{aligned}\Big(\frac{383}{443}\Big) &\underset{(7)}{=} -\Big(\frac{443}{383}\Big) \underset{(1)}{=} -\Big(\frac{60}{383}\Big) \underset{(2)}{=} -\Big(\frac{2^2}{383}\Big)\Big(\frac{15}{383}\Big)\\ &\underset{(3)}{=} -\Big(\frac{15}{383}\Big) \underset{(7)}{=} \Big(\frac{383}{15}\Big) \underset{(1)}{=} \Big(\frac{8}{15}\Big)\\ &\underset{(2)}{=} \Big(\frac{2^2}{15}\Big)\,\Big(\frac{2}{15}\Big) \underset{(3)}{=} \Big(\frac{2}{15}\Big) \underset{(6)}{=} 1,\end{aligned}$$

die Kongruenz $x^2 \equiv 383 \bmod 443$ ist somit lösbar.

4. Kapitel. Summen aus Quadraten und höheren Potenzen.

4.1. Def. und Satz.

(1) Ein Tripel $(x,y,z) \in \mathbb{N}^3$ heißt **pythagoräisches Tripel**, wenn es die Gleichung $x^2+y^2=z^2$ erfüllt. (Pythagoras, ca. 580–500 vor ZR).

(2) Man erhält alle pythagoräischen Tripel (x,y,z) mit $(x,y)=1,\ 2|x$ durch

$$x = 2ab, \quad y = a^2-b^2, \quad z = a^2+b^2,$$

wobei

$$(a,b)=1, \quad a>b>0 \quad \text{und} \quad a+b\equiv 1 \bmod 2.$$

Beweis. Die Einschränkungen $(x,y)=1$ und $2|x$ sind unerheblich, denn

a) aus $1<d|x,y$ folgt $d^2|z^2$, also $d|z$,

b) sind x und y beide ungerade, dann hat x^2+y^2 die Gestalt $4c+2$. Dies kann kein Quadrat sein.

Sei nun

$$(*) \qquad x^2+y^2=z^2$$

mit den obigen Bedingungen an x und y. Dann ist z ungerade, also $\frac{z-y}{2}$ und $\frac{z+y}{2}$ ganz und $\left(\frac{z-y}{2},\frac{z+y}{2}\right)=1$. Denn wenn ein p beiden Zahlen teilt, dann auch y und z, und somit x. Aus $(*)$ folgt

$$\left(\frac{x}{2}\right)^2=\frac{z+y}{2}\cdot\frac{z-y}{2}.$$

Wegen der Teilerfremdheit sind nach Satz 1.14 beide Zahlen selbst Quadrate.

$$\frac{z+y}{2}=a^2, \quad \frac{z-y}{2}=b^2, \quad x=2\,ab.$$

wobei

$$a>b>0, \quad (a,b)=1, \quad a+b\equiv a^2+b^2=z\equiv 1(2).$$

Sind umgekehrt a und b wie angegeben, dann gilt

$$x^2+y^2=4a^2b^2+(a^2-b^2)^2=(a^2+b^2)^2=z^2, \quad x,y,z>0, \quad 2|x.$$

Aus $(x,y)=d$ folgt $d|z=a^2+b^2$, $d|y=a^2-b^2$, also $d|2a^2, d|2b^2$. Wegen $(a,b)=1$ bedeutet das $d=1$ oder $d=2$. $d=2$ ist auf Grund von $d|z$ und $z\equiv 1(2)$ ausgeschlossen.

Die zulässigen Tripel (x,y,z) und Paare (a,b) sind einander bijektiv zugeordnet. □

Die ersten (a,b) ergeben folgende Tripel

a	b	x	y	z
2	1	4	3	5
3	2	12	5	13
4	1	8	15	17
4	3	24	7	25

Zu den ganz großen Problemen der Mathematik gehört die

4.2. Fermat–Vermutung (Pierre de F., 1601–1655).
Für $n>2$ besitzt die Gleichung $x^n+y^n=z^n$ keine Lösung $(x,y,z)\in\mathbb{N}^3$.

In den etwa 350 Jahren seiner Geschichte haben sich fast alle namhaften Mathematiker ernsthaft darum bemüht. Insbesondere die algebraische Zahlentheorie wurde durch die Arbeit am Fermatschen Problem entscheidend vorangetrieben. 1995 gelang Andrew Wiles

der vollständige Beweis der Vermutung.

Falls die Unlösbarkeit der Fermatschen Gleichung für ein $n > 2$ bewiesen ist, dann folgt sie wegen $a^{mn} = (a^m)^n$ für jedes Vielfache von n. Es ist somit für die Unlösbarkeit ausreichend, die Exponenten $n = 4$ und $n = p \geq 3$ zu untersuchen. Der Fall $n = 4$ ist nach Fermat elementar zugänglich, während $p \geq 3$ algebraische Hilfsmittel erfordert. Ein Hinweis:
Sei $\xi = \exp\left(\frac{2\pi i}{p}\right) = \cos\left(\frac{2\pi}{p}\right) + i \sin\left(\frac{2\pi}{p}\right)$, dann gilt

$$x^p + y^p = (x + y)(x + \xi y)(x + \xi^2 y) \cdots (x + \xi^{p-1} y).$$

Es wird daher nützlich sein, den „Kreisteilungskörper“ $\mathbb{Q}(\xi)$ zu untersuchen.

4.3. Satz (Fermat). Die Gleichung $x^4 + y^4 = z^4$ besitzt keine Lösung $(x, y, z) \in \mathbb{N}^3$.

Beweis: 1. Es reicht, die Unlösbarkeit der Gleichung

$$x^4 + y^4 = z^2 \tag{2}$$

in $x, y, z \in \mathbb{N}$ zu zeigen. Annahme, es gebe Lösungen. Sei z_0 die kleinste Zahl, zu der es x und y mit (1) gibt. Ein solches Paar (x, y) werde festgehalten. Es muß $(x, y) = 1$ gelten, da sonst in (1) gekürzt werden könnte. Insbesondere ist x oder y ungerade, also

$$z_0^2 = x^4 + y^4 \equiv 1 \quad \text{oder} \quad 2 \bmod 4.$$

$z_0^2 \equiv 2(4)$ tritt nicht ein. Bleibt

$$z_0 \equiv 1(2) \quad \text{und oBdA} \quad x \equiv 0, \quad y \equiv 1(2).$$

2. Auf (1) kann Satz 4.1. angewandt werden.

$$x^2 = 2ab,\ y^2 = a^2 - b^2,\ z_0 = a^2 + b^2 \tag{3}$$

mit $a, b > 0$, $(a, b) = 1$, $a + b \equiv 1(2)$ Aus $a \equiv 0(2)$, $b \equiv 1(2)$ folgte $y^2 \equiv -1(4)$, was nicht sein kann. Also bleibt

$$a \equiv 1(2), \quad b = 2c.$$

3. Aus 2. ergibt sich $\left(\frac{x}{2}\right)^2 = ac$, $(a, c) = 1$, also

$$a = z'^2, \quad c = d^2, \quad z' > 0, \quad d > 0, \quad (z', d) = 1, \quad y^2 = a^2 - b^2 = z'^4 - 4d^4,$$

und daher mit (2)

$$(2d^2)^2 + y^2 = (z'^2)^2, \tag{4}$$

wobei $2d^2, y$ und z'^2 paarweise teilerfremd sind.

4. Auf (3) wird erneut Satz 4.1. angewandt

$$2d^2 = 2a'b', \quad z'^2 = a'^2 + b'^2, \quad a' > 0, \quad b' > 0, \quad (a', b') = 1.$$

Wegen $d^2 = a'b',\ (a', b') = 1$ folgt $a' = x'^2,\ b' = y'^2,\ x' > 0,\ y' > 0$ und

$$x'^4 + y'^4 = z'^2. \tag{5}$$

Aber $0 < z' \leq z'^2 = a \leq a^2 < a^2 + b^2 = z_0$. Mit z' ist somit eine kleinere Zahl als z_0 gefunden, die (1) löst. Dies bedeutet einen Widerspruch. □

Die hier benutzte Methode, zu einer angenommenen Lösung eine kleinere zu konstruieren, geht auf Fermat zurück und wird nach ihm „descendente infinie“ genannt. Das Prinzip wird hier noch zweimal angewandt werden.

Als nächstes soll untersucht werden, welche Zahlen sich als Summe von zwei, drei oder mehr Quadraten schreiben lassen.

4.4. Satz von Euler. Für $n \in \mathbb{N}$ sind die folgenden Aussagen äquivalent

1) $n = x^2 + y^2 \quad$ mit $\quad x, y \in \mathbb{N}_0$,
2) in der Primfaktorzerlegung von n treten alle Primteiler p von n mit $p \equiv 3 \bmod 4$ in gerader Potenz auf.

Beweis. 1. Falls ein p mit $p \equiv 3(4)$ die Zahl n teilt, gibt es keine Zerlegung

$$n = x^2 + y^2 \quad \text{mit} \quad (x, y) = 1.$$

Falls es eine solche gibt, dann können wegen $n \geq 3$ x und y ≥ 1 gewählt werden. Wegen $p|n$ gilt $x, y \not\equiv 0(p)$. Nach Satz 3.2. existiert ein z mit $y \equiv zx\,(p)$, also $x^2(1 + z^2) = x^2 + y^2 \equiv 0(p)$ und somit $1 + z^2 \equiv 0(p)$. Damit ist $\left(\frac{-1}{p}\right) = 1$, also nach dem ersten Ergänzungsgesetz $p \equiv 1(4)$, was einen Widerspruch bedeutet.

2. Sei $p \equiv 3(4), \quad p^a|n, \quad p^{a+1} \nmid n, \quad a \geq 1,$ ungerade. Dann besitzt n keine Darstellung $n = x^2 + y^2$.

Angenommen, es existieren x und y mit $n = x^2 + y^2$. Sei $d = (x, y), \quad x = dx', \quad y = dy', \quad (x', y') = 1,$ also

$$(*) \qquad n = d^2(x'^2 + y'^2) = d^2 n'.$$

n' besitzt somit eine Darstellung $n' = x'^2 + y'^2$ mit $(x', y') = 1$.
Sei b der Exponent von p in der kanonischen Zerlegung von d. p kann nach 1. nicht in n' aufgehen. Dann teilt p nach $(*)$ die Zahl n in genau $2b$–ter Potenz, im Widerspruch zur Annahme.

3. Mit n_1 und n_2 ist auch $n_1 n_2$ darstellbar, wie die Identität

$$(x_1^2 + y_1^2)(x_2^2 + y_2^2) = (x_1 x_2 + y_1 y_2)^2 + (x_1 y_2 - x_2 y_1)^2$$

zeigt. Für die Richtung von 2) nach 1) reicht es danach aus, für jedes $p \equiv 1(4)$ die Lösbarkeit von $p = x^2 + y^2$ nachzuweisen.

Denn $2 = 1^2 + 1^2$ und für $p \equiv 3(4)$ ist $p^2 = 0 + p^2$.

4. Sei $p \equiv 1(4)$.
4.1. Es gibt x, y, m mit $p \nmid x, p \nmid y, 0 < m < p$ und

$$x^2 + y^2 = mp. \tag{4.1}$$

Denn wegen $p \equiv 1(4)$ ist $\left(\frac{-1}{p}\right) = 1$, also existiert ein x mit $0 < x \leq \frac{p-1}{2}$ und $x^2 + 1 = mp$. Wegen $0 < 1 + x^2 < p^2$ ist $0 < m < p$.

4.2. Sei $m_0 \in (0, p)$ das kleinste m, für das (4.1) lösbar ist. Es kommt darauf an, $m = 1$ zu zeigen. Nach der Fermatschen Idee wird im Fall $m_0 > 1$ ein kleineres m_1 mit (4.1) konstruiert. Werde also $m_0 > 1$ angenommen. Es gilt in (4.1)

$$m_0 \nmid x \ \vee \ m_0 \nmid y. \tag{4.2}$$

Denn aus $m_0 | x$ und $m_0 | y$ folgt $m_0^2 \,|\, x^2 + y^2 = m_0\, p$, $m_0 | p$, was wegen $1 < m_0 < p$ nicht eintreten kann.

4.3 Da $m_0 > 1$, lassen sich ganze a und b finden, so daß für

$$x_1 \underset{\text{Df}}{=} x - a\, m_0, \quad y_1 \underset{\text{Df}}{=} y - b\, m_0$$

$|x_1|, |y_1| \leq \frac{m_0}{2}$ gilt. Also ist

$$0 < x_1^2 + y_1^2 \leq 2\Big(\frac{m_0}{2}\Big)^2 < m_0^2 \quad \text{und} \quad x_1^2 + y_1^2 \equiv x^2 + y^2 \equiv 0(m_0),$$

$$x_1^2 + y_1^2 = m_1 m_0 \quad \text{mit} \quad 0 < m_1 < m_0 \tag{4.3}$$

4.4. 3. (4.1) und (4.3) ergeben

$$m_0^2 m_1 p = (x^2 + y^2)(x_1^2 + y_1^2) = (x x_1 + y y_1)^2 + (x y_1 - x_1 y)^2.$$

Wegen

$$x x_1 + y y_1 = x(x - a\, m_0) + y(y - b\, m_0) = m_0\, x'$$
$$x y_1 - x_1 y = x(y - b\, m_0) - y(x - a\, m_0) = m_0\, y'$$

folgt daraus

$$m_1 p = x'^2 + y'^2.$$

Wegen $0 < m_1 < m_0$ steht dies im Widerspruch zur Minimalität von m_0. □

Der Fall dreier Summanden ist wesentlich schwieriger und kann hier nur knapp diskutiert werden. Der Hauptgrund dafür ist, daß keine „Multiplikationsformel“ der Art

$$(x_1^2 + y_1^2 + z_1^2)(x_2^2 + y_2^2 + z_2^2) = L_1^2(x_1, \ldots, z_2) + \cdots + L_3^2(x_1, \ldots, z_2)$$

(L_1, L_2, L_3 Polynome zweiten Grades in den sechs Variablen) existiert. Adolf Hurwitz (1859–1919) hat gezeigt, daß es solche Formeln nur für 1,2,4 oder 8 Summanden gibt.

Satz 4.5. (Legendre) Für $n \in \mathbb{N}$ sind folgende Aussagen äquivalent

1) $n = x_1^2 + x_2^2 + x_3^2$ mit $x_1, x_2, x_3 \in \mathbb{N}_0$,
2) n hat nicht die Gestalt $n = 4^a(8b+7)$ $(a, b \in \mathbb{N}_0)$.

Die Richtung von 2) nach 1) erfordert einiges aus der Theorie der ternären quadratischen Formen

$$Q(x_1, x_2, x_3) = \sum_{j,k=1}^{3} a_{jk}\, x_j x_k \quad (a_{jk} \in \mathbb{Z})$$

sowie den Satz von Dirichlet (1805–1859), daß in jeder reduzierten Restklasse $a + k\mathbb{Z}$ $\big((a,k) = 1\big)$ unendlich viele Primzahlen liegen.

Die Richtung 1) $\Rightarrow$ 2) ist einfach. Sei

(1) $$n = 4^a(8b+7), \quad a, b \in \mathbb{N}_0 \quad \text{und} \quad n = x_1^2 + x_2^2 + x_3^2,$$

(2) $$x_j = 2^{a_j} y_j, \quad 2 \nmid y_j, \quad a_1 \le a_2 \le a_3.$$

Für jedes ungerade $y \in \mathbb{Z}$ gilt $y^2 \equiv 1\,(8)$. Dies sieht man durch Ausrechnen in den zwei Fällen $y \equiv 1\,(4)$ bzw. $y \equiv 3\,(4)$. Also ist $y_j^2 \equiv 1\,(8)$ $(j = 1, 2, 3)$.
Dann folgt

(3) $$0 \le a = a_1 \le a_2 \le a_3.$$

Denn aus $a_1 < a$ ergibt sich mit (1)

$$\frac{n}{4^{a_1}} = 4^{a-a_1}(8b+7) = y_1^2 + 2^{2(a_2-a_1)} y_2^2 + 2^{2(a_3-a_1)} y_3^2.$$

Die linke Seite ist $\equiv 0$ oder 4 mod 8. Der zweite und dritte Summand rechts sind $\equiv 0, 1$ oder 4 mod 8, die rechte Seite also $\equiv 1, 2, 3, 5$ oder 6 mod 8, was nicht zusammenpaßt.

Die Annahme $a_1 > a$ ergibt unmittelbar einen Widerspruch zu (1).

Aus (1), (2) und (3) erhält man

$$n_1 = 8b + 7 = y_1^2 + 2^{b_2} y_2^2 + 2^{b_3} y_3^2$$

mit $2 \nmid y_1, y_2, y_3$ und $0 \le b_2 \le b_3$. Man überzeugt sich durch Verfolgen aller Möglichkeiten, daß die rechte Seite nicht $\equiv 7$ mod 8 sein kann. □

Am Beispiel

$$3 \cdot 5 = (1^2 + 1^2 + 1^2) \cdot (2^2 + 1^2 + 0^2) = 15 = 8 + 7$$

sieht man, daß die Eigenschaft, Summe dreier Quadrate zu sein, nicht „multiplikativ“ ist.

Das Problem mit vier oder mehr Summanden hat eine einfache Lösung.

4.6. Satz von Lagrange. Jedes $n \in \mathbb{N}$ besitzt eine Darstellung

$$n = x_1^2 + x_2^2 + x_3^2 + x_4^2 \quad (x_1, \ldots, x_4 \in \mathbb{N}_0).$$

Beweis. 1. Die wichtige Multiplikationsformel (**Lagrangesche Identität**) lautet hier

$$\begin{aligned}(x_1^2 + \cdots + x_4^2)(y_1^2 + \cdots + y_4^2) \\ = (x_1y_1 + \cdots + x_4y_4)^2 + (x_1y_2 - x_2y_1 + x_3y_4 - x_4y_3)^2 \\ + (x_1y_3 - x_3y_1 + x_4y_2 - x_2y_4)^2 + (x_1y_4 - x_4y_1 + x_2y_3 - x_3y_2)^2.\end{aligned}$$

Wegen $2 = 1^2 + 1^2 + 0^2 + 0^2$ reicht es aus, den Beweis für ungerade p zu führen.

2. Für $p > 2$ sind die Zahlen

$$\begin{aligned} a^2 \quad & \Big(0 \le a \le \frac{p-1}{2}\Big) \quad \text{und} \\ -1 - b^2 \quad & \Big(0 \le b \le \frac{p-1}{2}\Big)\end{aligned}$$

jeweils paarweise inkongruent mod p. Es muß also, da insgesamt $2\Big(1 + \frac{p-1}{2}\Big) > p$ Zahlen zur Verfügung stehen, eine aus der ersten zu einer aus der zweiten Menge mod p kongruent sein (Schubfachschluß!)

$$1 + a^2 + b^2 = mp \quad \text{mit} \quad 0 < mp \le 1 + 2\Big(\frac{p-1}{2}\Big)^2 < p^2,$$

also $0 < m < p$. Es existieren demnach $m \in (0, p)$, so daß

$$x_1^2 + \cdots + x_4^2 = mp, \quad \text{nicht alle} \quad x_j \equiv 0(p) \tag{2}$$

lösbar ist. Sei wieder m_0 das kleinste solche m. Es werde $m_0 > 1$ angenommen. $x_1, \ldots, x_4$ werden gemäß (2) zu m_0 gewählt.

3. Angenommen, m_0 sei gerade. Dann sind

a) alle x_j gerade oder
b) oBdA x_1, x_2 gerade und x_3, x_4 ungerade oder
c) alle x_j ungerade.

Im Fall b) sind

$$y_1 = x_1 + x_2, \quad y_2 = x_1 - x_2, \quad y_3 = x_3 + x_4, \quad y_4 = x_3 - x_4$$

sämtlich gerade, desgleichen bei a) und c).
Aus (2) folgt

$$\frac{1}{2}\, m_0\, p = \Big(\frac{y_1}{2}\Big)^2 + \cdots + \Big(\frac{y_4}{2}\Big)^2,$$

im Widerspruch zur Minimalität von m_0. Also ist m_0 ungerade und ≥ 3.

4. Nicht alle x_j sind durch m_0 teilbar, denn andernfalls folgte aus (2) $m_0|p$. Zur Konstruktion eines kleineren m_1 mit der Eigenschaft (2) werde wie im Beweis zu 4.3.

$$y_j = x_j - a_j\, m_0, \quad |y_j| < \frac{m_0}{2}$$

gesetzt (das strenge $<$ ist wegen $m \equiv 1(2)$ möglich).

$$0 < y_1^2 + \cdots + y_4^2 < 4\Big(\frac{m_0}{2}\Big)^2 = m_0^2.$$

Aus (2) ergibt sich

$$y_1^2 + \cdots + y_4^2 = m_0 m_1 \quad \text{mit} \quad 0 < m_1 < m_0. \tag{4}$$

5. Multiplikation von (2) und (4) gemäß 1. führt zu

$$m_0^2 m_1\, p = z_1^2 + \cdots + z_4^2, \tag{5}$$

wobei die z_j wie angegeben aus den x_j und y_j berechnet werden. Zum Beispiel

$$\begin{aligned} z_1 &= \sum_{1\le j\le 4} x_j y_j = \sum_{1\le j\le 4} x_j(x_j - a_j m_j) \\ &\equiv \sum_{1\le j\le 4} x_j^2 \equiv 0(m_0). \end{aligned}$$

Ebenso für z_2, z_3, z_4, d.h.

$$z_j = m_0\, c_j.$$

In (5) eingesetzt, ergibt das

$$m_1\, p = c_1^2 + \cdots + c_4^2,$$

was der Minimalität von m_0 widerspricht. □

Der Satz von Lagrange kann als Spezialfall eines allgemeineren Problems aufgefaßt werden.

4.7. Waringsches Problem (Edmund W., 1734–1798).
Existiert zu jedem $k \ge 2$ ein $\ell \in \mathbb{N}$, so daß jedes n als Summe von ℓ k–ten Potenzen x_j^k $(x_j \in \mathbb{N}_0)$ dargestellt werden kann?

Der erste allgemeine Beweis wurde 1909 von David Hilbert (1862–1943) gegeben. Unter den verschiedenen, sämtlich nicht elementaren Lösungswegen hat sich folgender als ergiebigsten erwiesen. Sei für $n \in \mathbb{N}$ und $\alpha \in \mathbb{R}$

$$S_k(\alpha) = \sum_{m\le n^{1/k}} e(\alpha\, m^k), \quad e(\beta) = e^{2\pi i\beta}.$$

Dann gilt

$$\begin{aligned} R_{l,k}(n) &= \#\big\{(m_1,\dots,m_l) \in \mathbb{N}^l,\ m_1^k + \cdots + m_l^k = n\big\} \\ &= \int_0^1 \big(S_k(\alpha)\big)^l\, e(-\alpha n)\, d\alpha. \end{aligned}$$

Eine genaue Analyse des Integrals führt für hinreichend großes $l \geq l_0(k)$ zu einer Näherungsformel für $R_{l,k}(n)$, und damit zur Lösung des Waringschen Problems.

5. Kapitel. Zahlentheoretische Funktionen.

5.1. Def. (1) Eine Abbildung $f : \mathbb{N} \to \mathbb{C}$ heißt **zahlentheoretische Funktion (zF).**

(2) Eine zF heißt **multiplikativ**, wenn

a) $f(1) = 1$ und
b) $\forall\, m, n : (m, n) = 1 \Rightarrow f(mn) = f(m)f(n)$.

f heißt **vollständig multiplikativ**, wenn stets $f(mn) = f(m)f(n)$ gilt.

(3) Eine zF heißt **additiv**, wenn

$$\forall\, m, n : (m, n) = 1 \Rightarrow f(mn) = f(m) + f(n).$$

Analog **vollständig additiv.**

Bemerkungen. (1) Bedingung (2) a) kann auch durch

a') $\exists\, n_0 : f(n_0) \neq 0$ ersetzt werden. Denn mit
b) folgt daraus $f(n_0) = f(n_0 \cdot 1) = f(n_0) \cdot f(1)$, also $f(1) = 1$.

Während es sich bei der Multiplikativität als günstig erweist, die Null–Funktion auszuschließen, ist dies bei der Additivität nicht nötig.

(2) Multiplikative Funktionen sind wegen

$$f\left(p_1^{a_1} \cdots p_k^{a_k}\right) = f\left(p_1^{a_1}\right) \cdots f\left(p_k^{a_k}\right) \quad (p_1 < \cdots < p_k)$$

durch ihre Werte auf den Primzahlpotenzen vollständig bestimmt, vollständig multiplikative durch ihre Werte an den Primzahlen.

(3) Ist g additiv, dann ist $f = e^g$ multiplikativ.

5.2. Beispiel. Für $\alpha \in \mathbb{R}$ wird

$$\sigma_\alpha(n) = \sum_{d|n} d^\alpha$$

als **Teilersummen–Funktion** bezeichnet. $\sum\limits_{d|n}$ bedeutet Summation über alle natürlichen Teiler d von n. Insbesondere

$$\sigma(n) \underset{\text{Df}}{=} \sigma_1(n) = \sum_{d|n} d,$$

$$d(n) = \sigma_0(n) = \sum_{d|n} 1.$$

Folgerung. σ_α ist multiplikativ.

Dies ergibt sich unmittelbar aus dem späteren Satz 5.8. Zur Übung werde der Beweis hier ausgeführt.

Sei $(m,n)=1$. Dann sind die Paare (d,k) mit $d|m,\ k|n$ und die Teiler $l=dk$ von mn einander bijektiv zugeordnet.

$$\begin{aligned}\sigma_\alpha(mn) &= \sum_{l|mn} l^\alpha = \sum_{d|m}\ \sum_{k|n} (dk)^\alpha \\ &= \sum_{d|m} d^\alpha \sum_{k|n} k^\alpha = \sigma_\alpha(m)\sigma_\alpha(n).\end{aligned}$$

Es ist

$$\sigma_\alpha(p^k) = 1 + p^\alpha + p^{2\alpha} + \cdots + p^{k\alpha} \neq 0.$$

Wegen

$$\begin{aligned}&\sigma_\alpha(p^2) = 1 + p^\alpha + p^{2\alpha} \quad \text{und} \\ &\sigma_\alpha(p)\cdot\sigma_\alpha(p) = (1+p^\alpha)(1+p^\alpha) = 1 + 2p^\alpha + p^{2\alpha}\end{aligned}$$

sieht man, daß σ_α nicht vollständig multiplikativ ist. □

Die folgenden Bezeichnungen gehen auf die alten Griechen zurück.

5.3. Def. (1) n heißt **vollkommen** (oder perfekt), wenn

$$\sigma(n) = \sum_{d|n} d = 2n.$$

(2) Zwei verschiedene natürliche Zahlen n und m heißen **befreundet**, wenn $\sigma(n)-n=m$ und $\sigma(m)-m=n$.

5.4. Satz. (1) (Euklid–Euler) Die folgenden Eigenschaften sind äquivalent

(a) n ist gerade und vollkommen,

(b) $n = 2^k(2^{k+1}-1)$ und $2^{k+1}-1$ ist Primzahl.

(2) (Marin Mersenne, 1588–1648).
Falls 2^m-1 Primzahl, ist m prim. Die Zahlen $M_p = 2^p-1$ heißen **Mersenne–Zahlen**.

Bemerkung. Für $p = 2,3,5,7,13,17,19,31,61,89,107,127$ sind die M_p prim, für die übrigen $p<500$ ist M_p zusammengesetzt. Man kennt bis heute (Februar 2005) 42 prime M_p. Das größte solche M_p gehört zu $p = 25\,964\,951$ und hat $7\,816\,230$ Dezimalstellen. Es ist zugleich die zur Zeit größte berechnete Primzahl. Man vermutet, daß es unendlich viele prime und zusammengesetzte M_p gibt. Ob ungerade vollkommene Zahlen existieren, ist ein offenes Problem. Falls es welche gibt, müssen sie größer als 10^{150} sein. Ebenso ist unbekannt, ob es unendlich viele Paare befreundeter Zahlen gibt.

Beweis zu 5.4.
1. zu (1), (b) $\Rightarrow$ (a). Im Fall $2^{k+1}-1=p$ sieht man

$$\begin{aligned}\sigma(n)=\sigma\big(2^k(2^{k+1}-1)\big) &= 1+2+\cdots+2^k+(1+2+\cdots+2^k)p\\ &= (2^{k+1}-1)(p+1)=p\cdot 2^{k+1}=2n.\end{aligned}$$

Die ist die Euklidsche Feststellung.

2. zu (1), (a) $\Rightarrow$ (b). Eine Zweierpotenz $n=2^k$ ist wegen $\sigma(n)=1+\cdots+2^k=2^{k+1}-1$ nicht vollkommen. Sei also

$$n=2^k u,\quad k\geq 1,\quad u\geq 3,\quad 2\nmid u$$

vollkommen. Dann folgt mit der Multiplikativität von σ

$$2^{k+1}u=2n=\sigma(n)=\sigma(2^k)\sigma(u)=(2^{k+1}-1)\sigma(u),$$

also

$$(*)\qquad \sigma(u)=2^{k+1}u(2^{k+1}-1)^{-1}=u+\frac{u}{2^{k+1}-1}.$$

Da u und $\sigma(u)$ ganz sind, ist es auch $u(2^{k+1}-1)^{-1}$, das heißt $2^{k+1}-1$ und $u(2^{k+1}-1)^{-1}$ sind Teiler von u. Aus der Identität $(*)$ entnimmt man, daß u und $u(2^{k+1}-1)^{-1}$ die einzigen Teiler von u sind. Also ist u Primzahl und $u(2^{k+1}-1)^{-1}=1$.

3. Zu (2). Sei $m=kl$ mit $k,l>1$. Dann ist

$$2^{kl}-1=(2^k-1)\big(2^{k(l-1)}+2^{k(l-2)}+\cdots+2^k+1\big)$$

zusammengesetzt. □

5.5. Definition.

$$\mu(n)\underset{\text{Df}}{=}\begin{cases}(-1)^r, & \text{falls}\quad n=p_1\dots p_r,\\ & \qquad 2\leq p_1<\cdots<p_r,\ r\geq 0\\ 0 & \text{sonst.}\end{cases}$$

Möbius–Funktion (August Ferdinand M., 1790–1868).

Diese merkwürdige Funktion, deren fundamentale Bedeutung bald klar sein wird, ist multiplikativ, aber nicht vollständig multiplikativ. Zahlen, in deren kanonischer Zerlegung keine Primzahlen in zweiter oder höherer Potenz auftreten, heißen **quadratfrei** (square-free).

$$\begin{aligned}&n\quad\text{quadratfrei}\ \Leftrightarrow |\mu(n)|=\mu^2(n)=1.\\ &n\quad\text{quadrathaltig (d.h. }\exists\, p: p^2|n)\ \Leftrightarrow \mu(n)=0.\end{aligned}$$

Sei $(n,m)=1$. Falls beide quadratfrei sind, ist es auch nm und

$$\mu(nm)=\mu(p_1\dots p_a\, q_1\dots q_b)=(-1)^{a+b}=(-1)^a(-1)^b=\mu(n)\mu(m).$$

Falls nm quadrathaltig ist, dann wegen der Teilerfremdheit auch n oder m, also

$$0 = \mu(nm) = \mu(n)\mu(m).$$

Wegen $\mu(p^2) = 0$, aber $\mu(p)\mu(p) = 1$ liegt keine vollständige Multiplikativität vor.

5.6. Definition. $n > 1$ habe die kanonische Zerlegung $p_1^{a_1} \dots p_k^{a_k}$. Dann wird

$$\begin{aligned} &\omega(n) = k && (\textbf{Primteiler–Anzahl}) \\ &\Omega(n) = a_1 + \dots + a_k && (\textbf{Primfaktoren–Anzahl}) \\ &\omega(1) = \Omega(1) \underset{\text{Df}}{=} 0 \end{aligned}$$

gesetzt.

ω ist additiv, aber nicht vollständig, Ω ist vollständig additiv.

Die folgende Funktion spielt in der Primzahltheorie eine wichtige Rolle.

5.7. Definition.

$$\Lambda(n) \underset{\text{Df}}{=} \begin{cases} \ln p, & \text{falls} \quad n = p^k \\ 0 & \text{sonst.} \end{cases}$$

von–Mangoldt–Funktion (Hans Karl Friedrich von M., 1854–1925).

Λ ist weder additiv noch multiplikativ.

Folgerung. $\displaystyle\sum_{d|n} \Lambda(d) = \ln n.$

Die Aussage ist richtig für $n = 1$. Sei $n = p_1^{a_1} \dots p_k^{a_k} > 1$. Nach Definition von Λ tragen zur Teilersumme nur die d etwas bei, die die Gestalt $p_\nu^{b_\nu}$ $(1 \le \nu \le k,\ 1 \le b_\nu \le a_\nu)$ haben.

$$\begin{aligned} \sum_{d|n} \Lambda(d) &= \sum_{1\le\nu\le k} \sum_{1\le b_\nu\le a_\nu} \ln\, p_\nu \\ &= \sum_{1\le\nu\le k} \ln(p_\nu^{a_\nu}) = \ln(p_1^{a_1} \dots p_k^{a_k}) = \ln n. \end{aligned}$$

□

5.8. Definition und Satz.
(1) Für zwei zahlentheoretische Funktionen f und g wird das **Falt–Produkt** $f * g$ definiert durch

$$(f * g)(n) = \sum_{d|n} f(d)\, g\Big(\frac{n}{d}\Big).$$

(2) Falls f und g multiplikativ sind, ist auch $f * g$ multiplikativ.

Bemerkung. Die vollständige Multiplikativität bleibt bei der Faltung nicht immer erhalten, wie das Beispiel

$$d = \underline{1} * \underline{1} \qquad \big(\underline{1}(n) = 1 \;\forall\, n\big) \quad \text{zeigt.}$$

Beweis zu (2). Ist $(n_1, n_2) = 1$ und durchlaufen d_1 und d_2 unabhängig voneinander alle Teiler von n_1 und n_2, so durchläuft $d = d_1 d_2$ alle Teiler von $n = n_1 n_2$. Umgekehrt läßt sich jeder Teiler d von n eindeutig als $d_1 d_2$ $(d_j | n_j)$ schreiben. Für $d_j | n_j$ ist wegen $(n_1, n_2) = 1 \quad (d_1, d_2) = \Big(\dfrac{n_1}{d_1}, \dfrac{n_2}{d_2}\Big) = 1$, also

$$\begin{aligned}(f * g)(n_1 \cdot n_2) &= \sum_{d_1 | n_1, d_2 | n_2} f(d_1 d_2) g\Big(\frac{n_1}{d_1}\ \frac{n_2}{d_2}\Big) \\ &= \sum_{d_1 | n_1} f(d_1) g\big(\frac{n_1}{d_1}\big) \sum_{d_2 | n_2} f(d_2)\, g\Big(\frac{n_2}{d_2}\Big) \\ &= (f * g)(n_1) \cdot (f * g)(n_2).\end{aligned}$$

□

Die Folgerung zu 5.2. ist hierin wegen

$$\sigma_\alpha = P_\alpha * \underline{1} \qquad \big(P_\alpha(n) = n^\alpha\big)$$

offenbar enthalten.

5.9. Satz. (1) Die Menge der zahlentheoretischen Funktionen f mit $f(1) \neq 0$, versehen mit $*$ als Verknüpfung, bildet eine abelsche Gruppe. Neutrales Element ist die Funktion ε,

$$\varepsilon(n) = \begin{cases} 1, & n = 1 \\ 0, & n > 1 = \big[\frac{1}{n}\big]. \end{cases}$$

(2) Die Möbius–Funktion ist das Faltungs–Inverse der Funktion $\underline{1}$ $(\underline{1}(n) = 1 \forall\, n)$.

(3) Die Menge der multiplikativen Funktionen bildet eine Untergruppe.

Beweis. 1. Daß $*$ auf $\mathcal{Z} = \{f,\ f \text{ zF}\}$ eine Verknüpfung bildet, ist klar. Daß $*$ nicht aus $\mathcal{M} = \{f,\ f \text{ multiplikative zF}\}$ hinausführt, wurde gerade gezeigt.

2. Zur Assoziativität. Es kann $(f*g)(n)$ auch als $\sum_{\substack{d_1,d_2|n\\ d_1d_2=n}} f(d_1)\,g(d_2)$ geschrieben werden, also

$$\begin{aligned}\big(f_1*(f_2*f_3)\big)(n) &= \sum_{\substack{d_1,d'|n\\ d_1d'=n}} f_1(d_1)(f_2*f_3)(d')\\ &= \sum_{\substack{d_1,d'|n\\ d_1d'=n}} f_1(d_1)\sum_{\substack{d_2,d_3|d'\\ d_2d_3=d'}} f_2(d_2)\,f_3(d_3)\\ &= \sum_{d_1,d_2,d_3|n,d_1d_2d_3=n} f_1(d_1)f_2(d_2)f_3(d_3).\end{aligned}$$

Denselben Ausdruck erhält man für $\big((f_1*f_2)*f_3\big)(n)$.

3. Die Kommutativität sieht man ähnlich und einfacher.

4. Die ε–Funktion ist ersichtlich multiplikativ, und es gilt für beliebiges $f\in\mathcal{Z}$

$$\begin{aligned}(f*\varepsilon)(n)=(\varepsilon*f)(n)&=\sum_{d|n}\varepsilon(d)f\Big(\frac{n}{d}\Big)\\ &= f\Big(\frac{n}{1}\Big)+\sum_{d|n,d>1} 0\cdot f\Big(\frac{n}{d}\Big)=f(n).\end{aligned}$$

5. Zum Nachweis des Inversen zu $f\in\mathcal{Z}$ mit $f(1)\neq 0$ wird die Gleichung

$$\text{(I)}\qquad g*f=\varepsilon$$

rekursiv nach g aufgelöst.

a) Für $n=1$ lautet (I) $g(1)f(1)=1$, also $g(1)=\big(f(1)\big)^{-1}$.

b) Für $n\geq 1$ seien $g(1),\dots,g(n)$ so bestimmt, daß (I) für die Werte $m=1,\dots,n$ erfüllt ist.

Die Gültigkeit für $n+1$ besagt

$$0=\varepsilon(n+1)=g(n+1)f(1)+\sum_{d|n+1,d<n+1} g(d)\,f\Big(\frac{n+1}{d}\Big).$$

Dies kann nach $g(n+1)$ aufgelöst werden.

6. Zur Untergruppen–Eigenschaft von $\mathcal{M}$ ist zu zeigen, daß mit f auch das durch $g*f=\varepsilon$ eindeutig festgelegte g multiplikativ ist. Für $n>1$ sei schon gezeigt, daß für alle d_1,d_2 mit $(d_1,d_2)=1$ und $d_1d_2<n$ $g(d_1d_2)=g(d_1)g(d_2)$ erfüllt ist. Nach 5. ist, wegen $f(1)=1$,

$$g(n)=-\sum_{d|n,d<n} g(d)\,f\Big(\frac{n}{d}\Big).$$

Sei $n = n_1 n_2,\ (n_1, n_2) = 1,\ 1 < n_1, n_2 < n$. Dann ergibt sich mit der Induktionsvoraussetzung

$$\begin{aligned} g(n_1 n_2) &= - \sum_{d \mid n_1 n_2, d < n_1 n_2} g(d)\, f\Big(\frac{n_1 n_2}{d}\Big) \\ &= - \sum_{\substack{d_1 \mid n_1, d_2 \mid n_2 \\ d_1 < n_1 \vee d_2 < n_2}} g(d_1 d_2)\, f\Big(\frac{n_1}{d_1}\,\frac{n_2}{d_2}\Big) \\ &= - \sum_{d_1 \mid n_1} \sum_{d_2 \mid n_2} g(d_1)\, g(d_2)\, f\Big(\frac{n_1}{d_1}\Big)\ f\Big(\frac{n_2}{d_2}\Big) + g(n_1)\, g(n_2)\, f(1)\, f(1) \\ &= -(g * f)(n_1) \cdot (g * f)(n_2) + g(n_1)\, g(n_2) \\ &= -\varepsilon(n_1)\, \varepsilon(n_2) + g(n_1)\, g(n_2) = g(n_1)\, g(n_2). \end{aligned}$$

7. Zu (2). Mit $\underline{1}$ und μ ist auch $\underline{1} * \mu$ multiplikativ. Für Primzahlpotenzen gilt jedoch

$$\begin{aligned} (\underline{1} * \mu)(p^k) = (\mu * \underline{1})(p^k) &= \sum_{d \mid p^k} \mu(d)\, \underline{1}\Big(\frac{p^k}{d}\Big) \\ &= \sum_{d \mid p^k} \mu(d) = \mu(1) + \mu(p) = 0, \end{aligned}$$

also $(\underline{1} * \mu)(n) = 0$ für $n > 1$. □

5.10. Möbiussche Umkehrformel. Für zwei zahlentheoretische Funktionen f und F sind äquivalent

$a)$ $\quad F = f * \underline{1}$ und $\quad b)$ $\quad f = F * \mu$.

Ausgeschrieben:

$a')$ $\quad \forall\, n : F(n) = \sum_{d \mid n} f(d)$ und

$b')$ $\quad \forall\, n : f(n) = \sum_{d \mid n} F(d) \mu\Big(\frac{n}{d}\Big)$.

Die Umkehrformel erlaubt es also, eine Teilersummen–Identität $F(n) = \sum_{d \mid n} f(d)$ nach f hin aufzulösen.

Der Beweis ist mit den Rechenregeln in $(\mathcal{Z}, *)$ unmittelbar klar. Die Einschränkung $f(1) \neq 0$ ist hier nicht erforderlich, da kein $*$–Inverses zu f benötigt wird.

5.11. 1. Beispiel zur Umkehrformel.

(1) $\forall\, n : \sum_{d|n} \varphi(d) = n \quad (\varphi * \underline{1} = Id)$

(2) $\forall\, n : \varphi(n) = \sum_{d|n} \mu(d)\frac{n}{d} \quad (\varphi = \mu * Id).$

Beweis zu (1). Mit d durchläuft n/d alle Teiler von n. Zu $d|n$ sei

$$k_d = \left\{1 \le a \le n, (a,n) = \frac{n}{d}\right\}$$

Hierdurch wird die n–elementige Menge $\{1, 2, \ldots, n\}$ in disjunkte Klassen eingeteilt, insbesondere

$$n = \sum_{d|n} \# k_d.$$

Wegen

$$\begin{aligned} k_d &= \left\{1 \le a'\frac{n}{d} \le n, \left(a'\frac{n}{d}, n\right) = \frac{n}{d}\right\} \\ &= \left\{1 \le a' \le d, (a', d) = 1\right\} \end{aligned}$$

gilt $\#k_d = \varphi(d)$, also die Behauptung.

(2) folgt aus (1) mit der Umkehrformel. Damit ist erneut die Multiplikativität von φ gezeigt. (2) ergibt obendrein für $n = p_1^{a_1} \cdots p_k^{a_k}$

$$\begin{aligned} \frac{\varphi(n)}{n} &= \sum_{d|n} \frac{\mu(d)}{d} = \left(1 + \frac{\mu(p_1)}{p_1} + \ldots\right) \ldots \left(1 + \frac{\mu(p_k)}{p_k} + \ldots\right) \\ &= \left(1 - \frac{1}{p_1}\right) \ldots \left(1 - \frac{1}{p_k}\right) = \prod_{p|n} \left(1 - \frac{1}{p}\right). \end{aligned}$$

□

5.12. 2. Beispiel zur Umkehrformel.

$$\Lambda(n) = \sum_{d|n} \mu(d) \ln \frac{n}{d} = -\sum_{d|n} \mu(d) \ln d.$$

Die erste Aussage entsteht durch Umkehrung der Folgerung zu Def. 5.7., und daraus

$$\begin{aligned} \Lambda(n) &= \ln n \sum_{d|n} \mu(d) - \sum_{d|n} \mu(d) \ln d \\ &= \ln n \cdot \varepsilon(n) - \sum_{d|n} \mu(d) \ln d = -\sum_{d|n} \mu(d) \ln d. \end{aligned}$$

5.13. Legendresche Formel.
Sei $\mathcal{A} \subseteq \mathbb{N},\ \#\mathcal{A} < \infty,\ k \in \mathbb{N}$, und f eine zahlentheoretische Funktion. Dann gilt

$$\sum_{n\in\mathcal{A},(n,k)=1} f(n) = \sum_{d|k} \mu(d) \sum_{n\in\mathcal{A},n\equiv 0(d)} f(n).$$

Die Aussage ist klar, wenn man bedenkt, daß die Summationsbedingung $(n,k) = 1$ ersetzt werden kann durch den Faktor $\varepsilon\big((n,k)\big) = \sum_{d|n,d|k} \mu(d)$ in der Summe.

5.13. entspricht dem **Inklusion–Exklusion–Prinzip** in der Kombinatorik. Sei M eine endliche Menge, seien $E_1, \dots, E_L$ Eigenschaften oder Merkmale, die die Elemente von M besitzen können. Für $r \geq 1$ und $1 \leq \nu_1 < \dots < \nu_r \leq L$ sei $M_{\nu_1}, \dots, \nu_r$ die Teilmenge der Elemente von M, die die Eigenschaften $E_{\nu_1}, \dots, E_r$ haben. Bezeichne M' die Menge aller Elemente von M, die keine der Eigenschaften E_j haben. Dann gilt

$$\text{(IE)} \quad \#M' = \#M - \sum_{1\leq\nu_1\leq L} \#M_{\nu_1} + \sum_{1\leq\nu_1<\nu_2\leq L} \#M_{\nu_1\nu_2} - + \dots + (-1)^L \#M_{1,\dots,L}.$$

Im Hinblick auf 5.13. sei

$$f = \underline{1}, \quad k = p_1^{a_1} \dots p_L^{a_L} \quad (\text{OBdA } L \geq 1).$$

E_j bedeutet: $p_j | n$. Dann ergibt (IE) mit $M = \mathcal{A}$

$$\begin{aligned}
\sum_{n\in\mathcal{A},(n,k)=1} 1 &= \#M' \\
&= \#\mathcal{A} - \sum_{1\leq\nu_1\leq L} \#\{n \in \mathcal{A}, n \equiv 0(p_{\nu_1})\} \\
&\quad + \sum_{1\leq\nu_1<\nu_2\leq L} \#\{n \in \mathcal{A}, n \equiv 0(p_{\nu_1}), n \equiv 0(p_{\nu_2})\} - + \dots \\
&= \sum_{d|k} \mu(d) \#\{n \in \mathcal{A}, n \equiv 0(d)\}.
\end{aligned}$$

□

Die gängigen zahlentheoretischen Funktionen wie σ_α, φ und μ weisen ein sehr sprunghaftes Verhalten auf. Insbesondere ist es unmöglich, sie durch vertraute stetige Funktionen zu approximieren, so wie es beispielsweise in der Stirlingschen Formel mit $f(n) = n!$ geschieht. Betrachtet man hingegen die Summenfunktion

$$F(x) \underset{\text{Df}}{=} \sum_{n\leq x} f(n) \quad (x \in \mathbb{R}, \geq 1),$$

so läßt sich in vielen Fällen ein Verhalten der Art

$$F(x) = H(x) + R(x)$$

feststellen. Dabei bedeutet H (= Hauptglied) eine „glatte“ Funktion, während das im allgemeinen nicht genau angebbare R (= Restglied) von geringerer Größenordnung ist als H. Hierzu hat sich eine - nicht auf die Zahlentheorie beschränkte - Schreibweise als

sehr nützlich erwiesen.

5.14. Bachmann–Landau–Symbolik (Paul B., 1837–1920; Edmund L., 1877–1938).

$$f, f_1, f_2 : [1,\infty) \to \mathbb{C},$$
$$g : [1,\infty) \to \mathbb{R}^+ = \{t \in \mathbb{R}, t > 0\}$$

(Statt des Definitionsbereiches $[1,\infty)$ kann auch $[x_0,\infty)$ mit einem $x_0 > 1$ vorliegen).

(1) $\quad f = \mathrm{O}(g) \underset{\text{Df}}{\Leftrightarrow} \exists\, C > 0\, \forall\, x \geq 1 : |f(x)| \leq C\, g(x)$

(bzw. $|f|/g$ ist beschränkt).
Gesprochen: ***f* gleich Groß O von *g***. Oder: f höchstens von der Ordnung g.

$$f_1 = f_2 + \mathrm{O}(g) \underset{\text{Df}}{\Leftrightarrow} f_1 - f_2 = \mathrm{O}(g).$$

(2) $\quad f = \mathrm{o}(g) \underset{\text{Df}}{\Leftrightarrow} \lim\limits_{x\to\infty} \dfrac{f(x)}{g(x)}$ existiert und ist $= 0$

(bzw. $f(x) = \varepsilon(x)\, g(x)$, wobei $\varepsilon(x) \to 0$ für $x \to \infty$)
Gesprochen: ***f* gleich klein o von *g***. Oder: f ist von kleinerer Ordnung als g.

$$f_1 = f_2 + o(g) \underset{\text{Df}}{\Leftrightarrow} f_1 - f_2 = o(g).$$

Beispiele und Bemerkungen.

1) $\ln x = \mathrm{O}(x^\varepsilon)$ für jedes $\varepsilon > 0$ (wobei die „O– Konstante“ C von ε abhängt).
2) x nicht $= \mathrm{O}(\ln x)$, kurz: $x \neq \mathrm{O}(\ln x)$.
3) $\sin x = \mathrm{O}(1)$, aber $\neq \mathrm{o}(1)$. $f(x) = \mathrm{O}(1)$ besagt nicht, daß f konstant ist, sondern nur, daß es beschränkt ist.
4) Eine „asymptotische Formel“ $F(x) = H(x) + \mathrm{O}\big(R(x)\big)$ macht nur Sinn, wenn R von geringerer Ordnung als H ist. Z.B. ist $F(x) = x + \mathrm{O}(x^2)$ nicht aussagekräftiger als $F(x) = \mathrm{O}(x^2)$.
5) Bei konkurrierenden O–Termen reicht es, den größten zu behalten.

$$\mathrm{O}(x) + \mathrm{O}(x^2) + \mathrm{O}(e^x) = \mathrm{O}(e^x).$$

6) Aus $f(x) = \mathrm{o}(x)$ folgt $f(x) = \mathrm{O}(x)$. Die Umkehrung gilt i.a. nicht.
7) $[x] = x + \mathrm{O}(1)$, aber nicht $[x] = x + \mathrm{o}(1)$.
8) $d(n) = \mathrm{O}(n^\varepsilon)$ für jedes $\varepsilon > 0$, aber $d(n) \neq \mathrm{o}(\ln n)$, da $d(2^k) = k + 1 > k = \ln(2^k)/\ln 2$.
9) $\varphi(n) = \mathrm{O}(n)$, aber $\varphi(n) \neq \mathrm{o}(n)$, da $\varphi(p) = p - 1 \geq p/2$.

5.15. Hilfssatz (Partielle oder **abelsche Summation**; Niels Henrik A., 1802–29). Sei

$$f : \mathbb{N} \to \mathbb{C}, \quad F(x) \underset{\text{Df}}{=} \sum_{n\leq x} f(n);$$
$$g : [1,\infty) \to \mathbb{C}, \quad g \text{ stetig differenzierbar.}$$

Dann gilt für $x \geq 1$

$$\sum_{n \leq x} f(n)\, g(n) = F(x)\, g(x) - \int_1^x F(t)\, g'(t)\, dt.$$

Beweis.

$$\int_1^x F(t)\, g'(t)\, dt = \int_1^x \Big(\sum_{n \leq t} f(n) \Big) g'(t)\, dt$$

$$= \sum_{n \leq x} f(n) \int_n^x g'(t)\, dt = F(x)\, g(x) - \sum_{n \leq x} f(n)\, g(n).$$

□

5.16. Hilfssatz.

(1) $\displaystyle\sum_{n \leq x} \frac{1}{n} = \ln x + \gamma + \mathrm{O}\Big(\frac{1}{x}\Big)$ $\qquad (\gamma = 0,5772\ldots,$ Euler–Konstante).

(2) $\displaystyle\sum_{n \leq x} \ln n = x \ln x + \mathrm{O}(x)$

Beweis zu (1). Es wird 5.15. auf $f(n) = 1$ und $g(t) = \dfrac{1}{t}$ angewandt.
$F(x) = [x] = x - \{x\} = x + \mathrm{O}(1)$, also

$$\sum_{n \leq x} \frac{1}{n} = \frac{x - \{x\}}{x} + \int_1^x \big(t - \{t\}\big) t^{-2}\, dt$$

$$= 1 + \mathrm{O}\Big(\frac{1}{x}\Big) + \ln x - \int_1^x \{t\}\, t^{-2}\, dt.$$

Das letzte Integral konvergiert. Der Rest bis ∞ läßt sich im Betrag abschätzen durch

$$\leq \int_x^\infty t^{-2}\, dt = \frac{1}{x}.$$

Setzt man $\gamma = 1 - \displaystyle\int_1^\infty \{t\}\, t^{-2}\, dt$, ergibt sich die Behauptung.

Zu (2). Da ln monoton wächst, gilt für $2 \le n \le x$ die Ungleichung $\ln \frac{x}{n} \le \int\limits_{n-1}^{n} \ln \frac{x}{t}\, dt$, also für $x \ge 2$

$$\sum_{2\le n\le x} \ln \frac{x}{n} \le \int\limits_{1}^{[x]} \ln \frac{x}{t}\, dt \le \int\limits_{1}^{x} \ln \frac{x}{t}\, dt$$
$$= x \int\limits_{1}^{x} v^{-2} \ln v\, dv < x \int\limits_{1}^{\infty} v^{-2}\, \ln v\, dv = \mathrm{O}(x).$$

Damit hat man

$$(*) \qquad \sum_{n\le x} \ln \frac{x}{n} = \mathrm{O}(x) \quad \text{für } x \ge 2$$

und offensichtlich auch für $1 \le x < 2$. Hiermit ist (2) sofort einzusehen

$$\sum_{n\le x} \ln n = \sum_{n\le x} \ln x - \sum_{n\le x} \ln \frac{x}{n}$$
$$= \big(x + \mathrm{O}(1)\big) \ln x + \mathrm{O}(x) = x \ln x + \mathrm{O}(x).$$

□

Den Abschluß des Kapitels bilden einige Beispiele von asypmtotischen Formeln für Summen über multiplikative Funktionen.

5.17. Satz von Dirichlet.

$$\sum_{n\le x} d(n) = x \ln x + (2\gamma - 1)x + \mathrm{O}(x^{1/2}).$$

Beweis.
1. Eine schwächere Aussage kann mit 5.16 (1) hergeleitet werden.

$$\sum_{n\le x} d(n) = \sum_{d,k,dk\le x} 1 = \sum_{d\le x} \sum_{k\le x/d} 1$$
$$= \sum_{d\le x} \Big[\frac{x}{d}\Big] = x \sum_{d\le x} \frac{1}{d} + \mathrm{O}\Big(\sum_{d\le x} 1\Big)$$
$$= x\big(\ln x + \gamma + \mathrm{O}(x^{-1})\big) + \mathrm{O}(x) = x \ln x + \mathrm{O}(x).$$

2. Ist n keine Quadratzahl, so ist für $d|n$ eine der Zahlen d und $n|d$ kleiner als $n^{1/2}$ und die andere größer, also

$$d(n) = 2 \sum_{d|n, d<\sqrt{n}} 1 + \begin{cases} \mathrm{O}, & \text{falls} \quad n \ne m^2, \\ \mathrm{O}(1), & \text{sonst.} \end{cases}$$

Somit folgt

$$\begin{aligned}
\sum_{n\le x} d(n) &= 2\sum_{n\le x}\sum_{d\mid n, d<\sqrt{n}} 1 + \mathrm{O}(x^{1/2}) \\
&= 2\sum_{d<\sqrt{x}}\sum_{\substack{d^2<n\le x\\ n\equiv 0(d)}} 1 + \mathrm{O}(x^{1/2}) \\
&= 2\sum_{d<\sqrt{x}}\sum_{d<m\le x/d} 1 + \mathrm{O}(x^{1/2}) \\
&= 2\sum_{d<\sqrt{x}}\Big(\Big[\frac{x}{d}\Big] - d\Big) + \mathrm{O}(x^{1/2}) \\
&= 2x\sum_{d<\sqrt{x}}\frac{1}{d} + \mathrm{O}(x^{1/2}) - 2\sum_{d<\sqrt{x}} d + \mathrm{O}(x^{1/2}) \\
&= 2x\sum_{d\le\sqrt{x}}\frac{1}{d} - 2\sum_{d\le[\sqrt{x}]} d + \mathrm{O}(x^{1/2}) \\
&= 2x\big(\ln x^{1/2} + \gamma + \mathrm{O}(x^{-1/2})\big) - [x^{1/2}]\big([x^{1/2}]+1\big) + \mathrm{O}(x^{1/2}) \\
&= x(\ln x + 2\gamma - 1) + \mathrm{O}(x^{1/2}).
\end{aligned}$$

□

5.18. Satz.

(1) $\displaystyle\sum_{n\le x}\mu^2(n) = \frac{6x}{\pi^2} + \mathrm{O}(x^{1/2}).$

(2) $\displaystyle\sum_{n\le x}\varphi(n) = \frac{3x^2}{\pi^2} + \mathrm{O}(x\ln x + 1)$

Bemerkungen.
1. Da mit Hilfe von μ^2 die quadratfreien Zahlen gezählt werden, kann (1) auch so gelesen werden. Für $x = N \in \mathbb{N}$ wird

$$\frac{1}{N}\#\{n \le N,\ n \text{ quadratfrei}\} = \frac{6}{\pi^2} + \mathrm{O}(N^{-1/2}) \to \frac{6}{\pi^2} \quad \text{für} \quad N \to \infty.$$

Dies heißt, daß die relative Häufigkeit der quadratfreien unter den natürlichen Zahlen $\le N$ mit $N \to \infty$ gegen $6/\pi^2$ strebt. Grob: Etwa zwei Drittel aller natürlichen Zahlen sind quadratfrei.

2. Man betrachte die Paare natürlicher $m, n \le N$. (Um Verwechslungen mit dem ggT zu vermeiden, wird auf die Paar–Klammern verzichtet.) Wieviele davon sind relativ prim?

$$\begin{aligned} T(N) &\underset{\text{Df}}{=} \#\{m,n \le N,\ (m,n)=1\} \\ &= \#\{m,n \le N,\ m \le n,\ (m,n)=1\} + \#\{m,n \le N,\ n \le m,\ (m,n)=1\} \\ &\quad - \#\{m \le N,\ (m,m)=1\} \\ &= \sum_{n \le N} \varphi(n) + \sum_{m \le N} \varphi(m) - 1 \\ &= \frac{6}{\pi^2}\, N^2 + \mathrm{O}(N \ln N + 1) \end{aligned}$$

(Die Eins im Fehler soll nur die Positivität der Funktion in der O–Klammer bewirken). Daraus erhält man

$$\frac{1}{N^2} \#\{m,n \le N,\ (m,n)=1\} \to \frac{6}{\pi^2} \quad \text{für} \quad N \to \infty,$$

das heißt grob: Etwa $\frac{2}{3}$ aller Paare natürlicher Zahlen sind relativ prim.

Beweis zu (1). Es besteht die Identität

$$\mu^2(n) = \sum_{d, d^2/n} \mu(d). \tag{1.1}$$

Mit der linken ist auch die rechte Seite multiplikativ. Sie ist für $n = p^a$ $(a \in \mathbb{N}_0)$ gleich 1, falls $a = 0$ oder $= 1$, und gleich 0 für $a \ge 2$. Das stimmt mit $\mu^2(p^a)$ überein. Damit folgt

$$\begin{aligned} \sum_{n \le x} \mu^2(n) &= \sum_{n \le x} \sum_{d \le x^{1/2}, d^2/n} \mu(d) \\ &= \sum_{d \le x^{1/2}} \mu(d) \sum_{n \le x, n \equiv 0 (d^2)} 1 \\ &= \sum_{d \le x^{1/2}} \mu(d) \Big(\frac{x}{d^2} + \mathrm{O}(1) \Big) \\ &= \sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} + \mathrm{O}\Big(x \sum_{d > x^{1/2}} \frac{1}{d^2} \Big) + \mathrm{O}(x^{1/2}) \\ &= x \sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} + \mathrm{O}(x^{1/2}). \end{aligned} \tag{1.2}$$

Zur Berechnung der letzten Reihe benutzt man die Eulersche Formel $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$.

Da beide Reihen absolut konvergieren, kann ausmultipliziert und beliebig angeordnet

werden.

$$\sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} \cdot \frac{\pi^2}{6} = \sum_{d,k\in\mathbb{N}} \frac{\mu(d)}{(dk)^2}$$
$$= \sum_{n=1}^{\infty} \frac{1}{n^2} \sum_{d|n} \mu(d) = 1,$$

das heißt, die Reihe in (1.2) hat den Wert $6/\pi^2$.

2. Beweis zu (2). Mit 5.11(2) sieht man

$$\sum_{n\le x} \varphi(n) = \sum_{n\le x} n \sum_{d|n} \frac{\mu(d)}{d} = \sum_{d\le x} \frac{\mu(d)}{d} \sum_{\substack{n\le x \\ n\equiv 0(d)}} n$$
$$= \sum_{d\le x} \frac{\mu(d)}{d} \sum_{m\le x/d} md = \sum_{d\le x} \mu(d)\ \frac{1}{2}\ \left[\frac{x}{d}\right]\left(\left[\frac{x}{d}\right]+1\right)$$
$$= \frac{1}{2} \sum_{d\le x} \mu(d) \left(\frac{x^2}{d^2} + \mathrm{O}\Big(\frac{x}{d}\Big)\right) = \frac{x^2}{2} \sum_{d\le x} \frac{\mu(d)}{d^2} + \mathrm{O}\Big(x \sum_{d\le x} \frac{1}{d}\Big)$$
$$= \frac{x^2}{2} \sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} + \mathrm{O}(x\ln x) = \frac{3}{\pi^2}\ x^2 + \mathrm{O}(x\ln x).$$

□

Das Verhalten der Summe $\sum_{n\le x} \mu(n)$, also insbesondere die Häufigkeit quadratfreier Zahlen mit geradzahlig bzw. ungeradzahlig vielen Primfaktoren, ist wesentlich schwieriger zu studieren. Jedenfalls ist das hier mehrfach benutzte Prinzip, eine zahlentheoretische Funktion f als Faltung zu schreiben, und in der entstehenden Doppelsumme die Reihenfolge richtig zu wählen, bei μ nicht ohne weiteres anwendbar.

6. Kapitel. Elementare Primzahltheorie.

x bezeichnet eine reelle Zahl ≥ 1 (evtl. auch $\ge x_0 > 1$)

In diesem Abschnitt soll die Verteilung der Primzahlen, insbesondere ihre Häufigkeit innerhalb der natürlichen Zahlen, näher untersucht werden.

6.1. Def.

$$\pi(x) \underset{\text{Df}}{=} \#\{p \le x\},$$
$$\psi(x) \underset{\text{Df}}{=} \sum_{n\le x} \Lambda(n) = \sum_{p^k\le x} \ln p.$$

6.2. Hilfssatz. Sei $n! = \prod_{p \le n} p^{k_{p,n}}$.

Dann gilt

$$k_{p,n} = \sum_{m \in \mathbb{N}} \left[\frac{n}{p^m} \right].$$

Bemerkung. Die Summe wird nur formal bis unendlich erstreckt, da für $p^m > n$ der Summand $= 0$ ist.

Beweis. Sei $j \ge 0$. Ein $r \le n$, das p in genau j–ter Potenz enthält, liefert zu $k_{p,n}$ den Beitrag j. $p^j \| r$ heiße: $p^j | r$, $p^{j+1} \nmid r$. Damit wird

$$\begin{aligned} k_{p,n} &= \sum_{j \ge 0} j \#\{r \le n,\ p^j \| r\} \\ &= \sum_{j \ge 0} j \big(\#\{r \le n,\ p^j | r\} - \#\{r \le n,\ p^{j+1} | r\} \big) \\ &= \sum_{j \ge 0} j \Big(\Big[\frac{n}{p^j} \Big] - \Big[\frac{n}{p^{j+1}} \Big] \Big) \\ &= \sum_{j \ge 0} j \Big[\frac{n}{p^j} \Big] - \sum_{j \ge 1} (j-1) \Big[\frac{n}{p^j} \Big] \\ &= \sum_{m \ge 1} \Big[\frac{n}{p^m} \Big]. \end{aligned}$$

□

Numerische Untersuchungen brachten Mathematiker wie Euler, Legendre und Gauss zu der Vermutung, daß $\pi(x)$ sich näherungsweise wie $x / \ln x$ verhält. Das erste in diese Richtung führende Ergebnis ist der

6.3. Satz von Tschebytschev (1850, Pafnuti Lwowitsch T., 1821–1894).
Es existieren $C_1, \dots, C_4 > 0$, so daß für $x \ge 2$

(1) $\quad C_1 \dfrac{x}{\ln x} \le \pi(x) \le C_2 \dfrac{x}{\ln x}$,

(2) $\quad C_3 x \le \psi(x) \le C_4 x$ gilt.

Bemerkung. Auf die Werte der Konstanten wird hier nicht geachtet. Der angegebene Beweis führt beispielsweise zu $C_1 = \dfrac{1}{8}$, $C_2 = 12$.

Beweis.
1. Es wird sich als günstig erweisen, die Ungleichungen

(1.1) $$C_5 x \le \vartheta(x) \underset{\text{Df}}{=} \sum_{p \le x} \ln p \le C_6 x$$

herzuleiten. Aus (1.1) folgt (1). Denn mit der oberen Abschätzung in (1.1) ergibt sich

$$\begin{aligned}\pi(x) &= \pi\Big(\frac{x}{\ln x}\Big) + \sum_{\frac{x}{\ln x}<p\le x} \frac{\ln p}{\ln p} \\ &\le \frac{x}{\ln x} + \frac{1}{\ln(\frac{x}{\ln x})} \sum_{\frac{x}{\ln x}<p\le x} \ln p \\ &\le \frac{x}{\ln x} + \frac{C_7}{\ln x}\ \vartheta(x) \le (1 + C_6C_7)\frac{x}{\ln x}.\end{aligned}$$

Und umgekehrt

$$\pi(x) = \sum_{p\le x} \frac{\ln p}{\ln p} \ge \frac{\vartheta(x)}{\ln x} \ge C_5\ \frac{x}{\ln x}.$$

Analog folgt (2) aus (1.1).

$$\psi(x) \ge \vartheta(x) \ge C_5 x.$$

$$\psi(x) = \vartheta(x) + \sum_{p^k\le x, k\ge 2} \ln p. \tag{1.2}$$

In der letzten Summe treten nur $p \le x^{1/2}$ auf, also

$$\begin{aligned}\sum_{p^k\le x, k\ge 2} \ln p &\le \sum_{p\le x^{1/2}} \ln p \sum_{2\le k\le \frac{\ln x}{\ln p}} 1 \\ &\le \sum_{p\le x^{1/2}} \ln x \le x^{1/2} \ln x \le C_8 x,\end{aligned}$$

und somit

$$\psi(x) \le (C_6 + C_8)x.$$

2. Die entscheidende Idee zum hier geschilderten Beweis von (1.1) wurde 1932 vom damals 19–jährigen Paul Erdös (1913–1997) gefunden.
Es werde

$$B_n = \binom{2n}{n} = \frac{(2n)!}{(n!)^2}$$

betrachtet. B_n genügt den Ungleichungen

$$B_n < \sum_{\nu=0}^{2n} \binom{2n}{\nu} = (1+1)^{2n} = 4^n, \tag{2.1}$$

$$B_n = \frac{n+1}{1} \cdot \frac{n+2}{2} \cdots \frac{2n}{n} \ge 2^n. \tag{2.2}$$

3. Für $n < p \le 2n$ teilt p den Zähler $(2n)!$ von B_n, aber nicht den Nenner, also

$$P \underset{\text{Df}}{=} \prod_{n<p\le 2n} p \quad \text{teilt} \quad B_n, \quad \text{d.h.} \quad P \le B_n.$$

Wegen

$$P = \exp\Big(\sum_{n<p\le 2n} \ln p\Big) = \exp\big(\vartheta(2n) - \vartheta(n)\big)$$

ergibt sich daraus mit (2.1)

$$\vartheta(2n) - \vartheta(n) \le \ln B_n < n \cdot \ln 4. \tag{3.1}$$

Sei $x < 2^k \le 2x$. (3.1), angewandt auf $n = 2^k, 2^{k-1}, \dots,$ liefert

$$\begin{aligned}\vartheta(x) \le \vartheta(2^k) &= \big(\vartheta(2^k) - \vartheta(2^{k-1})\big) + \big(\vartheta(2^{k-1}) - \vartheta(2^{k-2})\big) + \dots \\ &< \ln 4 \cdot \big(2^{k-1} + 2^{k-2} + \dots\big) \le \ln 4 \cdot 2^k \le \ln 16 \cdot x.\end{aligned}$$

4. Bei der linken Ungleichung in (1.1) muß man etwas sorgfältiger vorgehen. Sei $B_n = \prod\limits_{p\le 2n} p^{k_p}$. Hilfssatz 6.2 ergibt

$$0 \le k_p = \sum_m \Big(\Big[\frac{2n}{p^m}\Big] - 2\Big[\frac{n}{p^m}\Big]\Big).$$

Hier brauchen nur die $m \le \ln(2n)/\ln p$ berücksichtigt zu werden. Die Klammer hat die Gestalt

$$\begin{aligned}&\frac{2n}{p^m} - \xi_1 - 2\Big(\frac{n}{p^m} - \xi_2\Big) \quad (0 \le \xi_\nu < 1) \\ &= 2\xi_2 - \xi_1 \le 1,\end{aligned}$$

da der Wert ganzzahlig ist. Somit erhält man $0 \le k_p \le [\ln(2n)/\ln p]$ und

$$\begin{aligned}\ln B_n = \ln\Big(\prod_{p\le 2n} p^{k_p}\Big) &\le \sum_{p\le 2n} \Big[\frac{\ln(2n)}{\ln p}\Big] \ln p \\ &= \sum_{p\le 2n} \ln p \sum_{k,p^k\le 2n} 1 = \sum_{m\le 2n} \Lambda(m) = \psi(2n).\end{aligned}$$

Mit (2.2) führt das zu

$$\psi(2n) \ge \ln 2 \cdot n.$$

Wie in 1. sieht man hiermit $\vartheta(2n) \ge C_9 n$ und daher

$$\vartheta(x) \ge C_5 x \qquad (x \ge 2).$$

□

6.4. Satz.

(1) $\displaystyle\sum_{n\le x} \frac{\Lambda(n)}{n} = \ln x + \mathrm{O}(1),$

(2) $\displaystyle\sum_{p\le x} \frac{\ln p}{p} = \ln x + \mathrm{O}(1),$

(3) $\displaystyle\sum_{p\le x} \frac{1}{p} = \ln\ln x + C + \mathrm{O}\Big(\frac{1}{\ln x}\Big)$ für $x \ge 3$ $(C = 0{,}2615\dots).$

Beweis zu (1). Mit der Folgerung zu 5.7. sieht man

$$\sum_{d\le x}\ln d=\sum_{d\le x}\sum_{n\mid d}\Lambda(n)=\sum_{n\le x}\Lambda(n)\Big[\frac{x}{d}\Big]$$
$$=x\sum_{n\le x}\frac{\Lambda(n)}{n}+\mathrm{O}\big(\psi(x)\big).$$

6.3(2) und 5.16(2) ergeben die Behauptung.

Zu (2).

$$0\le\sum_{n\le x}\frac{\Lambda(n)}{n}-\sum_{p\le x}\frac{\ln p}{p}$$
$$=\sum_{p^k\le x,k\ge 2}\frac{\ln p}{p^k}=\sum_{p\le\sqrt{x}}\ln p\sum_{k\ge 2}\frac{1}{p^k}$$
$$\le\sum_{p\le\sqrt{x}}\ln p\,\frac{1}{p^2}\cdot\frac{1}{1-\frac{1}{p}}=\mathrm{O}(1).$$

Mit (1) ergibt das Aussage (2).

(3) folgt aus (2) mit partieller Summation. Man nimmt

$$f(n)=\begin{cases}p^{-1}\ln p & \text{für}\quad n=p,\quad n\ge 2\\ 0 & \text{sonst,}\end{cases}$$
$$g(t)=(\ln t)^{-1}\quad\text{für}\quad t\ge 2$$
und stetig differenzierbar fortgesetzt bis $t=1$.

Nach (2) ist

$$F(x)=\begin{cases}0 & \text{für}\quad 1\le x<2,\\ \ln x+R(x) & \text{mit}\quad R(x)=\mathrm{O}(1)\quad\text{für}\quad x\ge 2.\end{cases}$$
$$\sum_{p\le x}\frac{1}{p}=\sum_{n\le x}f(n)\,g(n)$$
$$=\big(\ln x+R(x)\big)\frac{1}{\ln x}+\int_2^x\big(\ln t+R(t)\big)\frac{dt}{t\ln^2 t}$$
$$=1+\frac{R(x)}{\ln x}+\int_2^x\frac{dt}{t\ln t}+\int_2^x\frac{R(t)dt}{t\ln^2 t}.$$

Da wegen $R(t)=\mathrm{O}(1)$ das letzte Integral konvergiert, ist es $=C'+\mathrm{O}\Big(\frac{1}{\ln x}\Big)$, also

$$\sum_{p\le x}\frac{1}{p}=\ln\ln x+C+\mathrm{O}\Big(\frac{1}{\ln x}\Big).\qquad\square$$

Hinweis. Die Divergenz der Summe $\sum_{p\le x}\frac{1}{p}$ erfolgt außerordentlich langsam. Zum Beispiel wird der Wert 4 erst etwa bei $1,79\cdot 10^{18}$ erreicht.

6.5. Satz. Für $x\ge 3$ gilt

(1) $\displaystyle\sum_{n\le x}\omega(n) = x\ln\ln x + C + \mathrm{O}\big(x(\ln x)^{-1}\big)$

(2) $\displaystyle\sum_{n\le x}\Omega(n) = x\ln\ln x + C' + \mathrm{O}\big(x(\ln x)^{-1}\big)$

(3) $\displaystyle\sum_{n\le x}\big(\omega(n)-\ln\ln x\big)^2 = \mathrm{O}(x\ln\ln x).$

Bemerkung. (1) und (2) besagen, daß die $n\le x$ im Mittel etwa $\ln\ln x$ Primteiler bzw. Primfaktoren besitzen. Wegen des langsamen Wachstums des iterierten Logarithmus ist dies eine überraschend niedrige Anzahl. Ein Mittelwert kann dadurch erreicht werden, daß viele Werte wesentlich darunter und viele wesentlich darüber liegen. So ist es bei multiplikativen Funktionen oft der Fall. Bei additiven Funktionen ist vielfach eine Versammlung der Werte nahe dem Mittelwert zu beobachten. (3) kann als Varianz–Abschätzung gedeutet werden. Aus (3) folgt insbesondere für jedes $\varepsilon>0$

$$\begin{aligned}&\#\{n\le x,\ |\omega(n)-\ln\ln x|>\varepsilon\ln\ln x\}\\ &\le(\varepsilon\ln\ln x)^{-2}\sum_{n\le x}\big(\omega(n)-\ln\ln x\big)^2\\ &=\mathrm{o}(x),\end{aligned}$$

das heißt, „für die meisten“ $n\le x$ liegt $\omega(n)$ sehr dicht beim Mittelwert $\ln\ln x$. Mit Methoden der analytischen Zahlentheorie und der Stochastik zeigten **Erdös** und **Kac** 1940, daß ω dem „zentralen Grenzwertsatz“ genügt:

$$\frac{1}{x}\#\Big\{n\le x,\ \frac{\omega(n)-\ln\ln x}{\sqrt{\ln\ln x}}\le t\Big\}\to\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{t}e^{-y^2/2}dy$$

für jedes $t\in\mathbb{R}$ und $x\to\infty$.

Ein so regelmäßiges Verhalten zeigen multiplikative Funktionen im allgemeinen nicht.

Beweis zu (1).

$$\begin{aligned}\sum_{n\le x}\omega(n) &= \sum_{n\le x}\sum_{p\mid n}1 = \sum_{p\le x}\Big[\frac{x}{p}\Big]\\ &= x\sum_{p\le x}\frac{1}{p}+\mathrm{O}\big(\pi(x)\big)\\ &= x\big(\ln\ln x + C + \mathrm{O}\big((\ln x)^{-1}\big)+\mathrm{O}\big(x(\ln x)^{-1}\big)\big)\end{aligned}$$

nach 6.4(3) und 6.3(1).

(2) ergibt sich analog mit

$$\sum_{n\le x}\Omega(n)=\sum_{p^k\le x}\Big[\frac{x}{p^k}\Big]=\sum_{p\le x}\frac{1}{p}+\mathrm{O}\big(\pi(x)\big)$$
$$+x\sum_{p^k\le x,k\ge 2}\frac{1}{p^k}+\mathrm{O}\Big(\sum_{p^k\le x,k\ge 2}1\Big).$$

Wie schon mehrfach ausgeführt, erweist sich der Beitrag der p^k mit $k\ge 2$ als

$$x\Big(\sum_{p^k,k\ge 2}\frac{1}{p^k}+\mathrm{O}\big(\frac{1}{\ln x}\big)\Big)+\mathrm{O}\big(x^{1/2}\big)$$
$$=C''x+\mathrm{O}\big(x(\ln x)^{-1}\big).$$

Beweis zu (3). Sei $y=x^{1/4}$,

$$\tilde\omega(n)=\#\{p|n,\ p\le y\}.$$

Dann ist für $n\le x$

$$0\le\omega(n)-\tilde\omega(n)\le 3$$

da n höchstens drei Primteiler zwischen $x^{1/4}$ und x besitzt. Mit $\ln\ln x-\ln\ln y=\ln 4$ und der Ungleichung $(a+b)^2\le 4(a^2+b^2)$ sieht man – wenn $\ln\ln y$ mit L abgekürzt wird –

$$\begin{aligned}S&\underset{\text{Df}}{=}\sum_{n\le x}\big(\omega(n)-\ln\ln x\big)^2=\sum_{n\le x}\big(\tilde\omega(n)-L+\mathrm{O}(1)\big)^2\\&\le 4\sum_{n\le x}\big(\tilde\omega(n)-L\big)^2+\mathrm{O}(x)\\&=4\Big(\sum_{n\le x}\big(\tilde\omega(n)\big)^2-2L\sum_{n\le x}\tilde\omega(n)+L^2x\Big)+\mathrm{O}(x).\end{aligned}\tag{3.1}$$

$$\begin{aligned}\sum_{n\le x}\tilde\omega(n)^2&=\sum_{p_1,p_2\le y}\#\{n\le x,\ p_1|n,p_2|n\}\\&=\sum_{\substack{p_1,p_2\le y\\ p_1\ne p_2}}\Big[\frac{x}{p_1p_2}\Big]+\sum_{p\le y}\Big[\frac{x}{p}\Big]\\&=\sum_{p_1,p_2\le y}\Big[\frac{x}{p_1p_2}\Big]-\sum_{p\le y}\Big[\frac{x}{p^2}\Big]+\sum_{p\le y}\Big[\frac{x}{p}\Big]\\&=x\Big(\sum_{p\le y}\frac{1}{p}\Big)^2+\mathrm{O}\big(y^2\big)-x\sum_{p\le y}\frac{1}{p^2}+\mathrm{O}(y)+x\sum_{p\le y}\frac{1}{p}+\mathrm{O}(y)\\&=x\,L^2+\mathrm{O}(xL),\quad\text{mit 6.4.(3).}\end{aligned}$$

Wegen

$$\sum_{n\le x}\tilde\omega(n) = xL + \mathrm{O}(x) \quad \text{wird aus (3.1)}$$

$S \le 4\cdot \mathrm{O}(xL) + \mathrm{O}(x)$, also, da $S \ge 0$,

$$S = \mathrm{O}(x \ln\ln x),$$

wie behauptet. □

Mit zusätzlichem Aufwand kann man

$$\sum_{n\le x}\big(\omega(n) - \ln\ln x\big)^2 = x\ln\ln x + \mathrm{O}(x)$$

zeigen.

Numerischer Vergleich von $\pi(x)$ und $\dfrac{x}{\ln x}$ legt die Vermutung nahe, daß $\pi(x)/(x/\ln x)$ für $x\to\infty$ gegen Eins konvergiert. Dies ist der Inhalt des berühmten Primzahlsatzes. Das Problem besteht hier darin, die Existenz des Limes zu zeigen.

6.6. Satz (Tschebytschev).
Falls $\lim\limits_{x\to\infty} \pi(x)/(x/\ln x)$ existiert, hat er den Wert Eins.

Beweis. Sei A der angenommene Grenzwert, das heißt

$$\pi(x) = A\,\frac{x}{\ln x} + \varepsilon(x)\,\frac{x}{\ln x}$$

mit einer Funktion $\varepsilon(x)$, für die $\lim\limits_{x\to\infty}\varepsilon(x) = 0$ gilt. Partielle Summation mit

$$f(n) = \begin{cases} 1, & \text{falls} \quad n = p \\ 0 & \text{sonst,} \end{cases}$$

$F(x) = \pi(x)$ und $g(t) = t^{-1}$ ergibt für $x \ge 3$

$$\sum_{p\le x}\frac{1}{p} = \sum_{n\le x} f(n)\,g(n) = \frac{\pi(x)}{x} + \int_2^x \pi(t)\,\frac{dt}{t^2}$$

$$= \mathrm{O}\big((\ln x)^{-1}\big) + A\int_2^x \frac{dt}{t\ln t} + \int_2^x \frac{\varepsilon(t)}{t\ln t}\,dt. \tag{1}$$

Sei $\delta > 0$. Für $x \ge x_0(\delta)$ ist $|\varepsilon(x)| \le \delta$ und für $2 \le x \le x_0$ gilt $|\varepsilon(x)| \le C_1$. Also wird für $x \ge x_0$

$$\Big|\int_2^x \frac{\varepsilon(t)}{t\ln t}\,dt\Big| \le C_1\int_2^{x_0}\frac{dt}{t\ln t} + \delta\int_{x_0}^x \frac{dt}{t\ln t}$$

$$\le C_1\ln\ln x_0 + \delta\ln\ln x$$

$$\le 2\delta\ln\ln x, \quad \text{falls} \quad x \ge x_1(\delta).$$

Aus (1) erhält man daher

$$\sum_{p\le x}\frac{1}{p} = A\ln\ln x + \mathrm{o}(\ln\ln x),$$

was nach 6.4(3) nur für $A = 1$ richtig sein kann. □

Den entscheidenden Anstoß zum Beweis des Primzahlsatzes gab 1859 Bernhard Riemann (1826–1866) durch das Studium der nach ihm benannten **Riemannschen Zeta–Funktion**

$$\zeta(s) = \sum_{n=1}^{\infty}\frac{1}{n^s} \qquad (\mathrm{Res}\ > 1).$$

Ihre Bedeutung für die Primzahlverteilung wird sichtbar durch die im gleichen Bereich gültigen Formeln

$$\zeta(s) = \prod_p \Big(1-\frac{1}{p^s}\Big)^{-1}, \quad -\frac{\zeta'}{\zeta}(s) = \sum_n \frac{\Lambda(n)}{n^s}.$$

Nach der von Riemann vorgeschlagenen Methode konnten 1896 erstmals Jaques **Hadamard** (1866–1963) und Charles de la **Vallée–Poussin** (1866-1962) den Primzahlsatz beweisen. Einen elementaren Zugang, der ganz ohne komplexe Funktionentheorie auskommt, fanden 1948 Paul **Erdös** und Atle **Selberg**.

6.7. Primzahlsatz.
Es gelten die asymptotischen Formeln

(1) $$\pi(x) = \frac{x}{\ln x} + \mathrm{o}\Big(\frac{x}{\ln x}\Big), \quad \text{d.h.} \quad \pi(x)/(x/\ln x) \to 1 \quad \text{für} \quad x\to\infty,$$

(2) $$\psi(x) = x + \mathrm{o}(x) \quad \big(\psi(x)/x \to 1\big),$$

(3) $$M(x) \underset{\mathrm{Df}}{=} \sum_{n\le x}\mu(n) = \mathrm{o}(x), \quad \big(M(x)/x \to \mathrm{o}\big).$$

Ein analytischer Beweis zu (1) und (2) – deren Äquivalenz leicht einzusehen ist – wird in der Fortsetzungsvorlesung gegeben. Der elementare erfordert zwar keine weitgehenden Hilfsmittel, ist aber extrem verwickelt. Als Beispiel für kunstvolle elementare Umformungen, insbesondere mit Hilfe der Möbius–Funktion, soll hier die Implikation (2) ⇒ (3) gezeigt werden. Es sei also

(a) $$\psi(x) = \sum_{n\le x}\Lambda(n) = x + \varepsilon(x)\,x \quad \text{mit} \quad \lim_{x\to\infty}\varepsilon(x) = 0$$

Mit Hilfssatz 5.16(2) sieht man

$$\begin{aligned} M(x)\ln x &= \sum_{n\le x}\mu(n)\ln\Big(\frac{x}{n}\Big) + \sum_{n\le x}\mu(n)\ln n \\ &= \sum_{n\le x}\mu(n)\ln n + \mathrm{O}\Big(\sum_{n\le x}\ln\,\frac{x}{n}\Big) \\ &= \sum_{n\le x}\mu(n)\ln n + \mathrm{O}(x). \qquad \text{(b)} \end{aligned}$$

Wegen

$$\begin{aligned} \sum_{d|n}\mu\Big(\frac{n}{d}\Big)\Lambda(d) &= -\sum_{d|n}\mu\Big(\frac{n}{d}\Big)\sum_{k|d}\mu(k)\ln k \\ &= -\sum_{k|n}\mu(k)\ln k\sum_{d|n,d=k_1k}\mu\Big(\frac{n}{d}\Big) \\ &= -\sum_{k|n}\mu(k)\ln k\sum_{k_1|(n/k)}\mu\Big(\frac{n}{kk_1}\Big) \\ &= -\sum_{k|n}\mu(k)\ln k\;\varepsilon\Big(\frac{n}{k}\Big) = -\mu(n)\ln n \end{aligned}$$

erhält man

$$\begin{aligned} -\sum_{n\le x}\mu(n)\ln n &= \sum_{n\le x}\sum_{d|n}\mu\Big(\frac{n}{d}\Big)\,\Lambda(d) \\ &= \sum_{d,k,dk\le x}\mu(k)\,\Lambda(d) = \sum_{k\le x}\mu(k)\,\psi(x/k) \\ &= \sum_{k\le x}\mu(k)\frac{x}{k} + \sum_{k\le x}\mu(k)\,\varepsilon\Big(\frac{x}{k}\Big)\;\frac{x}{k} \\ &\underset{\text{Df}}{=} S_1(x)+S_2(x). \qquad \text{(c)} \end{aligned}$$

$$\begin{aligned} S_1(x) &= \sum_{k\le x}\mu(k)\Big(\Big[\frac{x}{k}\Big]+\mathrm{O}(1)\Big) = \sum_{k\le x}\mu(k)\sum_{d\le\frac{x}{k}}1+\mathrm{O}(x) \\ &= \sum_{k,d,dk\le x}\mu(k)+\mathrm{O}(x) = \sum_{n\le x}\sum_{k|n}\mu(k)+\mathrm{O}(x) \\ &= 1+\mathrm{O}(x)=\mathrm{O}(x) \qquad \text{(d)} \end{aligned}$$

Nach (a) existiert zu vorgegebenem $\delta>0$ ein $x_0=x_0(\delta)$, so daß für $x\ge x_0$ $|\varepsilon(x)|\le\delta$ erfüllt ist. Für $x\le x_0$ hat man $|\varepsilon(x)|\le C_1$. Damit ergibt sich für $x\ge x_0$

$$|S_2(x)| \le x\sum_{k\le x/x_0}\delta/k + x\sum_{x/x_0<k\le x}C_1/k.$$

Anwendung von Hilfssatz 5.16 (1) führt zu

$$\begin{aligned} |S_2(x)| &\le \delta x \sum_{k\le x} \frac{1}{k} + C_1 x\big(\ln x - \ln(x/x_0) + \mathrm{O}(1)\big) \\ &\le \delta x \ln x + C_2 x \quad \text{mit} \quad C_2 = C_2(\delta). \end{aligned}$$

Faßt man das Vorige zusammen, dann ergibt sich für $x \ge x_0(\delta)$

$$\begin{aligned} |M(x)| &\le \frac{1}{\ln x}\Big|\sum_{n\le x} \mu(n)\ln n\Big| + C_3 x/\ln x \\ &\le \delta x + C_4 \frac{x}{\ln x} \le 2\delta x \quad \text{für} \quad x \ge x_1(\delta). \end{aligned}$$

Dies besagt aber $M(x) = \mathrm{o}(x)$, wie behauptet. □

In ähnlicher Weise kann auch die Umkehrung (3) ⇒ (2) gezeigt werden. Insofern ist es gleichgültig, ob man (1), (2) oder (3) ansteuert. Dementsprechend gibt es elementare Beweise von vergleichbarem Schwierigkeitsgrad zu (2) oder (3). (1) wird seltener direkt gezeigt, da die Indikatorfunktion zur Menge der Primzahlen nicht so günstige Summationseigenschaften hat wie Λ.