3 Algebra und Zahlentheorie

3.1 Grundlegende Sätze der elementaren Zahlentheorie Mit a, b usw. bezeichnen wir, solange nichts anderes gesagt wird, immer ganze Zahlen. Wir schreiben $a \mid b$, wenn a ein Teiler von b ist, wenn also $b = aq$ für eine ganze Zahl q.

$p \in \mathbb{N}$ heißt *Primzahl* genau dann, wenn p genau zwei Teiler hat, nämlich 1 und p. Damit ist 1 keine Primzahl, weil sie nur einen Teiler besitzt.

Aus der Schule ist der folgende Satz über die eindeutige Zerlegung einer natürlichen Zahl in ihre Primfaktoren bekannt.

Satz 3.1 (Fundamentalsatz der Arithmetik) *Ist $p_1 = 2$, $p_2 = 3$, $p_3 = 5, \ldots$ die Folge der Primzahlen, so gibt es zu jeder natürlichen Zahl $a > 1$ eindeutige Exponenten $r_1, \ldots, r_k \in \mathbb{N}_0$ mit*

$$a = p_1^{r_1} p_2^{r_2} \ldots p_k^{r_k}, \quad r_k > 0.$$

Der Satz ist hoffentlich intuitiv klar. Wenn wir vor der rechten Seite ein Minuszeichen setzen dürfen, gilt er auch für alle $a \in \mathbb{Z}$ mit $|a| > 1$.

Unter den vielen Anwendungen dieses Satzes erwähnen wir: Ist eine Primzahl p Teiler von ab, so ist $p \mid a$ oder $p \mid b$. Schauen wir uns nämlich die Primfaktorzerlegungen von a und b an, so muss p in einer der beiden vorkommen.

Wir sagen, a ist *kongruent zu b modulo m*, wenn die natürliche Zahl m ein Teiler von $b - a$ ist, und schreiben dafür $a \equiv b \mod m$. Die Zahl m heißt *Modul* der Kongruenz. Die Differenz zweier gerader Zahlen ist gerade, sie sind daher kongruent modulo 2. Ebenso sind zwei ungerade Zahlen kongruent modulo 2, weil ihre Differenz ebenfalls geradzahlig ist.

Sind zwei Zahlen kongruent modulo m, so muss die Differenz der beiden Zahlen ein ganzzahliges Vielfaches von m sein, daher

$$a \equiv b \mod m \quad \Leftrightarrow \quad m \mid a - b \quad \Leftrightarrow \quad a = b + qm \text{ für ein } q \in \mathbb{Z}. \tag{3.6}$$

Ist a eine natürliche Zahl, so hinterlässt sie beim Teilen durch m einen Rest in der Menge $\{0, 1, \ldots, m-1\}$. Zwei natürliche Zahlen a, b sind genau dann kongruent modulo m, wenn sie beim Teilen durch m den gleichen Rest besitzen, denn dieser Rest fällt ja in $b - a$ heraus. Dieses Prinzip lässt sich auch auf negative Zahlen ausdehnen, wenn wir m auf den Rest addieren.

Aus (3.6) entnimmt man direkt die Rechenregeln

$$a \equiv b \mod m, \ \ c \equiv d \mod m \quad \Rightarrow \quad a \pm c \equiv b \pm d \mod m \ \text{ und } \ ac \equiv bd \mod m,$$

insbesondere auch

$$a \equiv b \mod m \quad \Rightarrow \quad a^k \equiv b^k \mod m.$$

Diese Regeln lassen sich folgendermaßen zusammenfassen: Ist $p(x)$ ein Polynom mit ganzzahligen Koeffizienten, so gilt

$$a \equiv b \mod m \ \Rightarrow \ p(a) \equiv p(b) \mod m.$$

Zwei natürliche Zahlen heißen *teilerfremd*, wenn sie nur 1 als gemeinsamen Teiler besitzen.

Vorsicht ist bei der Division in der Kongruenzrelation geboten. Es gilt $m \equiv 2m \mod m$, aber $1 \not\equiv 2 \mod m$. Daher

$$ac \equiv bc \mod m, \ c \text{ und } m \text{ teilerfremd} \ \Rightarrow \ a \equiv b \mod m.$$

Man beweist diese Regel, indem man für $ac \equiv bc \mod m$ die äquivalente Form $m \mid (b - a)c$ betrachtet. Sind m und c teilerfremd, so kommen in den Primfaktorzerlegungen von m und c nur verschiedene Primzahlen vor. Damit muss m ein Teiler von $b - a$ sein.

Satz 3.2 (Kleiner Satz von Fermat) *Sei a positiv und p eine Primzahl. Dann gilt*

$$a^p \equiv a \mod p.$$

Ist p kein Teiler von a, folgt hieraus

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

Beweis: Wir verwenden vollständige Induktion über a. Für $a = 1$ ist $p \mid 1^p - 1$ richtig. Als Induktionsvoraussetzung nehmen wir an, dass die Behauptung für a richtig ist, dass also $p \mid a^p - a$. Wir müssen zeigen, dass

$$p \mid (a+1)^p - (a+1).$$

Mit der binomischen Formel Satz 2.7 erhalten wir

$$\begin{aligned}(a+1)^p - (a+1) &= \sum_{i=0}^{p} \binom{p}{i} a^i - (a+1) \\ &= a^p + 1 + \sum_{i=1}^{p-1} \binom{p}{i} a^i - (a+1) \\ &= a^p - a + \sum_{i=1}^{p-1} \binom{p}{i} a^i. \qquad (3.7)\end{aligned}$$

Auf der rechten Seite ist $a^p - a$ aufgrund der Induktionsvoraussetzung durch p teilbar. Die Binomialkoeffizienten

$$\binom{p}{i} = \frac{p!}{i!(p-i)!}$$

sind ganzzahlig. Ist p eine Primzahl, so kann der Faktor p im Zähler für $i \neq 0$ und $i \neq p$ nicht herausgekürzt werden. Da die Binomialkoeffizienten in (3.7) durch p teilbar sind, ist auch die linke Seite von (3.7) durch p teilbar. □

Beispiel 3.3 Zeigen Sie, dass für jede positive Zahl n gilt $30 \mid n^5 - n$.

Lösung: Die Teilbarkeit durch 5 folgt aus dem Fermatschen Satz. Wegen

$$n^5 - n = (n-1)n(n+1)(n^2+1)$$

ist $n^5 - n$ außerdem durch 2 und durch 3 teilbar, denn beide Zahlen müssen Teiler einer Zahl in der Folge $n-1, n, n+1$ sein. □

Aus der Elementarmathematik gut bekannt ist die „Division mit Rest“: Sind $a \in \mathbb{N}_0$ und $b \in \mathbb{N}$, so gibt es eindeutig bestimmte Zahlen $m, r \in \mathbb{N}_0$ mit

$$a = mb + r, \quad 0 \leq r < b. \qquad (3.8)$$

Dazu überlegt man sich, dass jede nichtnegative ganze Zahl in genau einem Intervall $[0, b), [b, 2b), \ldots$ liegen muss. Daher sind sowohl m als auch r eindeutig bestimmt.

Wir können (3.8) auch für ganzzahliges a übernehmen. In diesem Fall existieren eindeutige $b \in \mathbb{Z}$ und $0 \leq r < m$ mit

$$a = mb + r. \qquad (3.9)$$

Wir bekommen damit die *Ganzzahldivision* $a \operatorname{div} m = b$, die auch in den meisten Programmiersprachen implementiert ist. Man beachte $15 \operatorname{div} 7 = 2$, aber, weil $r \geq 0$ gefordert wird, $-15 \operatorname{div} 7 = -3$.

Den in (3.9) auftretenden Rest $r \in \{0, \ldots, m-1\}$ bezeichnen wir als *Rest von a modulo m* und schreiben dafür $a \bmod m$. Mit diesen Bezeichnungen können wir (3.9) in der Form

$$a = m \cdot (a \operatorname{div} m) + (a \bmod m)$$

schreiben.

d heißt *größter gemeinsamer Teiler* von $a \in \mathbb{N}$ und $b \in \mathbb{N}$, wenn $d \,|\, a, b$ und wenn aus $t \,|\, a$ und $t \,|\, b$ folgt, dass $t \,|\, d$. Wir schreiben dafür $d = ggT(a, b)$. Für teilerfremde Zahlen gilt $ggT(a, b) = 1$. Den größten gemeinsamen Teiler kann man aus den Primfaktorzerlegungen der Zahlen a und b bestimmen, indem man das Produkt der gemeinsamen Primfaktoren bildet. Dieses Verfahren ist allerdings sehr langsam, so dass man besser auf den im Beweis des nächsten Satzes dargestellten *erweiterten Euklidischen Algorithmus* zurückgreift.

Satz 3.4 (Satz vom größten gemeinsamen Teiler, Lemma von Bézout) *Für $a, b \in \mathbb{N}$ existiert genau ein größter gemeinsamer Teiler $d \in \mathbb{N}$. Ferner gibt es Zahlen $\alpha, \beta \in \mathbb{Z}$ mit*

$$d = \alpha a + \beta b.$$

Beweis: Wir dürfen $a > b$ annehmen. Wir wenden fortgesetzte Division mit Rest nach folgendem Schema solange an, bis der Rest 0 entsteht:

$$\begin{aligned}
a &= b \cdot q_1 + r_1, && 0 < r_1 < b,\\
b &= r_1 \cdot q_2 + r_2, && 0 < r_2 < r_1,\\
r_1 &= r_2 \cdot q_3 + r_3, && 0 < r_3 < r_2,\\
&\vdots\\
r_{k-4} &= r_{k-3} \cdot q_{k-2} + r_{k-2}, && 0 < r_{k-2} < r_{k-3},\\
r_{k-3} &= r_{k-2} \cdot q_{k-1} + r_{k-1}, && 0 < r_{k-1} < r_{k-2},\\
r_{k-2} &= r_{k-1} \cdot q_k + r_k, && 0 < r_k < r_{k-1},\\
r_{k-1} &= r_k \cdot q_k.
\end{aligned}$$

Da die Folge der Reste nichtnegativ und streng monoton fallend ist, kommen wir nach endlich vielen Schritten zum Rest 0. Wir zeigen nun, dass die Zahl r_k der größte gemeinsame Teiler von a und b ist. Liest man nämlich die Gleichungen von unten nach oben, so kommt man auf die Beziehungen

$$r_k \,|\, r_{k-1},\ r_k \,|\, r_{k-2}, \ldots\ r_k \,|\, b,\ r_k \,|\, a,$$

womit r_k ein gemeinsamer Teiler von b und a ist. Für einen beliebigen gemeinsamen Teiler t von a und b kommt man, wenn man die Gleichungen von oben nach unten liest, auf

$$t \,|\, r_1,\ t \,|\, r_2, \ldots,\ t \,|\, r_k.$$

Damit ist in der Tat $r_k = ggT(a, b)$.

Zum Nachweis von $r_k = \alpha a + \beta b$ gehen wir die obigen Gleichungen nochmals von unten nach oben durch. Aus der vorletzten Gleichung ergibt sich

$$r_k = r_{k-2} - r_{k-1} q_k$$

und mit der darüberstehenden Gleichung folgt

$$r_k = (1 + q_{k-1} q_k) r_{k-2} - q_k r_{k-3}.$$

Auf die gleiche Weise kann man hier r_{k-2} durch eine Kombination von r_{k-4} und r_{k-3} darstellen und verbleibt am Ende mit

$$r_k = \alpha a + \beta b.$$

□

Beispiel 3.5 Das im letzten Beweis dargestellte Verfahren ist deshalb so effektiv, weil sich die r_i in jedem Schritt mindestens halbieren. Für $a = 38$ und $b = 10$ erhält man

$$38 = 10 \cdot 3 + 8$$
$$10 = 8 \cdot 1 + 2$$
$$8 = 2 \cdot 4,$$

also $ggT(38, 10) = 2$. α und β bestimmt man aus

$$\begin{aligned} 2 &= 10 - 1 \cdot 8 \\ &= 10 - 1 \cdot (38 - 10 \cdot 3) = 4 \cdot 10 - 1 \cdot 38, \end{aligned}$$

also $\alpha = -1$ und $\beta = 4$. □

3.2 Stellenwertsysteme Sei $g \in \mathbb{N} \setminus \{1\}$. Die *g-adische Darstellung* einer natürlichen Zahl n ist von der Form

$$n = a_0 \cdot g^0 + a_1 \cdot g^1 + \ldots + a_s g^s = \sum_{k=0}^{s} a_k g^k \tag{3.10}$$

mit „Ziffern“ $a_k \in \{0, 1, \ldots, g-1\}$. Für die *Basis* g hat sich im täglichen Gebrauch $g = 10$ durchgesetzt, wir schreiben ja $a_s \ldots a_0$ für eine solche Dezimalzahl. Relikte anderer Basen sind bei uns noch erkennbar: Stunden, Minuten und Sekunden sind im 60er System strukturiert, das Dutzend und das Gros erinnern an die Basis 12.

Für die Darstellung in (3.10) schreiben wir

$$n = a_s a_{s-1} \ldots a_{0g}.$$

Diese Darstellung ist offenbar eindeutig und es gilt

$$a_k = (n \operatorname{div} g^k) \mod g \quad \text{für } k = 0, 1, \ldots.$$

Für die praktische Rechnung dividiert man fortgesetzt ganzzahlig durch g und nimmt anschließend die Ergebnisse modulo g. Z.B. für $n = 50$ und $g = 2$ gilt

$$50 \operatorname{div} 1 = 50,\ 50 \operatorname{div} 2 = 25,\ 25 \operatorname{div} 2 = 12,\ 12 \operatorname{div} 2 = 6,\ 6 \operatorname{div} 2 = 3,\ 3 \operatorname{div} 2 = 1,$$

daher $50_{10} = 110010_2$.

Kommen wir nun zur Darstellung ganzer Zahlen im Rechner. Stehen uns $s+1$ Bits im Binärsystem $g = 2$ zur Verfügung, so geht ein Bit für das Vorzeichen verloren. Es ist aber ungünstig, explizit das Vorzeichen zu codieren, weil das bei der Addition zu Fallunterscheidungen führt, denn die Vorzeichen der beiden zu addierenden Zahlen entscheiden darüber, ob addiert oder subtrahiert wird. Besser ist es daher, *Zweierkomplemente* zu verwenden, nämlich

$$[a_s a_{s-1} \ldots a_0]_2 = a_{s-1} \ldots a_{02} - 2^s a_s.$$

Bei $a_s = 0$ werden die nichtnegativen ganzen Zahlen von $0 = [0 \ldots 0]_2$ bis $2^s - 1 = [011 \ldots 1]_2$ codiert. Die negativen Zahlen laufen von $-1 = [11 \ldots 1]_2$ bis $-2^s = [10 \ldots 0]_2$. Bei der Addition solcher Zahlen führt man eine normale binäre Addition durch, ohne die besondere Bedeutung der Stelle s zu berücksichtigen. Allerdings fällt ein Übertrag von der Stelle s unter den Tisch. Solange sich die Zahlen im angegebenen Bereich bewegen, ist diese Addition korrekt:

Beispiel 3.6 Für $s = 3$ können die Zahlen von $-2^3 = -8$ bis $2^3 - 1 = 7$ dargestellt werden. Es ist klar, dass zwei nichtnegative Zahlen korrekt addiert werden, solange die 7 nicht überschritten wird. Andernfalls erhalten wir z.B. $4 + 4 = [0100]_2 + [0100]_2 \stackrel{?}{=} [1000]_2 = -8$. Bei der Summe zweier negativer Zahlen darf die Summe nicht kleiner als -8 werden, z.B $-1 - 1 = [1111]_2 + [1111]_2 = [1110]_2 = -2$, aber $-4 - 5 = [1100]_2 + [1011]_2 \stackrel{?}{=} [0111]_2 = 7$. □

3.3 Untergruppen und der Satz von Lagrange Wir hatten $\mathbb{G} = (G, e, \circ)$ eine Gruppe (siehe Abschnitt 1.5) genannt, wenn die zweistellige Operation $\circ$ assoziativ ist, das neutrale Element e besitzt und es zu jedem x ein x^{-1} gibt mit $x \circ x^{-1} = x^{-1} \circ x = e$.

$U \subset G$ heißt *Untergruppe* von G, wenn $(U, e, \circ)$ ebenfalls eine Gruppe ist mit der gleichen Operation „$\circ$“ eingeschränkt auf $U \times U$. In diesem Fall schreiben wir $U \leq G$ und, falls $U \neq G$, $U < G$.

Satz 3.7 (Untergruppenkriterium) *$(G, e, \circ)$ sei eine Gruppe. $U \subset G$ ist genau dann eine Untergruppe von G, wenn*

(a) *$U \neq \emptyset$,*

(b) *Mit $x, y \in U$ ist auch $x \circ y \in U$.*

(c) *Zu jedem $x \in U$ existiert $x^{-1} \in U$.*

Beweis: Eine Untergruppe erfüllt (a),(b),(c). Wegen (a) gibt es ein $x \in U$, das nach (c) ein inverses Element $x^{-1} \in U$ besitzt. Nach (b) ist dann auch $x \circ x^{-1} = e \in U$. Das Assoziativgesetz gilt in U, weil es in G gilt. □

$U = \{e\}$ und $U = G$ sind immer Untergruppen einer Gruppe G, man nennt sie die *trivialen Untergruppen*. Weitere Beispiele:

- Die ganzen Zahlen sind mit der üblichen Addition eine Untergruppe der rationalen Zahlen. Die geraden Zahlen sind wiederum eine Untergruppe der ganzen Zahlen.
- Die Menge der Permutationen von $A_n = \{1, 2, \ldots, n\}$ mit $p(1) = 1$ ist eine Untergruppe, die die gleiche Struktur wie die Permutationen der Menge A_{n-1} besitzt.

Satz 3.8 (Satz von Lagrange) *Sei G eine endliche Gruppe. Ist U eine Untergruppe von G, so ist ihre Kardinalität $|U|$ ein Teiler von $|G|$.*

Beweis: Sei U Untergruppe der endlichen Gruppe G. Für jedes $x \in G$ betrachten wir die *Nebenklasse*

$$xU = \{xy : y \in U\}.$$

Ist $xy_1 = xy_2$ für $y_1, y_2 \in U$, so folgt $y_1 = y_2$. Damit sind alle Nebenklassen gleich groß und haben $|U|$ viele Elemente. Haben zwei Nebenklassen x_1U, x_2U ein Element $x_1y_1 = x_2y_2$ gemeinsam, so sind die Nebenklassen gleich wegen

$$x_1U = x_1(y_1U) = (x_1y_1)U = x_2y_2U = x_2U.$$

Wegen $x = xe \in xU$ kommt jedes $x \in G$ in einer Nebenklasse vor. Daher unterteilen die Nebenklassen die Menge G in endlich viele disjunkte Teilmengen mit $|U|$ Elementen, womit $|G|$ ein ganzzahliges Vielfaches von $|U|$ sein muss. □

3.4 Restklassenkörper und der Satz von Wilson Wir hatten in Abschnitt 1.6 $\mathbb{K} = (K, 0, 1, +, \cdot)$ Körper genannt, wenn $(K, 0, +)$ und $(K \setminus 0, 1, \cdot)$ abelsche Gruppen sind und das Distributivgesetz $a \cdot (b + c) = a \cdot b + a \cdot c$ gilt. Das inverse Element von a bezüglich der Addition schreiben wir als $-a$, das der Multiplikation als a^{-1}. Üblicherweise verwendet man $a - b$ statt $a + (-b)$ und ab statt $a \cdot b$. Weiter gilt die bekannte Regel „Punktrechnung geht vor Strichrechnung“.

Hieraus lassen sich alle Rechenregeln ableiten, die wir von den reellen Zahlen kennen:

Satz 3.9 *Sei $(K, 0, 1, +, \cdot)$ ein Körper. Dann gilt:*

(a) *Die neutralen Elemente der Addition und der Multiplikation sind eindeutig bestimmt.*

(b) *Das inverse Element $-a$ der Addition und das inverse Element a^{-1}, $a \neq 0$, der Multiplikation sind eindeutig bestimmt.*

(c) *Es gilt $a \cdot 0 = 0$, $(-1)a = -a$, $(-a)b = -ab$.*

(d) *Ist $a \neq 0$, so folgt aus $ab = ac$, dass $b = c$.*

(e) *Ein Körper ist* nullteilerfrei*, d.h. aus $ab = 0$ folgt $a = 0$ oder $b = 0$.*

Beweis: (a) und (b) folgen aus Satz 1.6.

(c) Aus $a0 = a(0 + 0) = a0 + a0$ folgt $a0 = 0$. Aus $0 = 0a = (1 + (-1))a = a + (-1)a$ folgt $(-1)a = -a$. Mit $(-1)a = -a$ folgt $(-a)b = (-1)ab = (-1)(ab) = -ab$.

(d) Dies ist wieder Satz 1.6.

(e) Ist $ab = 0$ und $b \neq 0$, so $a = abb^{-1} = 0b^{-1} = 0$ wegen (c). □

Sei $n > 1$ eine natürliche Zahl. Dann ist die auf $\mathbb{Z} \times \mathbb{Z}$ erklärte Relation $a \equiv b \mod n$ eine Äquivalenzrelation. Denn sie ist reflexiv und symmetrisch sowie transitiv wegen

$$a \equiv b \mod n,\ b \equiv c \mod n \quad \Rightarrow \quad a = b + qm,\ b = c + q'm \quad \Rightarrow \quad a = c + (q + q')m.$$

Zwei ganze Zahlen sind daher äquivalent, wenn sie bei der Division durch n den gleichen Rest modulo n besitzen. Die zugehörigen Äquivalenzklassen besitzen daher die natürlichen Vertreter $0, 1, \ldots, n-1$. Die Menge

$$\mathbb{Z}_n = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$$

bildet eine Partition von $\mathbb{Z}$.

Auf $\mathbb{Z}_n$ können wir die Operationen

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

definieren. Wir beweisen die Korrektheit dieser Definitionen, also die Unabhängigkeit von den Vertretern der jeweiligen Äquivalenzklasse. Ist $a' \in \overline{a}$, $b' \in \overline{b}$, so $a' = a + pn$, $b' = b + qn$. Dann

$$a' + b' = a + b + (p + q)n \in \overline{a + b}, \quad a' \cdot b' = ab + aqn + bpn + pqn^2 \in \overline{a \cdot b}.$$

Alternativ wird auch

$$\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$$

geschrieben. Man addiert und multipliziert diese Zahlen „normal“ in $\mathbb{N}_0$ und ordnet das Ergebnis der zugehörigen Äquivalenzklasse beziehungsweise ihrem Vertreter in $\mathbb{Z}_n$ zu. Um nicht in Konfusion mit den üblichen Operationen zu kommen, schreiben wir dann $+_n$ und $\cdot_n$ für die so definierten Operationen. Beispielsweise gilt in $\mathbb{Z}_4$ $2 \cdot 3 = 6 \equiv 2 \mod 4$, daher $2 \cdot_4 3 = 2$.

Wir erhalten für $n = 2$ die Tafeln

$+_2$	0	1
0	0	1
1	1	0

$\cdot_2$	0	1
0	0	0
1	0	1

Für $n = 4$:

$+_4$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$\cdot_4$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

Welche algebraischen Eigenschaften haben die so definierten Operationen? Zunächst ist klar, dass beide Operationen assoziativ und kommutativ sind. Ferner ist 0 neutral bezüglich der Addition. Zu $a \in \mathbb{Z}_n$ ist $n-a$ das inverse Element bezüglich der Addition, denn es gilt $a+(n-a) = n \equiv 0 \mod n$. Damit ist $(\mathbb{Z}_n, 0, +)$ eine kommutative Gruppe. Das Distributivgesetz wird von der Rechnung mit ganzen Zahlen geerbt und ist daher ebenfalls gültig. 1 ist neutrales Element der Multiplikation, was $\mathbb{Z}_n$ zu einem kommutativen Ring macht (siehe Abschnitt 1.6).

Wie die Tafel oben rechts zeigt, gibt es für die 2 bei $n = 4$ kein inverses Element. Allgemein ist für zusammengesetztes $n = kl$ die Struktur kein Körper wegen $k \cdot_n l = n \equiv 0 \mod n$, sie ist damit nicht nullteilerfrei.

Bei Primzahlen p haben wir dagegen:

Satz 3.10 *Ist p eine Primzahl, so ist $\mathbb{Z}_p$ zusammen mit den Operationen $+_p$ und $\cdot_p$ ein Körper, der* Restklassenkörper modulo p *genannt wird. Für $a \neq 0$ gilt $-a = p - a$ sowie $a^{-1} \equiv a^{p-2} \mod p$. Genau die Elemente 1 und $p-1$ sind zu sich selbst invers bezüglich der Multiplikation $\cdot_p$, alle anderen Elemente $\neq 0$ lassen sich zu Paaren a, a', $a \neq a'$, zusammenfassen mit $a \cdot_p a' = 1$.*

Beweis: Nach dem kleinen Satz von Fermat 3.2 gilt $a^{p-1} \equiv 1 \mod p$ für alle $a \in \{1, \ldots, p-1\}$. Somit $a \cdot a^{p-2} \equiv 1 \mod p$ und die Restklasse modulo p von a^{p-2} ist das inverse Element von a bezüglich $\cdot_p$.

Aus $a^2 \equiv 1 \mod p$ folgt $(a-1)(a+1) \equiv 0 \mod p$, was genau für $a = 1$ oder $a = p-1$ erfüllt ist. □

Satz 3.11 (Wilson) *Für jede Primzahl p gilt*

$$(p-2)! \equiv 1 \mod p, \quad (p-1)! \equiv -1 \mod p.$$

Beweis: Es gilt $(p-2)! = 2 \cdot \ldots \cdot (p-2)$. Nach dem letzten Satz wird dieses Produkt von Paaren mit $aa' \equiv 1 \mod p$ gebildet, daher $(p-2)! \equiv 1 \mod p$. Wir multiplizieren dies mit $p-1$ und erhalten den zweiten Teil der Behauptung. □

Es gilt auch die Umkehrung dieses Satzes: Ist $(p-1)! \equiv -1 \mod p$, so ist p eine Primzahl.

3.5 Geheimcodes dienen dazu, Nachrichten so zu verschlüsseln, dass sie nur vom Empfänger lesbar gemacht werden können. Wir untersuchen zunächst klassische Verschlüsselungen und behandeln dann die moderne RSA-Technik.

Die Substitution besteht darin, jeden Buchstaben eines Textes durch einen anderen zu ersetzen. Im Folgenden verwenden wir kleine Buchstaben für den zu verschlüsselnden Text (=Klartext) und große Buchstaben für die verschlüsselte Nachricht (=Geheimtext). Verwenden wir die Zuordnung

Klartextalphabet:	a b c d e f g h i j k l m n o p q r s t u v w x y z
Geheimtextalphabet:	J L P A W I Q B C T R Z Y D S K E G F X H U O N V M

so erhalten wir beispielsweise

Klartext:	gehen wir aus?
Geheimtext:	QWBWD OCG JHF?

Die Anzahl der auf diese Weise erzeugten Geheimcodes ist gleich der Anzahl der Permutationen der 26 Buchstaben, das sind $26! \sim 4 \cdot 10^{26}$. Obwohl diese Zahl zu groß ist, um alle Möglichkeiten auch mit Hilfe eines Rechners durchzuprobieren, sind solche Codes leicht zu entschlüsseln, wenn nur der Text genügend lang ist. Man macht sich dabei die Tatsache zu Nutze, dass in jeder Sprache die Buchstaben unterschiedlich häufig vorkommen. Im Deutschen gilt für die Häufigkeiten:

Buchstabe	Häufigkeit in %	Buchstabe	Häufigkeit in %
a	6,51	n	9,78
b	1,89	o	2,51
c	3,06	p	0,79
d	5,08	q	0,02
e	17,40	r	7,00
f	1,66	s	7,27
g	3,01	t	6,15
h	4,76	u	4,35
i	7,55	v	0,67
j	0,27	w	1,89
k	1,21	x	0,03
l	3,44	y	0,04
m	2,53	z	1,13

Neben dem im Deutschen leicht zu identifizierendem e kann man sich an den Wörtern mit drei Buchstaben orientieren: Sie bezeichnen meist einen der Artikel der, die, das, ein.

Vignère-Verschlüsselungen Bei der Vignère-Verschlüsselung nimmt man für jeden Buchstaben in Abhängigkeit seiner Position im Klartext einen anderen Schlüssel. Im einfachsten Fall vereinbart man ein Schlüsselwort, beispielsweise `LICHT`, das wiederholt über den Klartext geschrieben wird.

Schlüsselwort	`LICHTLICHTLICHTLICHTL`
Klartext	`truppenabzugnachosten`
Geheimtext	`EZWWIPVCISFOEHVSWUAXY`

Der Buchstabe des Schlüsselworts gibt an, wieweit der Buchstabe des Klartextes im Alphabet verschoben werden muss. Im obigen Beispiel ist L der 12. Buchstabe des Alphabets und man verschiebt das t des Klartexts um $12 - 1 = 11$ Positionen nach rechts modulo 26, das ist gerade E. Der nächste Buchstabe r wird wegen des an 9. Position stehenden I um 8 Positionen nach rechts verschoben, das ist Z.

Damit wird jeder Buchstabe auf 5 verschiedene Arten verschlüsselt, eine Häufigkeitsanalyse der Buchstaben ist zur Entschlüsselung nicht mehr möglich. Allerdings kann bei kurzen Schlüsselwörtern eine Häufigkeitsanalyse nach Sequenzen vorgenommen werden wie etwa nach dem häufigsten dreibuchstabigen Wort „die“. Auch nach Verschlüsselung werden die zugehörigen verschlüsselten Sequenzen immer noch häufig sein und führen somit auf das Schlüsselwort.

Man kann die Vignère-Verschlüsselung dahingehend verbessern, dass an Stelle eines Schlüsselwortes ein ganzer Text vereinbart wird, beispielsweise ein Abschnitt eines Romans. In diesem Fall muss der Entschlüssler den Text kennen. Eine moderne Version dieser Technik verwendet einen Zufallsgenerator an Stelle eines Textes. Vor der Verschlüsselung müssen daher nur die Daten des Generators festgelegt werden.

Eine Variante der Vignére-Verschlüsselung wurde von Deutschland mit dem Enigma-Gerät im 2. Weltkrieg verwendet. In der einfachsten Version besteht die Enigma aus einer Tastatur und mindestens drei Rotoren sowie einigen Steckverbindungen, die eine involutorische Permutation der Buchstaben bewirken. Dabei heißt eine Permutation p *involutorisch*, wenn $p^2 = p \circ p = id$. Wird eine Buchstabentaste gedrückt, so fließt ein Strom durch die Steckverbindungen und Rotoren, der den zugehörigen Buchstaben des Geheimtextes erscheinen lässt. Die Rotoren haben jeweils 26 Positionen, die mit den Buchstaben des Alphabets beschriftet sind und sich nach jeder Eingabe eines

Buchstabens ändern. Zudem fließt der Strom nach Durchlaufen der Rotoren auf eine „Umkehrwalze“, die eine involutorische Permutation u darstellt. Anschließend fließt der Strom durch die Rotoren und die Steckverbindungen zurück. Insgesamt entsteht bei einem Zustand z der Rotoren eine Permutation der 26 Buchstaben der Form

$$b_i = R_z(a_i) = r^{-1} \circ p_z^{-1} \circ u \circ p_z \circ r(a_i), \quad i = 1, \dots 26,$$

wobei p_z die Permutation bezeichnet, die von den Rotoren herrührt und vom aktuellen Zustand z der Rotoren abhängt. r ist die Permutation, die aus den Steckverbindungen hervorgeht und sich während der Nachrichtenübermittlung nicht ändert. Wegen $u^2 = id_{A_{26}}$ gilt

$$(r^{-1} \circ p_z^{-1} \circ u \circ p_z \circ r) \circ (r^{-1} \circ p_z^{-1} \circ u \circ p_z \circ r) = r^{-1} \circ p_z^{-1} \circ u \circ u \circ p_z \circ r = id_{A_{26}}.$$

R_z ist dadurch ebenfalls involutorisch, was den Vorteil hat, dass das Dechiffrieren mit dem selben Gerät erfolgen kann, wenn die Anfangsstellung der Rotoren und der Steckverbindungen bekannt ist. Insgesamt kommt man auf eine Verschlüsselung mit einer Bitlänge von etwa 70 – ein auch für die heutige Zeit kaum knackbarer Code. Diese anscheinend hohe Zahl relativiert sich aber, denn die Rotoren liefen nach einem teilweise bekannten Algorithmus und bestimmte Wörter wie oberkommandoderwehrmacht oder wetterbericht kamen in fast jeder Nachricht vor. Jedenfalls konnten die Alliierten die meiste Zeit alle Funksprüche der Deutschen dechiffrieren, was den 2. Weltkrieg sicherlich abkürzte.

Alle diese Verschlüsselungsmethoden eignen sich nicht für eine moderne Kommunikation zwischen wechselnden Partnern über Handy oder Internet, da zuvor der Schlüssel ausgetauscht werden muss. Dies geschieht unverschlüsselt und kann daher abgehört werden.

Die RSA-Verschlüsselung beruht auf zwei Sätzen, die mit den uns zur Verfügung stehenden Methoden leicht bewiesen werden können.

Satz 3.12 (Existenz der modularen Inversen) *Sind a und n teilerfremde natürliche Zahlen, so gibt es eine ganze Zahl b mit der Eigenschaft*

$$ab \equiv 1 \mod n.$$

Beweis: Nach dem Satz über den größten gemeinsamen Teiler 3.4 gibt es Zahlen $\alpha, \beta \in \mathbb{Z}$ mit

$$1 = ggT(a, n) = \alpha a + \beta n,$$

also $\alpha a \equiv 1 \mod n$. Die Zahl $b = \alpha$ erfüllt daher die Behauptung. □

Satz 3.13 *Seien p und q zwei verschiedenen Primzahlen und sei a teilerfremd zu pq. Dann gilt*

$$a^{(p-1)(q-1)} \equiv 1 \mod pq.$$

Beweis: Mit a teilerfremd zu q ist auch a^{p-1} teilerfremd zu q. Mit dem kleinen Satz von Fermat 3.2 folgt

$$a^{(p-1)(q-1)} \equiv 1 \mod q \;\Leftrightarrow\; a^{(p-1)(q-1)} = kq + 1$$

Auf die gleiche Weise folgt $a^{(p-1)(q-1)} = lp + 1$, daher $kq = lp$. Also ist $kq = lp$ sowohl durch q als auch durch p teilbar. Somit $kq = lp = mqp$ und $a^{(p-1)(q-1)} = mpq + 1$ oder $a^{(p-1)(q-1)} \equiv 1 \mod pq$. □

Die RSA-Verschlüsselung ist asymmetrisch. Wer mir eine verschlüsselte Nachricht senden will, verschlüsselt sie mit einem öffentlichen Schlüssel, den ich beispielsweise im Internet zur Verfügung stelle. Das Entschlüsseln geschieht mit einer geheimen Zahl, die nicht versendet werden muss und auch dem Sender der Nachricht unbekannt ist. Genauer geht man folgendermaßen vor:

- Es werden zwei verschiedene Primzahlen p und q gewählt und $n = pq$ berechnet.
- Mit einer weiteren frei gewählten Zahl e, die teilerfremd zu $(p-1)(q-1)$ ist, wird d so berechnet, dass

  $$ed \equiv 1 \mod (p-1)(q-1) \text{ oder } ed = 1 + k(p-1)(q-1).$$

  Dies ist die modulare Inverse aus Satz 3.12. d kann mit Hilfe des erweiterten euklidischen Algorithmus aus Satz 3.4 effektiv berechnet werden.
- Öffentlicher Schlüssel: e und n.
- Privater Schlüssel: d (kann größer als Null gewählt werden).
- p, q und $(p-1)(q-1)$ werden nicht mehr benötigt und sollten sicherheitshalber vernichtet werden.

Nun gibt man die Zahlen n und e öffentlich bekannt. Die „geheime“ Zahl d wird nicht bekannt gegeben. Will jemand eine Nachricht $m < n$ an uns senden, so übermittelt er

$$c \equiv m^e \mod n.$$

Die Zahl c wird entschlüsselt durch

$$m' \equiv c^d \mod n.$$

Satz 3.14 (Korrektheit der RSA-Verschlüsselung) *Mit obigem Verschlüsselungsverfahren gilt $m' = m$.*

Beweis: Aus $a \equiv r \mod n$ folgt $a^d \equiv r^d \mod n$. Für $a = m^e$ ergibt das

$$m^e \equiv c \mod n \Leftrightarrow m^{ed} \equiv c^d \equiv m' \mod n.$$

Wir müssen daher zeigen, dass $m^{ed} \equiv m \mod n$ gilt. Nach Definition von e und d ist $ed = 1 + k(p-1)(q-1)$. Daraus folgt

$$m^{ed} = m^{1+k(p-1)(q-1)} \equiv m \cdot m^{k(p-1)(q-1)} \mod n.$$

Wegen $m < n$ gilt $m \equiv m \mod n$ und wegen Satz 3.13 $m^{k(p-1)(q-1)} \equiv 1 \mod n$. Bilden wir das Produkt dieser Kongruenzen, so

$$m \cdot m^{k(p-1)(q-1)} \equiv m \cdot 1 \equiv m \mod n.$$

Somit ergibt sich nach dem Dechiffrieren mit dem privaten Schlüssel tatsächlich wieder m. □

Im Gegensatz zu den in den vorigen Abschnitten beschriebenen Verfahren werden keine Schlüssel ausgetauscht. Jeder kann mir eine verschlüsselte Nachricht senden, wenn er sich die von mir bekannt gegebenen Zahlen n und e verschafft. Das Verfahren ist daher abhörsicher.

Die RSA-Verschlüsselung beruht auf dem Glauben, dass aus den öffentlichen Zahlen n und e der Schlüssel d nicht in vernünftiger Zeit rekonstruiert werden kann, wenn n genügend groß gewählt wurde. In der Tat kann d nur über die Faktoren in $n = pq$ bestimmt werden. Man ist sich ziemlich sicher, dass diese Faktorisierung nicht „schnell“ gelingt.