Diskrete Mathematik für Informatiker

Universität Leipzig

WS 2007 / 08

Claus Diem

Inhaltsverzeichnis

Kapitel 1

Algebraische Strukturen

1.1 Boolesche Ringe und boolesche Algebren

Boolesche Ringe

Definition Sei X eine Menge und $\circ$ eine Verknüpfung auf X. Sei nun $x \in X$. Dann heißt x *idempotent* (bez. $\circ$), wenn $x \circ x = x$. Die Verknüpfung $\circ$ heißt *idempotent*, wenn $x \circ x = x$ für alle $x \in X$.

Folgerung 1.1 *Sei $(H, \cdot)$ eine abelsche Halbgruppe mit einer idempotenten Verknüpfung. Wir definieren wie folgt eine Relation $\leq$ auf H:*

$$a \leq b :\Longleftrightarrow ab = a$$

Dann ist $\leq$ eine Ordnungsrelation.

Beweis.
(R) Nach Definition ist $a^2 = a$ für alle $a \in R$.
(A) Seien $a, b \in B$ mit $a \leq b$ und $b \leq a$. Dann ist $a = ab = ba = b$.
(T) Seien $a, b, c \in B$ mit $a \leq b, b \leq c$. Dann ist $a = ab = a(bc) = (ab)c = ac$. □

Definition Sei R ein Ring.[1] Dann heißt R *boolesch*, wenn die Multiplikation auf R idempotent ist.

Beispiel 1.2 Die Standardbeispielklasse zu booleschen Ringen ist wie folgt: Sei S eine Menge. Nun ist die Potenzmenge $\mathcal{P}(S)$ mit der symmetrischen Differenz Δ als Addition und dem Durchschnitt ($\cap$) als Multiplikation ein

[1] Ringe haben bei uns immer eine 1 ($(R, \cdot)$ ist ein Monoid). Dies wird in der Literatur nicht immer vorausgesetzt.

boolescher Ring. Die Null ist die leere Menge und die Eins ist S. Wir nennen diesen Ring den *vollen Mengenring* auf S.

Sei im Folgenden R ein boolescher Ring.

Folgerung 1.3 *R ist kommutativ (d.h. die Multiplikation ist kommutativ). Außerdem gilt: $a + a = 0$ für alle $a \in R$ (d.h. $a = -a$).*

Bemerkung Für $a \in R$ und $n \in \mathbb{N}$ setzen wir $na := \overbrace{a + \cdots + a}^{n \text{ mal}}$ (siehe [LA, S. 27])[2]. Die Aussage ist also, dass in R stets $2a = 0$ gilt.

Beweis der Folgerung. Sei zunächst $a \in R$. Dann gilt $2a = (2a)^2 = 4a^2 = 4a$. Damit ist also $2a = 0$.

Seien $a, b \in R$. Dann gilt $a+b = (a+b)^2 = a^2+ab+ba+b^2 = a+ab+ba+b$. Somit gilt $0 = ab + ba$. Da $ba = -ba$, folgt $ab = ba$. □

Betrachten wir die Abbildung $\varphi : \mathbb{F}_2 \longrightarrow R, 0 \mapsto 0_R, 1 \mapsto 1_R$. Diese Abbildung ist ein Ringhomomorphismus. Die einzige nicht-triviale Aussage ist, dass $\varphi(1+1) = \varphi(1) + \varphi(1)$. Dies stimmt aber, da $\varphi(1+1) = \varphi(0) = 0_R$ und $\varphi(1) + \varphi(1) = 1_R + 1_R = 0_R$ ist (s.o.).

Insbesondere definiert ein boolescher Ring "in offensichtlicher Weise" einen $\mathbb{F}_2$-Vektorraum. (Wenn immer ein Körper K, ein Ring R und ein Ringhomomorphismus $\varphi : K \longrightarrow R$ gegeben ist, ist R mittels der Skalarmultiplikation $\cdot : K \times R \longrightarrow R\,, r := \varphi(a) \cdot r$ für $a \in K, r \in R$ ein K-Vektorraum.)

Somit haben wir insbesondere:

Folgerung 1.4 *Wenn R endlich ist, dann ist $\#R$ ist eine Zweierpotenz.*

Wie in Folgerung 1.1 definieren wir für $a, b \in R$:

$$a \leq b :\Longleftrightarrow ab = a$$

Nach Folgerung 1.1 ist $\leq$ eine Ordnungsrelation auf R.

Definition Ein bez. $\leq$ minimales Element in $R - \{0\}$ heißt *Atom*.

Mit anderen Worten: Ein Atom ist ein Element $a \in R$ mit $a \neq 0$ und

$$\forall b \in R - \{0\} : b \leq a \longrightarrow b = a\ .$$

Beachten Sie hier: $b \leq a$ bedeutet $ab = b$.

Beispiel 1.5 Im vollen Mengenring $\mathcal{P}(S)$ gilt $X \subseteq Y \longleftrightarrow X \leq Y$, und die Atome in $\mathcal{P}(S)$ sind die ein-elementigen Mengen.

[2] Mit [LA] ist das Skript zur meiner Vorlesung Lineare Algebra gemeint.

Lemma 1.6

- *Sei a ein Atom. Dann gilt für alle $b \in R$: $a \leq b$ oder $ab = 0$.*
- *Für zwei verschiedene Atome a, b gilt: $ab = 0$.*

Beweis. Sei a ein Atom und $b \in R$. Dann gilt $(ab) \cdot a = a^2 b = ab$, und somit $ab \leq a$. Da a ein Atom ist, impliziert dies: $ab = a$ oder $ab = 0$. Die Aussage $ab = a$ bedeutet gerade, dass $a \leq b$.

Sei nun b ein von a verschiedenes Atom. Wenn nun $a \leq b$ gelten würde, würde $a = b$ gelten nach Definition. Also gilt $ab = 0$. □

Lemma 1.7 *Sei R endlich. Dann gilt:*

- *Zu jedem Element $x \in R$ gilt es Atom $a \in R$ mit $a \leq x$.*
- *Es gilt die Gleichung $1 = \sum_{a\ Atom} a$.*

Beweis. Die erste Aussage ist leicht.

Zur zweiten Aussage: Wir nehmen an, dass $1 - \sum_{a\ \mathrm{Atom}} a = 1 + \sum_{a\ \mathrm{Atom}} a \neq 0$. Dann gibt es ein Atom a_0 mit $a_0 \leq 1 + \sum_{a\ \mathrm{Atom}} a$. Somit gilt $a_0 = a_0(1 + \sum_{a\ \mathrm{Atom}} a) = a_0 + \sum_{a\ \mathrm{Atom}} (a_0 a) = a_0 + a_0$ nach dem obigen Lemma. Dies ist ein Widerspruch. □

Dies impliziert:

Folgerung 1.8 *Sei $b \in R$ und sei $X := \{a \in R |\ a \leq b\,, a\ Atom\}$. Dann gilt:*

$$b = \sum_{a \in X} a\ .$$

Ferner gilt für $Y \subseteq R$: Wenn $b = \sum_{a \in Y} a$, dann ist $X = Y$.

Beweis. Es ist $b = 1 \cdot b = (\sum_{a\ \mathrm{Atom}} a) b = \sum_{a\ \mathrm{Atom}} (ab) = \sum_{a \in X} a$.

Seien nun $Y, Z \subseteq R$ mit $b = \sum_{a \in Y} a = \sum_{a \in Z} a$. Wir nehmen an, dass $Y \neq Z$. Dann gibt OBdA es ein Atom $a_0 \in Z - Y$. Nun haben wir einerseits $ba_0 = \sum_{a \in Z} a a_0 = a_0$ und andererseits $ba_0 = \sum_{a \in Y} a a_0 = 0$, ein Widerspruch. □

Wir haben nun den folgenden Satz:

Satz 1.1 *Sei R ein endlicher boolescher Ring und A die Menge der Atome in R. Dann ist die Abbildung*

$$\varphi : \mathcal{P}(A) \longrightarrow R\ ,\ X \mapsto \sum_{a \in X} a$$

ein Isomorphismus von Ringen, wobei wir $\mathcal{P}(A)$ wie in Beispiel 1.2 als booleschen Ring auffassen. Die Umkehrabbildung ist

$$R \longrightarrow \mathcal{P}(A)\ ,\ b \mapsto \{a \in R \mid a \leq b\,, a\ Atom\ \}\ .$$

Beweis. Wir wissen schon, dass wir eine Bijektion haben. Ferner gilt für $X, Y \subseteq A$: $\varphi(X) + \varphi(Y) = (\sum_{a \in X} a) + (\sum_{a \in Y} a) = \sum_{a \in X \Delta Y} a = \varphi(X \Delta Y)$ und $\varphi(X) \cdot \varphi(Y) = (\sum_{a \in X} a) \cdot (\sum_{a \in Y} a) = \sum_{(a,b) \in X \times Y} ab = \sum_{a \in X \cap Y} a = \varphi(X \cap Y)$, d.h. φ ist ein Ringhomomorphismus. □

Bemerkung Sei $A = \{a_1, \ldots, a_n\}$ mit paarweise verschiedenen a_i. Dann gilt also: Für alle $b \in R$ gibt es eindeutig bestimmte Elemente $c_1, \ldots, c_n \in \mathbb{F}_2$ mit $b = \sum_{i=1}^{n} c_i a_i$. Mit anderen Worten: $a_1, \ldots, a_n$ ist eine $\mathbb{F}_2$-Vektorraumbasis von A. (Man kann auch sagen: A ist eine $\mathbb{F}_2$-Vektorraumbasis von R.)

Boolesche Algebren

Boolesche Ringe stehen in enger Beziehung zu *booleschen Algebren.*

Definition Eine *boolesche Algebra* ist eine Menge B zusammen mit zwei Verknüpfungen $\wedge$ und $\vee$ so dass das gilt:

- $(B, \vee)$ ist ein abelsches Monoid.
- $(B, \wedge)$ ist ein abelsches Monoid.
- Es gelten die *Distributivgesetze*

  $$\forall a, b, c : (a \vee b) \wedge c = (a \vee c) \wedge (b \vee c) \ , \quad (a \wedge b) \vee c = (a \wedge c) \vee (b \wedge c) \ .$$

- Sei 0 das neutrale Element von $(B, \vee)$ und 1 das neutrale Element von $(B, \wedge)$. Dann gilt: Für alle $a \in B$ gibt es ein $b \in B$ mit

  $$a \vee b = 1 \text{ und } a \wedge b = 0 \ .$$

Gegeben ein Element $a \in B$ heißt ein Element b wie im letzten Punkt ein *Komplement* von a.

Bemerkung Eine etwas formalere Definition (mit demselben Inhalt) ist: Eine boolesche Algebra ist ein Tupel $(B, \wedge, \vee)$, wobei B eine Menge, $\wedge$ und $\vee$ Verknüpfungen auf B sind, so dass die obigen Eigenschaften gelten. Die Rollen von $\wedge$ und $\vee$ bzw. 0 und 1 sind nun symmetrisch. D.h. wenn $(B, \wedge, \vee)$ eine boolesche Algebra ist, dann ist auch $(B, \vee, \wedge)$ eine.

Beispiel 1.9 Die Standardbeispielklasse zu booleschen Algebren ist: Sei S eine Menge. Dann ist $\mathcal{P}(S)$ mit der Vereinigung als Operation $\vee$ und dem Durchschnitt als Operation $\wedge$ eine boolesche Algebra. Dabei ist $0 = \emptyset$ und $1 = S$. Wir nennen $\mathcal{P}(S)$ die *volle Mengenalgebra* auf S.

Wie immer definieren wir gleich auch die entsprechenden Morphismen.

Definition Seien B, B' zwei boolesche Algebren. Dann ist ein *Homomorphismus von booleschen Algebren* von B nach B' eine Abbildung $\varphi : B \longrightarrow B'$, die ein Homomorphismus von Monoiden von $(B, \vee)$ nach $(B', \vee')$ sowie von $(B, \wedge)$ nach $(B, \wedge')$ ist. Mit anderen Worten: Es muss gelten: $\varphi(a \wedge_B b) = \varphi(a) \wedge_{B'} \varphi(b'), \varphi(0_B) = 0_{B'}, \varphi(a \vee_B b) = \varphi(a) \vee_{B'} \varphi(b), \varphi(1_B) = 1_{B'}$ für alle $a, b \in B$.

Sei nun B eine boolesche Algebra. Wir leiten einige "Rechenregeln" ab und beweisen unter anderen, dass Komplemente stets eindeutig sind.

Lemma 1.10 *Für $a, b \in B$ gilt:*

- $a \vee a = a\,, \qquad a \wedge a = a$
- $a \vee 1 = 1\,, \qquad a \wedge 0 = 0$
- $a \vee (a \wedge b) = a\,, \qquad a \wedge (a \vee b) = 0$

Beweis. Aufgrund der Symmetrie von $\vee$ und $\wedge$ müssen wir jeweils nur eine der drei Aussagen beweisen.

Sei $c \in B$ ein Komplement von a. Dann ist

$$a = a \vee 0 = a \vee (a \wedge c) = (a \vee a) \wedge (a \vee c) = (a \vee a) \wedge 1 = a \vee a$$
$$a \vee 1 = a \vee (a \vee c)) = (a \vee a) \vee c = a \vee c = 1$$
$$a \vee (a \wedge b) = (a \wedge 1) \vee (a \wedge b) = a \wedge (1 \vee b) = a \vee 1 = a\,.$$

□

Lemma 1.11 *Sei $a \in B$, und seien $b, c \in B$ Komplemente zu a. Dann gilt $b = c$.*

Beweis. Es ist $b = 1 \wedge b = (a \vee c) \wedge b = (a \wedge b) \vee (c \wedge b) = 0 \vee (c \wedge b) = c \wedge b$. Analog gilt $c = b \wedge c = c \wedge b$. Also gilt $b = c$. □

Definition Sei $a \in B$, und sei $b \in B$ das nach Lemma 1.11 eindeutig bestimmte Element mit $a \vee b = 1, a \wedge b = 0$. Dann setzen wir $a^* := b$.

Lemma 1.12 *Seien $a, b \in B$. Dann sind äquivalent:*

- $a \vee b = b$
- $a \wedge b = a$

Beweis. Aufgrund der Symmetrie müssen wir nur zeigen, dass die erste Aussage die zweite impliziert.

Sei also $a \vee b = b$. Dann ist $a \wedge b = a \wedge (a \vee b) = (a \wedge a) \vee (a \wedge b) = a \vee (a \wedge b) = a$. □

Definition Wir definieren $a \leq b :\Longleftrightarrow a \wedge b = a$.

Nach Folgerung 1.1 ist $\leq$ wiederum eine Ordnungsrelation.

Normalerweise bezeichnen wir eine Menge "mit Zusatzstruktur" genau wie die Menge selbst. (Z.B. machen wir keinen Unterschied zwischen dem booleschen Ring R und der zugehörigen Menge). Dies ist jedoch für das Folgende unpraktisch.

Die folgende Folgerung kann man mit relativ einfachen aber teilweise länglichen Rechnungen beweisen. Die Idee ist, dass man den offensichtlichen Zusammenhang zwischen $(\mathcal{P}(S), \Delta, \cap)$ und $(\mathcal{P}(S), \cup, \cap)$ verallgemeinern kann.

Folgerung 1.13 *Sei M eine Menge.*

- *Seien $+, \cdot$ zwei Verknüpfungen auf M so dass $(M, +, \cdot)$ ein boolescher Ring ist. Wir definieren eine Verknüpfung $\vee$ auf M durch*

$$a \vee b := a + b + ab$$

  für $a, b \in M$. Dann ist $(M, \vee, \cdot)$ eine boolesche Algebra. (Es ist naheliegend, die Verknpfung $\cdot$ dann $\wedge$ zu nennen.) Außerdem gilt dann $a^* = 1 - a = 1 + a$ für alle $a \in M$.

- *Seien umgekehrt Verknüpfungen $\wedge$, $\vee$ auf M gegeben so dass $(M, \vee, \wedge)$ eine boolesche Algebra ist. Wir definieren eine Verknüpfung $+$ auf M durch*

$$a + b := (a \wedge b^*) \vee (a^* \wedge b) \ .$$

  Dann ist $(M, +, \wedge)$ ein boolescher Ring. (Es ist naheliegend, die Verknüpfung $\wedge$ dann $\cdot$ zu nennen.)

- *Sei $\mathfrak{R}$ die Menge der booleschen Ringe auf M (d.h. $\mathfrak{R}$ ist die Menge der booleschen Ringe, deren unterliegende Menge M ist), und sei $\mathfrak{B}$ die Menge der booleschen Algebren auf M. Dann definieren die obigen Zuordnungen $(M, +, \cdot) \mapsto (M, \wedge, \cdot)$ und $(M, \vee, \wedge) \mapsto (M, +, \wedge)$ zueinander inverse Bijektionen zwischen $\mathfrak{R}$ und $\mathfrak{B}$.*

(Beweis Übungsaufgabe.)

Gegeben ein boolescher Ring, kann man also von der *zugehörigen booleschen Algebra* sprechen, und gegeben eine boolesche Algebra, kann man vom *zugehörigen booleschen Ring* zu sprechen. Beachten Sie außerdem, dass die Relation $\leq$ für einen booleschen Ring mit der Relation $\leq$ für die zugehörige boolesche Algebra identisch ist.

Folgerung 1.14 *Seien nun M, M' zwei Mengen. Sei R ein boolescher Ring auf M, R' ein boolescher Ring auf M, und seien B bzw. B' die entsprechenden booleschen Algebren. Sei $\varphi : M \longrightarrow M'$ eine Abbildung. Dann ist φ genau dann ein Homomorphismus von booleschen Ringen von R nach R', wenn φ ein Homomorphismus von B nach B' ist.*

(Beweis Übungsaufgabe).

Sei nun B eine boolesche Algebra. Die Relation $\leq$ ist genau so definiert wie für boolesche Ringe. (Mit $\wedge$ anstelle von $\cdot$.) (Mit anderen Worten: Die Relation $\leq$ auf B ist die Relation $\leq$ des zugehörigen booleschen Rings.) Atome sind dann genau so definiert wie für boolesche Ringe. (D.h. die Menge der Atome in der booleschen Algebra und im zugehörigen booleschen Ring sind identisch.)

Sei $+$ die Addition des zugehörigen booleschen Rings. Dann gilt für Atome $a, b : a \vee b = a + b + ab = a + b$. Allgemeiner gilt für eine endliche Menge X von Atomen: $\sum_{a \in X} a = \bigvee_{a \in X} a$.

Wir erhalten das folgende Korollar zu Satz 1.1:

Korollar 1.15 *Sei B eine endliche boolesche Algebra, und sei A die Menge der Atome von B. Dann ist*

$$\varphi : \mathcal{P}(A) \longrightarrow B \, , X \mapsto \bigvee_{a \in X} a$$

ein Isomorphismus von booleschen Algebren, wobei wir die Potenzmenge $\mathcal{P}(A)$ wie in Beispiel 1.9 als boolesche Algebra auffassen.

Beweis. Die Abbildung φ ist genau die in Satz 1.1 definierte Abbildung. Nach Satz 1.1 wissen also schon, dass die Abbildung ein Isomorphismus der zugehörigen booleschen Algebren ist. Somit ist die Abbildung insbesondere bijektiv. Es ist ein Homomorphismus nach Folgerung 1.14. Alternativ kann man auch leicht direkt nachrechnen, dass es sich um einen Homomorphismus handelt. □

1.2 Allgemeine Algebra

Motivation

Wir setzen uns das Ziel, möglichst formal zu definieren, was wir unter einer Halbgruppe, einem Monoid und einer Gruppe meinen. Die Idee ist, dass die Strukturen jeweils aus einer Menge und einer Verknüpfung bestehen so dass gewissen Eigenschaften gelten. Wir wollen diese Eigenschaften mit Aussagen der Prädikatenlogik beschreiben.

Definition (alt)

- Eine Halbgruppe ist ein Tupel $(H, \cdot)$, wobei H eine Menge und $\cdot$ eine Verknüpfung auf H ist so dass gilt:

$$\forall a, b, c \in H : (a \cdot b) \cdot c = a \cdot (b \cdot c)$$

- Ein Monoid ist ein Tupel $(M, \cdot)$, wobei M eine Menge und $\cdot$ eine Verknüpfung auf M ist so dass gilt:

$$\forall a, b, c \in M : (a \cdot b) \cdot c = a \cdot (b \cdot c)$$
$$\exists e \in M : \forall a \in M : e \cdot a = e \wedge a \cdot e = a$$

- Eine Gruppe ist ein Tupel $(G, \cdot)$, wobei G eine Menge und $\cdot$ eine Verknüpfung auf G ist so dass gilt:

$$\forall a, b, c \in G : (a \cdot b) \cdot c = a \cdot (b \cdot c)$$
$$\exists e \in G : ((\forall a \in G : e \cdot a = e \wedge a \cdot e = a) \wedge (\forall a \in G \, \exists b \in G : a \cdot b = e \wedge b \cdot a = e))$$

Ist es auch möglich die Objekte durch, "durch Identitäten" zu definieren, wobei man auf den Existenz-Quantor verzichtet?

Für Halbgruppen ist dies kein Problem. Bez. Monoide erinnern wir uns, dass das neutrale Element stets eindeutig bestimmt ist. Anstatt des Tupels $(M, \cdot)$ kann man auch das Tupel $(M, \cdot, e)$ betrachten, wobei e das neutrale Element ist. (Formal ist dieses 3-er-Tupel ein anderes Objekt als das ursprüngliche Zweiertupel, aber dieser Unterschied ist "vernachlässigbar".) Wir erhalten also neue Definition eines Monoids:

Definition (neu) Ein Monoid ist ein Tupel $(M, \cdot, e)$, wobei M eine Menge, $\cdot$ eine Verknüpfung auf M und $e \in M$ ist so dass gilt:

$$\forall a, b, c \in M : (a \cdot b) \cdot c = a \cdot (b \cdot c)$$
$$\forall a \in M : e \cdot a = e \wedge a \cdot e = a$$

Für Gruppen geht es ähnlich, nur müssen wir nun auch die Inversion betrachten. Wir erhalten:

Definition (neu) Eine Gruppe ist ein Tupel $(G, \cdot, \iota, e)$, wobei G eine Menge, $\cdot$ eine Verknüpfung auf G, $\iota : G \longrightarrow G$ eine Abbildung und $e \in G$ ist so dass gilt:

$$\forall a, b, c \in G : (a \cdot b) \cdot c = a \cdot (b \cdot c)$$
$$\forall a \in G : a \cdot \iota(a) = e \wedge \iota(a) \cdot a = e$$
$$\forall a \in G : e \cdot a = e \wedge a \cdot e = a \, .$$

Natürlich schreibt man wieder $a^{-1} = \iota(a)$ für alle $a \in G$.

Die *Allgemeine Algebra*[3] ist eine Theorie, die eine solche Art von Beschreibungen algebraischer Objekte in einen allgemeinen Rahmen gibt. Aussagen der Allgemeinen Algebra ergeben dann als Spezialfälle Aussagen über die verschiedensten algebraischen Objekte wie z.B. Halbgruppen, Monoide, Gruppen, Ringe, Vektorräume.

Erste Definitionen

Bisher haben wir die folgende Definition für eine "Verknüpfung" benutzt:

Definition Sei X eine Menge. Dann ist eine Verknüpfung auf X eine Abbildung $X \times X \longrightarrow X$.

In der Theorie der "Allgemeinen Algebra" verallgemeinert man diese Definition.

Sei im Folgenden X eine beliebige Menge. Erinnern Sie sich, dass wir für $n \in \mathbb{N}$ $X^n := X^{\{1,\dots,n\}}$ definiert haben (die Menge der Abbildungen von $\{1, \dots, n\}$ nach X).[4]

Man kann somit $X^0 := X^{\emptyset}$ definieren. Diese Menge hat genau ein Element, welches wir mit $*$ bezeichnen.

Definition Sei $n \in \mathbb{N}_0$. Dann ist eine *n-stellige Verknüpfung* (oder *Operation*) auf X eine Abbildung $X^n \longrightarrow X$. Eine *(endlichstellige) Verknüpfung* ist eine n-stellige Verknüpfung für ein n. Wenn σ eine n-stellige Verknüpfung ist, setzen wir $\mathrm{ar}(\sigma) := n$.[5]

Eine Verknüpfung im Sinne der bisherigen Definition ist nach dieser Definition also eine 2-stellige Verknüpfung, die man auch *binäre Verknüpfung* nennt. Was den Sprachgebrauch betrifft, muss man etwas aufpassen: Normalerweise meint man mit einer Verknüpfung immer eine 2-stellige Verknüpfung. In der Theorie der Allgemeinen Algebra versteht man unter einer Verknüpfung allerdings eine beliebige endlichstellige Verknüpfung, und so verfahren wir auch in diesem Abschnitt.

Auch in der Allgemeinen Algebra bezeichnet man eine binäre Verknüpfung oft mittels eines "Verknüpfungssymbols" (z.B. "$\circ$", "$\cdot$", "$+$"), dass man zwischen die zu verknüpfenden Elemente schreibt. Man schreibt also nicht

[3] "Allgemeine Algebra" wird traditioneller Weise "Universelle Algebra" genannt. Im Englischen wird traditionell von "Universal Algebra" gesprochen, aber es kommt immer mehr auch die Bezeichnung "General Algebra" auf.

[4] Wie in der Vorlesung Lineare Algebra, bezeichne ich 0 nicht als natürliche Zahl.

[5] "ar" bezieht sich auf "arity" (Englisch für "Stelligkeit").

$+(x, y)$ sondern (wie gewohnt) $x + y$ für die Addition zweier Zahlen (z.B. in $\mathbb{Z}$).

Eine 0-stellige Verknüpfung auf X ist per Definition eine Abbildung $\{*\} \longrightarrow X$. Diese Abbildung ist durch die Angabe des Bildes von X eindeutig bestimmt, und umgekehrt kann man jedem $x \in X$ die Abbildung $\{*\} \longrightarrow X$, $* \mapsto x$ zuordnen. Mit anderen Worten, wir haben einen eine Bijektion

$$X \xrightarrow{\sim} \mathrm{Abb}(\{*\}, X),\ x \mapsto (* \mapsto x)$$

(siehe auch [LA, Beispiel 1.4].)

Man kann eine nullstellige Verknüpfung mit dem Bild von $*$ in X "identifizieren", und so verfahren wir auch im Folgenden. Eine nullstellige Verknüpfung (bzw. die Elemente von X) heißen auch *Konstanten*.

Beachten Sie ferner, dass eine 1-stellige Verknüpfung nichts anderes als eine Abbildung $X \longrightarrow X$ ist. Man spricht auch von einer *unären Verknüpfung*.

Definition Ein *Typ von Algebren* ist eine Menge I zusammen mit Familie $(n_i)_{i\in I} \in \mathbb{N}_0^I$. (Zur Erinnerung: Eine "Familie" ist nichts anderes als eine Abbildung.)

Wir nennen die Menge I *(Verknüpfungs-)indexmenge* und die Familie $(n_i)_{i\in I} \in \mathbb{N}_0^I$ *Familie von Stelligkeiten*. Oftmals ist die Menge I gleich $\{1, \ldots, \ell\}$ für ein $\ell \in \mathbb{N}$. In diesem Fall kann man I in der Beschreibung weglassen und einfach das Tupel $(n_1, \ldots, n_\ell)$ angeben. Normalerweise wählt man die Stelligkeiten so, dass dann $n_1 \geq n_2 \geq \cdots \geq n_\ell$.

Es sind jedoch auch unendliche Indexmengen möglich.

Sei nun $\Omega = (I, (n_i)_{i\in I})$ ein Typ.

Definition Eine Ω-*Algebra* ist eine Menge A zusammen mit einer Familie $(\sigma_i)_{i\in I}$ von Verknüpfungen auf A so dass $\mathrm{ar}(\sigma_i) = n_i$ für alle $i \in I$.

Eine Ω-Algebra wird auch *Algebra* vom *Typ* Ω genannt. Wiederum muss man aufpassen, was den Sprachgebrauch betrifft: Außerhalb der Allgemeinen Algebra hat das Wort "Algebra" eine andere Bedeutung.

Wir bezeichnen eine Ω-Algebra $(A, (\sigma_i)_{i\in I})$ wieder mit A.Wie immer definieren wir gleich, was die entsprechenden Homomorphismen sind:

Definition Seien A, A' zwei Ω-Algebren. Dann ist ein *Homomorphismus von Ω-Algebren* von A nach A' eine Abbildung $\varphi : A \longrightarrow A'$ so dass für alle $i \in I$ gilt:

$$\forall a_1, \ldots, a_{n_i} \in A : \varphi(\sigma_i^A(a_1, \ldots, a_{n_i})) = \sigma_i^{A'}(\varphi(a_1), \ldots, \varphi(a_{n_i}))$$

Bemerkung Die Bedingung an φ kann man auch so formulieren: Für alle $i \in I$ gilt $\varphi \circ \sigma_i^A = \sigma_i^{A'} \circ (\overbrace{\varphi \times \cdots \times \varphi}^{n_i \text{ mal}}) : A^{\mathrm{ar}(i)} \longrightarrow A$.

Definition Seien A, A' wiederum Ω-Algebren, $\varphi : A \longrightarrow A'$ ein Homomorphismus. Dann ist φ ein *Isomorphismus*, wenn es einen Homomorphismus $\psi : A' \longrightarrow A$ mit $\psi \circ \varphi = \mathrm{id}_A$ und $\varphi \circ \psi = \mathrm{id}_{A'}$ gibt.

Bemerkung Ein Homomorphismus von Ω-Algebren ist genau dann ein Isomorphismus, wenn er bijektiv ist (Übungsaufgabe).

Beispiel 1.16 Nach den ursprünglichen Definitionen sind Halbgruppen, Monoide und Gruppe Algebren vom Typ (2). Nach der erneuerten Definition "im Sinne der allgemeinen Algebra" sind es jeweils Algebren vom Typ $(2), (2,0)$ bzw. $(2,1,0)$.

Bemerkung / Definition Sei A eine Ω-Algebra, und sei $B \subseteq A$ *abgeschlossen* unter allen Verknüpfungen. D.h.: Für alle $i \in I$ und alle $a_1, \ldots, a_{n_i} \in B$ gilt $\sigma_i(a_1, \ldots, a_{n_i}) \in B$. Dann ist B mit den *induzierten Verknüpfungen* $\sigma_i|_{B^{n_i}}$ eine Ω-Algebra, genannt eine Ω-*Unteralgebra* von A.

Bemerkung Sei σ_i eine 0-stellige Verknüpfung. Dann ist eine Teilmenge $B \subseteq A$ genau dann abgeschlossen unter σ_i, wenn $\sigma_i(*) \in B$.

Beispiel 1.17 Sei M ein Monoid. Wenn wir nun wie oben M als eine Algebra vom Typ $(2,1)$ betrachten, ergibt sich die Definition eines Untermonoids als ein Spezialfall der obigen Definition (siehe auch [LA]).

Dies gilt auch für Halbgruppen und Gruppen, aber bei Gruppen ist es egal, von welcher Definition man startet.

Bemerkung Sei A eine Ω-Algebra, und sei $\mathcal{B}$ eine Menge von Ω-Unteralgebren von A. Dann ist $\bigcap_{B \in \mathcal{B}} B$ eine Ω-Unteralgebra von A.

Bemerkung / Definition Sei A eine Ω-Algebra, und sei $X \subseteq A$ eine Teilmenge. Sei

$$\mathcal{B} := \{B \subseteq A \mid B \text{ ist } \Omega\text{-Unteralgebra von A mit } X \subseteq B\} .$$

Dann ist $\bigcap_{B \in \mathcal{B}} B$ die kleinste Ω-Unteralgebra von A, die X enthält; sie wird mit $\langle X \rangle$ bezeichnet und heißt die *von X erzeugte* Unteralgebra von A. Falls $\langle X \rangle = A$, sagen wir, dass A von X *erzeugt* wird.

(*Begründung*: Nach der obigen Bemerkung ist es eine Ω-Unteralgebra, und offensichtlich enthält sie auch alle anderen.)

Wir geben nun eine explizite Beschreibung der von einer Menge erzeugten Unteralgebra.

Folgerung 1.18 *Sei A eine Ω-Algebra und sei $X \subseteq A$. Sei für $j \in \mathbb{N}_0$:*

$$S^{(0)} := X$$

$$S^{(j+1)} := S^{(j)} \cup \{\sigma_i(a_1, \ldots, a_{n_i}) \mid i \in I, a_1, \ldots, a_{n_i} \in S^{(j)}\} .$$

Dann ist $\langle X \rangle = \bigcup_{j=0}^{\infty} S^{(j)}$.

Beweis. Wir zeigen, dass $\bigcup_{j=0}^{\infty} S^{(j)}$ die kleinste Ω-Unteralgebra von A ist, die X enthält. Sicher enthält sie X. Außerdem ist sie abgeschlossen unter allen Verknüpfungen, also ist eine Ω-Unteralgebra.

Sei nun B eine beliebige Ω-Unteralgebra von A, die X enthält. Mit Induktion nach j sieht man sofort, dass $S^{(j)} \subseteq B$ für alle $j \in \mathbb{N}_0$. □

Folgerung 1.19 *Sei A eine Ω-Algebra, die von $X \subseteq A$ erzeugt wird, und sei B eine zweite Ω-Algebra. Sei $f : X \longrightarrow B$ eine Abbildung. Dann gibt es höchstens einen Homomorphismus $\varphi : A \longrightarrow B$ mit $\varphi(x) = f(x)$ für alle $x \in X$.*

Beweis. Seien φ, φ' zwei Homomorphismen wie in der Behauptung. Dann gilt per Induktion nach j: $\varphi(a) = \varphi'(a)$ für alle $x \in S^{(j)}$.

Dies ist richtig für $j = 0$ nach Voraussetzung. Der Induktionsschluss $j - 1 \curvearrowright j$ ist wie folgt:

Für $a \in S^{(j-1)}$ ist die Behauptung klar nach Induktionsvoraussetzung. Sei also $a \in S^{(j)} - S^{(j-1)}$. Dann gilt $a = \sigma_i(a_1, \ldots, a_{n_i})$ für ein $i \in I$ und $a_1, \ldots, a_{n_i} \in S^{(j-1)}$. Damit ist $\varphi(a) = \varphi(\sigma_i(a_1, \ldots, a_{n_i})) = \sigma_i(\varphi(a_1), \ldots, \varphi(a_{n_i})) = \sigma_i(\varphi'(a_1), \ldots, \varphi'(a_{n_i})) = \varphi'(\sigma_i(a_1, \ldots, a_{n_i}))$. □

Termalgebren

Definition Sei X eine Menge. Eine Ω-*Termalgebra* auf X ist eine Ω-Algebra $(T, (\sigma_i)_{i \in I})$ zusammen mit einer Inklusion $\iota : X \hookrightarrow T$ so dass gilt:

- T ist von $\iota(X)$ erzeugt.
- Für alle $t \in T$ gilt: *Entweder* es ist $t = \iota(x)$ für ein $x \in X$ *oder* es gibt ein $i \in I$ und $t_1, \ldots, t_{n_i} \in t$ mit $t = \sigma_i(t_1, \ldots, t_{n_i})$. Im zweiten Fall sind i und $t_1, \ldots, t_{n_i}$ *eindeutig* (durch t) *bestimmt.*

Wir zeigen nun, dass Termalgebren immer existieren und in einer gewissen Hinsicht "im Wesentlichen" eindeutig bestimmt sind.

Definition Sei weiterhin X eine Menge. Eine *freie* Ω-Algebra auf X ist eine Ω-Algebra $(F, (\sigma_i)_{i\in I})$ zusammen mit einer Abbildung $u : X \longrightarrow F$ so dass gilt:

Für alle Ω-Algebren A und alle Abbildungen $f : X \longrightarrow A$ gibt es genau einen Homomorphismus von Ω-Algebren $\varphi : F \longrightarrow A$ mit $\varphi \circ u = f$.

Bemerkung Eine Ω-Algebra F zusammen mit einer Abbildung $u : X \longrightarrow F$ ist also genau dann eine freie Ω-Algebra, wenn gilt: Für alle Ω-Algebren A ist der Homomorphismus

$$\mathrm{Hom}(F, A) \longrightarrow A^X = \mathrm{Abb}(X, A)\ ,\ \varphi \mapsto \varphi \circ u$$

bijektiv.

Bemerkung Die Eigenschaft in der Definition heißt auch *universelle Eigenschaft*. Hiernach ist es naheliegend, freie Ω-Algebren "universelle Ω-Algebren" zu nennen. Leider wird aber in vielen Büchern über Allgemeine Algebra ("Universelle Algebra") *jede* Ω-Algebra auch "universelle Algebra" genannt.

Lemma 1.20 *Seien (F, u), (F, u') zwei freie Ω-Algebren auf X. Dann ist der eindeutig bestimmte Homomorphismus $\varphi : F \longrightarrow F'$ mit $\varphi \circ u = u'$ ein Isomorphismus.*

Beweis. Sei φ wie im Lemma, und sei $\psi : F' \longrightarrow F$ der eindeutig bestimmte Homomorphismus mit $\psi \circ u' = u$. Dann gilt $\psi \circ \varphi \circ u = u$ und natürlich auch $\mathrm{id}_F \circ u = u$. Somit ist (nach der definierenden Eigenschaft von F und u) $\psi \circ \varphi = \mathrm{id}_F$. Analog gilt $\varphi \circ \psi \circ u' = u'$ und somit $\varphi \circ \psi = \mathrm{id}_{F'}$. □

Satz 1.2 *Sei X eine Menge. Dann gilt:*

a) Es gibt eine Ω-Termalgebra auf X.

b) Jede Ω-Termalgebra auf X ist eine freie Algebra auf X.

c) Seien $(T, \iota), (T', \iota')$ zwei Ω-Termalgebren auf X. Dann gibt es einen eindeutig bestimmten Isomorphismus von Ω-Algebren $\varphi : T \longrightarrow T$ mit $\varphi \circ \iota = \iota'$.

Beweis. a) Wir setzen für $j \in \mathbb{N}_0$:

$$T^{(0)} := \{(*, x) \mid x \in X\}$$
$$T^{(j+1)} := T^{(j)} \cup \{(i, (t_1, \ldots, t_{n_i})) \mid i \in I, t_1, \ldots, t_{n_i} \in S^{(j)}\}$$

Wir setzen nun $T := \bigcup_{j=0}^{\infty} T^{(j)}$ und $\iota : X \longrightarrow T$ $x \mapsto (*, x)$, $\sigma_i : T^{n_i} \longrightarrow T$, $(t_1, \ldots, t_{n_i}) \mapsto (i, (t_1, \ldots, t_{n_i}))$.

Nun ist $(T, (\sigma_i)_{i \in I}$ eine Ω-Algebra. Es ist offensichtlich, dass die von $\iota(X)$ erzeugt wird. (Wir haben $S^{(j)} = T^{(j)}$ mit den Bezeichnungen von Folgerung 1.18.) Die zweite Bedingung ist offensichtlich auch erfüllt.

Sei nun T eine Termalgebra. (Beachten Sie: T muss nicht wie im Beweis von a) "konstruiert" sein.)

Sei A eine Ω-Algebra, und sei $f : X \longrightarrow A$ eine Abbildung. Nach Folgerung 1.19 wissen wir schon, dass es höchstens einen Homomorphismus $\varphi : T \longrightarrow A$ mit $\varphi \circ \iota = f$ gibt. Wir wollen einen Homomorphismus $\varphi : T \longrightarrow A$ mit $\varphi \circ \iota = f$ finden. Seien $S^{(j)} \subseteq T$ wie in Folgerung 1.18 für die Ω-Algebra T und die Menge X. Beachten Sie: Für $j > 0$ gilt nach Definition von $S^{(j)}$ und Voraussetzung an T: Für alle $t \in S^{(j)} - S^{(j-1)}$ gibt ein eindeutig bestimmtes $i \in I$ sowie eindeutig bestimmte $t_1, \ldots, t_{n_i} \in S^{(j-1)}$ mit $t = \sigma_i(t_1, \ldots, t_{n_i})$.

Wir definieren nun $\varphi_j : S^{(j)} \longrightarrow A$ $(j \in \mathbb{N}_0)$ induktiv wie folgt: $\varphi_0(\iota(x)) := f(x)$ für $x \in X$, sowie für $j > 0$: $\varphi_j(t) := \varphi_{j-1}(t)$ falls $t \in S^{(j-1)}$ sowie $\varphi_j(t) := \sigma_j(\varphi_{j-1}(t_1), \ldots, \varphi_{j-1}(t_{n_i}))$, falls $t \notin S^{(j-1)}$ und $t = \sigma_j(t_1, \ldots, t_{n_i})$.

Für $j < k$ gilt nun $\varphi_j|_{S^{(j)}} = \varphi_k$. Wir definieren nun $\varphi : T \longrightarrow A$ durch $\varphi(t) := \varphi_j(t)$ falls $t \in S^{(j)}$. (Dies hängt wie gesagt nicht von j ab.)

Der Definition ist φ nun ein Homomorphismus mit $\varphi \circ \iota = f$. □

Nach dem Satz sind Ω-Termalgebren "so eindeutig wie möglich". Man definiert entsprechend:

Definition Sei X eine Menge. Dann bezeichnen wir eine Ω-Termalgebra auf X mit $T_\Omega(X)$.

Bemerkung Sei X endlich, sagen wir $X = \{x_1, \ldots, x_n\}$ mit x_i paarweise verschieden. Dann werden die Elemente in $T_\Omega(X)$ mit $t = t(x_1, \ldots, x_n)$ bezeichnet. Allgemeiner: Sei I eine "Indexmenge" und $I \longrightarrow X$, $i \mapsto x_i$ eine Bijektion. Dann ist es sinnvoll, $t = t((x_i)_{i \in I})$ zu schreiben. (Vergleiche die Ähnlichkeit zur Schreibweise von Polynomen.)

Durch Gleichungen definierte Objekte

Sei nach wie vor Ω ein Typ, und sei X eine Menge.

Definition Sei A eine Ω-Algebra. Sei $t \in T_\Omega(X)$, und sei $(a_x)_{x \in X} \in A^X$. Sei $\varphi : T_\Omega(X) \longrightarrow A$ die eindeutig bestimmte Abbildung mit $\varphi(x) = a_x$ für alle $x \in X$. Dann setzen wir $t((a_x)_{x \in X}) := \varphi(t)$.

Wenn $X = \{x_1, \ldots, x_n\}$ mit paarweise verschiedenen x_i, kann man auch definieren: Seien $a_1, \ldots, a_n \in A$, und sei $\varphi : T_\Omega(X) \longrightarrow A$ die eindeutig bestimmte Abbildung mit $\varphi(x_i) = a_i$ für alle $i = 1, \ldots, n$. Dann setzen wir $t(a_1, \ldots, a_n) := \varphi(t)$.

Definition Unter einer *Gleichung* bez. Ω über X verstehen wir ein Tupel $(s, t) \in T_\Omega(X)^2$. Dies schreiben wir auch in der Form $s \stackrel{\circ}{=} t$. (Wir führen das neue Symbol "$\stackrel{\circ}{=}$" ein, um zu verhindern, dass man die "Gleichung" "$s \stackrel{\circ}{=} t$" (d.h. das Tupel (s, t)) mit der Aussage verwechselt, dass der Term s gleich dem Term t ist). (Eine andere sinnvolle Schreibweise wäre $s \stackrel{!}{=} t$; man findet auch $s \approx t$ oder auch einfach $s = t$ in der Literatur.)

Definition Seien $s, t \in T_\Omega(X)$, und sei $(a_x)_{x \in X} \in A^X$. Dann sagen wir "$(a_x)_{x \in X} \in A^X$ erfüllt die Gleichung $s \stackrel{\circ}{=} t$", wenn $s((a_x)_{x \in X}) = t((a_x)_{x \in X})$. Die obigen Bemerkungen falls $X = \{x_1, \ldots, x_n\}$ gelten auch hier.

Beispiel 1.21 Sei $X = \{x_1, x_2, x_3\}$ eine Menge mit drei Elementen, und sei Ω gegeben durch das Tupel (2). (Jede Ω-Algebra hat genau eine Verknüpfung, und diese ist binär.) Wir schreiben die Verknüpfung σ_1 auf $T_\Omega(X)$ wie gewohnt mittels des Symbols "$\circ$". Sei nun A eine Ω-Algebra. Dann ist A genau dann eine Halbgruppe, wenn für alle $a_1, a_2, a_3 \in A$ gilt: (a_1, a_2, a_3) erfüllt die Gleichung $(x_1 \circ x_2) \circ x_3 \stackrel{\circ}{=} x_1 \circ (x_2 \circ x_3)$.

Bemerkung Man sollte allgemein die "Verknüpfungssymbole" für A und in $T_\Omega(X)$ "kompatibel" wählen. Ansonsten können die Aussagen sehr verwirrend sein.

Definition Sei nun A eine Ω-Algebra. Dann sagen wir, dass A "die Gleichung $t \stackrel{\circ}{=} s$ erfüllt", wenn gilt: Alle $(a_x)_{x \in X} \in A^X$ erfüllen die Gleichung $s \stackrel{\circ}{=} t$. Wir schreiben dann

$$A \vDash s \stackrel{\circ}{=} t \; .$$

Wenn Σ eine Menge von Gleichungen ist (beachten Sie die formale Definition von "Gleichung"), dann sagen wir, "A erfüllt Σ", wenn A alle Gleichungen in Σ erfüllt. Wir schreiben dann

$$A \vDash \Sigma \; .$$

Da Ω-Algebren frei sind, ist jeder Homomorphismus $\varphi : T_\Omega(X) \longrightarrow A$ von der Form $t \mapsto t((a_x)_{x \in X})$ für ein eindeutig bestimmtes $(a_x)_{x \in X}$. Damit gilt:

Lemma 1.22 *Es gilt genau dann $A \vDash s \mathrel{\hat{=}} t$, wenn für alle Morphismen $\varphi : T_\Omega(X) \longrightarrow A$ gilt: $\varphi(s) = \varphi(t)$.*

Beispiel 1.23 Sei Ω gegeben durch das Tupel $(2, 1, 0)$. Wir schreiben $\circ$ für die zweistellige Verknüpfung, ι für die einstellige Verknüpfung und e für die "Konstante" zur 0-stelligen Verknüpfung in $T_\Omega(X)$.

Sei nun A eine Ω-Algebra. Dann ist A genau dann eine Gruppe, wenn A die Gleichungen

$$(x \circ y) \circ z \mathrel{\hat{=}} x \circ (y \circ z)$$
$$e \circ x \mathrel{\hat{=}} x$$
$$\iota(x) \circ x \mathrel{\hat{=}} e$$

erfüllt. (Hier steckt eine nicht-triviale Aussage drin, nämlich, dass es auch-reicht, dass ein links-neutrales Element und links-Inverse zu fordern. Können Sie das beweisen?)

Mit anderen Worten: Sei

$$\Sigma := \{(x \circ y) \circ z \mathrel{\hat{=}} x \circ (y \circ z),\ e \circ x \mathrel{\hat{=}} x,\ \iota(x) \circ x \mathrel{\hat{=}} e\} \ .$$

Dann ist A genau dann eine Gruppe, wenn $A \vDash \Sigma$.

Definition Sei nun $\mathcal{K}$ eine Klasse[6] von Ω-Algebren, und sei Σ eine Menge von Gleichungen in $T_\Omega(X)$. Dann sagen wir, "$\mathcal{K}$ erfüllt Σ", wenn für alle A aus $\mathcal{K}$ gilt: $A \vDash \Sigma$. Wir schreiben dann

$$\mathcal{K} \vDash \Sigma \ .$$

Wenn Σ nur aus einer Gleichung $s \mathrel{\hat{=}} t$ besteht, schreiben wir dann auch

$$\mathcal{K} \vDash s \mathrel{\hat{=}} t \ .$$

Definition Sei $\mathcal{K}$ eine Klasse. Die *Menge der Identitäten* von $\mathcal{K}$ über X ist dann die Menge der Gleichungen, welche von $\mathcal{K}$ erfüllt werden:

$$\Sigma_{\mathcal{K}}(X) := \{s \mathrel{\hat{=}} t \in T_\Omega(X) \mid \mathcal{K} \vDash s \mathrel{\hat{=}} t\}$$

Bemerkung Per Definition ist $\Sigma_{\mathcal{K}}(X) \subseteq T_\Omega(X)^2$. Offensichtlich definiert $\Sigma_{\mathcal{K}}(X)$ eine Äquivalenzrelation. Lassen Sie uns hierfür $\mathrel{\hat{=}}_{\mathcal{K}}$ zu schreiben. Dann gilt also: $s \mathrel{\hat{=}}_{\mathcal{K}} t \longleftrightarrow \forall A \in \mathcal{K} : A \vDash s \mathrel{\hat{=}} t$.

[6]Der Begriff der "Klasse" ist allgemeiner als der der Menge. Z.B. sollte man nicht von der "Menge aller Mengen" reden (siehe [LA]), aber man kann von der "Klasse aller Mengen" zu reden.

Definition Sei Σ eine Menge von Gleichungen. Dann ist die *durch* Σ *definierte Klasse* die Klasse der Ω-Algebren A mit $A \vDash \Sigma$. Die Klasse $\mathcal{K}$ heißt dann auch *durch* Σ *definiert* oder *axiomatisiert*. Man nennt $\mathcal{K}$ auch die *Varietät* zu Σ, Bezeichnung: $V(\Sigma)$. Wenn $\mathcal{K}$ durch irgendein Σ definiert ist, nennt man $\mathcal{K}$ auch einfach eine Varietät oder "durch Gleichungen definiert".

Satz 1.3 *(Birkhoff) Sei $\mathcal{K}$ eine Klasse von Ω-Algebren. Dann ist $\mathcal{K}$ genau dann eine Varietät, wenn gilt:*

- *Seien A und B Ω-Algebren mit $B \in \mathcal{K}$ und sei $\iota : A \longrightarrow B$ ein injektiver Homomorphismus. Dann ist auch $A \in \mathcal{K}$.*
- *Seien A und B Ω-Algebren mit $A \in \mathcal{K}$ und sei $p : A \longrightarrow B$ ein surjektiver Homomorphismus. Dann ist auch $B \in \mathcal{K}$.*
- *Sei I eine Menge und sei $(A_i)_{i \in I}$ eine Familie von Ω-Algebren in $\mathcal{K}$. Dann ist auch $\prod_{i \in I} A_i$ in $\mathcal{K}$.*

Zum Beweis. Es ist leicht zu zeigen, dass jede Varietät diese Eigenschaften hat. Die Umkehrung zeigen wir nicht.

Definition Sei weiterhin $\mathcal{K}$ eine Klasse von Ω-Algebren. Eine *freie Algebra* in $\mathcal{K}$ über X ist eine Ω-Algebra F in $\mathcal{K}$ zusammen mit einer Abbildung $u : X \longrightarrow F$ so dass gilt:

Für alle Ω-Algebren A und alle Abbildungen $f : X \longrightarrow A$ gibt es genau einen Homomorphismus von Ω-Algebren $\varphi : F \longrightarrow A$ mit $\varphi \circ u = f$.

Satz 1.4 *In jeder Varietät gibt es freie Algebren über X. Zwei solche Algebren sind "im wesentlichen eindeutig" wie in Lemma 1.20 beschrieben. Wenn (F, u) eine solche Algebra ist, dann ist der eindeutig bestimmte Homomorphismus $\varphi : T_\Omega(X) \longrightarrow F$ mit $\varphi \circ \iota = u$ surjektiv.*

Diese Aussage zeigen wir nicht.

Definition Wir bezeichnen eine freie Algebra von $\mathcal{K}$ über X mit $F_\mathcal{K}(X)$ und den eindeutig bestimmten Homomorphismus $\varphi : T_\Omega(X) \longrightarrow F_\mathcal{K}(X)$ mit $\varphi \circ \iota = u$ mit π. Wenn t ein Term über X ist, setzen wir $\overline{t} := \pi(t)$.

Definition Sei Σ eine Menge von Gleichungen (wie immer über X), und seien s, t Terme. Dann sagen wir, dass $s \stackrel{\circ}{=} t$ aus Σ *folgt*, wenn $V(\Sigma) \vDash s \stackrel{\circ}{=} t$, d.h. wenn für alle $A \in V(\Sigma)$ gilt: $A \vDash s \stackrel{\circ}{=} t$. Wir schreiben dann:

$$\Sigma \vDash s \stackrel{\circ}{=} t$$

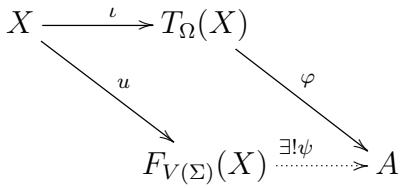
Folgerung 1.24 *Sei Σ eine Menge von Gleichungen, und sei $s \stackrel{\circ}{=} t$ eine Gleichung. Dann sind äquivalent:*

- $\Sigma \vDash s \stackrel{\circ}{=} t$
- $\overline{s} = \overline{t} \in F_{V(\Sigma)}(X)$

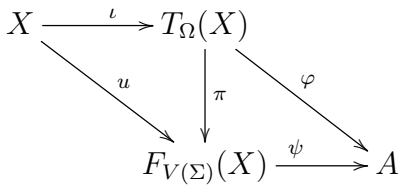
Beweis. Es gelte $\Sigma \vDash s \stackrel{\circ}{=} t$. Dann gilt insbesondere $F_{V(\Sigma)}(X) \vDash s \stackrel{\circ}{=} t$. Mit anderen Worten: Für alle Morphismen $\varphi : T_\Omega(X) \longrightarrow F_{V(\Sigma)}(X)$ gilt: $\varphi(s) = \varphi(t)$. Damit gilt insbesondere $\overline{s} = \overline{t}$.

Es gelte nun $\overline{s} = \overline{t} \in F_{V(\Sigma)}(X)$. Sei A eine beliebige Ω-Algebra in $\mathcal{K}$. Wir müssen zeigen: Für alle Morphismen $\varphi : T_\Omega(X) \longrightarrow A$ gilt: $\varphi(s) = \varphi(t)$.

Sei also $\varphi : T_\Omega(X) \longrightarrow A$ so ein Homomorphismus. Dann gibt es per Definition von $F_{V(\Sigma)}(X)$ genau einen Homomorphismus $\psi : F_{V(\Sigma)}(X) \longrightarrow A$ mit $\psi \circ u = \varphi \circ \iota$.



Ich behaupte, dass nun auch $\varphi = \psi \circ \pi$ gilt. Hierzu: Es gilt $\varphi \circ \iota = \psi \circ u = \psi \circ (\pi \circ \iota) = (\psi \circ p) \circ \iota$. Damit ist, da $T_\Omega(X)$ frei auf X ist, $\varphi = \psi \circ \pi$.



Wir haben also $\varphi(s) = \psi(\overline{s}) = \psi(\overline{t}) = \varphi(t)$. □

Beispiele 1.25

- Die freie Algebra über X zu allen Ω-Algebren ist die Term-Algebra $T_\Omega(X)$ (s.o.).
- Das freie Monoid über X ist die Menge der "Strings" X^* (mit der offensichtlichen Inklusion) (das neutrale Element ist das "leere Wort" $\Box$).
- Die freie Halbgruppe über X ist $X^* - \{\Box\}$.

- Die freie Gruppe über X kann man explizit wie folgt beschreiben: Wir fixieren eine von X disjunkte Menge, die wir mit X^{-1} bezeichnen und eine Bijektion $X \longrightarrow X^{-1}$, $x \mapsto x^{-1}$. Dann betrachten wir die Teilmenge von $(X \cup X^{-1})^*$, die aus Strings besteht so dass gilt: Es kommt niemals ein Element x direkt neben einen Element x^{-1} for. Zwei Elemente werden verknüpft, indem man sie hintereinanderschreibt und dann in offensichtlicher Weise "kürzt". Diese Teilmenge ist dann die freie Gruppe auf X. (Das neutrale Element ist wieder das leere Wort.) (Wie lauten die inversen Elemente?)

Sei ab jetzt $X = \{x_1, x_2, x_3, \ldots\}$ mit paarweise verschiedenen x_i.

Wenn nun Σ endlich ist, könnte man versuchen, die Menge der Identitäten auf $V(\Sigma)$ algorithmisch zu bestimmen. (Diese Idee ist noch sehr allgemein und muss präzisiert werden.) Hierzu ist es naheliegend, formale Beweisschemata zu betrachten.

Definition Sei Σ eine Menge von Gleichungen, und sei $s \stackrel{\circ}{=} t$ eine Gleichung über X vom Typ Ω. Dann sagen wir, dass $s \stackrel{\circ}{=} t$ *in einem Schritt aus* Σ *(formal) abgeleitet* oder *in einem Schritt aus* Σ *(formal) abgeleitet hergeleitet* oder *in einem Schritt aus* Σ *deduziert* werden kann, wenn (mindestens) eine der folgenden Bedingungen erfüllt ist:

- $s = t \in T_\Omega(X)$ (gemeint ist hier wirklich die Gleichheit von Termen)
- $t \stackrel{\circ}{=} s \in \Sigma$
- Es gibt ein $r \in T_\Omega(X)$ mit $s \stackrel{\circ}{=} r \in \Sigma$ und $r \stackrel{\circ}{=} t \in \Sigma$.
- Es gibt $p \stackrel{\circ}{=} q \in \Sigma$, einen Term r und ein $i \in \mathbb{N}$ so dass gilt: Wenn man in r die Variable x_i durch p ersetzt, erhält man s, wenn man in r die Variable x_i durch q ersetzt, erhält man t.
- Es gibt $p \stackrel{\circ}{=} q \in \Sigma$, einen Term r und ein $i \in \mathbb{N}$ so dass gilt: Wenn man in p die Variable x_i durch r ersetzt, erhält man s, wenn man in q die Variable x_i durch r ersetzt, erhält man t.

Definition Sei Σ eine Menge von Gleichungen über X, und sei $s \stackrel{\circ}{=} t$ eine Gleichung. Dann sagen wir, dass $s \stackrel{\circ}{=} t$ aus Σ *(formal) abgeleitet* oder *hergeleitet* oder *deduziert* werden kann, wenn es eine endliche Folge

$$s_1 \stackrel{\circ}{=} t_1,\ s_2 \stackrel{\circ}{=} t_2, \ldots, s_n \stackrel{\circ}{=} t_n \tag{1.1}$$

von Gleichungen mit $s_n = s$ und $t_n = t$ gibt, so dass jede der Gleichungen $s_i \stackrel{\circ}{=} t_i$ aus der Menge $\Sigma \cup \{s_1 \stackrel{\circ}{=} t_1, \ldots, s_{i-1} \stackrel{\circ}{=} t_{i-1}\}$ in einem Schritt formal hergeleitet werden kann. Die endliche Folge (1.1) nennen wir dann auch eine *formale Herleitung* oder eine *formale Ableitung* oder einen *formalen Beweis* von $s \stackrel{\circ}{=} t$ aus Σ.

In diesem Fall schreiben wir

$$\Sigma \vdash s \stackrel{\circ}{=} t \ .$$

Nun gilt der folgende Satz:

Satz 1.5 *(Vollständigkeitssatz) Sei Σ eine Menge von Gleichungen auf $X = \{x_1, x_2, x_3, \ldots\}$, und sei $s \stackrel{\circ}{=} t$ eine Gleichung auf X. Dann sind äquivalent:*

- $\Sigma \models s \stackrel{\circ}{=} t$
- $\Sigma \vdash s \stackrel{\circ}{=} t$
- $\overline{s} = \overline{t} \in F_{V(\Sigma)}(X)$

Zum Beweis. Man sieht schnell, dass alle Gleichungen, die aus Σ formal abgeleitet werden können, auch aus Σ folgen. (Mit anderen Worten: Wenn $s \stackrel{\circ}{=} t$ aus Σ formal abgeleitet werden kann und A die Gleichungsmenge Σ erfüllt, dann erfüllt A auch $s \stackrel{\circ}{=} t$.) Der schwierige Teil des Beweises ist, die Umkehrung dieser Aussage zu zeigen. Aber diesen Beweis führen wir nicht. Die Äquivalenz des ersten und des dritten Punktes ist Folgerung 1.24. □

Hieraus folgt relativ schnell:

Korollar 1.26 *Es gibt eine "Prozedur" (eine Turing Maschine), die unter Eingabe eines "endlichen" Typs, einer endlichen Menge von Gleichungen Σ auf X bezüglich Ω alle Gleichungen $s \stackrel{\circ}{=} t$, mit $\Sigma \models s \stackrel{\circ}{=} t$ und nur solche ausgibt. (Die Maschine terminiert nicht.)*

Mit anderen Worten: Es gibt eine "Prozedur" (eine Turing Maschine), die unter Eingabe eines "endlichen" Typs Ω, einer endlichen Menge von Gleichungen Σ auf X bezüglich Ω und einer Gleichung $s \stackrel{\circ}{=} t$ genau dann terminiert, wenn $\Sigma \models s \stackrel{\circ}{=} t$.

Es gibt aber keinen Algorithmus, der dieses Problem entscheidet, d.h. der immer terminiert und die korrekte Antwort "Ja", "Nein" ausgibt. Es gilt sogar der folgende berühmte Satz:

Satz 1.6 *Es gibt einen endlichen Typ Ω und eine endliche Menge von Gleichungen Σ so dass es keinen Algorithmus (keine Turing Maschine) gibt, welche unter Eingabe einer Gleichung $s \stackrel{\circ}{=} t$ mit Termen in $T_\Omega(\emptyset)$ ausgibt, ob $\Sigma \models s \stackrel{\circ}{=} t$.*

Stichwort: "Wortproblem" für endlich dargestellte Gruppen oder für endlich dargestellte Halbgruppen.

Literatur Wenn Sie sich wirklich für das Themengebiet interessieren, finden Sie vielleicht die folgenden Texte interessant:

Bergman: An Invitation to General Algebra and Universal Constructions (http://math.berkeley.edu/∼gbergman/245)

Burris, Sankappanavar: A Course in Universal Algebra (Signatur Alg 1 82 in der Bibliothek)

Grätzer: Universal Algebra (Signatur Alg 1 39 in der Bibliothek)

Wechler: Universal Algebra for Computer Scientists (Signatur ST 120 W386 in der Bibliothek)

Ich persönlich finde das Buch von Herrn Bergman am besten. Ich möchte noch anmerken, das die Definition von "Typ" in der Vorlesung formal ein Spezialfall der Definition von Herrn Bergman ist. Aber die Darstellung ist in der Vorlesung eine leicht andere ...

Kapitel 2

Endliche Körper und Anwendungen

2.1 Körpererweiterungen

Definition Sei L ein Körper und K ein Unterkörper von L. Dann sagen wir, dass L ein *Erweiterungskörper* von K ist. Wir sagen dann auch: $K \subseteq L$ (oder $L|K$ oder L/K) ist eine *Körpererweiterung*.

Häufig ist ein Homomorphismus $\iota : K \longrightarrow L$ von Körpern gegeben. Dann sagen wir wiederum auch, dass $L|K$ eine Körpererweiterung ist.

Bemerkung Wenn $L|K$ eine Körpererweiterung ist, dann ist L "in offensichtlicher Weise" ein K-Vektorraum.

Definition Der *Grad* einer Körpererweiterung $L|K$ ist die Dimension von L als K-Vektorraum. Bezeichnung: $[L : K] := \dim_K(L)$.

Insbesondere ist $[L : K] = 1 \longleftrightarrow L = K$.

Folgerung 2.1 *Seien $K \subseteq L \subseteq M$ Körpererweiterungen. Dann ist $[M : K] = [M : L] \cdot [L : K]$. (Falls eine der Grade $[M : L]$ oder $[L : K]$ unendlich ist, ist dies im offensichtlichen Sinn zu verstehen.)*

Beweis. Wenn eine der beiden Erweiterungen $M|L$ oder $L|K$ unendlichen Grad hat, so hat offensichtlich auch $M|K$ unendlichen Grad. Seien also die Grade endlich. Sei $b_1, \ldots, b_m$ eine Basis von L über K, und sei $c_1, \ldots, c_n$ eine Basis von M über L. Ich behaupte, dass dann die Elemente $b_j c_i$ mit $i = 1, \ldots, n, j = 1, \ldots, m$ eine Basis von M über K bilden.

Dazu: Sei $a \in M$. Dann gibt es also $a_1, \ldots, a_n \in L$ mit $a = \sum_{i=1}^{n} a_i c_i$. Nun gibt es für jedes $i = 1, \ldots, n$ $a_{i,1}, \ldots, a_{i,m} \in K$ mit $a_i = \sum_{j=1}^{m} a_{i,j} b_j$. Damit

ist also $a = \sum_{i,j} a_{i,j}\, c_i b_j$. Somit bilden die Elemente $b_j c_i$ mit $i = 1, \ldots, n, j = 1, \ldots, m$ ein Erzeugendensystem von M über K.

Zur linearen Unabhängigkeit: Seien $a_{i,j} \in K$ mit $\sum_{i,j} a_{i,j}\, c_i b_j = 0$. Wir haben dann also $\sum_i (\sum_j a_{i,j} b_j) c_i = 0$. Dies ist eine Linearkombination mit Koeffizienten in L. Da die c_i eine Basis von M über L bilden, gilt somit $\sum_j a_{i,j} b_j = 0$ für alle i. Da die b_j eine Basis von L über K bilden, gilt also $a_{i,j} = 0$ für alle i, j. □

Definition Sei $K \subseteq R$, wobei R ein kommutativer Ring ist, und seien $r_1, \ldots, r_r \in R$. Dann ist *der von $r_1, \ldots, r_n$ über K erzeugte Unterring* von R der kleinste Unterring von R, der K und $r_1, \ldots, r_n$ enthält. Dieser Ring wird mit $k[r_1, \ldots, r_n]$ bezeichnet.

Lemma 2.2 *Sei $r \in R$. Dann ist $K[r] := \{f(r) \in R \mid f \in K[X]\}$.*

Der *Beweis* ist einfach.

Beispiel 2.3 Der Polynomring $K[X]$ selbst wird von X über K erzeugt.

Lemma 2.4 *Sei $\varphi : k[X] \longrightarrow R\ , f \mapsto f(r)$. Dann ist φ ein K-linearer Ringhomomorphismus, und $k[r] = \mathrm{Bild}(\varphi)$. Genauer gilt: φ induziert einen K-linearen Isomorphismus von Ringen*

$$k[X]/(\mathrm{Kern}(\varphi)) \longrightarrow k[r]\ .$$

Beweis. Die erste Aussage ist eine Umformulierung des obigen Lemmas. Für den Isomorphismus siehe Ü1 auf Übungsblatt 7 zur Linearen Algebra. □

Definition Sei $K \subseteq L$ eine Körpererweiterung, und seien $\alpha_1, \ldots, \alpha_n \in L$. Dann ist der von $\alpha_1, \ldots, \alpha_n$ über K *erzeugte Körper* der kleinste Unterkörper von L, der K und $\alpha_1, \ldots, \alpha_n$ enthält. Dieser Körper wird mit $K(\alpha_1, \ldots, \alpha_n)$ bezeichnet. (Es ist offensichtlich, dass dieser Körper existiert.)

Lemma 2.5 *Sei $\alpha \in K$. Dann ist $K(\alpha) = \{\frac{p(\alpha)}{q(\alpha)} \mid p(X), q(X) \in K[X], q(\alpha) \neq 0\}$.*

Beispiel 2.6 Man kann aus dem Ring $\mathbb{Z}$ der ganzen Zahlen den Körper $\mathbb{Q}$ rationalen Zahlen "konstruieren" (siehe [LA, §1.7]). Analog kann man aus dem Polynomring $K[X]$ über dem Körper K einen Körper "konstruieren", dessen Elemente durch Brüche $\frac{p(X)}{q(X)}$ von Polynomen über K mit $q(X) \neq 0$ gegeben sind. Hierbei gilt wiederum die übliche "Erweiterungsregel" $\frac{p(X)}{q(X)} = \frac{p(X)\cdot r(X)}{q(X)\cdot r(X)}$ für jedes Polynom $r(X) \neq 0$. Dieser Körper wird offensichtlich von

X erzeugt. Deshalb bezeichnet man ihn mit $K(X)$. Dieser Körper heißt der *Körper der rationalen Funktionen* über K.

Sei nun $L|K$ eine Körpererweiterung.

Definition Sei $\alpha \in L$. Dann heißt α *algebraisch* über K, wenn α eine nicht-triviale "algebraische Gleichung"

$$c_n X^n + c_{n-1} X^{n-1} + \cdots + c_0 \stackrel{\circ}{=} 0$$

mit "Koeffizienten" $c_i \in K$ erfüllt. ("Nicht-trivial" heißt hier, dass nicht alle Koeffizienten gleich Null sind.) Ist dies nicht der Fall, so heißt α *transzendent* über K. L heißt *algebraisch über* K, wenn jedes Element von L algebraisch über K ist.

Beispiele 2.7

- Die "n-ten Einheitswurzeln" $\zeta_n^j := e^{\frac{2\pi \mathrm{i} \cdot j}{n}} \in \mathbb{C}$ $(n, j \in \mathbb{N})$ sind algebraisch über $\mathbb{Q}$. (Sie erfüllen die Gleichung $X^n - 1 \stackrel{\circ}{=} 0$.)
- Die Zahlen π und e sind transzendent über $\mathbb{Q}$. (Das sind sehr schwierige Resultate.)

Lemma 2.8 *Jede endliche Körpererweiterung ist algebraisch.*

Beweis. Sei $L|K$ so eine Körpererweiterung, und sei $\alpha \in L$. Sei $n := \dim_K(L)$. Dann ist sind die Elemente $1, \alpha, \ldots, \alpha^n$ linear abhängig. Somit erfüllen sie eine lineare Gleichung über K. Dies bedeutet gerade, dass α algebraisch über K ist. □

Achtung Die Umkehrung dieser Aussage gilt nicht!

Bemerkung Das Element α ist genau dann algebraisch über K, wenn es ein Polynom $f \in K[X], f \neq 0$, mit $f(\alpha) = 0$ gibt. Sei $\varphi : K[X] \longrightarrow L$, $\varphi(p) := f(\alpha)$. Dann ist φ ein K-linearer Ringhomomorphismus. Dann ist genau dann $f(\alpha) = 0$, wenn $f \in \mathrm{Kern}(\varphi)$. Insbesondere ist α genau dann algebraisch über K, wenn $\mathrm{Kern}(\varphi) \neq \{0\}$.

Beachten Sie ferner: In jedem Ideal $I \neq \{0\}$ von $K[X]$ gibt es ein eindeutig bestimmtes normiertes Polynom $f(X)$ mit $I = (f(X))$ (siehe [LA, Lemma 1.70]); $f(X)$ ist das eindeutig bestimmte normierte Polynom in I kleinsten Grades.

Definition Sei $\alpha \in L$ algebraisch über K. Dann heißt das eindeutig bestimmte normierte Polynom $f \in K[X]$ kleinstes Grades mit $f(\alpha) = 0$ das *Minimalpolynom* von α. Die Bezeichung ist μ_α.

Sei weiterhin α algebraisch, und sei – wie gerade definiert – sein Minimalpolynom μ_α. Sei auch weiterhin φ wie oben definiert. Dann ist also $\mathrm{Kern}(\varphi) = (f(X))$. Damit induziert φ einen K-linearen Isomorphismus von Ringen

$$k[X]/(f(X)) \longrightarrow k[\alpha] \ , \ [f]_{(\mu_\alpha(X))} \mapsto f(\alpha) \ .$$

Insbesondere ist also $[K[\alpha] : K] = \mathrm{Grad}(\mu_\alpha)$.

Folgerung 2.9 *Das Minimalpolynom μ_α ist irreduzibel in $K[X]$. Die Ringe $K[X]/(f(X))$ und $K[\alpha]$ sind Körper.*

Beweis. Annahme: $f = f_1 \cdot f_2$ mit nicht-konstanten $f_1, f_2 \in K[X]$. Dann ist also $f_1(\alpha) = 0$ oder $f_2(\alpha) = 0$ (oder beides). Das widerspricht aber der Definition von f.

Hieraus folgt mit [LA, Satz 1.4], dass $K[X]/(f(X))$ ein Körper ist. Nach dem obigen Isomorphismus ist somit $K[\alpha]$ auch ein Körper. □

Sei nun α wieder ein beliebiges Element von L. Dann haben wir die folgende Charakterisierung von "α ist algebraisch über K":

Folgerung 2.10 *Die folgenden Aussagen sind äquivalent:*

a) α ist algebraisch über K.

b) $K[\alpha]$ ist ein Körper.

c) $K[\alpha] = K(\alpha)$.

Beweis. Wir zeigen zunächst, dass b) und c) äquivalent sind.

Offensichtlich impliziert c) b). Es gelte also b). Sicherlich ist $K[\alpha] \subseteq K(\alpha)$. Nach Voraussetzung gilt aber nun auch $K(\alpha) \subseteq K[\alpha]$. Damit gilt also c).

Wir haben bereits gesehen, dass a) b) impliziert. Wir müssen noch zeigen, dass b) a) impliziert. Sei also $K[\alpha]$ ein Körper. Dann ist der (surjektive) Homomorphismus $\varphi : K[X] \longrightarrow K[\alpha] \ , \ f \mapsto f(\alpha)$ nicht injektiv. (Andernfalls wäre φ ein Isomorphismus und somit $K[\alpha] \simeq K[X]$ kein Körper.) Damit ist α Nullstelle eines nicht-trivialen Polynoms in $K[X]$, d.h. α ist algebraisch über K. □

Satz 2.1 *Sei K ein Körper, und sei $f \in K[X]$. Dann gibt es einen Körper L und einen injektiven Homomorphismus $\iota : K \longrightarrow L$ so dass gilt: Das Polynom f (genauer: das Bild von f unter der offensichtlichen Inklusion $K[X] \longrightarrow L[X]$) zerfällt in L in Linearfaktoren, und L wird über K (genauer: über $\iota(K)$) von den Nullstellen von f erzeugt.*

Außerdem gilt: Seien L mit ι und L' mit ι' zwei solche Körper. Dann gibt es einen Isomorphismus $\varphi : L \longrightarrow L'$ mit $\varphi \circ \iota = \iota' : K \longrightarrow L'$.

Bemerkung Im Gegensatz zu ähnlichen früheren Aussagen ist der Isomorphismus φ nicht immer eindeutig bestimmt.

Definition Man nennt einen Körper L wir im Satz (zusammen mit ι) einen *Zerfällungskörper* von f über K.

Beweis. Wir beweisen zunächst die Existenz. Wenn f konstant ist, ist nichts zu zeigen. Sei also f nicht konstant. Ich gebe zunächst einen etwas informalen Beweis und dann noch einen genaueren Beweis.

Sei $f = f_1 \cdots f_k \in K[X]$ mit irreduziblen f_i. Nun ist $L_1 := K[X]/(f_1(X))$ ein Körper, in dem $f(X)$ eine Nullstelle hat, nämlich $\alpha_1 := [X]_{(f_1(X))}$.

Wir betrachten nun das Polynom $\frac{f}{X-\alpha_1} \in L_1[X]$. Wenn dieses Polynom konstant ist, ist wiederum nichts zu zeigen. Anderenfalls kann man analog wie zuvor einen Körper L_2 finden, in dem $\frac{f}{X-\alpha_1}$ eine Nullstelle hat. Wenn man diese "Prozedur" iteriert, erhält man schließlich einen Körper L mit den gewünschten Eigenschaften.

Genauer kann man wie folgt vorgehen: Wir beweisen per vollständiger Induktion nach n, dass die folgende Aussage richtig ist: Für alle Körper K und alle Polynome $f \in K[X]$ mit $\mathrm{Grad}(f) = n$ gibt es einen Zerfällungskörper von f.

Der Induktionsanfang $n = 1$ ist trivial.

Es gelte die Aussage also für n. Sei also K ein Körper und $f \in K[X]$ ein Polynom vom Grad $n + 1$.

Wir definieren L_1 und α_1 wie oben. Nun wenden wir die Induktionsvoraussetzung auf $\frac{f}{X-\alpha_1} \in L_1[X]$ an und erhalten einen Zerfällungskörper L von $\frac{f}{X-\alpha_1}$ über L_1.

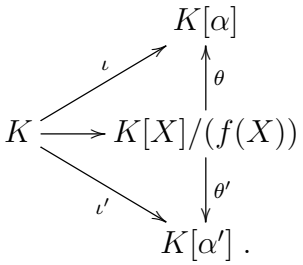
Ich behaupte, dass L ein Zerfällungskörper von f über K ist. Wir wissen schon, dass f über L in Linearfaktoren zerfällt. Wir müssen zeigen, dass L von den Nullstellen von f erzeugt wird. Sei hierzu Z ein Unterkörper von L, der K und alle Nullstellen von f enthält. Dann enthält Z insbesondere K und α_1, und somit auch L_1. Da L auch alle Nullstellen von $\frac{f}{X-\alpha_1}$ enthält, ist nach Definition von L $Z = L$.

Seien nun L mit ι und L' mit ι' zwei Zerfällungskörper von f über K. Die Idee ist, φ "induktiv" auf Teilkörpern zu definieren, die von Nullstellen von f erzeugt werden.

Wiederum nehmen wir an, dass f nicht-konstant ist. Sei wiederum f_1 ein irreduzibler Teiler von f, und seien $\alpha_1 \in L$ und $\alpha_1' \in L'$ Nullstellen von f_1. Wir haben dann also K-lineare Isomorphismen von Körpern $\theta : K[X]/(f(X)) \longrightarrow K[\alpha_1], \theta' : K[X]/(f(X)) \longrightarrow K[\alpha']$ wie oben. Damit ist also

$$\varphi_1 : (\theta')^{-1} \circ \theta : K[\alpha_1] \longrightarrow K[\alpha_1'],\ f(\alpha) \mapsto f(\alpha')$$

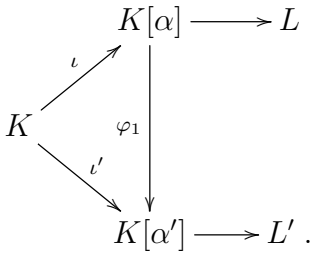
ein Isomorphismus von Körpern. Offensichtlich gilt $\varphi_1 \circ \iota = \iota'$.



Wiederum zeigen wir die Aussage nun der Induktion nach dem Grad von f über beliebigen Körpern.

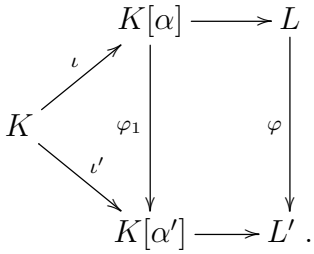
Die Induktionsbasis ist wiederum trivial. Die Aussage gelte für alle Polynome vom Grad n, und f sei ein Polynom vom Grad $n+1$.

Seien f_1, α_1 und α_1' wie zuvor. Dann haben wir also ein kommutatives Diagramm



Nun ist L ein Zerfällungskörper von $\frac{f(X)}{X-\alpha_1}$ über $K[\alpha]$. Analog ist L' ein Zerfällungskörper von $\frac{f(X)}{X-\alpha_1'}$ über $K[\alpha_1']$. Dies bedeutet aber, dass L' mit der Inklusion $\varphi_1 : K[\alpha] \longrightarrow L'$ auch ein Zerfällungskörper von von $\frac{f(X)}{X-\alpha_1}$ über $K[\alpha]$ ist. Nach Induktionsvoraussetzung gibt es somit einen Homomorphismus von

Körpern $L \longrightarrow L'$ so dass die rechte Seite des Diagramms



kommutiert. Dies bedeutet insbesondere, dass $\varphi\iota = \iota'$ ist. □

Ich erwähne noch:

Definition Ein Körper K so dass jedes Polynom in $K[X]$ in Linearfaktoren zerfällt, heißt *algebraisch abgeschlossen.*

Satz 2.2 *Sei K ein Körper. Dann gibt es einen Erweiterungskörper $\overline{K}$ von K, der algebraisch über K und algebraisch abgeschlossen ist.*

Definition Ein Körper wie im Satz heißt ein *algebraischer Abschluss* von K.

Beispiele 2.11

- Der Körper $\mathbb{C}$ der komplexen Zahlen ist algebraisch abgeschossen, und $\mathbb{C}$ ist "der" algebraische Abschluss von $\mathbb{R}$.

- Der Menge aller komplexer Zahlen, die algebraisch über $\mathbb{Q}$ sind, ist ein algebraischer Abschluss von $\mathbb{Q}$. Es ist $[\overline{\mathbb{Q}} : \mathbb{Q}] = \infty$.

Wenn man die Existenz von algebraischen Abschlüßen voraussetzt, ist es sehr einfach, die Existenz von Zerfällungskörpern zu zeigen. Es gilt die folgende offensichtliche Aussage:

Folgerung 2.12 *Sei K ein Körper und $f \in K[X]$. Sei $L|K$ eine Körpererweiterung, so dass $f \in L[X]$ in Linearfaktoren zerfällt (z.B. sei L ein algebraischer Abschluss von K). Sei $f = c \cdot (X - \alpha_1) \cdots (X - \alpha_k)$ mit $c \in K$ und $\alpha_i \in L$. Dann ist $K(\alpha_1, \ldots, \alpha_k)$ ein Zerfällungskörper von f über K.*

2.2 Endliche Körper

Existenz und Eindeutigkeit

Ich erinnere, wie die *Charakteristik* eines Körpers definiert ist:

Sei K ein Körper. Wir betrachten den Ringhomomorphismus $\varphi : \mathbb{Z} \longrightarrow K, n \mapsto n \cdot 1_K$. Die *Charakteristik* von K, $\mathrm{char}(K)$, ist nun die eindeutig bestimmte Zahl $n \in \mathbb{N}_0$ mit $\mathrm{Kern}(\varphi) = (n)$.

Mit anderen Worten: Wenn es eine Zahl $n \in \mathbb{N}$ mit $n \cdot 1_K = 0 \in K$ gibt, dann ist $\mathrm{char}(K)$ die kleinste solche Zahl. Wenn es keine solche Zahl gibt, dann ist $\mathrm{char}(K) = 0$.

Wenn die Charakteristik von K gleich 0 ist, dann haben wir also einen injektiven Homomorphismus $\mathbb{Z} \hookrightarrow K$. Man zeigt dann leicht, dass man auch einen injektiven Homomorphismus $\mathbb{Q} \hookrightarrow K$ hat. Also: Jeder Körper der Charakteristik 0 "enthält" $\mathbb{Q}$.

Sei nun $\mathrm{char}(K) = p > 0$. Dann haben wir also einen injektiven Homomorphismus $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z} \hookrightarrow K$. Also: Jeder Körper der Charakteristik $p > 0$ "enthält" $\mathbb{F}_p$. (Wenn wir $\mathbb{F}_p$ mit seinem Bild in K "identifizieren".)

Insbesondere sehen wir: Jeder Körper der Charakteristik 0 ist "in offensichtlicher Weise" ein $\mathbb{Q}$-Vektorraum, und jeder Körper der Charakteristik $p > 0$ ist "in offensichtlicher Weise" ein $\mathbb{F}_p$-Vektorraum.

Ein *endlicher Körper* ist, wie der Name schon sagt, ein Körper mit endlich vielen Elementen. Offensichtlich ist die Charakteristik eines endlichen Körpers positiv.

Sei K ein endlicher Körper der Charakteristik p. Dann ist K offensichtlich ein endlich erzeugter $\mathbb{F}_p$-Vektorraum. Sei $e := [K : \mathbb{F}_p]$. Dann ist $\#K = p^e$. Wir erhalten somit:

Folgerung 2.13 *Sei K ein endlicher Körper. Dann ist die Charakteristik von K positiv; sei $\mathrm{char}(K) = p$ und $e := [K : \mathbb{F}_p]$. Dann ist $\#K = p^e$.*

Folgerung 2.14 *Sei K ein endlicher Körper der Charakteristik p. Dann ist die Abbildung $\sigma : K \longrightarrow K\ , a \mapsto a^p$ ein Automorphismus des Körpers K.*

Beweis. Es ist klar, dass $0^p = 0$ und $(ab)^p = a^p b^p$ für alle $a, b \in K$. Um zu zeigen, dass K ein Endomorphismus von K (Homomorphismus von K nach K) ist, müssen wir zeigen, dass $(a+b)^p = a^p + b^p$ für alle $a, b \in K$.

Seien also $a, b \in K$ beliebig. Dann ist $(a+b)^p = \sum_{i=0}^{p} \binom{p}{i} a^i b^{p-i}$.

Sei nun $i = 1, \ldots, p-1$. Dann gilt $p | \binom{p}{i}$. Denn: $\binom{p}{i} = \frac{p!}{i! \cdot (p-i)!}$, $p | p!, p \nmid i!$ und $p \nmid (p-i)!$.

Damit ist also $(a+b)^p = a^p + b^p$.

Somit ist die Abbildung σ ein Endomorphismus von K. (Dies gilt allgemeiner für beliebige Körper der Charakteristik $p > 0$.)

Wir wissen, dass Homomorphismen von Körpern immer injektiv sind. Da K endlich ist, ist σ auch surjektiv und somit bijektiv. □

Definition Sei K ein endlicher Körper der Charakteristik p. Dann heißt der Körperautomorphismus $\sigma : K \longrightarrow K\,, a \mapsto a^p$ der *Frobeniusautomorphismus* von K.

Wir wollen nun zeigen, dass es zu jeder "Primpotenz" $q := p^e$ (p prim und $e \in \mathbb{N}$) "bis auf Isomorphie" genau einen endlichen Körper mit q Elementen gibt. Diese Aussagen folgen aus dem folgenden Satz (und der Existenz und Eindeutigkeit von Zerfällungskörpern).

Satz 2.3 *Sei p eine Primzahl, $e \in \mathbb{N}$ und $q = p^e$. Dann gilt:*

a) Sei K ein Zerfällungskörper von $X^q - X$ über $\mathbb{F}_p$. Dann hat K genau q Elemente, und es gilt $X^q - X = \prod_{a \in K} X - a$.

b) Sei K ein endlicher Körper mit q Elementen. Dann ist K ein Zerfällungskörper von $X^q - X$ über $\mathbb{F}_p$.

Zum Beweis benötigen wir die folgende Definition und die folgenden Lemmata.

Definition Sei K ein Körper und $f = \sum_{i=0}^{d} a_i X^i \in K[X]$. Dann ist die *formale Ableitung* von f definiert als

$$f' := \sum_{i=1}^{d} a_i \, i \, X^{i-1} \; .$$

Lemma 2.15 *Die Abbildung $K[X] \longrightarrow K[X], f \mapsto f'$ ist K-linear und erfüllt die* Leibniz-Regel *(*Produkt-Regel*)* $(fg)' = f'g + fg'$ *und die* Ketten-Regel $(f(g(X)))' = f'(g(X)) \cdot g(X)'$ *für $f, g \in K[X]$.*

Der *Beweis* ist eine Übungsaufgabe.

Bemerkung Wenn K Charakteristik $p > 0$ hat, gilt die überraschende Identität $(X^p)' = p \cdot X^{p-1} = 0$.

Lemma 2.16 *Seien $g, h \in K[X]$ und $f := g^2 h$. Dann gilt $g | \operatorname{ggT}(f, f')$.*

Beweis. Sicher gilt $g | f$. Es ist $f' = 2gg'h + g^2 h'$, und somit gilt auch $g | f'$. □

Bemerkung Eine Umformulierung dieses Lemmas ist: Sei $f \in K[X]$ normiert, $f = f_1^{e_1} \cdots f_k$ die Zerlegung von f in irreduzible Polynome f_i mit $e_i \geq 1$. Es gelte $\mathrm{ggT}(f, f') = 1$. Dann sind alle Indizes $e_i = 1$.

Beweis des Satzes. Sei K ein Zerfällungskörper von $X^q - X$ über $\mathbb{F}_p$. Es gilt $(X^q - X)' = -1$, insbesondere ist also $\mathrm{ggT}(X^q - X, (X^q - X)') = 1$. Damit zerfällt $X^q - X$ über K also in ein Produkt von *verschiedenen* normierten Linearfaktoren $X - a$. Mit anderen Worten: $X^q - X$ hat in K genau (verschiedene) q Nullstellen. Sei U die Menge der Nullstellen von $X^q - X$ in K. Dann hat U also q Elemente. Ich behaupte, dass U ein Körper ist.

Dazu: Es ist $1 \in U$. Seien $a, b \in U$ beliebig. Dann ist also $a^q = a$ und $b^q = b$. Damit ist $(ab)^q = ab$ und $(a+b)^q = (a+b)^{(p^e)} = a^q + b^q = a + b$. Es ist $(-a)^q = (-1)^q a^q = (-1)^q a = -a$, denn wenn $p = 2$ ist, ist $-1 = 1 \in K$, und wenn p ungerade ist, ist $(-1)^q = -1$. Damit ist also $ab \in U, a+b \in U, -a \in U$. Sei $a \neq 0$. Dann ist $(\frac{1}{a})^q = \frac{1}{a^q} = \frac{1}{a}$, also auch $\frac{1}{a} \in U$.

U ist nun ein Körper, der K enthält so dass $X^q - X$ in $U[X]$ in Linearfaktoren zerfällt. Da K ein Zerfällungskörper, enthält (per Definition von "Zerfällungskörper") K keinen echten Teilkörper[1] mit dieser Eigenschaft. Somit gilt $U = K$. Insbesondere hat K also q Elemente. Es gilt nun auch $X^q - X = \prod_{a \in K} X - a$

Sei nun K ein endlicher Körper mit q Elementen. Ich behaupte, dass für jedes Element $a \in K$ $a^q = a$ gilt. Dies ist trivial für $a = 0$. Sei also $a \neq 0$, d.h. $a \in K^*$ (die multiplikative Gruppe von K). Die Gruppe K^* hat $q - 1$ Elemente. Hieraus folgt $a^{q-1} = 1$ (siehe "Wiederholung"), insbesondere also $a^q = a$.

Nun ist jedes Element von K also eine Nullstelle von $X^q - X$. Allerdings hat K q Elemente, und $X^q - X$ hat Grad q, also höchstens q Nullstellen. Damit ist $X^q - X = \prod_{a \in K}(X - a)$. Damit ist offensichtlich K ein Zerfällungskörper von $X^q - X$ über $\mathbb{F}_p$. □

Wiederholung (siehe Aufgaben H2 und H3 von Übungsblatt 5 zur Linearen Algebra)

Sei G eine (multiplikativ geschriebene) endliche Gruppe, und sei $a \in G$. Dann ist die *Ordnung* von a, $\mathrm{ord}(a)$, ist die kleinste Zahl $n \in \mathbb{N}$ mit $a^n = e$. Man sieht leicht, dass für alle $n \in \mathbb{Z}$ gilt: $a^n = e \longleftrightarrow \mathrm{ord}(a) | n$. Sei $\langle a \rangle$ die von a erzeugte Untergruppe. Dann ist $\mathrm{ord}(a) = \#\langle a \rangle$. Man kann relativ leicht zeigen (siehe H2), dass $\mathrm{ord}(a) | \#G$. Damit gilt also insbesondere $a^{\#G} = e$.

Wir haben gesehen, dass es zu jeder Primpotenz q bis auf Isomorphie genau einen Körper mit q Elementen gibt.

[1] Ein *echter Teilkörper* von K ist ein Teilkörper M von K, der nicht gleich K ist.

Definition Sei q eine Primpotenz. Den (bis auf Isomorphie eindeutig bestimmten) Körper mit q Elementen bezeichnen wir mit $\mathbb{F}_q$.

Seien nun $d, e \in \mathbb{N}$. Wir fragen uns, wann $\mathbb{F}_{p^d}$ ein Unterkörper von $\mathbb{F}_{p^e}$ ist, genauer, wann es einen Homomorphismus $\mathbb{F}_{p^d} \hookrightarrow \mathbb{F}_{p^e}$ gibt.

Wenn dies der Fall ist, ist $d = [\mathbb{F}_{p^d} : \mathbb{F}_p]$ ein Teiler von $e = [\mathbb{F}_{p^e} : \mathbb{F}_p]$ nach Folgerung 2.1.

Sei umgekehrt d ein Teiler von e. Ich behaupte, dass nun die Menge

$$M := \{a \in \mathbb{F}_q : a^{p^d} = a\}$$

ein Körper mit p^d Elementen ist. Hierzu betrachten wir das Polynom $X^{p^d} - X$. Dieses Polynom ist ein Teiler von $X^{p^e} - X$.

Denn: p^e ist eine Potenz von p^d, und somit ist $p^d - 1$ ein Teiler von $p^e - 1$. (Für alle $a \in \mathbb{Z}$ und $n \in \mathbb{N}$ ist $a - 1$ ein Teiler von $a^n - 1$, da $a^n - 1 = (a - 1) \cdot (a^{n-1} + \cdots + a + 1)$.) Deshalb ist $X^{p^d-1} - 1$ ein Teiler von $X^{p^e-1} - 1$. (Für alle Polynome $g \in R[X]$, R ein beliebiger kommutativer Ring, und alle $n \in \mathbb{N}$ gilt $g^n - 1 = (g - 1) \cdot (g^{n-1} + \cdots + g + 1)$.)

Da $\mathbb{F}_q$ der Zerfällungskörper von $X^q - X$ ist, zerfällt das Polynom $X^{p^d} - X$ in $\mathbb{F}_q[X]$ in Linearfaktoren; M ist die Menge der Nullstellen von $X^{p^d} - X$. Damit hat M genau p^d Elemente. Wie im Beweis des obigen Satzes sieht man nun, dass M ein Körper ist. Offensichtlich ist M der einzige Unterkörper von $\mathbb{F}_q$ mit p^d Elementen. (Wenn N ein anderer solcher Körper ist, erfüllen die Elemente von N auch alle die Gleichung $X^{p^d} \mathrel{\mathring{=}} X$.)

Wir haben bewiesen:

Folgerung 2.17 *Seien $d, e \in \mathbb{N}$. Dann ist $\mathbb{F}_{p^d}$ genau dann in $\mathbb{F}_{p^e}$ enthalten (genauer: kann in $\mathbb{F}_q$ eingebettet werden), wenn $d|e$. In diesem Fall ist (das eindeutig bestimmte Bild von) $\mathbb{F}_{p^d}$ gleich der Menge $\{a \in \mathbb{F}_q : \sigma^d(a) = a\}$.*

(Hier ist $\sigma : \mathbb{F}_q \longrightarrow \mathbb{F}_q\,, a \mapsto a^p$ der Frobeniusautomorphismus.)

Die multiplikative Gruppe

Sei G eine multiplikativ geschriebene Gruppe.

Definition Ein Element $g \in G$ mit $\langle g \rangle = G$ heißt ein *Erzeugendes* von G. Wenn G ein Erzeugendes hat, dann heißt G *zyklisch*.

Bemerkung Das Element $g \in G$ ist genau dann ein Erzeugendes von G, wenn gilt: $\forall a \in G \exists n \in \mathbb{Z} : g^n = a$. Wenn G endlich ist, ist dies äquivalent zu $\forall a \in G \exists n \in \mathbb{N}_0 : g^n = a$.

In jeden Fall ist g genau dann ein Erzeugendes von G, wenn der Gruppenhomomorphismus $\varphi : \mathbb{Z} \longrightarrow G, n \mapsto g^n$ surjektiv ist. Falls dies der Fall ist, erhalten wir einen induzierten Isomorphismus

$$\mathbb{Z}/\operatorname{Kern}(\varphi) \longrightarrow G \; . \tag{2.1}$$

Wenn G endlich ist, ist $\operatorname{Kern}(\varphi) = (\operatorname{ord}(\#G)) \; (\stackrel{\text{Def}}{=} \operatorname{ord}(\#G)\cdot\mathbb{Z})$. Die endlichen zyklischen Gruppen sind also genau diejenigen Gruppen, die isomorph zu $\mathbb{Z}/n\mathbb{Z}$ für ein $n \in \mathbb{N}$ sind. Die einzige unendliche zyklische Gruppe ist $\mathbb{Z}$.

Lemma 2.18 *Sei $\#G = n < \infty$ und g ein Erzeugendes von G, sei $j \in \mathbb{N}$. Dann ist g^j genau dann ein Erzeugendes von G, wenn $\operatorname{ggT}(n, j) = 1$, d.h. wenn $[j]_n \in (\mathbb{Z}/n\mathbb{Z})^*$.*

Beweis. Wir wissen schon, dass $\operatorname{ggT}(n, j) = 1$ genau dann wenn $[j]_n \in (\mathbb{Z}/n\mathbb{Z})^*$.

Aufgrund des Isomorphismus (2.1) müssen wir für die erste Behauptung nur zeigen, dass $[j]_n$ genau dann Erzeugendes von $\mathbb{Z}/n\mathbb{Z}$ ist, wenn $[j]_n \in (\mathbb{Z}/n\mathbb{Z})^*$.

Nun ist offensichtlich $[j]_n$ genau dann ein Erzeugendes von $\mathbb{Z}/\mathbb{Z}_n$, wenn $1 = [1]_n$ ein Vielfaches von $[j]_n$ ist. Dies bedeutet aber gerade, dass $[j]_n$ invertierbar ist, d.h. $[j]_n \in (\mathbb{Z}/n\mathbb{Z})^*$. □

Wenn also G eine endliche zyklische Gruppe mit n Elementen ist, dann hat G genau $\varphi(n)$ Erzeugende.

Definition Die *Eulersche φ-Funktion* ist die Funktion $\varphi : \mathbb{N} \longrightarrow \mathbb{N}$ mit $\varphi(1) := 1$ und $\varphi(n) := \#(\mathbb{Z}/n\mathbb{Z})^*$ für $n \geq 2$.

Lemma 2.19 *Es gilt $\sum_{d|n} \varphi(d) = n$.*

Beweis. Sei für $d|n$ $S_d := \{j \in \{0, \ldots, n-1\} : \operatorname{ggT}(j, n) = \frac{n}{d}\}$. Offensichtlich ist $\{0, \ldots, n-1\} = \dot{\bigcup}_{d|n} S_d$. Wir wollen zeigen, dass $\#S_d = \varphi(d)$. Daraus folgt dann die Behauptung. Beachten Sie, dass offensichtlich $\#S_n = \varphi(n)$.

Die Menge S_d besteht offensichtlich aus den Elementen $i \cdot \frac{n}{d}$ für $i = 1, \ldots, d-1$ und $\operatorname{ggT}(i \cdot \frac{n}{d}, n) = \frac{n}{d}$. Es ist $\operatorname{ggT}(i \cdot \frac{n}{d}, n) = \operatorname{ggT}(i \cdot \frac{n}{d}, d \cdot \frac{n}{d}) = \frac{n}{d} \cdot \operatorname{ggT}(i, d)$. Somit besteht S_d also aus den Elementen $i \cdot \frac{n}{d}$ für $i = 1, \ldots, d-1$ und $\operatorname{ggT}(i, d) = 1$. Wir haben also eine Bijektion

$$\{1, \ldots, d \mid \operatorname{ggT}(i, d) = 1\} \longrightarrow S_d \; , i \mapsto i \cdot \frac{n}{d} \; ,$$

und die linke Seite hat per Definition $\varphi(d)$ Elemente. □

Sei im Folgenden $q = p^e$, wobei p eine Primzahl ist.

Satz 2.4 *Die Gruppe $\mathbb{F}_q^*$ hat ein Erzeugendes, ist also zyklisch.*

Beweis. Ich wiederhole noch einmal, dass die Ordnung eines Elements von $\mathbb{F}_q^*$ immer $q-1$ teilt.

Sei d ein Teiler von $q-1$. Beachten Sie, dass die Menge derjenigen Elemente, deren Ordnung d teilt, genau die Menge der Nullstellen des Polynoms $X^d - X$ in $\mathbb{F}_q$ ist.

Ich behaupte, dass es entweder kein oder genau $\varphi(d)$ Elemente von Ordnung d gibt. (Es wird sich dann herausstellen, dass der erste Fall nicht eintreten kann.)

Wir nehmen an, dass es ein Element a von Ordnung d gibt. Dann sind die Elemente $1, a, a^2, \ldots, a^{d-1}$ alle verschieden. Somit haben wir also $X^d - X = \prod_{j=0}^{d-1}(X - a^j)$, und jedes Element, dessen Ordnung d teilt, ist eines der Elemente a^j. Wir wissen schon, dass a^j genau dann Ordnung d hat, wenn $\mathrm{ggT}(j,d) = 1$ ist. Damit sind genau die $\varphi(d)$ vielen Elemente a^j mit $\mathrm{ggT}(j,d) = 1$ die Elemente von Ordnung d.

Wir haben nun einerseits $q - 1 = \sum_{d|q-1} \varphi(d)$ nach dem obigen Lemma, andererseits ist die Anzahl aller Elemente von $\mathbb{F}_q^*$ auch $q-1$. Das bedeutet, dass es für jedes $d|q-1$ genau $\varphi(d)$ (und niemals 0) Elemente von Ordnung $\varphi(d)$ geben muss.

Die Behauptung folgt nun mit $d = q - 1$. □

Definition Ein Erzeugendes der Gruppe $\mathbb{F}_q^*$ heißt auch *multiplikatives Erzeugendes* von $\mathbb{F}_q$. Ein Element $\alpha \in \mathbb{F}_q$ mit $\mathbb{F}_q = \mathbb{F}_p[\alpha]$ heißt ein *primitives Element* von $\mathbb{F}_q$.

Jedes multiplikative Erzeugende von $\mathbb{F}_q$ ist ein primitives Element. Damit gilt insbesondere:

Korollar 2.20 *(Satz vom primitiven Element für endliche Körper) Es gibt ein Element $\alpha \in \mathbb{F}_q$ mit $\mathbb{F}_q = \mathbb{F}_p[\alpha]$. Genauer gibt es sogar mindestens $\varphi(q)$ solche α.*

Automorphismen endlicher Körper

Sei wie oben $\sigma : \mathbb{F}_q \longrightarrow \mathbb{F}_q\,, a \mapsto a^q$ der Frobeniusautomorphismus.

Satz 2.5 *Sei $\alpha \in \mathbb{F}_q$, sei $\mu_\alpha \in \mathbb{F}_p[X]$ das Minimalpolynom von α über $\mathbb{F}_p$, und sei $d := \mathrm{Grad}(\mu_\alpha)$. Dann sind alle Elemente $\sigma^i(\alpha) = \alpha^{p^i}$ mit $i = 0, \ldots, d-1$ paarweise verschieden, und es gilt $\mu_\alpha = (X - \alpha)(X - \sigma(\alpha)) \cdots (X - \sigma^{d-1}(\alpha))$.*

Beweis. Sei $\mu_\alpha = \sum_i a_i X^i$ mit $a_i \in \mathbb{F}_p$. Dann gilt $\sigma(a_i) = a_i$ für alle i. Somit gilt $\mu_\alpha(\sigma(\alpha)) = \sum_i a_i \sigma(\alpha) = \sum_i \sigma(a_i \alpha) = \sigma(\sum_i a_i \alpha) = \sigma(\mu_\alpha(\alpha)) = 0$. Indem man dies iteriert, sieht man, dass $\mu_\alpha(\sigma^i(\alpha)) = 0$ für alle $i \in \mathbb{N}$.

Ich behaupte, dass alle $\sigma^i(\alpha)$ mit $i = 0, \ldots, d-1$ paarweise verschieden sind.

Angenommen, dies ist nicht der Fall. Sei $\sigma^i(\alpha) = \sigma^j(\alpha)$ für $0 \leq i < j < d$. Wenn wir auf beide Seiten der Gleichung $\sigma^j(\alpha) = \sigma^i(\alpha)$ den Automorphismus σ^{d-j} anwenden, erhalten wir: $\alpha = \sigma^d(\alpha) = \sigma^{d-j+i}(\alpha) = \sigma^{d-(j-i)}(\alpha)$. Nun gilt $0 < d-(j-i) < e$. Es gibt also ein $k \in \mathbb{N}$ mit $0 < k < d$ und $\alpha^{p^k} = \sigma^k(\alpha) = \alpha$.

Wir betrachten den Körper $\mathbb{F}_p[\alpha]$. Dieser Körper hat p^d Elemente. Sei nun $a \in \mathbb{F}_p[\alpha]$ beliebig. Dann gibt es also ein $g \in \mathbb{F}_p[X]$ mit $g(\alpha) = a$. Damit gilt $\mu_\alpha(\alpha)^{p^k} = \sigma^k(\mu_\alpha(\alpha)) = \mu_\alpha(\sigma^k(\alpha)) = \mu_\alpha(\alpha)$. Somit ist *jedes* Element von $\mathbb{F}_p[\alpha]$ eine Nullstelle des Polynoms $X^{p^k} - X$. Dieses Polynom hat aber höchstens $p^k < p^d$ verschiedene Nullstellen. Dies ist ein Widerspruch. □

Wir betrachten nun die Automorphismen der endlichen Körper.

Lemma 2.21 *Der einzige Automorphismus des Körpers $\mathbb{F}_p$ ist die Identität.*

Beweis. Sei τ ein Automorphismus von $\mathbb{F}_p$. Dann ist $\tau(1) = 1$. Aufgrund der Additivität ist damit $\tau(a) = a$ für alle $a \in \mathbb{F}_p$. □

Korollar 2.22 *Sei τ ein Automorphismus des Körpers $\mathbb{F}_q$. Dann ist $\tau = \sigma^i$ für ein $i \in \mathbb{N}$. Mit anderen Worten: Die Gruppe $\mathrm{Aut}(\mathbb{F}_q)$ ist eine zyklische Gruppe mit Erzeugendem σ, und die Ordnung von σ (die Anzahl der Elemente von $\mathrm{Aut}(\mathbb{F}_q)$) ist e.*

Beweis. Sei τ ein Automorphismus von $\mathbb{F}_q$. Sei nun $\alpha \in \mathbb{F}_q$ ein primitives Element, d.h. $\mathbb{F}_q = \mathbb{F}_p$. Sei (entsprechend der allgemeinen Notation) $\mu_\alpha \in \mathbb{F}_p[X]$ das Minimalpolynom von α (über $\mathbb{F}_p[X]$). Dann gilt also $\mu_\alpha = (X - \alpha)(X - \sigma(\alpha)) \cdots (X - \sigma^{n-1}(\alpha)) \in \mathbb{F}_q[X]$. Nun gilt $\mu_\alpha(\alpha) = 0$. Damit gilt auch $\tau(\mu_\alpha(\alpha)) = 0$, und daraus folgt (mit dem obigen Lemma), dass $\mu_\alpha(\tau(\alpha)) = 0$. Damit ist $\tau(\alpha)$ eine der Nullstellen von μ_α, d.h. $\tau(\alpha) = \sigma^i(\alpha)$ für ein $i = 1, \ldots, e-1$. Damit ist die wesentliche Aussage gezeigt. Wir wissen nun, dass die Automorphismengruppe von $\mathrm{Aut}(\mathbb{F}_q))$ von σ erzeugt wird. Außerdem wissen wir schon, dass $\sigma^e = \mathrm{id}$ und $\sigma^{e-1} \neq \mathrm{id}$. □

Irreduzible Polynome über $\mathbb{F}_p$ und endliche Körper

Sei im Folgenden $\mathbb{F}_q$ ein fester Körper mit q Elementen, und sei wie immer $q = p^e$.

Sei $f \in \mathbb{F}_p[X]$ ein nicht-konstantes, irreduzibles, normiertes Polynom, und sei $d := \mathrm{Grad}(f)$. Wir fragen uns, ob f in $\mathbb{F}_q$ eine Nullstelle hat. Wir

setzen zunächst voraus, dass f eine Nullstelle α in $\mathbb{F}_q$ hat. Wir haben dann einen einen Körperhomomorphismus $\mathbb{F}_p[X]/(f) \hookrightarrow \mathbb{F}_q$, der durch $[X]_{(f)} \mapsto \alpha$ gegeben ist. Nun ist $[\mathbb{F}_p[X]/(f) : \mathbb{F}_p] = d$, und damit muss $d|e$ gelten.

Es gelte andererseits $d|e$. Nach Folgerung 2.17 enthält damit $\mathbb{F}_{p^e}$ einen Unterkörper mit p^d Elementen, und nach der Eindeutigkeit endlicher Körper ist dieser Körper ist isomorph zu $\mathbb{F}_p[X]/(f)$. Damit haben wir einen Körperhomomorphismus $\mathbb{F}_p[X]/(f) \hookrightarrow \mathbb{F}_q$. Nun hat f in $\mathbb{F}_q$ eine Nullstelle α, nämlich das Bild von $[X]_{(f)}$. Per Definition ist nun $\mathbb{F}_p[\alpha]$ der eindeutig bestimmte Unterkörper von $\mathbb{F}_q$ mit p^d Elementen. Wir wenden nun Satz 2.5 an und erhalten: Die Elemente $\alpha, \sigma(\alpha), \ldots, \sigma^{d-1}(\alpha)$ sind alle verschieden, und wir haben $f = \prod_{i=0}^{d-1}(X - \sigma^i(\alpha))$.

Wir haben also den folgenden Satz bewiesen:

Satz 2.6 *Sei $f \in \mathbb{F}_p[X]$ ein nicht-konstantes, irreduzibles, normiertes Polynom von Grad d. Dann hat f in $\mathbb{F}_q$ genau dann eine Nullstelle, wenn $d|e$. Wenn dies der Fall ist, und α eine solche Nullstelle ist, sind die Elemente $\alpha, \sigma(\alpha), \ldots, \sigma^{d-1}(\alpha)$ paarweise verschieden, und es gilt $f = \prod_{i=0}^{e-1}(X - \sigma^i(\alpha))$. Insbesondere zerfällt f in $\mathbb{F}_q$ also in ein Produkt von verschiedenden normierten Linearfaktoren.*

Man kann diesen Satz benutzen, um Aussagen über irreduzible Polynome über $\mathbb{F}_p$ zu erhalten.

Ich schildere zunächst die Idee: Sei zunächst ein irreduzibles Polynom von Grad $d|e$ gegeben. Dann kann man diesem Polynom die Menge der Nullstellen in $\mathbb{F}_q$ zuordnen. Diese Menge ist $\{\alpha, \sigma(\alpha), \ldots, \sigma^{d-1}(\alpha)\}$, wobei α eine Nullstelle ist. Sei anderseits $\alpha \in \mathbb{F}_q$. Dann ist α Nullstelle genau eines irreduziblen normierten Polynoms, nämlich des Minimalpolynoms. Wenn das Minimalpolynom μ_α Grad d hat, sind die Elemente $\alpha, \sigma(\alpha), \ldots, \sigma^{d-1}(\alpha)$ alle verschieden und haben dasselbe Minimalpolynom; wir wissen dass f genau die Nullstellen $\alpha, \sigma(\alpha), \ldots, \sigma^{e-1}(\alpha)$ hat. Jedes andere Element von $\mathbb{F}_q$ hat ein von μ_α verschiedenes Minimalpolynom.

Folgerung 2.23 *Es gilt $X^q - X$ ist das Produkt über alle normierten irreduziblen Polynome von $\mathbb{F}_p[X]$, deren Grad e teilt. Also*

$$X^q - X = \prod_{f \in \mathbb{F}_p[X] \text{normiert, irreduzibel}, \operatorname{Grad}(f)|e} f \, .$$

Beweis. Wir wissen schon, dass $X^q - X = \prod_{a \in \mathbb{F}_q}(X - a)$. Wir müssen also zeigen, dass das Produkt über alle normierten irreduziblen Polynome, deren Grad e teilt, auch gleich $\prod_{a \in \mathbb{F}_q}(X - a)$ ist. Wir wissen:
- Jedes normierte irreduzible Polynom von $\mathbb{F}_p[X]$ zerfällt in $\mathbb{F}_q[X]$ in ein

Produkt von verschiedenen normierten Linearfaktoren.
- Jedes Element von $\mathbb{F}_q$ ist Nullstelle genau eines normierten irreduziblen Polynoms (des Minimalpolynoms).

Die zweite Aussage bedeutet: Jedes normierte lineare Polynom $X-a$ teilt genau ein normiertes irreduzibles Polynom, und dieses teilt es genau einmal. Damit teilt $X-a$ also das Produkt $\prod_{f\in\mathbb{F}_p[X]\text{normiert, irreduzibel}} f$ genau einmal. Zusammen mit der ersten Aussage folgt daraus die Behauptung. □

Definition Sei für $n \in \mathbb{N}$ $\Phi_{p,n}$ das Produkt aller normierten irreduziblen Polynome in $\mathbb{F}_p[X]$. Sei $I_p(n)$ die Anzahl aller normierten irreduziblen Polynome von Grad n in $\mathbb{F}_p[X]$.

Folgerung 2.23 können wir auch mittels der Gleichung

$$X^q - X = \prod_{d|e} \Phi_{p,d} \tag{2.2}$$

formulieren.

Offensichtlich haben wir $n \cdot I_p(n) = \text{Grad}(\Phi_{p,n})$. Somit erhalten wir die Formeln

$$\sum_{d|e} d \cdot I_p(d) = p^e \ . \tag{2.3}$$

Mit diesen Formeln kann man alle $I_p(n)$ im Prinzip explizit berechnen.

Es ist $I_p(1) = p$. Sei nun n eine Primzahl. Dann ist $n \cdot I_p(n) + p = p^n$, also $I_p(n) = \frac{p^n-p}{n} = \frac{p^n}{n} \cdot (1 - \frac{1}{p^{n-1}})$.

Man kann zeigen:

Satz 2.7 *Es gilt stets*

$$\frac{1-2p^{-n/2}}{n} \leq \frac{I_p(n)}{p^n} \leq \frac{1}{n} \ .$$

Zum Vergleich: Es gibt genau p^n normierte irreduzible Polynome von Grad n in $\mathbb{F}_p[X]$. Damit ist die Zahl $\frac{I_p(n)}{p^n}$ die Wahrscheinlichkeit, dass ein uniform gewähltes normiertes Polynom von Grad n in $\mathbb{F}_p[X]$ irreduzibel ist.

Wir wissen schon, dass $I_p(2)$ immer positiv ist. Für $n \geq 3$ ist $2 \cdot p^{-n/2} \leq 2 \cdot p^{-3/2} \leq 2^{-1/2}$. Somit ist $I_p(n)$ stets positiv. Mit anderen Worten: Gegeben eine Primzahl p und ein $n \in \mathbb{N}$, gibt es immer ein irreduzibles Polynom von Grad n. Wenn f so ein Polynom ist, ist natürlich $\mathbb{F}_p[X]$ ein endlicher Körper mit p^n Elementen.

(Wir wussten eigentlich vorher schon, dass es immer so ein Polynom gibt. Warum?)

Außerdem sieht man:

Korollar 2.24 *Für festes p gilt* $\frac{\frac{I_p(n)}{p^n}}{\frac{1}{n}} \longrightarrow 1$ *für* $n \longrightarrow \infty$.

Bemerkung Beachten Sie, dass $I_p(n)$ eine Funktion ist, die von zwei Parametern abhängt: p und n.

Bemerkung Das obige Korollar kann man auch so ausdrücken: Für festes p und $n \longrightarrow \infty$ sind $\frac{I_p(n)}{p^n}$ und $\frac{1}{n}$ *asymptotisch äquivalent*. Man schreibt: $\frac{I_p(n)}{p^n} \sim \frac{1}{n}$ für $n \longrightarrow \infty$ und p fest.

Wir kommen zur Frage, wie man in endlichen Körpern explizit rechnet. Hierzu muss man zunächst die Elemente explizit (durch Bit-Strings) darstellen. Sagen wir, dass wir in $\mathbb{F}_q$ mit $q = p^e$ explizit rechnen wollen. Eine einfache Methode, um die Elemente explizit darzustellen, ist: Man wählt ein normiertes irreduzibles Polynom $f \in \mathbb{F}_p[X]$ von Grad e. Dann ist $\mathbb{F}_p[X]/(f)$ ein Körper mit q Elementen, also kann man $\mathbb{F}_q := \mathbb{F}_p[X]/(f)$ setzen. Sei (wie schon oft) $\alpha := [X]_{(f)}$. Dann ist $1, \alpha, \ldots, \alpha^{e-1}$ eine Basis über $\mathbb{F}_p$. Jedes Element von $\mathbb{F}_q$ kann dann durch den Koordinatenvektor bezüglich der Basis dargestellt werden.

Das Rechnen in $\mathbb{F}_q$ kann man dann wie folgt algorithmisch umsetzen: Addieren und Subtrahieren erfolgt wie in der Linearen Algebra. Für das Multiplizieren und Dividieren kann man den Euklidischen Algorithmus benutzen (siehe [LA]).

Es bleibt die Frage, wie man ein irreduzibles Polynom in $\mathbb{F}_p[X]$ eines vorgegebenen Grades findet.

Wir wissen bereits, dass ein uniform zufällig gewähltes Polynom von Grad n in $\mathbb{F}_p[X]$ mit ungefähr einer Wahrscheinlichkeit von ungefähr $\frac{1}{n}$ irreduzibel ist (siehe Satz 2.7).

Wir benötigen also nur einen effizienten Test auf Irreduzibilität, um ein geeignetes Polynom schnell zu finden. Es gibt einen solchen Test. Hier ist die Idee:

Sei zunächst $f \in \mathbb{F}_p[X]$ ein Polynom von Grad n. Sei $e < n$. Wir betrachten den ggT von f und $X^{p^e} - X$. Es gilt

$$\mathrm{ggT}(f, X^{p^e} - X) = \prod_{g \in \mathbb{F}_p[X] \text{ normiert, irreduzibel}, \mathrm{Grad}(g) | e, g | f} g \; .$$

Insbesondere ist also genau dann $\mathrm{ggT}(f, X^q - X) = 1$, wenn f keinen Teiler hat, dessen Grad ein Teiler von e ist.

Jedes reduzible (nicht irreduzible) Polynom von Grad n hat einen Teiler, der Grad $\leq n/2$ hat. Somit erhalten wir:

Folgerung 2.25 *Sei $f \in \mathbb{F}_p[X]$ mit $\mathrm{Grad}(f) = n$. Dann ist f genau dann irreduzibel, wenn für alle $e \in \mathbb{N}$ mit $e \leq n/2$ gilt:* $\mathrm{ggT}(f, X^{p^e} - X) = 1$.

Man kann also das Polynom f auf Irreduzibilität überprüfen, indem man für alle $e \leq n/2$ testet, ob $\mathrm{ggT}(f, X^{p^e} - X) = 1$ ist. Die Rechnung könnte man "wie immer" mit dem Euklidischen Algorithmus durchführen. Man würde allerdings auf jeden Fall eine Laufzeit bekommen, die exponentiell und auf keinen Fall polynomiell in der Eingabelänge ist. Das ist katastrophal! Ein effizienter Algorithmus sollte eine Laufzeit haben, die polynomiell in der Eingabelänge ist (d.h. polynomiell in $\log(q)$).

Beachten Sie aber: Es ist $\mathrm{ggT}(f, X^{p^e} - X) = \mathrm{ggT}(f, X^{p^e} - X \bmod f)$, wobei für ein Polynom $g \in \mathbb{F}_p[X]$ $g \bmod f$ das eindeutig bestimmte Polynom kleinsten Grades ist, welches kongruent zu g modulo f ist.

Wenn man $\mathrm{ggT}(f, X^{p^e} - X)$ berechnet, kann man also "modulo f" rechnen. Mit anderen Worten: Man kann in dem Ring $\mathbb{F}_p[X]/(f)$ rechnen. Man geht also so vor: Sei $x := [X]_{(f)}$. Zunächst berechnet man $y := x^q \in \mathbb{F}_p[X]/(f)$ mittels "Quadrieren und Multiplizieren".[2]

Sei $y = [h]_{(f)}$ mit $\mathrm{Grad}(h) < \mathrm{Grad}(f)$. (Hier ist nichts zu rechnen.) Nun ist also $X^q \equiv h \mod f$. Dann berechnet man $\mathrm{ggT}(h - x, f)$.

Man sieht leicht, dass die Laufzeit der Methode polynomiell in $\log(p) \cdot n$ ist.

Man kann dies auch noch ein wenig optimieren: Wenn $\mathrm{ggT}(f, X^{p^e} - X) = 1$ ist und $d|e$, dann ist auch $\mathrm{ggT}(f, X^{p^d} - X) = 1$. Damit braucht man den zweiten ggT nicht zu berechnen.

Beispiel 2.26 Sei $f \in \mathbb{F}_p[X]$ ein Polynom von Grad 13. Nach Folgerung 2.25 sollten wir $\mathrm{ggT}(X^{p^e} - X, f)$ für $e = 1, \ldots, 6$ berechnen. Wir müssen aber den ggT nur für $e = 4, 5, 6$ berechnen.

Explizites Rechnen in endlichen Körpern

Ich gebe noch einige Informationen zum Rechnen in endlichen Körpern.

Oben sind wir von einem irreduziblen Polynom $f \in \mathbb{F}_p[X]$ ausgegangen, und wir haben $\mathbb{F}_q := \mathbb{F}_p[X]/(f)$ definiert. Hierin haben wir dann die Basis $1, \alpha, \ldots, \alpha^{e-1}$, wobei $\alpha := [X]_{(f)}$ und $e := \mathrm{Grad}(f)$.

Eine Basis der Form $1, \alpha, \ldots, \alpha^{e-1}$ mit $\alpha \in \mathbb{F}_q$ nennt man eine *Polynomialbasis*.

[2] Sei G eine explizit gegebene endliche Gruppe (oder allgemeiner ein endliches Monoid), und sei $g \in G$, $n \in \mathbb{N}$ mit $n = \sum_{i=0}^{e} n_i 2^i$ mit $n_i \in \{0, 1\}$ und $n_e = 1$. Dann ist $g^n = g^{\sum_{i=0}^{e} n_i 2^i} = \prod_{i=0}^{e} (g^{2^i})^{n_i}$. Beachten Sie dabei: $g^{2^{i+1}} = (g^{2^i})^2$. Also kann man g^n berechnen, indem man g wiederholt quadriert und dann ein Produkt berechnet. Die Laufzeit beträgt offensichtlich $\mathcal{O}(e) = \mathcal{O}(\log(n))$ Gruppenoperationen (oder Monoidoperationen).

Mittels des Euklidischen Algorithmus kann man nun beweisen:

Folgerung 2.27 *Wenn man die Elemente des endlichen Körpers $\mathbb{F}_q$ mittels einer Polynomialbasis darstellt, kann die Arithmetik in $\mathbb{F}_q$ (die vier "Grundrechenarten") in einer Zeit von $\mathcal{O}(\log(q)^2)$ ausgeführt werden.*

Beachten Sie hierzu: Wenn man die Aussagen von [LA, §§1.9 und 1.10] zugrundelegt, erhält man eine Laufzeit von $\mathcal{O}(\log(q)^3)$. Aber man kann die Analyse des Euklidischen Algorithmus verbessern.

Man kann jede Basis von $\mathbb{F}_q|\mathbb{F}_p$ für explizite Arithmetik benutzen. Dies stelle ich in einem allgemeineren Kontext dar.

Sei $L|K$ irgendeine Körpererweiterung von Grad n. Sei $b_1, \ldots, b_n$ eine Basis von L "über K" (d.h. eine Basis von L als K-Vektorraum). Dann können wir alle Produkte $b_i b_j$ in der Basis entwickeln: $b_i b_j = \sum_{k=1}^{n} m_{i,j,k} b_k$ mit eindeutig bestimmten $m_{i,j,k} \in \mathbb{F}_p$. Diese $m_{i,j,k}$ kann man in einer "Multiplikationstabelle" $M = (m_{i,j,k})_{i,j,k=1,\ldots,n}$ zusammenfassen. Dies ist eine "dreidimensionale Matrix", d.h. ein Objekt in $\mathbb{F}_p^{n\times n\times n}$.

Wir gehen nun davon aus, dass wir in K explizit rechnen können, und wir stellen die Elemente von L durch ihre Koordinatenvektoren bezüglich der Basis $b_1, \ldots, b_n$ dar. Dann ist Addition und Subtraktion in L wieder einfach. Die Multiplikation von zwei Elementen kann man mittels der Multiplikationstabelle ausrechnen. Invertieren kann man, indem man ein lineares Gleichungssystem löst (wie geht das genau?)

Ich erwähne noch, dass es neben Polynomialbasen noch eine weitere wichtige Klasse von Basen endlicher Körper gibt. Man kann beweisen:

Satz 2.8 *Sei $q = p^e$. Dann gibt es ein Element $b \in \mathbb{F}_q$ so dass die Elemente $b, \sigma(b), \ldots, \sigma^{e-1}(b)$ eine Basis von $\mathbb{F}_q$ über $\mathbb{F}_p$ bilden.*

Eine Basis wie im Satz heißt *Normalbasis* von $\mathbb{F}_q$.

Endliche Körper der Charakteristik 2 sind besonders interessant für industrielle Anwendungen. (Koordinatenvektoren über $\mathbb{F}_2$ kann man in offensichtlicher Weise mit Bit-Strings identifizieren.) In diesen Fall ist $\sigma(a) = a^2$. Wenn nun $b \in \mathbb{F}_q$ eine Normalbasis definiert, haben wir für $a_0, \ldots, a_{n-1} \in \mathbb{F}_2$: $(\sum_{i=0}^{n-1} a_i \sigma(b))^2 = \sum_{i=0}^{n-1} a_i \sigma^{i+1}(b) = \sum_{i=1}^{n} a_{i-1} \sigma^i(b) = \sum_{i=0}^{n-1} a_{i-1} \sigma^i(b)$, wobei $a_{-1} := a_{n-1}$. Mit anderen Worten: Quadrieren entspricht einem "zyklischen Shift" auf dem Koordinatenvektor. Die Berechnung hiervon ist (fast) vernachläßigbar.

Aus diesem Grund sind Normalbasen recht populär. Aber Normalbasen haben auch Nachteile gegenüber Polynomialbasen: Man muss die Multiplikation mittels einer Multiplikationstabelle beschreiben, und Dividieren ist im Vergleich besonders langsam.

"Gute" Basen für konkrete Anwendungen zu finden war lange Zeit ein lebhaftes Forschungsthema. (Inzwischen scheint mir die Forschung im Wesentlichen abgeschlossen. Nach meinen Informationen haben sich Polynomialbasen durchgesetzt.)

Literatur Ich kann die beiden folgenden Bücher empfehlen.

E. Bach und J. Shallit. Algorithmic Number Theory (Kapitel 1 - 6)

N. Koblitz. A Course in Number Theory and Cryptography (Kapitel I und Abschnitt II.1)

In den empfohlenen Kapiteln werden auch einige Konzepte aus der Vorlesung Lineare Algebra wiederholt. Im Buch von Koblitz werden auch Anwendungen in der so genannten Public Key Kryptographie behandelt, zu der wir auch noch kommen werden.

2.3 Kodierungstheorie

Die Zielsetzung der Kodierungstheorie ist, Daten bei Übertragung und Speicherung vor Ausfällen oder Veränderungen zu schützen. Hierzu werden die Daten in einer gewissen redundanten Weise gesendet bzw. gespeichert.

Übrigens sind Kodierungstheorie und Kryptographie verschiedene Sachen. Beide kommen bei der Datenübertragung zum Einsatz. Aber bei Ersterem geht es darum, sicherzustellen, dass Daten unversehrt vor Übertragungsfehlern ankommen, bei Zweiterem geht es um Geheimhaltung und verwandte Themen.

Kodierungstheorie kommt zum Beispiel beim Mobilfunk, beim Internet (IP-Protokoll), bei CDs und DVDs und bei der Erforschung des Sonnensystems zum Einsatz.

Allgemeine Codes

Vom mathematischen Gesichtspunkt können Codes so definiert werden:

Definition Sei A eine endliche Menge. Ein *(Block-)Code* über A besteht aus natürlichen Zahlen $m < n$ und einer injektiven Abbildung $f : A^m \longrightarrow A^n$. Die Elemente aus A heißen dann *Symbole*, die Menge A heißt *Alphabet*. Die Elemente aus $f(A^m) \subseteq A^n$ heißen *Codewörter*. Die Zahl n heißt die *Länge* des Codes, die Zahl m heißt auch die *Dimension* des Codes. Die Elemente aus A^m kann man als *Nachrichten* bezeichnen.

Ganz formal ist ein Code ein Tupel (A, m, n, f), wobei A eine Menge, $m, n \in \mathbb{N}$ und $f : A^m \longrightarrow A^n$ eine injektive Abbildung ist. Wir fixieren einen solchen Code, den wir mit $\mathcal{C}$ bezeichnen.

Wir definieren auch noch:

Definition Sei $p : A^n \longrightarrow A^m$ die Projektion auf die ersten m Koordinaten. Der Code $\mathcal{C}$ heißt *systematisch*, wenn $p \circ f = \mathrm{id}_{A^m}$.

Mit anderen Worten: Ein Code heißt systematisch, wenn der Beginn eines Codeworts zu einer Nachricht die Nachricht enthält.

Bei der Datenübertragung wird zu einer Nachricht $\underline{x} \in A^m$ das Codewort $f(\underline{x})$ berechnet und dieses übertragen. Dabei können Fehler auftreten, d.h. möglicherweise wird ein Wort $\underline{y} \neq f(\underline{x})$ empfangen. Das (möglicherweise fehlerhafte) empfangene Wort $\underline{y}$ will man dann wieder decodieren, d.h. man will eine "passende" Nachricht $\underline{x}' \in A^m$ berechnen. Was dabei eine "passende" Nachricht in A^m ist, lassen wir zunächst weitgehend offen. Eine sinnvolle Forderung ist aber sicherlich: Wenn $f(\underline{x})$ fehlerfrei übertragen wurde, soll die Dekodierung $\underline{x}$ ergeben.

Bei Anwendungen tritt oftmals die folgende Situation ein: Es treten Übertragungsfehler auf, wobei man allerdings weiß, an welcher Stelle Verlust eingetreten ist. Um dies zu modellieren, fixieren wir ein weiteres Element ϵ, und wir betrachten $(A \cup \{\epsilon\})^n$ anstatt von A^n. Wir definieren:

Definition Eine *Dekodierungsfunktion (zur Fehlerkorrektur und für Datenausfall)* ist eine Abbildung $g : (A \cup \{\epsilon\})^n \longrightarrow A^m$ mit $g \circ f = \mathrm{id}_{A^m}$.

Es ist naheliegend, ein $\underline{y} \in (A \cup \{\epsilon\})^n$ zu einer Nachricht $\underline{x} \in A^m$ zu dekodieren, deren Codewort $f(\underline{x})$ mit $\underline{y}$ "am besten übereinstimmt". Hierzu definieren wir:

Definition Sei X eine beliebige Menge, und seien $\underline{x}, \underline{y} \in X^n$. Dann ist der *Hamming-Abstand* von x und y definiert als

$$d(\underline{x}, \underline{y}) := \#\{1 \leq i \leq n | x_i \neq y_i\} \ .$$

Lemma 2.28 *Die Funktion $d : X^n \longrightarrow \mathbb{N}_0$ erfüllt die folgenden Eigenschaften:*

a) $\forall \underline{x}, \underline{y} \in X^n : d(\underline{x}, \underline{y}) = 0 \longleftrightarrow \underline{x} = \underline{y}$

b) $\forall \underline{x}, \underline{y}, \underline{z} \in X^n : d(\underline{x}, \underline{z}) \leq d(\underline{x}, \underline{y}) + d(\underline{y}, \underline{z})$ *("Dreiecksungleichung").*

Für $\underline{y} \in (A \cup \{\epsilon\})^n$ sind die Codewörter, die intuitiv mit $\underline{y}$ am besten übereinstimmen gerade diejenigen Codewörter, die minimalen Hamming-Abstand zu $\underline{y}$ haben.

Definition Eine Dekodierungsfunktion $g : (A \cup \{\epsilon\})^n \longrightarrow A^m$ mit

$$d(\underline{y}, f(g(\underline{y}))) = \min_{\underline{x} \in A^m} d(\underline{y}, f(\underline{x}))$$

für alle $\underline{y} \in (A \cup \{\epsilon\})^n$ heißt eine Dekodierungsfunktion nach der *Hamming-Regel*.

Definition Die *Distanz* des Codes ist der minimale Hamming-Abstand zwischen zwei Codewörtern. Die Distanz wird mit $d(\mathcal{C})$ bezeichnet.

Es ist offensichtlich, dass für feste Länge und feste Dimension eine große Distanz ein Merkmal der Güte eines Codes ist.

Wir definieren noch:

Definition Sei $\underline{x} \in X^n$ (X eine beliebige Menge). Die *Hamming-Kugel* um $\underline{x}$ mit Radius $e \in \mathbb{N}$ ist

$$U_e(\underline{x}) := \{\underline{y} \in X^n \mid d(\underline{x}, \underline{y}) \leq e\} .$$

Lemma 2.29 *Sei $g : (A \cup \{\epsilon\})^m \longrightarrow A^n$ eine Dekodierungsfunktion nach der Hamming-Regel. Sei $\underline{x} \in A^m$, und sei $\underline{y} \in A^n$.*

a) Wenn $d(f(\underline{x}), \underline{y}) \leq \lfloor \frac{d(\mathcal{C})-1}{2} \rfloor$, dann gilt $g(\underline{y}) = \underline{x}$.

b) Wenn für alle $i = 1, \ldots, n$ $x_i = y_i$ oder $x_i = \epsilon$ gilt und $d(f(\underline{x}), \underline{y}) \leq d(\mathcal{C}) - 1$, dann gilt $\underline{x} = g(\underline{y})$.

Beweis.

zu a) Sei $\underline{x}' \in A^m, \underline{x}' \neq \underline{x}$. Dann gilt also $d(f(\underline{x}), f(\underline{x}')) \geq d(\mathcal{C})$. Damit gilt $d(f(\underline{x}'), \underline{y}) \geq d(f(\underline{x}'), f(\underline{x})) - d(f(\underline{x}), \underline{y}) \geq d(\mathcal{C}) - \lfloor \frac{d(\mathcal{C})-1}{2} \rfloor = \lceil \frac{d(\mathcal{C})+1}{2} \rceil > \lfloor \frac{d(\mathcal{C})-1}{2} \rfloor \geq d(f(\underline{x}), \underline{y})$.

zu b) Offensichtlich ist für alle $\underline{x}' \in A^m$

$$d(f(\underline{x}'), \underline{y}) \geq d(\mathcal{C}) > d(f(\underline{x}), \underline{y}) .$$

□

Bemerkung Intuitiv kann man das Lemma so formulieren: Treten bei der Übertragung höchstens $\lfloor \frac{d(\mathcal{C})-1}{2} \rfloor$ Fehler oder Ausfälle auf, so kann man sie mit der Hamming-Regel korrekt dekodieren. Treten bei der Übertragung höchstens $d(\mathcal{C}) - 1$ Ausfälle aus, so kann man sie mit der Hamming-Regel ebenfalls korrekt dekodieren.

Definition Ein Code $\mathcal{C}$ bestehend aus A, n, m, f wie oben mit $\#A = q$ und $d(\mathcal{C}) = d$ heißt q-närer (n, m, d)-*Code*. Ist $t \leq d-1$, so wird $\mathcal{C}$ auch t-*ausfalltoleranter Code* genannt. Ist $t \leq \lfloor \frac{d-1}{2} \rfloor$, so wird $\mathcal{C}$ auch t-*fehlerkorrigierender Code* genannt.

Definition Mit den obigen Daten heißt $r(\mathcal{C}) := \frac{m}{n}$ die *Rate* und $\delta(\mathcal{C}) := \frac{d}{n}$ die *relative Distanz* von $\mathcal{C}$.

Folgerung 2.30 *(Singleton-Schranke) Es gilt* $d(\mathcal{C}) \leq n - m + 1$.

Beweis. Sei $d := d(\mathcal{C})$. Die Idee ist: Wenn die letzten $d-1$ Symbole eines Codewortes einer Nachricht aus A^m nicht übertragen werden, wird das empfangene Wort trotzdem richtig dekodiert. Insbesondere muss es deshalb mindestens $\#A^m$ Wörter in A^{n-d+1} geben.

Hier ist ein formaler Beweis.

Sei

$$h : A^n \longrightarrow (A \cup \{\epsilon\})^n, \underline{y} \mapsto (y_1, \ldots, y_{n-d+1}\ ,\ \epsilon_{n-d+2}, \ldots, \epsilon_n)\ .$$

(Die letzten $d-1$ Einträge werden durch ϵ ersetzt ("fallen aus").) Dann gilt für $\underline{x} \in A^m$ $d(f(\underline{x}), h(f(\underline{x}))) \leq d-1$. Damit wird mit der Hamming-Regel $h(f(\underline{x}))$ zu $\underline{x}$ dekodiert. D.h. wenn f eine Dekodierungsfunktion nach der Hamming-Regel ist, gilt $g \circ h \circ f = \mathrm{id}_{A^m}$. Damit ist die Funktion $h \circ f$ injektiv. Andererseits enthält das Bild von h (also auch das Bild von $h \circ f$) höchstens $\#A^{n-m+1}$ Elemente. Dies bedeutet aber insbesondere, dass $m \leq n - d + 1$ bzw. $d \leq n - m + 1$. □

Lineare Codes

Definition Ein *linearer Code* besteht aus einem endlichen Körper K, zwei natürlichen Zahlen m, n und einer linearen Abbildung $f : K^m \longrightarrow K^n$.

Definition Die Abbildungsmatrix von f (bezüglich der Standardbasis) heißt *Generatormatrix* des Codes.

Bemerkung Sei U ein linearer Unterraum von K^n. Wenn man nun eine Basis $B = (b_1, \ldots, b_m)$ von U wählt, erhält man eine lineare Abbildung $f : K^m \longrightarrow K^n, \underline{x} \mapsto B\underline{x}$. (Ich identifiziere B mit der entsprechenden Matrix.) Damit erhält man also einen Code dessen Codewortmenge U ist, und B ist die Generatormatrix des Codes. Auf diese Weise enthält man auch alle Codes mit Codewortmenge U in K^n.

Sei nun C ein linearer Unterraum von K^n der Dimension m. Durch Umordnen der Zeilen können wir erreichen, dass die ersten m Zeilen linear unabhängig sind. Wir nehmen dies an. Dann gibt es eine Basis so dass die entsprechende Matrix die Form $\begin{pmatrix} I \\ M \end{pmatrix}$ mit einer Matrix $M \in K^{n-m \times m}$ hat. (Beides folgt mittels des Gauß-Algorithmus.) Dies bedeutet aber gerade, dass der entsprechende Code systematisch ist. Wir sehen also, dass wir bis auf Vertauschung der Reihenfolge der gesendeten Bits stets einen linearen Unterraum von K^n zum Raum der Codewörter eines systematischen Codes machen können.

Sei nun $f : K^m \longrightarrow K^n$ ein linearer Code. Sei $C := f(K^m)$.

Definition Für $\underline{y} \in K^n$ ist das *Gewicht* von $\underline{x}$ definiert als die Anzahl der Koordinaten von $\underline{y}$, die ungleich 0 sind.

Lemma 2.31 *Die Distanz $d(\mathcal{C})$ ist gleich dem minimalen Gewicht eines Codewortes $c \neq 0$:*

$$d(\mathcal{C}) = \min_{\underline{c} \in C} w(c)$$

Beweis. Es ist $w(\underline{y}) = d(\underline{y}, 0)$, $d(\underline{x}, \underline{y}) = w(\underline{x} - \underline{y})$ und für $c, d \in C$ ist $c + d \in C$. □

Reed-Solomon Codes

Reed-Solomon Codes sind spezielle lineare Codes, die auf dem "Auswerten" von Polynomen beruhen. Diese Codes werden wirklich in der Praxis verwendet. Sie wurden in einer Arbeit von 1960 vorgeschlagen.

Sei hierzu K ein endlicher Körper mit q Elementen, seien $m \leq n \in \mathbb{N}$, und seien $\alpha_1, \ldots, \alpha_n \in K$. Der Code ist nun so gegeben: Wir ordnen einem Vektor $\underline{a} \in K^m$ das Polynom $a(X) := \sum_{i=0}^{m-1} a_{i+1} X^i \in K[X]$ zu. Dieses Polynom "werten wir an den α_i aus". Wir erhalten den Vektor

$$\begin{pmatrix} a(\alpha_1) \\ \vdots \\ a(\alpha_n) \end{pmatrix} \in K^n \, .$$

Offensichtlich ist die Abbildung

$$f : K^m \longrightarrow K^n \, , \; \underline{a} \mapsto \begin{pmatrix} a(\alpha_1) \\ \vdots \\ a(\alpha_n) \end{pmatrix}$$

linear. Es ist $f(\underline{e}_i) = \begin{pmatrix} \alpha_1^{i-1} \\ \alpha_2^{i-1} \\ \vdots \\ \alpha_n^{i-1} \end{pmatrix}$, und somit ist die Generatormatrix gleich

$$\begin{pmatrix} 1 & \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^{m-1} \\ 1 & \alpha_2 & \alpha_2^2 & \cdots & \alpha_2^{m-1} \\ \vdots & & & \ddots & \vdots \\ 1 & \alpha_n & \alpha_n^2 & \cdots & \alpha_n^{m-1} \end{pmatrix}.$$

Beachten Sie: Für $m = n$ erhalten wir die sogenannte *Vandermondesche Matrix*. Wir wissen schon, dass diese Matrix invertierbar ist (siehe S. 128- 129 aus [LA]). Insbesondere hat die obige Matrix deshalb immer vollen Spaltenrang, d.h. die Abbildung f ist injektiv. Wir haben also einen Code.

Folgerung 2.32 *Ein Reed-Solomon Code der Länge n und Dimension m über $\mathbb{F}_q$ hat die Distanz $n - m + 1$.*

Mit anderen Worten: Reed-Solomon Codes erfüllen die Singleton-Schranke.

Beweis. Ich zeige, dass das Gewicht eines nicht-trivialen Codewortes immer mindestens $n - m + 1$ ist. Dann ist $d(\mathcal{C}) \geq n - m + 1$. Da auch $d(\mathcal{C}) \leq n - m + 1$ nach der Singleton-Schranke, gilt damit $d(\mathcal{C}) = n - m + 1$.

Sei $\underline{a} \in K^m$ mit $w(f(\underline{a})) < n - m + 1$, d.h. $w(f(\underline{a})) \leq n - m$. Dann enthält die Menge der $i = 1, \ldots, n$ mit $a(\alpha_i) \neq 0$ höchstens $n - m$ Elemente. Damit enthält die Menge der $i = 1, \ldots, n$ mit $a(\alpha_i) = 0$ also mindestens m Elemente. Somit hat also das Polynom $a(X)$ mindestens m Nullstellen. Da aber $\mathrm{Grad}(a(X)) \leq m - 1$, ist somit $a(X)$ das Nullpolynom. Damit ist auch $\underline{a} = 0$. □

Wir kommen zu der Frage, wie man effizient dekodiert. Hierbei sind zwei Fälle zu unterscheiden:

1. Man geht davon aus, dass nur Ausfälle aufgetreten sind (oder es gibt Übertragungsfehler, deren Position man kennt).

2. Es gibt möglicherweise auch Übertragungsfehler (deren Position man nicht kennt).

Dekodieren ohne Übertragungsfehler Wir behandeln zuerst den ersten Fall. Dieser Fall ist wesentlich leichter als der allgemeinere zweite Fall. Man geht wie folgt vor: Gegeben ist $\underline{b} \in (K \cup \{\epsilon\})^n$. Sei $I \subseteq \{1, \ldots, n\}$ so dass für alle $i \in \{1, \ldots, n\}$ gilt: $\underline{b}_i \in K$ genau dann wenn $i \in I$. (Die Idee ist, dass

$\underline{b}$ die übertragene Nachricht ist und an den Stellen, die nicht in I enthalten sind, Ausfälle stattgefunden haben.) Wir nehmen an, dass es höchstens $d-1$ Ausfälle gibt ($d = d(\mathcal{C})$), d.h. $n - \#I \leq d-1$. Unter dieser Voraussetzung gibt es höchstens eine Nachricht $\underline{a}$ mit $a_i = b_i$ für alle $i \in I$ (siehe Lemma 2.29). Wir wollen herausfinden, ob es so ein $\underline{a}$ gibt und gegebenenfalls $\underline{a}$ berechnen.

Das gesuchte $\underline{a}$ korrespondiert zu einem Polynom $a(X)$ mit $a(\alpha_i) = b_i$ für alle $i \in I$. So ein Polynom zu berechnen ist genau die Fragestellung der "Lagrange-Interpolation". Der Vektor $\underline{a}$ ist die Lösung des linearen Gleichungssystems, dessen Gleichungen

$$A_1 + \alpha_i A_2 + \alpha_i^2 A_3 + \cdots + \alpha_i^{m-1} A_m \stackrel{\circ}{=} b_i$$

für $i \in I$ lauten, wobei $A_1, \ldots, A_{m-1}$ die Unbestimmten sind. Beachten Sie hier: Die Anzahl der Gleichungen ist $\#I \geq n-d+1 = m$, und die Anzahl der Unbestimmten ist m. Außerdem enthält die Matrix des zugehörigen homogenen LGS eine $m \times m$-Vandermonde-Matrix als Untermatrix. Somit ist der Rang gleich m. Wir sehen wieder, dass es höchstens eine Lösung gibt (was wir ja schon wußten).

Man kann die Lösungsmenge (die leer ist oder aus einem Element besteht) mittels des Gauß-Algorithmus berechnen. Die Laufzeit ist hierbei polynomiell in $\log(q^n)$.

Exkurs: Interpolation Die obigen Ausführungen motivieren, sich intensiver mit dem Lagrange Interpolationsproblem zu beschäftigen.

Für die folgenden Ausführungen fixieren wir einen Körper K, der nicht endlich sein muss. Das Interpolationsproblem besteht nun in der folgenden algorithmischen Fragestellung: Gegeben $m < n \in \mathbb{N}$, paarweise verschiedene $\alpha_1, \ldots, \alpha_n \in K$ und $b_1, \ldots, b_n \in K$, gibt es ein Polynom $a(X)$ von Grad $\leq m$ mit $a(\alpha_i) = b_i$ für alle $i = 1, \ldots, n$? Wenn ja, berechne man ein solches Polynom!

Es gibt ein eindeutig bestimmtes Polynom $a(X)$ von Grad $\leq m-1$ mit $a(\alpha_i) = b_i$ für alle $i = 1, \ldots, m$. Nehmen wir zunächst an, dass wir dieses Polynom gefunden haben. Dann können wir leicht testen, ob $a(X)$ das Interpolationsproblem löst: Wir müssen nur testen, ob $a(\alpha_i) = b_i$ für $i = m+1, \ldots, n$. An dieser Stelle eine Bemerkung zur Auswertung von Polynomen: Es ist vorteilhaft, das sogenannte *Horner-Schema* zu benutzen: Hier wird $a(\alpha)$ (für $\alpha \in K$) nach der Formel

$$a(\alpha) = a_0 + \alpha(a_1 + \alpha(\alpha_2 + \ldots))$$

berechnet.

Wir müssen jetzt noch $a(X)$ berechnen. Das Polynom $a(X)$ kann nun wie oben beschrieben mittels eines LGS mit m Gleichungen und m Unbestimmten gelöst werden. Es gibt aber noch einen anderen Ansatz:

Definition Das i-te *Lagrange-Polynom* zu $\alpha_1, \ldots, \alpha_m$ ist das Polynom

$$\ell_i(X) := \prod_{j=1,\ldots,m, j\neq i} \frac{X-\alpha_j}{\alpha_i-\alpha_j} .$$

Beachten Sie: Es gilt $\ell_i(\alpha_j) = 0$ für alle $j \neq i$ und $\ell_i(\alpha_i) = 1$.
Somit lautet die eindeutige Lösung des Interpolationsproblems

$$\sum_{i=1}^{m} b_i \ell_i(X) .$$

Beachten Sie noch die folgende Interpretation der Lagrange-Polynome: Sei für $d \in \mathbb{N}$ $K[X]_{\leq d}$ die Menge der Polynome in $K[X]$ von Grad $\leq d$. Wir haben einen Isomorphismus von K-Vektorräumen

$$K[X]_{\leq m-1} \longrightarrow K^m \ , \ a(X) \mapsto \begin{pmatrix} a(\alpha_1) \\ a(\alpha_2) \\ \vdots \\ a(\alpha_m) \end{pmatrix}$$

Die Abbildungsmatrix dieses Isomorphismus bezüglich der Basis $1, X, \ldots, X^{m-1}$ einerseits und der Standardbasis andererseits ist die Vandermondesche Matrix. Sei M das Inverse dieser Matrix, d.h. M ist die Abbildungsmatrix der Umkehrabbildung bezüglich denselben Basen. Dann ist die j-te Spalte von M der Koordinatenvektor von $\ell_j(X)$ bez. $1, X, \ldots, X^{m-1}$.

Die Lagrange-Polynome kann man wie folgt schnell berechnen:
Zuerst berechnet man

$$u(X) := \prod_{j=1}^{m} (X-\alpha_j) .$$

Dann berechnet man für alle $i = 1, \ldots, m$ $u_i(X) := \frac{u(X)}{(X-\alpha_i)}$ mit Polynomdivison. Schließlich berechnet man $\ell_i(X) = \frac{u_i(X)}{u_i(\alpha_i)}$. Man benötigt dann $\mathcal{O}(m^2)$ Körperoperationen (in K), um alle Lagrange-Polynome zu berechnen.

In der Anwendung der Kodierungstheorie ist dies allerdings nicht relevant, da hier die α_i fest gewählt sind. Man kann deshalb alle Rechnungen, die nur die α_i involvieren, schon im vorhinein durchführen. Dann benötigt man nur $\mathcal{O}(n \cdot m)$ Körperoperationen, um zu testen, ob ein Polynom $a(X)$ mit $a(\alpha_i) = b_i$ für alle $i = 1, \ldots, n$ existiert und ggf. dieses Polynom zu berechnen.

Dekodieren mit Übertragungsfehlern Wir wenden uns der zweiten Aufgabe zu: Wir wollen dekodieren, wobei auch möglicherweise Übertragungsfehler aufgetreten sind (deren Position wir nicht kennen).

Zunächst die folgende allgemeine Folgerung, die eine Verallgemeinerung des Lagrange-Interpolationsproblems behandelt.

Folgerung 2.33 *Sei K ein beliebiger Körper, seien $n, s, t \in \mathbb{N}$ mit $n < s+t$ und $s \leq n$. Seien $\alpha_1, \ldots, \alpha_s \in K$ paarweise verschieden und $b_1, \ldots, b_n \in K$. Dann gibt es Polynome $g(X)$ und $h(X) \in K[X]$ mit $\mathrm{Grad}(g(X)) \leq s-1$ und $\mathrm{Grad}(h(X)) \leq t-1$ und $h(X) \neq 0$ so dass $g(\alpha_i) = h(\alpha_i)\, b_i$ für alle $i = 1, \ldots, n$.*

Beweis. Betrachte das homogene LGS mit den n Gleichungen

$$1 \cdot G_1 + \alpha_i \cdot G_2 + \cdots + \alpha_i^{s-1} \cdot G_s - b_i \cdot H_1 - b_i \cdot \alpha_i \cdot H_2 - \cdots - b_i \cdot \alpha_i^{t-1} \cdot H_t \stackrel{\circ}{=} 0 \ ,$$

wobei die G_i und H_i die Unbestimmten sind. Dieses LGS hat mehr Variablen $(s+t)$ als Gleichungen (n), also hat es eine nicht-triviale Lösung. Diese Lösung definiert zwei Polynome $g(X)$ und $h(X)$ mit $g(\alpha_i) = h(\alpha_i)\, b_i$ für alle $i = 1, \ldots, n$, wobei nicht beide Polynome $= 0$ sind. Wenn nun $h(X) = 0$ wäre, dann würde für alle $i = 1, \ldots, n$ $g(\alpha_i) = 0$ gelten. Da $\mathrm{Grad}(g(X)) \leq s-1 \leq n-1$, wäre dann auch $g(X) = 0$, also wären beide Polynome $= 0$.□

Folgerung 2.34 *Sei K wieder ein beliebiger Körper. Seien $m, k \in \mathbb{N}$, $\alpha_1, \ldots, \alpha_{m+k} \in K$ paarweise verschieden, und sei $a(X) \in K[X]$ mit $\mathrm{Grad}(a(X)) \leq m-1$. Seien nun $g(X), h(X) \in K[X]$ mit $\mathrm{Grad}(g(X)) \leq m-1+k$ und $\mathrm{Grad}(h(X)) \leq k$, $h(X) \neq 0$ und $g(\alpha_i) = h(\alpha_i)\, a(\alpha_i)$ für alle $i = 1, \ldots, m+k$. Dann gilt*

$$a(X) = \frac{g(X)}{h(X)} \ .$$

Beweis. Wir haben $(h(X)\, a(X) - g(X))(\alpha_i) = h(\alpha_i)a(\alpha_i) - g(\alpha_i) = 0$ für alle $i = 1, \ldots, m+k$. Das Polynom $h(X)\, a(X) - g(X)$ hat Grad $\leq m-1+k$. Somit gilt $ha - g = 0$. Dies impliziert $a(X) = \frac{g(X)}{h(X)}$. □

Aufgrund von Folgerung 2.33 und der obigen Folgerung haben wir den folgenden Dekodieralgorithmus. Der Algorithmus dekodiert richtig, wenn man höchstens $\lfloor \frac{d(\mathcal{C})-1}{2} \rfloor = \lfloor \frac{n-m}{2} \rfloor$ Übertragungsfehler (oder Ausfälle) hat (s.u.).

Dekodierungsalgorithmus

1. Setze $k \longleftarrow \lceil \frac{n-m}{2} \rceil$.

2. Berechne $g(X), h(X) \in \mathbb{F}_q[X]$ mit $\mathrm{Grad}(g(X)) \leq m-1+k$, $\mathrm{Grad}(h(X)) \leq k$ und $h(X) \neq 0$ mit $g(\alpha_i) = h(\alpha_i)\, b_i$ für $i = 1, \ldots, n$ (z.B. mit dem Gauß-Algorithmus).

3. Setze $a^*(X) \longleftarrow \frac{g(X)}{h(X)}$.

4. Wenn $a^*(X)$ ein Polynom ist, gib den Koordinatenvektor $\underline{a}^*$ von a^* bez. $1, X, \ldots, X^{m-1}$ aus. Andernfalls gib aus "Zu viele Fehler".

Zur Korrektheit des Algorithmus.

Sei $\underline{a}$ eine Nachricht, $\underline{b} \in K^n$ ein Wort mit $d(f(\underline{a}), \underline{b}) \leq \lfloor \frac{d-1}{2} \rfloor = \lfloor \frac{n-m}{2} \rfloor$.

Es gilt $(k+m)+k = m+2k \geq n$. Deshalb gibt es die Polynome im 2. Schritt. (Das gilt ohne Voraussetzung an $\underline{b}$.)

Zum 3. Schritt: Es gibt nach Voraussetzung mindestens $n - \lfloor \frac{d-1}{2} \rfloor = n - \lfloor \frac{n-m}{2} \rfloor = m + \lceil \frac{n-m}{2} \rceil = m+k$ fehlerfrei übertragene Symbole. Mit anderen Worten: Es gibt eine Menge $I \subseteq \{1, \ldots, n\}$ mit mindestens $m+k$ Elementen so dass für $i \in I$ $g(\alpha_i) = h(\alpha_i)\, a(\alpha_i)$ gilt. Nach Satz 2.34 angewendet auf die Stellen α_i mit $i \in I$, das Polynom $a(X)$ und die Polynome $g(X)$ und $h(X)$ gilt nun $\frac{g(X)}{h(X)} = a(X)$. Also gilt nach Definition von $a^*(X)$: $a^*(X) = a(X)$.□

Die Laufzeit des Algorithmus ist offensichtlich polynomiell in $\log(q) \cdot n$. Wir erhalten also den folgenden Satz.

Satz 2.9 *Mit dem obigen Algorithmus kann man bei einem Reed-Solomon Code über $\mathbb{F}_q$ der Länge n und Dimension n bis zu $\lfloor \frac{n-m}{2} \rfloor$ Fehler korrigieren. Die benötigte Zeit (in Bit-Operationen) ist hierbei polynomiell in $\log(q) \cdot n$.*

Literatur Ich kann die Wikipedia-Artikel zur Kodierungstheorie (insbesondere zu Reed-Solomon Codes) empfehlen (auf Deutsch und Englisch). Der obige Text beruht auf dem folgenden Skript:

J. Blömer: Algorithmische Codierungstheorie (Universität Paderborn)

Das Skript ist im deutschen Wikipedia-Artikel zu Reed-Solomon Codes verlinkt.

2.4 Kryptographie

Das Diffie-Hellman Protokoll

Anton will Berta eine Nachricht schicken, die kein Unbefugter lesen können soll. Nehmen wir an, dass Anton und Berta ein "klassisches" Chiffriersystem zur Verfügung haben: Unter Verwendung eines Schlüssels S verschlüsselt das System eine Nachricht, die wiederum mit S entschlüsselt werden kann. In vielen Fällen werden sich Anton und Berta persönlich kennen, so dass sie sich im Vorhinein auf einen Schlüssel einigen können. In anderen Fällen werden sie einer Organisation angehören, die allen Teilnehmern Schüssel aushändigt. Aber nehmen wir an, dass dies nicht der Fall ist, dass sich z.B. Anton und Berta gerade "übers Internet" kennen gelernt haben. Sie wollen nun ad hoc und "in der Öffentlichkeit" einen gemeinsamen Schlüssel vereinbaren. Geht das?

Es geht, und zwar z.B. mit dem so genannten *Diffie-Hellman Protokoll*. Das Protokoll funktioniert so:

Anton und Berta einigen sich zunächst auf eine Primpotenz q. Alle folgenden Rechnungen finden in $\mathbb{F}_q$ statt. Hierbei sei eine ("vernünftige") Darstellung der Elemente von $\mathbb{F}_q$ mittels Bitstrings fixiert. Nun einigen sie sich auch noch auf ein Element $g \in \mathbb{F}_q^*$. Sei ℓ die Ordnung von g (die Größe der von g erzeugten zyklischen Untergruppe von $\mathbb{F}_q$); wir gehen davon aus, dass ℓ bekannt ist.

Nun wählt Anton zufällig ein Element $x \in \{0, \ldots, \ell-1\}$, und Berta wählt zufällig ein Element $y \in \{0, \ldots, \ell-1\}$. Nun berechnet Anton $X := g^x \in \mathbb{F}_q$, und Berta berechnet $Y := g^y \in \mathbb{F}_q$.[3] Anton schickt X an Berta, und Berta schickt Y an Anton. (Aber Anton hält x geheim, und Berta hält y geheim.) Abschließend berechnet Anton Y^x, und Berta berechnet X^y. Beachten Sie: Es ist $Y^x = g^{xy} = X^y$, also haben beide das gleiche Element berechnet. Dies ist der gemeinsame Schlüssel für die weitere Kommunikation.

Man kann das Protokoll wie folgt symbolisieren:

[3] Diese Rechnung geschieht mit "Quadrieren und Multiplizieren"; siehe Fußnote 2.

q Primpotenz, $g \in \mathbb{F}_q^*, \ell := \mathrm{ord}(g)$ öffentlich

Anton	**Berta**
Wähle $x \in \{0, \ldots, \ell - 1\}$	Wähle $y \in \{0, \ldots, \ell - 1\}$
$\xrightarrow{X \leftarrow g^x}$	
	$\xleftarrow[Y \leftarrow g^y]{}$
$S_A \longleftarrow Y^x$	$S_B \longleftarrow X^y$

Offensichtlich ist $S_A = g^{xy} = S_B$.

Wie sicher ist dieses Protokoll? Nehmen wir an, dass ein "Angreifer" Emil die Konversation abhört. Er kennt dann q und $g, X = g^x, Y = g^y \in \mathbb{F}_q$, und er möchte $X^y = Y^x$ berechnen. Das entsprechende algorithmische Problem, nämlich: gegeben eine Primpotenz q, sowie $g, g^x, g^y \in \mathbb{F}_q^*$, berechne g^{xy}, heißt *Diffie-Hellman Problem*.

Offensichtlich kann Emil insbesondere dann den gemeinsamen Schlüssel berechnen, wenn er das folgende Problem in $\mathbb{F}_q^*$ bezüglich g lösen kann: Gegeben $X = g^x$, berechne $x \in \{0, \ldots, \ell - 1\}$!

Die Zahl $x \in \{0, \ldots, \ell - 1\}$ mit $g^x = h$ heißt der *diskrete Logarithmus von h zur Basis g*. Man schreibt auch $x = \log_g(h)$.[4] Eine klassische Terminologie für x ist auch *Index*.

Das algorithmische Problem "gegeben q und $g, g^x \in \mathbb{F}_q^*$ berechne $x \in \{0, \ldots, \mathrm{ord}(g) - 1\}$" heißt (klassisches) *diskretes Logarithmus Problem (DLP)*.

Beachten Sie auch: Wir haben den Gruppenisomorphismus

$$\mathbb{Z}/\,\mathrm{ord}(g)\mathbb{Z} \longrightarrow \langle g \rangle \ , \ [x] \mapsto g^x \ .$$

Unter der Identifikation von $\{0, \ldots, \mathrm{ord}(g) - 1\}$ mit $\mathbb{Z}/\,\mathrm{ord}(g)\mathbb{Z}$ ist $\log_g : \langle g \rangle \longrightarrow \mathbb{Z}/\,\mathrm{ord}(g)\mathbb{Z}$ die Umkehrabbildung dieses Isomorphismus.

Beachten Sie, dass die Sicherheit des Protokolls offensichtlich von ℓ, der Ordnung von g, abhängt: Wenn ℓ zu klein ist, kann Emil ein x mit $g^x = X$ per Ausprobieren finden. Unten werden wir noch effizientere "Angriffe" kennen lernen.

Es stellen sich somit insbesondere die folgenden Fragen:

[4] Verwechseln Sie bitte nicht den diskreten Logarithmus mit dem "normalen Logarithmus"!

1. Wie groß muss der Parameter q sowie die Ordnung von g in $\mathbb{F}_q^*$ sein, damit ein Angreifer nicht den geheimen Schlüssel aus q, g, X, Y berechnen kann?

2. Wie sollten dann x und y gewählt werden?

3. Kann ein *aktiver Angreifer* das Protokoll überlisten?

Auf Fragen 1. und 2. komme ich weiter unten zurück. Ich zeige jetzt, dass man das Protokoll tatsächlich überlisten kann.

Gehen wir davon aus, dass Emil nicht nur Nachrichten abhören sondern auch Konversationen manipulieren kann. Er kann sich nun "in die Mitte stellen" und sich gegenüber Anton als Berta und gegenüber Berta als Anton ausgeben. Hierzu fängt er alle Nachrichten von Anton an Berta ab und gibt sich gegenüber Anton als Berta aus. Er fängt auch alle Nachrichten von Berta an Anton ab und gibt sich gegenüber Berta als Anton aus. Er führt nun an Stelle von Berta den Schlüsselaustausch mit Anton durch und an Stelle von Anton den Schlüsselaustausch mit Berta. Am Ende hat er einen gemeinsamen Schlüssel S_A mit Anton und einen gemeinsamen Schlüssel S_B mit Berta. Im Allgemeinen ist hier $S_A \neq S_B$.

Später schickt vielleicht Anton eine Nachricht an Berta. Nun kann Emil die Nachricht mit S_A entschlüsseln, lesen und vielleicht verändern, mit S_B verschlüsseln und dann an Berta schicken. Beide merken nichts von der Manipulation. Analog kann er verfahren, wenn Berta eine Nachricht an Anton schickt.

Wenn man sich das Protokoll nochmal anschaut, ist klar, dass so ein Angriff möglich ist: In dem Protokoll hat Anton gar keine Möglichkeit zu überprüfen, ob die Person, die sich als Berta ausgibt, wirklich Berta ist. Bei diesem Protokoll könnte es auch passieren, dass sich von vorne herein eine andere Person als Berta ausgibt.

Um einen "Mann in der Mitte Angriff" zu verhindern, muss also zumindest die *Authentizität* des Kommunikationspartners sicher gestellt werden. Wie dies möglich ist, behandeln wir nicht. Ich möchte hier die Warnung abgeben, dass es nicht einfach ist, Protokolle zu konstruieren, die sicher gegenüber allen möglichen Angriffen sind. Dies ist ein aktives Forschungsgebiet, und immer wieder werden Protokolle, die zuvor als besonders sicher galten, angegriffen, und zwar ohne, dass das unterliegende mathematische Problem – in diesem Fall das Diffie-Hellman Problem – gelöst wird.

Ich komme zur 2. Frage: Wie sollten x und y gewählt werden? Eine offensichtlich Antwort heißt: "so zufällig wie möglich". Mit anderen Worten: Man sollte versuchen, x und y uniform zufällig zu wählen. Außerdem sollten

x und y *unabhängig* von vorherigen Wahlen gewählt werden. (Wenn man das Protokoll öfters benutzt.)

Diese Forderung in der Praxis zu erreichen, ist nicht einfach. Es gibt in der Praxis im Wesentlichen zwei Möglichkeiten, um x (und y) zu wählen:

Die erste Möglichkeit ist, einen Pseudozufallsgenerator zu verwenden. Dann werden x und y auf deterministische Weise gewählt, aber wenn "mehrere x" hintereinander gewählt werden, "sieht die Folge der x zufällig aus". Man muss dann sicherstellen, dass die Folge der x wirklich keine naheliegenden statistischen Eigenschaften hat, die man für einen Angriff ausnutzen könnte.

Die zweite Möglichkeit ist, Zufälligkeit aus der "Umgebung" zu benutzen. Dies kann von Mausklicks bis zu radioaktiver Strahlung reichen.

Man kann auch diese zwei Möglichkeiten kombinieren.

Die praktische Umsetzung des Protokolls sollte man von einer mathematischen Beschreibung unterscheiden, die man als idealisierte Beschreibung der praktischen Umsetzung ansehen kann. In der mathematischen Beschreibung kann man oben beschriebene Wahl von x (und analog y) so ausdrücken: x ist eine uniform verteilte Zufallsvariable mit Werten in $\{0, \ldots, \ell - 1\}$; wenn das Protokoll des Öfteren angewandt wird, ist die Folge der x eine Folge von uniform unabhängig verteilten Zufallsvariablen.

Das diskrete Logarithmusproblem

Zur Beantwortung der 1. Frage liegt es nahe, das diskrete Logarithmusproblem in den Gruppen $\mathbb{F}_q^*$ näher zu untersuchen. Wir untersuchen das Problem dabei *vom asymptotischen Standpunkt* aus und verwenden insbesondere die $\mathcal{O}$-Notation. Beachten Sie hier: Per Definition werden Laufzeiten also nur bis auf multiplikative Konstanten bestimmt. Insbesondere kann man aus einer Laufzeit in $\mathcal{O}$-Notation als solche keine konkreten Empfehlungen für Schlüssellängen ableiten. Andererseits ist es in der Regel schon so, dass Algorithmen, die asymptotisch schneller sind, auch zu schnelleren Rechnungen in der Praxis führen.

Im Folgenden sei q eine Primpotenz, $g \in \mathbb{F}_q^*$ und $h \in \langle g \rangle \subseteq \mathbb{F}_q^*$. Das Ziel ist, ein $x \in \{0, \ldots, q-1\}$ mit $g^x = h$ zu berechnen.

Genauer wollen wir Algorithmen studieren, die unter Eingabe von q, g, h wie oben sowie der Ordnung ℓ von g in $\mathbb{F}_q^*$ ein x wie oben berechnen.

Wie schon gesagt, haben wir das folgende offensichtliche Resultat: *Man kann ein x mit $g^x = h$ in $\mathcal{O}(\mathrm{ord}(g))$ Gruppenoperationen in $\mathbb{F}_q^*$ berechnen.*

Dies kann man stark verbessern:

Satz 2.10 *Gegeben eine Primpotenz q, $g \in \mathbb{F}_q^*$, die Ordnung ℓ von g und*

$h \in \langle g \rangle$, kann man ein $x \in \{0, 1, \ldots, \ell - 1\}$ mit $g^x = h$ in einer Zeit von $\mathcal{O}(\sqrt{\mathrm{ord}(g)} \cdot \log(q)^2)$ berechnen.

Dieses Resultat erhält man mittels des so genannten *Baby-Step-Giant-Step* Algorithmus, den ich jetzt beschreibe.

Sei $N := \lfloor \sqrt{\ell} + 1 \rfloor$. Dann ist $N^2 > \ell$. Somit gibt es $a, b \in \mathbb{N}_0$, $a, b < N$ mit

$$g^{a+bN} = h \; . \tag{2.4}$$

Die Zahl $x := a + bN \bmod \ell$ ist dann die Lösung des diskreten Logarithmusproblems. Gleichung (2.4) kann man auch so ausdrücken:

$$(g^N)^b = h \cdot (g^{-1})^a \tag{2.5}$$

Ein passendes Paar (a, b) kann man finden, indem man alle Potenzen $h \cdot (g^{-1})^i$ und $(g^N)^j$ für $i, j = 0, \ldots, N - 1$ konstruiert und in geeigneter Weise einen Sortieralgorithmus benutzt.

Der Algorithmus ist wie folgt:

Baby-Step-Giant-Step Algorithmus

Eingabe: q Primpotenz, $g \in \mathbb{F}_q^*, h \in \langle g \rangle, \ell := \mathrm{ord}(g)$.
Ausgabe: $x \in \{0, \ldots, \ell - 1\}$ mit $g^x = h$.

1. Setze $N \longleftarrow \lfloor \sqrt{\ell} + 1 \rfloor$.

2. ("Baby Steps")
   Konstruiere eine Tabelle T_1 wie folgt:
   Für $i = 0, \ldots, N - 1$:
   Speichere $(1, i, h \cdot ((g^{-1})^i)$ als den $i + 1$-sten Eintrag von T_1.

3. ("Giant Steps")
   Konstruiere eine Tabelle T_2 wie folgt:
   Für $j = 0, \ldots, N - 1$:
   Speichere $(2, j, (g^N)^j)$ als den $j + 1$-sten Eintrag von T_2.

4. Setze T_3 als die Konkatenation von T_1 und T_2.

5. Sortiere T_3 nach dem letzten Eintrag.

6. Gehe T_3 durch, bis ein Paar von Einträgen gefunden wird, das denselben letzten Eintrag aber verschiedene erste Einträge hat. Sei $(1, a, \alpha)$, $(2, b, \alpha)$ so ein Paar.

7. Gib $x \longleftarrow a + bN$ aus.

Zur Analyse: Die Tabellen können jeweils in $\mathcal{O}(N) = \mathcal{O}(\sqrt{\text{ord}(g)})$ Multiplikationen konstruiert werden, das sind $\mathcal{O}(\sqrt{\text{ord}(g)} \cdot \log(q)^2)$ Bitoperationen. Das Sortieren kann z.B. mit Merge-Sort mit $\mathcal{O}(N \cdot \log(N)) = \mathcal{O}(\sqrt{\text{ord}(g)} \cdot \log(\text{ord}(g)))$ Vergleichen und Kopier-Operationen geschehen, das sind $\mathcal{O}(\sqrt{\text{ord}(g)} \cdot \log(\text{ord}(g)) \cdot \log(q))$ Bitoperationen. Die anderen Schritte sind nicht dominant. □

Bemerkung. Schritt 6 kann man in Schritt 5 integrieren: Man beendet das Sortieren vorzeitig, wenn man ein geeignetes Paar gefunden hat.

Wir gehen im Folgenden davon aus, dass wir auch die *Faktorisierung der Ordnung von g* kennen. Unter dieser Bedingung zeige ich jetzt, dass das diskrete Logarithmusproblem bezüglich g nicht schwerer ist als das diskrete Logarithmusproblem in allen Untergruppen von $\langle g \rangle$ von *primer Ordnung.*

Sei wie immer $\ell := \text{ord}(g)$, und sei zunächst $\ell = \ell_1 \ell_2$, wobei ℓ_1 und ℓ_2 teilerfremd sind. Die Idee ist, den *Chinesischen Restsatz* auszunutzen, der besagt: Die von den Projektionen induzierte Abbildung $\mathbb{Z}/\ell_1\ell_2\mathbb{Z} \longrightarrow \mathbb{Z}/\ell_1\mathbb{Z} \times \mathbb{Z}/\ell_2\mathbb{Z}$ ist ein Isomorphismus. Hiernach können wir zuerst den diskreten Logarithmus "modulo ℓ_1" und "modulo ℓ_2" berechnen und dann das Ergebnis zusammensetzen.

Genauer: Nach dem Euklidischen Algorithmus haben wir dann ganze Zahlen a, b mit $a\ell_1 + b\ell_2 = 1$, die wir effizient berechnen können.

Beachten Sie, dass g^{ℓ_1} ein Element von Ordnung ℓ_2 und g^{ℓ_2} ein Element von Ordnung ℓ_1 ist. Außerdem gilt $h^{\ell_1} \in \langle g^{\ell_1} \rangle$, $h^{\ell_2} \in \langle g^{\ell_2} \rangle$.

Seien nun $x_1 \in \{0, \ldots, \ell_2 - 1\}$ und $x_2 \in \{0, \ldots, \ell_1 - 1\}$ die diskreten Logarithmen von h^{ℓ_1} bezüglich g^{ℓ_1} bzw. von h^{ℓ_2} bezüglich g^{ℓ_2}. Dann ist

$$g^{a\ell_1 x_1 + b\ell_2 x_2} = h^{a\ell_1 + b\ell_2} = h \; .$$

Somit ist $x := a\ell_1 x_1 + b\ell_2 x_2 \bmod \ell_1\ell_2$ der gesuchte diskrete Logarithmus von h bezüglich g. Wir haben also die Berechnung des diskreten Logarithmus von h bezüglich g auf die Berechnung der beiden diskreten Logarithmen x_1 und x_2 zurückgeführt.

Die soeben beschriebene Methode zur Berechnung des diskreten Logarithmus auf Basis des Chinesischen Restsatzes heißt auch *Pohlig-Hellman Algorithmus.*

Mit dieser Methode kann man die Berechnung diskreter Logarithmen bezüglich beliebiger "Basen" $g \in \mathbb{F}_q^*$ auf die Berechnung diskreter Logarithmen bezüglich Elementen reduzieren, deren Ordnung eine *Primpotenz* ist und die Ordnung von g teilt. Ich zeige jetzt, wie man dann auch auf diskrete

Logarithmusprobleme bezüglich Elementen von primer Ordnung reduzieren kann.

Sei $\mathrm{ord}(g) = p^e$, wobei p eine Primzahl und $e \in \mathbb{N}$ ist. Die Idee ist nun wie folgt: Sei x der diskrete Logarithmus von h bezüglich g, d.h. $g^x = h$. Dann können wir x im "p-Zahlensystem entwickeln": Es gibt eindeutig bestimmte $x_0, \ldots, x_{e-1} \in \{0, \ldots, p-1\}$ mit $x = x_0 + x_1 p + \cdots + x_{e-1} p^{e-1}$. Diese Zahlen x_i kann man iterativ (oder rekursiv) berechnen.

Beachten Sie, dass g^p die Ordnung p^{e-1} hat und $h^p \in \langle g^p \rangle$. Sei nun $y \in \{0, \ldots, p^{e-1} - 1\}$ mit $g^{yp} = h^p$. (Man sieht übrigens leicht, dass $y = x_0 + x_1 p + \cdots + x_{e-2} p^{e-2}$.) Dann gilt $(hg^{-y})^p = 1$. Damit liegt hg^{-y} in einer Untergruppe von $\langle g \rangle$, die p Elemente hat. Es gibt aber nur eine solche Untergruppe, nämlich $\langle g^{p^{e-1}} \rangle$. Somit gibt es ein $x_{e-1} \in \{0, \ldots, p-1\}$ mit $(g^{p^{e-1}})^{x_{e-1}} = hg^{-y}$. Dann gilt $g^{y + x_{e-1} p^{e-1}} = h$. Somit ist $x := y + x_{e-1} p^{e-1}$ der gesuchte diskrete Logarithmus.

Wir haben also die Berechnung des diskreten Logarithmus x auf die Berechnung der beiden diskreten Logarithmen y und x_{e-1} zurückgeführt. Die Berechnung von y kann rekursiv erfolgen. Man muss insgesamt (höchstens) e diskrete Logarithmen bezüglich Elementen der Ordnung p berechnen.

Insgesamt haben wir das diskrete Logarithmusproblem bezüglich beliebiger "Basen" $g \in \mathbb{F}_q^*$ in diskrete Logarithmusprobleme bezüglich Elementen reduziert, deren Ordnung *prim* ist und die Ordnung von g teilt.

Wir können also die folgende Forderung für das Diffie-Hellman Protokoll (sowie für jedes kryptographische Protokoll, dessen Sicherheit auf dem diskreten Logarithmusproblem beruht) aufstellen:

Die Ordnung der "Basis" g sollte eine Primzahl sein.

Es gibt übrigens effiziente Algorithmen, um zu testen, ob eine Zahl eine Primzahl ist.

Die bisher beschriebenen Methoden zum Lösen des diskreten Logarithmusproblems funktionieren nicht nur in $\mathbb{F}_q^*$ sondern in allen endlichen Gruppen. Man muss nur voraussetzen, dass die Gruppe explizit gegeben ist und die Gruppenoperationen algorithmisch ausgeführt werden können. Man spricht deshalb von *generischen Algorithmen*.

Index Calculus

Ich beschreibe jetzt die Idee der so genannten *Index Calculus Algorithmen*. Diese Algorithmen sind nicht generisch. Der Einfachheit halber beschränken wir uns hierbei auf Primkörper. Es gibt viele Varianten des sogleich beschriebenen Algorithmus.

Sei also p eine Primzahl, und seien $g, h \in \mathbb{F}_p^*$ mit $h \in \langle g \rangle$. Zunächst einige allgemeine Vorbemerkungen: Sei ℓ die Ordnung von g in $\mathbb{F}_p^*$. Dann induziert das Potenzieren

$$\mathbb{Z} \longrightarrow \langle g \rangle \ , \ e \mapsto g^e$$

einen Isomorphismus

$$\mathbb{Z}/\ell\mathbb{Z} \longrightarrow \langle g \rangle \ , \ [e]_\ell \mapsto g^e \ .$$

Allgemeiner gilt: Für eine natürliche Zahl m mit $\ell | m$ haben wir eine Surjektion

$$\mathbb{Z}/m\mathbb{Z} \longrightarrow \langle g \rangle \ , \ [e]_m \mapsto g^e \ .$$

Wir können somit definieren (für $\ell | m$):

$$g^{[e]_m} := g^e \ .$$

Im Algorithmus rechnen wir im Exponenten "modulo $p-1$, was ja auch ein Vielfaches von ℓ ist. Wir suchen also ein $\xi \in \mathbb{Z}/(p-1)\mathbb{Z}$ mit $a^\xi = b$.

Im Folgenden bezeichnen wir natürliche Zahlen mit großen Buchstaben und Restklassen modulo p mit kleinen Buchstaben. Nun wählt man zuerst eine so genannte *Glattheitsschranke* S und bestimmt alle Primzahlen $< S$ (mit dem Sieb des Eratosthenes). Seien $P_1, P_2, \ldots, P_k$ diese Primzahlen (die Menge $\{P_1, P_2, \ldots, P_k\}$ heißt dann *Faktorbasis*). Wie man die Schranke S optimal wählt, erfordert eine eingehende Analyse, die wir nicht durchführen. Es stellt sich dabei heraus, dass S ungefähr $e^{C \cdot (\log(p) \cdot \log\log(p))^{1/2}}$ für eine Konstante $C > 0$ sein sollte.

Nun versucht man, *Relationen* der Form

$$\prod_j [P_j]_p^{r_{i,j}} = g^{\alpha_i} \cdot h^{\beta_i} \tag{2.6}$$

mit $\alpha_i, \beta_i, r_{i,j} \in \mathbb{Z}/(p-1)\mathbb{Z}$ zu erzeugen. (Beachten Sie hier nochmal: Exponenten können wir "modulo $p-1$" nehmen.) Hierzu geht man wie folgt vor: Man wählt zufällig $\alpha_i, \beta_i \in \mathbb{Z}/(p-1)$ und berechnet die eindeutig bestimmte Zahl $N < p$ mit $[N]_p = g^{\alpha_i} \cdot h^{\beta_i}$ (mittels "Quadrieren und Mulitplizieren" modulo p). Nun testet man, ob N über der Faktorbasis "zerfällt", d.h. ob man man N als ein Produkt $\prod_j P_j^{e_{i,j}}$ mit gewissen Exponenten $e_{i,j}$ schreiben kann. Wenn dies der Fall ist, gilt (2.6) mit $r_{i,j} := [e_{i,j}]_{p-1}$. (Eine Zahl, die über der Faktorbasis zerfällt, heißt *glatt*.). Falls dies der Fall ist, berechnet man dann die Exponenten (d.h. man faktorisiert N). Diese Tests und das Faktorisieren können mit Ausprobieren und Probedivision durchgeführt werden.

Wenn man eine Relation wie oben gefunden hat, speichert man den Vektor $(r_{i,j})_j$ als die i-te Zeiler einer Matrix R ab. Nehmen wir an, wir haben auf diese Weise mehr als k Relationen gefunden (sagen wir $k+1$ Relationen). Dann kann mittels Linearer Algebra die Faktorbasiselemente eliminieren und so eine Relation zwischen g und h erhalten. Beachten Sie hierzu: R hat mehr Zeilen als Spalten, d.h. die Zeilen sind linear abhängig. Man berechnet nun ein $\underline{\gamma} \in (\mathbb{Z}/(p-1)\mathbb{Z})^{k+1}$ mit $\underline{\gamma} \neq 0$ und $\underline{\gamma}R = 0$. Ein Problem ist hier, dass (für $p > 3$) $\mathbb{Z}/(p-1)\mathbb{Z}$ kein Körper sondern nur ein Ring ist ($p-1$ ist gerade also insbesondere keine Primzahl). In einer gewissen Weise kann man aber trotzdem den Gauß-Algorithmus anwenden, um so ein $\underline{\gamma}$ zu berechnen. Beachten Sie hier, dass der Algorithmus hier auf Spalten angewandt wird. Es gilt dann gilt für alle j:

$$\sum_i \gamma_i r_{i,j} = 0 \ .$$

Dies impliziert:

$$g^{\sum_i \gamma_i \alpha_i} h^{\sum_i \gamma_i \beta_i} = \prod_i g^{\gamma_i \alpha_i} h^{\gamma_i \beta_i} = \prod_i (g^{\alpha_i} h^{\beta_i})^{\gamma_i} = \prod_i (\prod_j [P_j]_p^{r_{i,j}})^{\gamma_i} =$$
$$\prod_{i,j} [P_j]_p^{\gamma_i r_{i,j}} = \prod_j [P_j]_p^{\sum_i \gamma_i r_{i,j}} = 1 \ .$$

Es ist also $g^{-\sum_i \gamma_i \alpha_i} = h^{\sum_i \gamma_i \beta_i}$. Nehmen wir nun an, dass $\sum_i \gamma_i \beta_i \in \mathbb{Z}/(p-1)\mathbb{Z}$ invertierbar ist. Mit

$$\xi := -(\sum_i \gamma_i \alpha_i)\,(\sum_i \gamma_i \beta_i)^{-1} \in \mathbb{Z}/(p-1)\mathbb{Z}$$

gilt dann $g^\xi = h$.

Sei nun wie immer $\ell := \mathrm{ord}(g)$, und sei $[\xi]_\ell$ das Bild von ξ unter der Projektion $\mathbb{Z}/(p-1)\mathbb{Z} \longrightarrow \mathbb{Z}/\ell\mathbb{Z}$. Dann ist $x \in \{0, \ldots, \ell-1\}$ mit $[x]_\ell = [\xi]_\ell$ der gesuchte diskrete Logarithmus. (Wir gehen wie immer davon aus, dass ℓ bekannt ist; dann ist die Berechnung von x aus ξ kein Problem.)

Mit der Index Calculus Methode kann man beweisen:

Satz 2.11 *Man kann das diskrete Logarithmusproblem in endlichen Körpern $\mathbb{F}_q^*$, wobei q eine beliebige Primpotenz ist, in einer erwarteten Zeit von $e^{\mathcal{O}(1)\cdot(\log(q)\cdot\log(\log(q)))^{1/2}}$ lösen.*

RSA

Mit dem Diffie-Hellman Protokoll kann man keine Nachrichten übertragen sondern nur einen Schlüssel für die weitere Kommunikation festlegen. Außerdem gibt es noch das Problem der fehlenden Authentisierung.

Mit dem so genannten *RSA-Verfahren* kann man beide Ziele erreichen. Das Verfahren beruht auf der Schwierigkeit, große Produkte zweier ganzer Zahlen zu faktorisieren. In diesem Verfahren hat jeder Teilnehmer ein Paar von Schlüsseln: einen privaten und einen öffentlichen Schlüssel. Der private Schlüssel sollte dabei wirklich sicher aufbewahrt werden. Am besten man benutzt eine Chipkarte, auf der auch gleich der gesamte Algorithmus ausgeführt wird so dass der private Schlüssel die Karte nie verlässt.

Nehmen wir zunächst an, dass Anton Berta eine geheime Nachricht schicken will. (Beachten Sie: Eine geheime Nachricht kann auch ein Schlüßel für ein klassisches Chiffrierverfahren sein.) Nun wählt Berta zwei große Primzahlen p, q und bildet das Produkt $n := pq$. Alle folgenden Rechnungen finden in $\mathbb{Z}/n\mathbb{Z}$ statt, und die Nachrichten, die verschickt werden, fassen wir auch als Elemente von $\mathbb{Z}/n\mathbb{Z}$ auf.

Beachten Sie: Für $a \in \mathbb{Z}$ mit $a \equiv 1 \mod \varphi(n) = (p-1)(q-1)$ und $m \in \mathbb{Z}/n\mathbb{Z}$ gilt $m^a = m$. (Dies ist klar, wenn m invertierbar ist, aber ansonsten ist es auch richtig. Warum?)

Nun wählt Berta eine natürliche Zahl $d \in \{2, \ldots, \varphi(n)\}$, die teilerfremd zu $\varphi(n)$ ist. Dann gibt es eine natürliche Zahl $e \in \{3, \ldots, \varphi(n)\}$ mit $de \equiv 1 \mod \varphi(n)$, und d kann man mit dem Euklidischen Algorithmus effizient berechnen.

Berta macht nun n und d öffentlich und hält p, q und e geheim (p und q kann sie vergessen). Nun will Anton Berta eine Nachricht $m \in \mathbb{Z}/n\mathbb{Z}$ schicken. Hierzu berechnet er $c := m^d$ und schickt dies an Berta. Berta kann dies mit $c^e = m^{de} = m$ entschlüsseln.

Das Verfahren kann man auch zum Signieren benutzen. Nehmen wir an, dass Berta einen (langen) Text T signieren will. Dann bildet sie zunächst mit einer *Hash-Funktion* (z.B. MD5, SHA1) eine Prüfsumme m, die wir als Element von $\mathbb{Z}/n\mathbb{Z}$ auffassen. Nun berechnet sie m^e und hängt dies an ihren Text an. Nehmen wir an, dass Anton wissen will, ob Berta den T signiert hat. Sei hierzu s die angehängte (angebliche) Signatur. Dann berechnet er auch die Prüfsumme m. Außerdem berechnet er s^d. Wenn wirklich $s = m^e$ ist, ist $s^d = m$. Ansonsten kommt etwas anderes heraus.

Mittels Signaturen kann man auch das Problem der Authentisierung lösen: Anton weist sich gegenüber einer "vertrauenswürdigen Instanz" aus, und diese Instanz signiert dann ihren öffentlichen Schlüssel.

Auch Mann-in-der-Mitte Angriffe lassen sich mit Signaturen umgehen.

Sicher kann man das RSA-Verfahren brechen, wenn man die Zahl $n = p \cdot q$ faktorisieren kann. Es gibt Faktorisierungsalgorithmen, die relativ ähnlich zu Index Calculus Algorithmen sind. Man kann dann beweisen:

Satz 2.12 *Eine Zahl n, die ein Produkt von zwei Primzahlen ist, kann man in einer erwarteten Zeit von $e^{\mathcal{O}(1)\cdot(\log(n)\cdot\log(\log(n)))^{1/2}}$ faktorisieren.*

Dieses Resultat ist natürlich auch für den Pohlig-Hellman Algorithmus relevant.

Elliptische Kurven Kryptographie

Wir haben gesehen, dass man das diskrete Logarithmusproblem in endlichen Körpern mittels Index Calculus angreifen kann, und ähnliche Methoden funktionieren auch für das Faktorisierungsproblem. Es stellt sich deshalb die folgende Frage: Gibt es Familien (explizit gegebener) endlicher Gruppen, für die nur generische Algorithmen zum Berechnen von diskreten Logarithmen bekannt sind? In der Tat gibt es solche Gruppen, nämlich (geeignet gewählte) *elliptische Kurven über endlichen Körpern*. Ich gehe hier nicht näher darauf ein. Es sei nur angemerkt, dass dies zur so genannten *Elliptischen Kurven Kryptographie (ECC)* führt. Diese Form der Kryptographie erlaubt wesentlich kürzere Schlüssellängen als Systeme, die auf dem klassischen diskreten Logarithmusproblem oder dem RSA-Verfahren beruhen. Im Allgemeinen wird z.B. davon ausgegangen, dass eine Schlüssellänge von 1024 bit RSA einer Schlüssellänge von etwa 160 bit ECC entspricht. Mit größeren Schlüssellängen werden die Vorteile von ECC immer deutlicher.

Die Vorteile von ECC auch durch die aktuellen Rekorde für Angriffe auf RSA und ECC deutlich: Der aktuelle Rekord für erfolgreiche Angriffe auf RSA liegt bei 664 bit, und der entsprechende Rekord für ECC liegt bei (nur) 97 bit.

Historische Anmerkungen

Die hier vorgestellten Protokolle fallen in das Gebiet der "Public-Key Kryptographie". Eine treibende Kraft hinter der Entwicklung war Ralf Merkle. Aufbauend auf seinen Ideen publizierten Whitfried Diffie und Martin Hellman das nach ihnen benannte Protokoll im Jahr 1976. Das RSA-Verfahren ist nach seinen Erfindern benannt: Ron Rivest, Adi Shamir und Leonard Adleman, die das Verfahren 1977 vorstellten.

Es ist inzwischen bekannt geworden, dass diese Verfahren schon vorher im britischen Geheimdienst DCHQ entwickelt wurden. Die Idee der Public-Key Kryptographie wurde von James Ellis ab 1970 propagiert. Chiffort Coocks

entwickelte im Wesentlichen RSA im Jahr 1973, und Malcolm Williams entwickelte kurz darauf das Diffie-Hellman Protokoll.

Literatur Ich empfehle noch einmal das Buch von N. Koblitz (siehe die Literatur zu endlichen Körpern)

Die weitere Vorlesung

Das Skript endet hier. In der Vorlesung wird noch das Thema "Graphen" behandelt. Hierbei folgen wir dem entsprechenden Kapitel aus dem Buch:

A. Steger. Diskrete Strukturen 1 (Signatur SK 110 S817 in der Bibliothek)