Freie algebraische Strukturen

Hartmut Laue

Mathematisches Seminar der Universität Kiel 2013

Inhaltsverzeichnis

Relief von der Insel Paros (ca. 570-560 v. Chr.), zeigt wahrscheinlich die drei Chariten Aglaia, Thalia, Euphrosyne. Glyptothek München, urheberrechtlich gestatteter Nachdruck.

Kapitel 1

Worte

Sind X, Y Mengen, so bedeute X^Y die Menge der Abbildungen von Y in X. Speziell besteht dann $X^\emptyset$ aus der „leeren Abbildung“, d.h. es gilt: $X^\emptyset = \{\emptyset\}$. Ein besonders wichtiger Fall ist der, in dem Y ein Anfangsstück von $\mathbb{N}(= \{1, 2, 3, 4, \dots\})$ ist: Für alle $n \in \mathbb{Z}$ setzen wir

$$\underline{n} := \{k | k \in \mathbb{N}, 1 \leq k \leq n\},$$

insbesondere gilt also $\underline{0} = \emptyset$. Die Elemente von $X^{\underline{n}}$ heißen n-Tupel über X. Mit $(x_1, \dots, x_n)$ wird dasjenige n-Tupel bezeichnet, das jedes $j \in \underline{n}$ auf x_j abbildet; das letztere Element heißt die j-te Komponente des n-Tupels. Wir setzen $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. Vermöge Induktion sieht man leicht:

1.0.1 *Für endliche Mengen X, Y gilt stets $|X^Y| = |X|^{|Y|}$, insbesondere $|X^{\underline{n}}| = |X|^n$ für alle $n \in \mathbb{N}_0$.* □

1.1 Definition Für jede Menge X setzen wir $\mathcal{T}(X) := \bigcup_{n \in \mathbb{N}_0} X^{\underline{n}}$ und nennen die Elemente von $\mathcal{T}(X)$ Tupel über X.

1.1.1 *Die Vereinigung $\bigcup_{n \in \mathbb{N}_0} X^{\underline{n}}$ ist disjunkt,*

denn für jedes $f \in \mathcal{T}(X)$ gibt es ein eindeutig bestimmtes $n \in \mathbb{N}_0$, so daß $\underline{n}$ der Definitionsbereich von f ist. Also liegt f in genau einer der Mengen $X^{\underline{n}}$. □

Für $f \in X^{\underline{n}}$ gilt (gemäß der Definition des Funktionsbegriffs) :

$$|f| = |\{(1, 1f), \dots, (n, nf)\}| = n.$$

Die Mächtigkeit n der Menge f wird aufgrund der üblichen Tupelschreibweise von f (s.o.) die Länge von f genannt und in natürlicher Weise mit $|f|$ bezeichnet. Auf der Menge $\mathcal{T}(X)$ definieren wir eine Verknüpfung (üblicherweise als Konkatenation bezeichnet) durch

$$(x_1, \ldots, x_n)(x'_1, \ldots, x'_m) := (x_1, \ldots, x_n, x'_1, \ldots, x'_m)$$

für alle m, $n \in \mathbb{N}_0$, $x_1 \ldots, x_n, x'_1, \ldots, x'_m \in X$.

1.2 Proposition *Sei X eine Menge.*

(1) *$\mathcal{T}(X)$ ist vermöge der Konkatenation ein Monoid, $\emptyset$ neutral.*

(2) *Ist U ein Untermonoid von $\mathcal{T}(X)$ mit $X^{\underline{1}} \subseteq U$, so ist $U = \mathcal{T}(X)$.*

(3) *Ist φ irgendeine Abbildung von $X^{\underline{1}}$ in ein Monoid M, so gibt es genau einen Monoid-Homomorphismus [1] $\bar{\varphi}$ von $\mathcal{T}(X)$ in M mit $\bar{\varphi}|_{X^{\underline{1}}} = \varphi$.*

Beweis. Seien f, g, $h \in \mathcal{T}(X)$, und seien n, m, $k \in \mathbb{N}_0$, $x_1, \ldots, x_n$, $x'_1, \ldots, x'_m$, $x''_1, \ldots, x''_k \in X$ mit $f = (x_1, \ldots, x_n)$, $g = (x'_1, \ldots, x'_m)$, $h = (x''_1, \ldots, x''_k)$.

(1) Es gilt:

$$\begin{aligned}(fg)h &= (x_1, \ldots, x_n, x'_1, \ldots, x'_m)(x''_1, \ldots, x''_k)\\ &= (x_1, \ldots, x_n, x'_1, \ldots, x'_m, x''_1, \ldots, x''_k)\\ &= (x_1, \ldots, x_n)(x'_1, \ldots, x'_m, x''_1, \ldots, x''_k)\\ &= f(gh),\end{aligned}$$

$f\emptyset = (x_1, \ldots, x_n) = \emptyset f$.

(2) Ist $X^{\underline{1}} \subseteq U$, so folgt durch Induktion nach n:[2]

$$X^{\underline{n}} = (X^{\underline{1}})^n \subseteq U \text{ für alle } n \in \mathbb{N}_0,$$

also $\mathcal{T}(X) \subseteq U$, d.h. $U = \mathcal{T}(X)$.

(3) Setzen wir $f\bar{\varphi} := (x_1)\varphi \ldots (x_n)\varphi$, so ist speziell $\emptyset\bar{\varphi}$ das leere Produkt in M, also gleich dem neutralen Element 1_M von M, und $(x)\bar{\varphi} = (x)\varphi$ für alle $x \in X$. Wegen

$$\begin{aligned}(fg)\bar{\varphi} &= (x_1, \ldots, x_n, x'_1, \ldots, x'_m)\bar{\varphi}\\ &= (x_1)\varphi \ldots (x_n)\varphi(x'_1)\varphi \ldots (x'_n)\varphi = f\bar{\varphi} \cdot g\bar{\varphi}\end{aligned}$$

[1] Ein Monoid-Homomorphismus ist eine verknüpfungstreue Abbildung eines Monoids in ein Monoid, die das neutrale Element des Urbild-Monoids auf das neutrale Element des Ziel-Monoids abbildet.

[2] Ist T eine Teilmenge eines Monoids, so bezeichnet T^n die Menge aller Produkte aus n in T liegenden Faktoren.

ist $\bar{\varphi}$, wie behauptet, eine Fortsetzung von φ zu einem Monoid-Homomorphismus von $\mathcal{T}(X)$ in M. Ist auch ψ eine solche, so gilt $f\psi = ((x_1)\dots(x_n))\psi \underset{\psi \text{ Hom.}}{=} (x_1)\varphi\dots(x_n)\varphi = f\bar{\varphi}$, also $\psi = \bar{\varphi}$. □

1.3 Definition Sei X eine Menge und N ein Monoid mit $X \subseteq N$. X heißt ein [Monoid-]Erzeugendensystem von N, wenn gilt: Ist N_0 ein Untermonoid von N mit $X \subseteq N_0$, so ist $N_0 = N$, d.h. wenn gilt:

$$\bigcap_{X \subseteq N_0 \leq N} N_0 = N.$$

(Hier bedeutet $\leq$: „ist Untermonoid von".) Für alle $n \in \mathbb{N}_0$ gilt offensichtlich $X^n \leq N$.

1.3.1 *X ist genau dann ein Erzeugendensystem von N, wenn* $\bigcup_{n\in\mathbb{N}_0} X^n = N$.

Denn $\bigcup_{n\in\mathbb{N}_0} X^n$ ist ein Untermonoid von N, das X enthält, und es liegt in jedem X enthaltenden Untermonoid von N. □

X heißt unabhängig, wenn gilt: Ist φ irgendeine Abbildung von X in ein Monoid M, so gibt es genau einen Homomorphismus $\bar{\varphi}$ von $\bigcap_{X\subseteq N_0\leq N} N_0$ in M mit $\bar{\varphi}|_X = \varphi$. Ist X ein unabhängiges Erzeugendensystem von N, so heißt N von X [als Monoid] frei erzeugt, kurz: frei über X. Ein Monoid heißt frei, wenn es eine Teilmenge besitzt, von der es frei erzeugt wird.

1.3.2 *Ist N frei über X, so gilt $1_N \notin X$.*

Gilt nämlich $1_N \in X$, so setzen wir $\varphi : X \to \mathbb{N}$, $x \mapsto \begin{cases} 1 & \text{für } x \neq 1_N \\ 2 & \text{für } x = 1_N \end{cases}$.
Dann gibt es keinen Homomorphismus $\bar{\varphi}$ von N in das multiplikative Monoid der natürlichen Zahlen mit $\bar{\varphi}|_X = \varphi$, da für jeden solchen $1_N\bar{\varphi} = 1_{\mathbb{N}}$ gelten muß. □

Eine Umformulierung von 1.2 ist

1.2' Für jede Menge X ist $\mathcal{T}(X)$ ein von $X^{\underline{1}}$ frei erzeugtes Monoid.

Streng genommen, ist natürlich X von $X^{\underline{1}}$ deutlich zu unterscheiden. Wir beleuchten im folgenden die beiden naheliegenden Fragen: Gibt es auch ein von X selbst frei erzeugtes Monoid? Wie hängen zwei von X frei erzeugte Monoide zusammen? Die letztere können wir sofort beantworten:

1.4 Proposition *Seien X, X' Mengen, und seien N bzw. N' von X bzw. X' frei erzeugte Monoide. Es gebe eine Bijektion φ von X auf X'. Dann gibt es einen eindeutig bestimmten Isomorphismus $\bar{\varphi}$ von N auf N' mit $\bar{\varphi}|_X = \varphi$.*

Folgerung. *Sind N, N' von X frei erzeugte Monoide, so gibt es einen eindeutig bestimmten Isomorphismus ψ von N auf N' mit $x\psi = x$ für alle $x \in X$.*

Beweis. Da N frei über X ist, gibt es eine Fortsetzung von φ zu einem Monoid-Homomorphismus $\bar{\varphi}$ von N in N'. Da N' frei über X' ist, läßt sich auch φ^{-1} zu einem Monoid-Homomorphismus $\overline{\varphi^{-1}}$ von N' in N fortsetzen. Für alle $x \in X$ gilt: $(x\varphi)\overline{\varphi^{-1}}\bar{\varphi} = x\bar{\varphi} = x\varphi$ und $x\bar{\varphi}\overline{\varphi^{-1}} = (x\varphi)\overline{\varphi^{-1}} = x$. Es folgt: $\overline{\varphi^{-1}}\bar{\varphi} = id_{N'}$, $\bar{\varphi}\overline{\varphi^{-1}} = id_N$, denn es handelt sich in beiden Fällen um Monoid-Endomorphismen, die ein Erzeugendensystem (nämlich X' bzw. X) elementweise festlassen. Also sind $\bar{\varphi}$, $\overline{\varphi^{-1}}$ zueinander inverse Isomorphismen: Es gilt $\bar{\varphi}^{-1} = \overline{\varphi^{-1}}$. – Die Folgerung ist der Spezialfall $X = X'$, $\varphi = id_X$. $\square$

Bis auf Isomorphie (in einem sehr strengen Sinne, nämlich sog. „X-Isomorphie“) gibt es also *höchstens* ein über X freies Monoid. Wie aber sieht es mit der Existenz aus? In 1.2 haben wir die Existenz eines über $X^{\underline{1}}$ freien Monoids nachgewiesen. Weiter gibt es eine kanonische Bijektion von X auf $X^{\underline{1}}$ (nämlich $x \mapsto (x)(= \{(1, x)\})$). Eine unscharfe, aber (leider) übliche Sprechweise läßt nun an dieser Stelle X mit X' „identifizieren“ und damit dann $\mathcal{T}(X)$ als über X freies Monoid „ansehen“. Diese – ohne nähere Erklärung durchaus mysteriösen – Ausdrucksweisen lassen sich aber vollkommen präzise fassen, was im folgenden geschehen soll. Hinter dem „Identifizieren“ steht der folgende rein mengentheoretische Satz, dessen (nicht tiefiegenden) Beweis wir hier nicht ausführen wollen:

Entgiftungssatz. *Seien A, X Mengen. Dann gibt es eine zu A gleichmächtige Menge A' mit $A' \cap X = \emptyset$.*

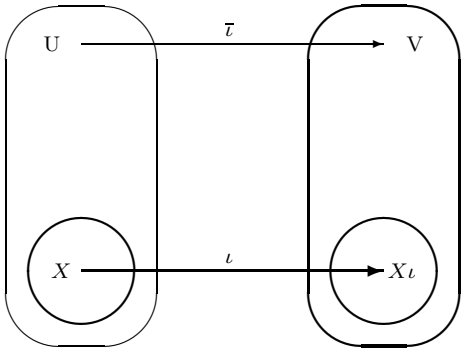
Als Folgerung erhalten wir das wichtige

Erweiterungsprinzip. *Seien X, V Mengen und ι eine injektive Abbildung von X in V. Dann gibt es eine Menge U mit $X \subseteq U$ und eine Bijektion $\bar{\iota}$ von U auf V mit $\bar{\iota}|_X = \iota$.*

Zusatz. *Ist $\bullet$ eine Verknüpfung auf V, so definiert die Setzung*

$$u \cdot u' := ((u\bar{\iota}) \bullet (u'\bar{\iota}))\bar{\iota}^{-1} \text{ für alle } u, u' \in U$$

eine Verknüpfung auf U, so daß $\bar{\iota}$ ein Isomorphismus von $(U, \cdot)$ auf $(V, \bullet)$ ist.



Beweis des Erweiterungsprinzips als Folgerung des Entgiftungssatzes: Sei $A := V \setminus X\iota$ und (nach dem Entgiftungssatz) A' eine zu A gleichmächtige Menge mit $A' \cap X = \emptyset$. Sei f eine Bijektion von A' auf A, $U := X \cup A'$ und $\bar{\iota} := \iota \cup f$. Die Behauptung des Erweiterungsprinzip folgt dann unmittelbar. Sind u, $u' \in U$, so folgt

$$(u \cdot u')\bar{\iota} = ((u\bar{\iota}) \bullet (u'\bar{\iota}))\bar{\iota}^{-1}\bar{\iota} = (u\bar{\iota}) \bullet (u'\bar{\iota}),$$

also gilt der Zusatz. □

Anwendung: *Zu jeder Menge X gibt es ein von X frei erzeugtes Monoid.*

Denn nach 1.2' gibt es ein von $X^{\underline{1}}$ frei erzeugtes Monoid, nämlich $\mathcal{T}(X)$. Sei ι die Injektion $X \to \mathcal{T}(X)$, $x \mapsto (x)$. Das Erweiterungsprinzip liefert unmittelbar die Behauptung. Die Fortsetzung $\bar{\iota}$ von ι bildet dabei die Menge X^n der Produkte der Länge n über X auf $X^{\underline{n}}$ ab. Ähnlich wie oben werden deswegen auch hier die Mengen X^n, $X^{\underline{n}}$ (meist ohne genaue Erläuterung des Wortsinnes) vielerorts „identifiziert“. Wegen der Eindeutigkeitsaussage aus der Folgerung zu 1.4 spricht man – unter Mißbrauch des bestimmten Artikels – von *dem* freien Monoid über X und bezeichnet es mit X^*. Wir schreiben $\imath$ für das neutrale Element von X^* und setzen $X^+ := X^* \setminus \{\imath\}$. Das Verknüpfungszeichen lassen wir, wie allgemein bei multiplikativ geschriebenen Verknüpfungen üblich, bei Produkten in X^* fort, wenn keine Verwechslungen zu befürchten sind. Gelegentlich hat die Juxtaposition von Elementen von X jedoch schon, durch den Zusammenhang vorgegeben, eine andere Bedeutung (z.B. im Falle, daß X Trägermenge einer Gruppe ist oder im Falle $X = \mathbb{N}$); dann verwenden wir zur Unterscheidung „.“ als Verknüpfungszeichen in X^*. Sofern in einem gegebenem Kontext kein spezielles X umfassendes Monoid gegeben ist, bedeutet X^n für jedes $n \in \mathbb{N}_0$ stets die Menge der Produkte der Länge n über X *im freien Monoid* X^*.

1.5 Proposition *Sei M ein Monoid, $X \subseteq M$. Es sind äquivalent:*

(i) *X ist unabhängig.*

(ii) *Für alle* $m, n \in \mathbb{N}_0$, $x_1, \ldots, x_n, x'_1, \ldots, x'_m \in X$ *gilt:*

$$x_1 \cdots x_n = x'_1 \cdots x'_m \Rightarrow (x_1, \ldots, x_n) = (x'_1, \ldots, x'_m)$$
$$(\textit{also } n = m,\ x_1 = x'_1,\ \ldots,\ x_n = x'_n).$$

Beweis. (i)⇒(ii): Seien $m, n \in \mathbb{N}_0$, $x_1, \ldots, x_n, x'_1, \ldots, x'_m \in X$ mit $x_1 \cdots x_n = x'_1 \cdots x'_m$, und sei $\varphi : X \to \mathcal{T}(X)$, $x \mapsto (x)$. Nach (i) hat φ eine homomorphe Fortsetzung $\bar{\varphi} : \bigcap_{X \subseteq N \leq M} N \to \mathcal{T}(X)$, und es gilt: $x_1 \cdots x_n,\ x'_1 \cdots x'_m \in \bigcap_{X \subseteq N \leq M} N$. Daraus folgt:

$$(x_1, \ldots, x_n) = (x_1) \cdots (x_n) = x_1\varphi \cdots x_n\varphi = x_1\bar{\varphi} \cdots x_n\bar{\varphi} = (x_1 \cdots x_n)\bar{\varphi}$$
$$= (x'_1 \cdots x'_m)\bar{\varphi} = x'_1\bar{\varphi} \cdots x'_m\bar{\varphi} = x'_1\varphi \cdots x'_m\varphi = (x'_1) \cdots (x'_m) = (x'_1, \ldots, x'_m).$$

(ii)⇒(i): Sei φ eine Abbildung von X in ein Monoid $\tilde{M}$. Zu jedem $a \in \bigcap_{X \subseteq N \leq M} N$ gibt es dann nach 1.3.1 ein $n \in \mathbb{N}_0$ und $x_1, \ldots, x_n \in X$ gibt mit $a = x_1 \ldots x_n$. Nach Voraussetzung ist dabei das Tupel $(x_1, \ldots, x_n)$ eindeutig bestimmt. Jeder Monoid-Homomorphismus von $\bigcap_{X \subseteq N \leq M} N$ in $\tilde{M}$ muß a auf das Produkt der Bilder der Faktoren $x_1, \ldots, x_n$ abbilden, so daß höchstens die Setzung $a\bar{\varphi} := x_1\varphi \cdots x_n\varphi$ als mögliche Fortsetzung von φ zu einem Monoid-Homomorphismus von $\bigcap_{X \subseteq N \leq M} N$ in $\tilde{M}$ in Frage kommt. Sicherlich gilt $\bar{\varphi}|_X = \varphi$. Ist $b \in \bigcap_{X \subseteq N \leq M} N$ und $(y_1, \ldots, y_k)$ *das* Tupel über X mit $b = y_1 \cdots y_k$, so ist $(x_1, \ldots, x_n, y_1, \ldots, y_k)$ ein, also *das* Tupel über X mit $ab = x_1 \cdots x_n\, y_1 \cdots y_k$. Es folgt:

$$(ab)\bar{\varphi} = x_1\varphi \cdots x_n\varphi\, y_1\varphi \cdots y_k\varphi = a\bar{\varphi}b\bar{\varphi}.$$

Ferner gilt $1_M\bar{\varphi} = 1_{\tilde{M}}$. □

1.6 Definition Sei X eine Menge, M ein von X erzeugtes Monoid, $m \in M$. Die Zahl

$$l_X(m) := \min\{n \mid n \in \mathbb{N}_0, \exists x_1, \ldots, x_n \in X \quad m = x_1 \cdots x_n\}$$

heißt die X-Länge von m. Offensichtlich gilt:

1.6.1 $\forall m, m' \in M \quad l_X(mm') \leq l_X(m) + l_X(m')$,

da ein Paar von Produktdarstellungen kürzester Länge von m, m' mit Faktoren aus X eine Produktdarstellung von mm' aus $l_X(m) + l_X(m')$ Faktoren ergibt. □

Ein Alphabet von M ist ein unabhängiges Erzeugendensystem von M. Aus 1.3.1 und 1.5 folgt:

1.6.2 *X Alphabet von $M \Leftrightarrow$ Für jedes $m \in M$ gibt es genau ein Tupel $(x_1, \ldots, x_n)$ über X mit $m = x_1 \cdots x_n$.* □

Ist X ein Alphabet von M, so lassen sich die Elemente von M nach 1.6.2 eindeutig durch Juxtaposition von Elementen von X darstellen. Deswegen heißen die Elemente von M auch Worte über X.

1.7 Proposition *Sei M ein Monoid.*

(1) *Ist X ein Alphabet von M, so gilt für alle m, $m' \in M$:*
$l_X(mm') = l_X(m) + l_X(m')$.

(2) *M hat höchstens ein Alphabet.*

Folgerung. *Alphabete freier Monoide sind eindeutig bestimmt.*

Ist daher M frei und X *das* Alphabet von M, so bezeichnet man die X-Länge eines Elements $w \in M$ einfach als die Länge von w. In einem freien Monoid (mit Alphabet X) soll daher „l“ stets „l_X“ bedeuten. Das Alphabet besteht genau aus den Worten der Länge 1, die man auch die Buchstaben von M nennt.

Beweis von 1.7: (1) Seien m, $m' \in M$. Seien $(x_1, \ldots, x_n)$, $(x'_1, \ldots, x'_k)$ die eindeutig bestimmten Tupel über X mit $m = x_1 \cdots x_n$, $m' = x'_1 \cdots x'_k$ (siehe 1.6.2). Dann ist $mm' = x_1 \cdots x_n\, x'_1 \cdots x'_k$, also nach 1.6.2 $l_X(mm') = n + k = l_X(m) + l_X(m')$.

(2) Seien X, Y Alphabete von M. Ist $x \in X$, so gibt es ein Tupel $(y_1, \ldots, y_n)$ über Y mit $x = y_1 \cdots y_n$. Es folgt:

$$1 = l_X(x) = l_X(y_1 \cdots y_n) \underset{(1)}{=} l_X(y_1) + \cdots + l_X(y_n) \geq n,$$

also $n = 1$, $l_X(y_1) = 1$ und damit $x = y_1 \in Y$. Es folgt: $X \subseteq Y$. Ebenso gilt $Y \subseteq X$. □

1.7.1 *Ist X Alphabet, Y irgendein Erzeugendensystem eines Monoids M, so gilt für alle $w \in M$: $l(w) \geq l_Y(w)$.*

Ist nämlich $w = y_1 \cdots y_k$ mit $y_1, \ldots, y_k \in Y$, so ist $l(y_j) \geq 1$ für alle $j \in \underline{k}$, also $l(w) \geq k$ nach 1.7(1). □

1.7.2 Beispiel Sei $X = \{x, y\}$, $U = \bigcap_{\{x, xy, yx\} \subseteq N \leq X^*} N$. Jedes Wort $\neq \iota$ in U enthält (bei seiner Darstellung als Wort über X) das Element x als Faktor, wie eine Anwendung von 1.3.1 auf das Erzeugendensystem $\{x, xy, yx\}$ des

Monoids U ergibt. Insbesondere gilt: $y \notin U$. Annahme, U sei frei. Dann sei Y das Alphabet von U. Es gilt: $l_X(x) = 1$, $l_X(xy) = 2 = l_X(yx)$, $y \notin Y$ und $l_X(w) \geq l_Y(w)$ für alle $w \in U$. Daraus folgt: $l_Y(x) = l_Y(xy) = l_Y(yx) = 1$, also x, xy, $yx \in Y$. Aber $(xy)x = x(yx)$; also ist Y nicht unabhängig, ein Widerspruch.

Das Beispiel lehrt insbesondere, daß Untermonoide freier Monoide nicht frei zu sein brauchen.

1.8 Definition Sei X eine Menge, $w \in X^*$. Ein Wort $v \in X^*$ heißt ein Rechtsfaktor (bzw. Linksfaktor) von w, wenn es ein $u \in X^*$ gibt mit $w = uv$ (bzw. $w = vu$). Wir verwenden dafür die Schreibweise $v \upharpoonright w$ (bzw. $v \upharpoonleft w$). Ist $w = x_1 \cdots x_n$ mit $x_1, \ldots, x_n \in X$, so sind also genau die Worte $\imath$ und $x_j x_{j+1} \cdots x_n$ mit $j \in \underline{n}$ die Rechtsfaktoren, die Worte $\imath$ und $x_1 \cdots x_j$ mit $j \in \underline{n}$ die Linksfaktoren von w. Ein Rechts- (oder Links-)Faktor $\neq w$ heißt echt.

Ein Wort $w' \in X^*$ heißt konjugiert zu w (Schreibweise: $w \sim w'$), wenn es Worte u, $v \in X^*$ gibt mit $w = uv$, $w' = vu$.

1.9 Proposition *Sei X eine Menge, und seien w, $w' \in X^*$. Es sind äquivalent:*

(i) $\exists z \in X^* \quad wz = zw'$

(ii) $w \sim w'$

Zusatz. *Für alle w, $w' \in X^+$, $z \in X^*$ gilt:*

$$wz = zw' \Leftrightarrow \exists u, v \in X^* : \quad w = uv,\ w' = vu,\ \exists j \in \mathbb{N}_0\ z = (uv)^j u.$$

Beweis. (ii)⇒(i): Gilt (ii), also $w = uv$, $w' = vu$ für geeignete u,$v \in X^*$, so setzen wir $z := u$, und wir erhalten (i).

(i)⇒(ii): Für $w = \imath$ oder $w' = \imath$ ist die Behauptung trivial. Für w, $w' \in X^+$ genügt es, den Zusatz zu beweisen: Gilt nämlich (i), so existieren nach dem Zusatz u, $v \in X^*$ mit $w = uv$, $w' = vu$, und damit gilt (ii). Damit geht es nun nur noch um den

Beweis des Zusatzes: Seien w, $w' \in X^+$, $z \in X^*$.

„⇐“: Sind u, $v \in X^*$, $j \in \mathbb{N}_0$ mit $w = uv$, $w' = vu$, $z = (uv)^j u$, so folgt:

$$wz = uv(uv)^j u = (uv)^j uvu = zw'.$$

„⇒“: Sei $wz = zw'$. Dann gilt: $l(w) = l(w')$ und $w^n z = zw'^n$ für alle $n \in \mathbb{N}$. Wegen $w \neq \imath$ existiert ein $n \in \mathbb{N}$ mit

$$l(w'^n) = n \cdot l(w') = n \cdot l(w) \geq l(z) \geq (n-1)l(w) = l(w^{n-1}).$$

Aus der Gleichung $w^{n-1}wz = zw'^n$ erhalten wir damit

$$z \upharpoonright w'^n,\ w^{n-1} \upharpoonright z.$$

Für geeignete u, $v \in X^*$ gilt also: $w'^n = vz$, $z = w^{n-1}u$. Es folgt:

$$w^{n-1}wz = zw'^n = w^{n-1}uvz,$$

also $w = uv$, und

$$w'^n = vz = vw^{n-1}u = v(uv)^{n-1}u = (vu)^n,$$

also $w' = vu$. Da auch $z = w^{n-1}u = (uv)^{n-1}u$ gilt, ist der Zusatz bewiesen.□

1.9.1 *Die Konjugiertheitsrelation $\sim$ ist eine Äquivalenzrelation auf X^*.*

Denn die Reflexivität und die Symmetrie ergeben sich unmittelbar aus der Definition, während die Transitivität bequem aus 1.9 folgt: Sind w, w', $w'' \in X^*$ mit $w \sim w'$, $w' \sim w''$, so gibt es nach 1.9 z, $y \in X^*$ mit $wz = zw'$, $w'y = yw''$. Es folgt: $wzy = zw'y = zyw''$, also $w \sim w''$ nach 1.9. □

Zu einem Wort w mit $l(w) = n$ gibt es höchstens n konjugierte Worte. Zwar sehen formal die Worte

$$\begin{gathered} x_1x_2\cdots x_n, \\ x_2\cdots x_nx_1, \\ \vdots \\ x_j\cdots x_nx_1\cdots x_{j-1} \\ \vdots \\ x_nx_1\cdots x_{n-1} \end{gathered}$$

paarweise verschieden aus, doch da über die Verschiedenheit der x_j nichts bekannt ist, können in der Liste durchaus Elemente mehrfach auftreten, z.B. gibt es im Fall $w = xxyxxy$ (wo $x \neq y$) genau *drei* Konjugierte – nämlich w, $xyxxyx$, $yxxyxx$ – und nicht etwa deren sechs.

Die Konjugierten eines Wortes entstehen durch spezielle Permutationen der Buchstaben. Dies legt allgemeiner die folgende Definition eines „Produktes“ zwischen Permutationen von $\underline{n}$ und Worten der Länge n fest:

1.10 Definition Sei X eine Menge, $n \in \mathbb{N}$. Für alle $\sigma \in S_n$ und $w = x_1 \cdots x_n \in X^n$ sei $\sigma w := x_{1\sigma} \cdots x_{n\sigma}$. Die Abbildung

$$\begin{aligned} S_n \times X^n &\to X^n \\ (\sigma, w) &\mapsto \sigma w \end{aligned}$$

heißt die Polya(-Weyl)-Aktion von S_n auf X^n. Wichtig zum Verständnis dieser Definition ist, daß die Indizes $1, \ldots, n$ nicht aus irgendeiner (vorher gegebenen) Numerierung von X, sondern aus *der* Darstellung von w als Produkt von Elementen von X stammen, also von w abhängen. Um z.B. die Permutation $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 4 & 5 & 6 & 1 \end{pmatrix}$ im Sinne der Polya-Aktion auf $w = xxyxxy$ anzuwenden, setzen wir „vorübergehend“ $x_1 := x_2 := x_4 := x_5 := x$, $x_3 := x_6 := y$, so daß $w = x_1x_2x_3x_4x_5x_6$ gilt, und erhalten

$$\sigma w = x_2x_3x_4x_5x_6x_1 = xyxxyx \sim w.$$

Ähnlich berechnen wir z.B.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 1 & 2 & 3 \end{pmatrix} w = x_4x_5x_6x_1x_2x_3 = xxyxxy = w,$$
$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 1 & 4 & 5 & 6 \end{pmatrix} w = x_3x_2x_1x_4x_5x_6 = yxxxxy \not\sim w.$$

1.10.1 $\forall \rho, \sigma \in S_n \; \forall w \in X^n \quad (\rho\sigma)w = \rho(\sigma w)$.

Beweis. Seien $\rho, \sigma \in S_n$, $w \in X^n$ und $x_1, \ldots, x_n \in X$ mit $w = x_1 \ldots x_n$. Sei $y_j := x_{j\sigma}$ für alle $j \in \underline{n}$. Es gilt dann:

$$\begin{aligned} \rho(\sigma w) &= \rho(x_{1\sigma} \cdots x_{n\sigma}) \\ &= \rho(y_1 \cdots y_n) \\ &= y_{1\rho} \cdots y_{n\rho} \\ &= x_{(1\rho)\sigma} \cdots x_{(n\rho)\sigma} \quad \text{nach Definition von } y_j \\ &= x_{1(\rho\sigma)} \cdots x_{n(\rho\sigma)} = (\rho\sigma)w. \end{aligned}$$

□

Unter Benutzung der Polya-Aktion können wir die Konjugierten von $x_1 \cdots x_n$ beschreiben als die Menge der Elemente

$$\tau^j x_1 \cdots x_n \quad (j \in \underline{n}), \quad \text{wobei } \tau = \begin{pmatrix} 1 & 2 & \ldots & n-1 & n \\ 2 & 3 & \ldots & n & 1 \end{pmatrix},$$

d.h. man benötigt zur Beschreibung der Konjugation nicht die Polya-Aktion der ganzen Gruppe S_n, sondern die der Untergruppe $\langle\tau\rangle$. Wir nennen die

Permutation τ den Standardzyklus der Länge n. Für $n = 6$ haben wir in 1.10 bereits zwei Elemente dieser Untergruppe betrachtet: $\tau(= \sigma)$ und τ^3: Unter dem letzteren war das dort betrachtete Wort w invariant ($\tau^3 w = w$). Allgemein sagt man, daß $\sigma \in S_n$ das Wort $w \in X^n$ stabilisiert, wenn gilt: $\sigma w = w$. Die Menge aller Elemente einer Untergruppe H von S_n, die w stabilisieren, wird der Stabilisator von w in H genannt und mit $Stab_H(w)$ bezeichnet. Aus 1.10.1 folgt leicht:

1.10.2 *$Stab_H(w)$ ist eine Untergruppe von S_n.* □

Weiter gilt:

1.10.3 $\forall \rho, \sigma \in S_n : \quad \rho w = \sigma w \Leftrightarrow \sigma^{-1}\rho \in Stab(w)$,

denn: $\rho w = \sigma w \Leftrightarrow \sigma^{-1}(\rho w) = \sigma^{-1}(\sigma w) \Leftrightarrow (\sigma^{-1}\rho)w = w$, nach 1.10.1. □

Folgerung: *Ist R ein Repräsentantensystem für die Restklassen von $Stab_{\langle\tau\rangle}(w)$ in $\langle\tau\rangle$, so ist $\{\rho w \,|\, \rho \in R\}$ die Konjugiertenklasse von w. Diese enthält genau $|R|$ Elemente.*

Beispiel: $Stab_{\langle\tau\rangle}(xxyxxy) = \{id, \begin{pmatrix}1&2&3&4&5&6\\4&5&6&1&2&3\end{pmatrix}\}$. Z.B. ist

$$\{id, \begin{pmatrix}1&2&3&4&5&6\\2&3&4&5&6&1\end{pmatrix}, \begin{pmatrix}1&2&3&4&5&6\\3&4&5&6&1&2\end{pmatrix}\}$$

ein Repräsentantensystem für die Restklassen von $\{id, \begin{pmatrix}1&2&3&4&5&6\\4&5&6&1&2&3\end{pmatrix}\}$ in $\langle\tau\rangle$. Wir haben oben bereits festgestellt, daß $xxyxxy$ genau drei Konjugierte besitzt, in Übereinstimmung mit der letzten Aussage der Folgerung in diesem Fall.

1.11 Definition Sei X eine Menge. Ein Wort $w \in X^*$ heißt primitiv, wenn die Zahl der Konjugierten von w mit $l(w)$ übereinstimmt. (Dies bedeutet, daß die Elemente in der vor 1.10 stehenden Auflistung paarweise verschieden sind.) Wegen $l(\imath) = 0$ ist $\imath$ nicht primitiv. Triviale Feststellungen sind ferner:

1.11.1 *Sei $n \in \mathbb{N}$, τ der Standardzyklus der Länge n. Genau dann ist ein Element $w \in X^n$ primitiv, wenn gilt: $Stab_{\langle\tau\rangle}(w) = \{id\}$.* □

1.11.2 *Ist $w \in X^*$ primitiv und $w \sim w'$, so ist auch w' primitiv.* □

1.12 Satz *Sei X eine Menge.*

(1) *Zu jedem $w \in X^+$ gibt es genau ein primitives Wort $v \in X^*$ mit $w = v^j$ für ein $j \in \mathbb{N}$.* (Im folgenden bezeichne $\underline{w}$ das primitive Wort, von dem w eine Potenz ist.) *Es gilt: $l(\underline{w})|l(w)$.*

(2) *Sind $w, w' \in X^+$ mit $l(w) = l(w')$, so gilt:*

$$w \sim w' \Leftrightarrow \underline{w} \sim \underline{w'}$$

(3) *Ist X endlich und (für beliebiges $n \in \mathbb{N}$) $\psi_X(n)$ die Zahl der Konjugiertenklassen primitiver Worte der Länge n über X, so gilt:*

$$\psi_X(n) = \frac{1}{n} \sum_{d|n} \mu(d) |X|^{\frac{n}{d}}.$$

(4) *Sind $X = \{x_1, \ldots, x_r\}$ (wobei $x_i \neq x_j$ für $i \neq j$), $k_1, \ldots, k_r \in \mathbb{N}_0$, $n := k_1 + \cdots + k_r$ und bezeichnet $\chi_X(k_1, \ldots, k_r)$ die Zahl der Konjugiertenklassen primitiver Worte der Länge n über X, die für jedes $i \in \underline{r}$ genau k_i-mal den Buchstaben x_i enthalten, so gilt:*

$$\chi_X(k_1, \ldots, k_r) = \frac{1}{n} \sum_{d|k_1, \ldots k_r} \mu(d) \binom{\frac{n}{d}}{\frac{k_1}{d}, \ldots, \frac{k_r}{d}}.$$

(Dabei ist μ die Möbius-Funktion $\mathbb{N} \to \{0, 1, -1\}$,

$$d \mapsto \begin{cases} (-1)^i & \text{wenn } d \text{ Produkt von } i \text{ verschiedenen Primzahlen ist} \\ 0 & \text{wenn } d \text{ durch eine Quadratzahl} \neq 1 \text{ teilbar ist} \end{cases}.)$$

Im Beweis von (3) und (4) werden wir den folgenden bekannten Satz verwenden:

1.13 Proposition (Möbius'sche Umkehrformel) *Seien f, $F \in \mathbb{Z}^{\mathbb{N}}$. Es sind äquivalent:*

(i) $F(n) = \sum_{d|n} f(d)$ *für alle* $n \in \mathbb{N}$,

(ii) $f(n) = \sum_{d|n} \mu(d) F(\frac{n}{d})$ *für alle* $n \in \mathbb{N}$.

Beweis. Die Hilfsaussage $\sum_{d|n} \mu(d) = \sum_{T \subseteq \underline{m}} (-1)^{|T|} = \begin{cases} 1 & \textit{falls } n = 1 \\ 0 & \textit{sonst} \end{cases}$ (wobei m die Anzahl der Primfaktoren von n ist) ermöglicht einen einfachen Beweis:

Gilt (i), so folgt für alle $n \in \mathbb{N}$

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

Gilt (ii), so auch $f(n) = \sum_{d|n} \mu(\frac{n}{d}) F(d)$ für alle $n \in \mathbb{N}$ und damit auch

$$\sum_{t|n} f(t) = \sum_{t|n} \sum_{d|t} \mu(\frac{t}{d}) F(d) = \sum_{d|n} \sum_{t^*|\frac{n}{d}} \mu(t^*) F(d) = F(n)$$

für alle $n \in \mathbb{N}$. □

Beweis von 1.12. (1): Sei $w \in X^+$, τ der Standardzyklus der Länge $l(w)$, $S := Stab_{\langle\tau\rangle}(w)$, $d := |\langle\tau\rangle : S|$. Es gilt $\tau^d \in S$, also auch $\tau^{dj} \in S$, d.h. $w = \tau^{dj} w$ für alle $j \in \mathbb{N}$. Seien $x_1, \ldots, x_n \in X$ mit $w = x_1 \cdots x_n$. Dann gilt: $x_i = x_j$, falls $i \equiv j$ modulo d, also

$$w = x_1 \cdots x_d x_1 \cdots x_d \ \cdots \ x_1 \cdots x_d = (x_1 \cdots x_d)^{\frac{n}{d}}.$$

Sei $v := x_1 \cdots x_d$. Wir zeigen, daß v primitiv ist:

Dazu sei $j \in \underline{d}$ mit $v = x_{j+1} \cdots x_d x_1 \cdots x_j$. Wir müssen zeigen: $j = d$. Es gilt:

$$\begin{aligned} w &= (x_1 \underbrace{\cdots x_d) \cdots (x_1 \cdots}_{\frac{n}{d}} x_d) \\ &= v^{\frac{n}{d}} = (x_{j+1} \underbrace{\cdots x_d x_1 \cdots x_j) \cdots (x_{j+1} \cdots x_d x_1 \cdots}_{\frac{n}{d}} x_j), \\ &= \tau^j w \end{aligned}$$

also $\tau^j \in S$. Wegen $d = |\langle\tau\rangle : S| = |\langle\tau\rangle / S|$ folgt daraus: $j \geq d$. Da $j \in \underline{d}$ gilt, erhalten wir $j = d$.

Also ist v primitiv und $w = v^{\frac{n}{d}}$. Zum Beweis der eindeutigen Bestimmtheit von v nehmen wir an, es sei u ein primitives Wort und $e \in \mathbb{N}$ mit $w = u^e$. Dann gilt: $\tau^{l(u)} \in S$, also $d | l(u)$, denn d ist die Ordnung der Faktorgruppe $\langle\tau\rangle / S$. Aus $u^e = w = v^{\frac{n}{d}}$ und $l(v) = d$ folgt, daß u eine Potenz von v, wegen der Primitivität von u also gleich v ist, und (1) ist bewiesen.

(2): Seien $w, w' \in X^+$ mit $l(w) = l(w')$, $j \in \mathbb{N}$ mit $w = \underline{w}^j$. Zunächst gelte $\underline{w} \sim \underline{w'}$. Dann gibt es $u, v \in X^*$ mit $\underline{w} = uv$, $\underline{w'} = vu$. Aus $l(w) = l(w')$ folgt nun: $(vu)^j = w'$, und damit

$$wu = (uv)^j u = u(vu)^j = uw',$$

woraus dank 1.9 folgt: $w \sim w'$.– Gilt umgekehrt $w \sim w'$, so existieren $u,v \in X^*$ mit $w = uv$, $w' = vu$. Seien $x_1, \ldots, x_d \in X$ mit $\underline{w} = x_1 \cdots x_d$. Dann gilt für ein geeignetes $k \in \underline{d}$:

$$w = \underbrace{(x_1 \cdots x_d) \cdots (x_1 \cdots x_{k-1})}_{=u} \underbrace{(x_k \cdots x_d) \cdots (x_1 \cdots x_d)}_{=v} = \underline{w}^j.$$

Es folgt:

$$\begin{aligned} w' = vu &= (x_k \cdots x_d)(x_1 \cdots x_d) \cdots (x_1 \cdots x_d)(x_1 \cdots x_{k-1}) \\ &= (x_k \cdots x_d x_1 \cdots x_{k-1})^j, \end{aligned}$$

und $x_k \cdots x_d x_1 \cdots x_{k-1}$ ist primitiv (siehe 1.11.2), da konjugiert zu $\underline{w}$. Damit gilt: $\underline{w'} = w_k \cdots w_d w_1 \cdots w_{k-1} \sim \underline{w}$.

(3): Die Abbildung $\pi : X^n \to X^*$, $w \mapsto \underline{w}$, ist injektiv und bildet nach (1) jede Konjugiertenklasse in X^n auf eine Konjugiertenklasse von primitiven Worten einer Länge ab, die n teilt. Ist v ein primitives Wort, deren Länge d ein Teiler von n ist, so ist $\underline{(v^{\frac{n}{d}})} = v$ und $l(v^{\frac{n}{d}}) = n$. Daher wird jede Konjugiertenklasse von primitiven Worten einer Länge d mit $d|n$ von π getroffen. Da X^n die disjunkte Vereinigung der Konjugiertenklassen von Worten der Länge n ist, folgt

$$|X|^n = \sum_{K \text{ Konj.-kl. in } X^n} |K| = \sum_{d|n} \sum_{K_d} |K_d| = \sum_{d|n} d \cdot \psi_X(d),$$

(wobei K_d die Konjugiertenklassen primitiver Worte der Länge d durchläuft); denn jede Konjugiertenklasse von primitiven Worten der Länge d enthält genau d Worte. Durch Anwendung der Möbius'schen Umkehrformel 1.13 ergibt sich daraus direkt die behauptete Gleichung.

(4) läßt sich nach einem Vorschlag von T. Bauer durch folgende Verfeinerung des Beweisgedankens von (3) einsehen: Für beliebige $l_1, \ldots, l_r \in \mathbb{N}_0$ sei $X^{(l_1,\ldots,l_r)}$ die Menge der Worte über X, die für jedes $i \in \underline{r}$ genau l_i-mal den Buchstaben x_i enthalten. Dann gilt:

$$|X^{(l_1,\ldots,l_r)}| = \binom{l}{l_1, \ldots, l_r},$$

wobei $l := l_1 + \cdots + l_r$. Seien nun $h_1, \ldots, h_r \in \mathbb{N}$ mit $\mathrm{ggT}(h_1, \ldots, h_r) = 1$, $h := h_1 + \cdots + h_r$. Ist $w \in X^{(jh_1,\ldots,jh_r)}$ für ein $j \in \mathbb{N}$, so gibt es einen eindeutig bestimmten Teiler d von j mit $\underline{w} \in X^{(dh_1,\ldots,dh_r)}$. Ist andererseits $v \in X^{(dh_1,\ldots,dh_r)}$ mit $d|j$ und v primitiv, so gilt $v^{\frac{j}{d}} \in X^{(jh_1,\ldots,jh_r)}$ und $\underline{v^{\frac{j}{d}}} = v$.

Also induziert π eine Bijektion von $X^{(jh_1,\ldots,jh_r)}$ auf die (disjunkte) Vereinigungsmenge der Mengen aller primitiven Worte von $X^{(dh_1,\ldots,dh_r)}$ für $d|j$. Setzen wir $f(k) := |\{v|v \in X^{(kh_1,\ldots,kh_r)}, v \text{ primitiv}\}|$ für alle $k \in \mathbb{N}$, so folgt

$$\binom{jh}{jh_1,\ldots,jh_r} = \sum_{d|j} f(d) \quad \text{für alle } j \in \mathbb{N}.$$

Durch Anwendung der Möbius'schen Umkehrformel 1.13 ergibt sich damit

$$f(k) = \sum_{d|k} \mu(d) \binom{\frac{k}{d}h}{\frac{k}{d}h_1,\ldots,\frac{k}{d}h_r} \quad \text{für alle } k \in \mathbb{N}.$$

Wir setzen nun speziell $k := \operatorname{ggT}(k_1,\ldots,k_r)$ und $h_i := \frac{k_i}{k}$ für alle $i \in \underline{r}$. Dann gilt $\operatorname{ggT}(h_1,\ldots,h_r) = 1$ und $h = \sum_{i=1}^{r} h_i = \frac{1}{k}\sum_{i=1}^{r} k_i = \frac{n}{k}$. Es folgt:

$$\chi_X(k_1,\ldots,k_r) = \frac{1}{n} f(k) = \frac{1}{n} \sum_{d|k} \mu(d) \binom{\frac{n}{d}}{\frac{k_1}{d},\ldots,\frac{k_r}{d}}.$$

□

Wir werden nun ein Repräsentantensystem für die Konjugiertenklassen primitiver Worte betrachten, das sich später algebraisch als von besonderer Wichtigkeit erweisen wird; es wurde 1958 von R. Lyndon eingeführt. Von vornherein bietet sich kein ausgezeichneter Repräsentant einer Konjugiertenklasse an, von dem man etwa sagen könnte, er sei etwa „der größte“ oder „der kleinste“. Dazu führen wir zunächst eine Ordnung auf der Menge X^* ein, die auf einer Ordnung auf X basiert:

1.14 Definition Sei X eine Menge und $\leq$ eine vollständige Ordnung auf X. Wir definieren dann induktiv (nach der Wortlänge) eine (strikte) Ordnung $\underset{lex}{<}$ auf X^* durch

$$\begin{aligned} \iota &\underset{lex}{<} w && \text{für alle } w \in X^+, \\ x_1\cdots x_n &\underset{lex}{<} y_1\cdots y_m && \text{falls } x_1 < y_1 \text{ oder: } x_1 = y_1 \text{ und } x_2\cdots x_n \underset{lex}{<} y_2\cdots y_m. \end{aligned}$$

(Den Nachweis, daß damit tatsächlich die Axiome einer vollständigen Ordnung erfüllt sind, führen wir hier nicht durch.) $\underset{lex}{<}$ heißt die (zu $<$ gebildete) lexikographische Ordnung auf X^*. Ist z.B. $X = \{a, b, c\}$ mit $a < b < c$, so gilt etwa

$$bb \underset{lex}{<} bbcac \underset{lex}{<} bca \underset{lex}{<} c.$$

Durch $\underset{lex}{<}$ wird X^* vollständig geordnet. Wie üblich bedeutet

$$v \underset{lex}{\leq} w: \quad v \underset{lex}{<} w \text{ oder } v = w.$$

Trivial ist die folgende Bemerkung:

1.14.1 *Ist $v \underset{lex}{\leq} w$ und v kein Linksfaktor von w, so gilt $vu \underset{lex}{<} w$ für alle $u \in X^*$.* □

Ein Element $w \in X^*$ heißt ein Lyndon-Wort, falls w primitiv ist und für alle $w' \in X^*$ gilt:

$$w \sim w' \Rightarrow w \underset{lex}{\leq} w'.$$

Mit $\mathfrak{L}^{X,<}$ bezeichnen wir die Menge aller Lyndon-Worte von X^*, mit $\mathfrak{L}_n^{X,<}$ die Menge der Lyndon-Worte der Länge n von X^*. Dann ist $\mathfrak{L}_n^{X,<}$ ein Repräsentantensystem für die Menge der Konjugiertenklassen primitiver Worte der Länge n. Den zweiten oberen Index „$<$“ erlauben wir uns wegzulassen, wenn über die Ausgangsordnung $<$ auf X Unklarheiten nicht zu befürchten sind.

1.14.2 *$\mathfrak{L}^X$ ist ein Repräsentantensystem für die Menge der Konjugiertenklassen primitiver Worte über der geordneten Menge X.* □

Als Beispiel betrachten wir $X = \{a, b\}$ mit $a < b$, $n = 3$. Die Konjugiertenklassen von Worten der Länge 3 sind (Lyndon-Worte jeweils unterstrichen):

$\{aaa\}$	nicht primitiv
$\{\underline{aab}, aba, baa\}$	primitiv
$\{\underline{abb}, bab, bba\}$	primitiv
$\{bbb\}$	nicht primitiv,

also $\mathfrak{L}_3^{\{a,b\}} = \{aab, abb\}$; die der Länge 4:

$\{aaaa\}$	nicht primitiv
$\{\underline{aaab}, baaa, abaa, aaba\}$	primitiv
$\{\underline{aabb}, baab, bbaa, abba\}$	primitiv
$\{abab, baba\}$	nicht primitiv
$\{\underline{abbb}, babb, bbab, bbba\}$	primitiv
$\{bbbb\}$	nicht primitiv,

also $\mathfrak{L}_4^{\{a,b\}} = \{aaab, aabb, abbb\}$. Aus 1.12(3) folgt:

1.14.3 $|\mathfrak{L}_n^X| = \frac{1}{n}\sum_{d|n} \mu(d)|X|^{\frac{n}{d}}$. □

Die folgenden drei Resultate befassen sich mit dem Buchstabenaufbau, also der „Innenwelt“ der Lyndon-Worte. Anschließend werden wir Aussagen zur strukturellen Bedeutung der Lyndon-Worte für X^*, also sozusagen zur „Außenwelt“ der Lyndon-Worte beweisen und damit das Kapitel beenden:

1.15 Proposition *Sei X eine geordnete Menge, $w \in X^+$. Es sind äquivalent:*

(i) $w \in \mathfrak{L}^X$,

(ii) $\forall v \in X^* \quad \iota \neq v \restriction w \neq v \Rightarrow w \underset{lex}{<} v$.

Zusatz. *Bei einem Lyndon-Wort w ist kein echter Rechtsfaktor $\neq \iota$ von w zugleich ein Linksfaktor von w.*

Beweis. (ii)⇒(i): Zunächst sehen wir, daß w primitiv sein muß, denn sonst gäbe es ein $v \in X^*$ und ein $j \in \mathbb{N}_{>1}$ mit $w = v^j$, damit $\iota \neq v \restriction w \neq v$, $v \underset{lex}{<} w$, im Widerspruch zu (ii). Sei weiter $w' \in X^*$ mit $w' \sim w$. Dann gibt es u, $v \in X^*$ mit $w = uv$, $w' = vu$. Ist $w' = w$, so $w \underset{lex}{\leq} w'$. Ist aber $w' \neq w$, so gilt $\iota \neq v \neq w$, also nach (ii) $w \underset{lex}{<} v \underset{lex}{\leq} vu = w'$.

(i)⇒(ii): Sei $w \in \mathfrak{L}^X$ und $\iota \neq v \restriction w \neq v$. Wir zeigen

$$(*) \qquad v \text{ ist kein Linksfaktor von } w,$$

d. h. wir beweisen den Zusatz. Daraus folgt nämlich (ii): Wäre $v \underset{lex}{\leq} w$, so nach $(*)$ und 1.14.1 sogar $vu \underset{lex}{<} w$, für alle $u \in X^*$. Ist nun $u \in X^*$ mit $w = uv$, so gilt $vu \underset{lex}{<} w$ und $vu \sim uv = w$ im Widerspruch zu (i). –

Beweis von $(*)$: Sei $u \in X^+$ mit $w = uv$. Angenommen, $v \upharpoonright w$, d.h. es gäbe ein $u' \in X^+$ mit $w = vu'$. Dann gälte $uv = vu'$, also nach dem Zusatz zu 1.9 $u = rs$, $u' = sr$, $v = (rs)^j r$ für geeignete r, $s \in X^*$, $j \in \mathbb{N}_0$. Es folgte:

$$w = rs\,(rs)^j r = (rs)^{j+1} r.$$

Wäre hierbei $r = \iota$, so $j = 0$ wegen der Primitivität von w und $\iota \neq v = r = \iota$, ein Widerspruch; also müßte $r \neq \iota$ gelten. Wegen $w \in \mathfrak{L}^X$ hätte man dann

$$r(sr)^{j+1} = (rs)^{j+1} r = w \underset{lex}{<} r(rs)^{j+1}$$

damit $(sr)^{j+1} \underset{lex}{<} (rs)^{j+1}$ nach Definition von $\underset{lex}{<}$ und folglich

$$w = r(sr)^{j+1} \sim (sr)^{j+1} r \underset{lex}{<} (rs)^{j+1} r = w$$

mit Widerspruch zur Lyndon-Wort-Eigenschaft von w. Damit gilt $(*)$, und 1.15 ist in allen Teilen bewiesen. □

1.16 Korollar *Sei X eine geordnete Menge und $w \in X^+$. Dann ist der längste zu $\mathfrak{L}^X$ gehörige Rechtsfaktor von w zugleich der lexikographisch kleinste Rechtsfaktor $\neq \imath$ von w.*

Beweis: Der lexikographisch kleinste Rechtsfaktor $u \neq \imath$ von w ist lexikographisch kleiner als jeder seiner eigenen echten Rechtsfaktoren $\neq \imath$, nach 1.15 also ein Lyndon-Wort über X. Damit ist u ein Rechtsfaktor des längsten zu $\mathfrak{L}^X$ gehörigen Rechtsfaktors v von w. Aus 1.15 folgte also im Falle $v \neq u$ der Widerspruch $v \underset{lex}{<} u$. Daher gilt $v = u$. □

Wir wenden nun 1.15 an auf echte Rechtsfaktoren v von $w \in \mathfrak{L}^X$, die selbst wieder Lyndon-Worte sind. Ist $w \notin X$, so gibt es solche immer, da ja zumindest der letzte Buchstabe von w von der Art ist.

1.17 Lemma *Sei X eine geordnete Menge und $w \in \mathfrak{L}^X \setminus X$. Sei v der längste zu $\mathfrak{L}^X$ gehörige echte Rechtsfaktor von w und $u \in X^*$ mit $w = uv$. Dann gilt: $u \in \mathfrak{L}^X$ und $u \underset{lex}{<} v$.*

Wir nennen das Paar (u, v) die Standard-Zerlegung des Lyndon-Wortes w.

Beweis. Trivialerweise gilt $u \underset{lex}{<} w$, $v \neq \imath$, und 1.15 ergibt: $w \underset{lex}{<} v$, da $w \in \mathfrak{L}^X$. Es folgt: $u \underset{lex}{<} v$. Es bleibt der erste Teil der Behauptung zu zeigen. Nach 1.15 genügt es dazu, für einen echten Rechtsfaktor $s \neq \imath$ von u nachzuweisen:

$$(*) \qquad u \underset{lex}{<} s.$$

Sei $r \in X^*$ mit $u = rs$, also

$$w = \underbrace{r\ s}_{=u}\ \underbrace{v}_{\in \mathfrak{L}^X}.$$

Nach Wahl von v gilt $sv \notin \mathfrak{L}^X$. Nach 1.15 gibt es also einen echten Rechtsfaktor t von sv mit $t \underset{lex}{<} sv$. Aus $sv \restriction w (= rsv)$ folgt $w \neq t \restriction w$, und wir erhalten, erneut mit 1.15:

$$u \underset{lex}{<} uv = w \underset{lex}{<} t.$$

Angenommen, es gälte $s \underset{lex}{<} t$. Aus $s \underset{lex}{<} t \underset{lex}{<} sv$ folgte dann die Existenz eines $z \in X^+$ mit $t = sz$, $z \underset{lex}{<} v$, damit $sv \neq sz \restriction sv$, $l(z) < l(v)$ und deswegen auch $v \neq z \restriction v$. Wegen $z \underset{lex}{<} v$ widerspräche das nach 1.15 der Voraussetzung, daß v ein Lyndon-Wort ist. Also gilt $s \underset{lex}{\geq} t$.

Es folgt: $u \underset{lex}{<} s$. Also ist $(*)$ bewiesen. □

1.18 Proposition *Sei X eine geordnete Menge, $w \in X^* \setminus X$. Dann gilt:*

$$w \in \mathfrak{L}^X \Leftrightarrow \exists u, v \in \mathfrak{L}^X \quad w = uv,\ u \underset{lex}{<} v.$$

Beweis. Da die Implikation „$\Rightarrow$“ unmittelbar aus 1.17 folgt, geht es im folgenden nur noch um die Implikation „$\Leftarrow$“. Offenbar gilt $w \neq \imath$. Wir wollen zeigen, daß w lexikographisch kleiner als jeder Rechtsfaktor $\neq \imath$, w von w ist, denn dann sind wir mit 1.15 fertig. Davon überlegen wir uns zunächst den folgenden Spezialfall:

$$(*) \qquad uv \underset{lex}{<} v.$$

Beweis von $(*)$: Es ist $u \in \mathfrak{L}^X$, also jedenfalls $u \neq \imath$.

1. Fall: $u \upharpoonright v$. Sei also $s \in X^*$ mit $v = us$. Wegen $u \neq v$ gilt $s \neq \imath$. Nach 1.15 gilt $v \underset{lex}{<} s$, da $v \in \mathfrak{L}^X$. Es folgt: $uv \underset{lex}{<} us = v$.

2. Fall: $\neg\, u \upharpoonright v$. Die Voraussetzung $u \underset{lex}{<} v$ impliziert dann sogar $uz \underset{lex}{<} v$ für alle $z \in X^*$ (siehe 1.14.1).

Also gilt $(*)$. Jeder echte Rechtsfaktor $\neq \imath$ von w ist *a)* ein Rechtsfaktor von v, oder *b)* er hat die Form sv für einen echten Rechtsfaktor s von u.

a) Sei $\imath \neq r \upharpoonright v$. Mit $(*)$ erhalten wir dann: $w = uv \underset{lex}{<} v \underset{lex}{\leq} r$, letzteres nach 1.15 wegen $v \in \mathfrak{L}^X$.

b) Sei $\imath \neq s \upharpoonright u$, $s \neq u$. Nach 1.15 gilt dann $u \underset{lex}{<} s$, also $w = uv \underset{lex}{<} sv$.

Also ist w lexikographisch kleiner als jeder echte Rechtsfaktor $\neq \imath$ von w. Aus 1.15 folgt: $w \in \mathfrak{L}^X$. □

Nun können wir auch charakterisieren, wann eine Produktzerlegung eines Lyndon-Worts in zwei Lyndon-Worte die Standardzerlegung ist:

1.19 Proposition *Seien X eine geordnete Menge und u, v, $w \in X^+$ mit $uv = w$. Es sind äquivalent:*

(i) *$w \in \mathfrak{L}^X$, und (u, v) ist die Standardzerlegung von w.*

(ii) *u, $v \in \mathfrak{L}^X$, $u \underset{lex}{<} v$, und für jeden zu $\mathfrak{L}^X$ gehörigen echten Rechtsfaktor t von u gilt: $t \underset{lex}{\geq} v$.*

(iii) *u, $v \in \mathfrak{L}^X$, $u \underset{lex}{<} v$, und im Falle $u \notin X$ gilt für den längsten zu $\mathfrak{L}^X$ gehörigen echten Rechtsfaktor t von u: $t \underset{lex}{\geq} v$.*

Beweis. (i)⇒ (ii): Sei (u, v) die Standardzerlegung von $w \in \mathfrak{L}^X$. Nach 1.17 gilt dann $u,\ v \in \mathfrak{L}^X$, $u \underset{lex}{<} v$. Gäbe es einen zu $\mathfrak{L}^X$ gehörigen echten Rechtsfaktor t von u mit $t \underset{lex}{<} v$, so folgte $tv \in \mathfrak{L}^X$ nach 1.18 mit Widerspruch dazu, daß v der längste zu $\mathfrak{L}^X$ gehörige Rechtsfaktor von w ist. Also gilt (ii).

(ii)⇒(iii) ist trivial.

(iii)⇒(i): Aus (iii) folgt zunächst $w \in \mathfrak{L}^X$ nach 1.18. Sei v' der längste zu $\mathfrak{L}^X$ gehörige echte Rechtsfaktor von w. Wir haben $v = v'$ zu zeigen. Wegen $v \in \mathfrak{L}^X$ gibt es jedenfalls ein $s \in X^*$ mit $v' = sv$. Um $s = \imath$ einzusehen, machen wir die Annahme $s \neq \imath$. Da s ein echter Rechtsfaktor von u ist, folgt dann $s \neq u \notin X$. Nach Voraussetzung gilt dann $v \underset{lex}{\leq} t$, und nach 1.16 (angewandt auf den maximalen echten Rechtsfaktor von u) ist t der lexikographisch kleinste echte Rechtsfaktor $\neq \imath$ von u, insbesondere $t \underset{lex}{\leq} s$. Damit folgt der Widerspruch

$$v \underset{lex}{\leq} t \underset{lex}{\leq} s \underset{lex}{<} sv = v' \underset{lex}{<} v,$$

wobei am Ende 1.15 auf $v' \in \mathfrak{L}^X$ angewandt wurde. □

Es entpuppt sich ein interessantes „Innenleben“ der Lyndon-Worte, wie die folgenden Beispiele andeuten mögen: Sei $X = \{a, b, c\}$ und $a < b < c$. Durch iteriertes Bilden der Standardzerlegung (d.h. zunächst für das Ausgangswort, dann für die aufgetretenen Faktoren usw., hier angedeutet durch entsprechendes Einklammern) erhalten wir

$$(abbac) = \big((abb)(ac)\big) = \Big(\big((ab)b\big)(ac)\Big) = \Big(\Big(\big((a)(b)\big)(b)\Big)\big((a)(c)\big)\Big),$$

wobei aus der Beklammerung hervorgeht, auf welchem Wege das Ausgangswort so lange zerlegt wurde, bis sämtliche Faktoren Buchstaben waren. Alle auftretenden Worte sind Lyndon-Worte, und 1.17 gewährleistet, daß dieser Prozeß tatsächlich erst dann abbricht, wenn Buchstaben erreicht sind. Spaltet man als Rechtsfaktor ein Lyndon-Wort kürzerer Länge ab als es in der Standardzerlegung der Fall ist, so kann man nicht erwarten, daß der dann zugehörige Linksfaktor wieder ein Lyndon-Wort ist: Spaltet man etwa das „zu kurze“ Lyndon-Wort c von $abbac$ ab, $abbac = (abba)c$, so gehört dazu der Linksfaktor $abba$, der kein Lyndon-Wort ist. Andererseits kann es durchaus iterierte Zerlegungen vermöge Faktoren geben, die stets Lyndon-Worte sind, ohne daß es sich durchgehend um die Standardzerlegung handelt. Ein weiteres Beispiel illustriert die beschriebenen Phänomene:

$ababc = (ab)(abc) = (ab)(a(bc)) = ((a)(b))((a)((b)(c)))$
(vollständige „Innenzerlegung“ durch Iteration der Standardzerlegung)
$= (abab)c$ (c ist Lyndon-Wort, aber $abab$ nicht!)
$= (aba)(bc)$ (bc ist Lyndon-Wort, aber aba nicht!)
$= (ab)(abc) = (ab)((ab)c) = ((a)(b))(((a)(b))(c))$
(eine vollständige „Nicht-Standard-Innenzerlegung“ (siehe 2. Schritt!))

1.20 Definition Sei X eine geordnete Menge, $w \in X^*$. Ein Worte-Tupel

$$\mathfrak{z} = (z^{(1)}, \ldots, z^{(k)}) \in \mathcal{T}(X^+)$$

heißt eine Zerlegung von w, wenn gilt: $w = z^{(1)} \cdots z^{(k)}$. Die Tupellänge k heißt dabei die Faktorenzahl der Zerlegung $\mathfrak{z}$. Wir nennen $\mathfrak{z}$ monoton fallend, wenn gilt: $z^{(1)} \underset{lex}{\geq} \cdots \underset{lex}{\geq} z^{(k)}$.

1.20.1 *Ist w ein Lyndon-Wort, so ist (w) die einzige monoton fallende Zerlegung von w.*

Gäbe es nämlich eine monoton fallende Zerlegung $(z^{(1)}, \ldots, z^{(k)})$ von w mit $k \geq 2$, so folgte, da $z^{(k)}$ echter Rechtsfaktor von w wäre, nach 1.15

$$w \underset{lex}{<} z^{(k)} \underset{lex}{\leq} z^{(k-1)} \underset{lex}{\leq} \ldots \underset{lex}{\leq} z^{(1)} \underset{lex}{<} w,$$

ein Widerspruch. □

Gilt $z^{(1)}, \ldots, z^{(k)} \in \mathfrak{L}^X$, so heißt $\mathfrak{z}$ eine Zerlegung von w in Lyndon-Worte. Jedes Wort besitzt trivialerweise, nämlich vermöge seiner Zerlegung in Buchstaben, eine Zerlegung in Lyndon-Worte; bei dieser ist die Faktorenzahl gleich der Länge des Wortes und damit die maximal mögliche.

Eine monoton fallende Zerlegung eines Wortes w in Lyndon-Worte nennen wir eine Lyndon-Zerlegung von w. Als Vorbereitung des darauffolgenden Hauptresultates zeigen wir:

1.21 Lemma *Sei X eine geordnete Menge. Jedes Element von X^* hat höchstens eine Lyndon-Zerlegung.*

Beweis. Durch Induktion nach $\min\{k, l\}$ beweisen wir, daß zwei Lyndon-Zerlegungen $(z^{(1)}, \ldots, z^{(k)})$, $(y^{(1)}, \ldots, y^{(l)})$ eines Wortes $w \in X^*$ notwendig übereinstimmen: Ist $\min\{k, l\} \leq 1$, also $w = \imath$ bzw. $w \in \mathfrak{L}^X$, so ist $\emptyset$ bzw. – nach 1.20.1 – (w) die einzige monoton fallende Zerlegung von w, somit die Behauptung trivial. Es gelte also k, $l \geq 2$, und o.B.d.A. $l(z^{(1)}) \geq l(y^{(1)})$. Wegen $z^{(1)} \cdots z^{(k)} = w = y^{(1)} \cdots y^{(l)}$ gilt dann $y^{(1)} \upharpoonright z^{(1)}$. Wäre $z^{(1)} \neq y^{(1)}$,

so gäbe es ein größtes $j \in \underline{l-1}$, zu dem es ein $u \in X^*$ gibt mit $z^{(1)} = y^{(1)} \cdots y^{(j)} u$. Dann wäre $u \restriction z^{(1)}$, $z^{(1)} \neq u \underset{1.20.1}{\neq} \imath$, und $u \restriction y^{(j+1)}$. Mit Hilfe von 1.15 folgte

$$z^{(1)} \underset{lex}{<} u \underset{lex}{\leq} y^{(j+1)} \underset{lex}{\leq} \ldots \underset{lex}{\leq} y^{(1)} \underset{lex}{\leq} z^{(1)},$$

ein Widerspruch. Also gilt: $z^{(1)} = y^{(1)}$. Dann sind aber $(z^{(2)}, \ldots, z^{(k)})$, $(y^{(2)}, \ldots, y^{(l)})$ Lyndon-Zerlegungen des zu dem Linksfaktor $z^{(1)}$ gehörigen Rechtsfaktors von w, stimmen also nach Induktionsvoraussetzung überein. Es folgt die Behauptung. □

1.22 Satz *Sei X eine geordnete Menge, $w \in X^*$, $k \in \mathbb{N}_0$, $\mathfrak{z} = (z^{(1)}, \ldots, z^{(k)}) \in \mathcal{T}(X^+)$. Es sind äquivalent:*

(i) *$\mathfrak{z}$ ist eine Zerlegung von w in Lyndon-Worte mit minimaler Faktorenzahl.*

(ii) *$\mathfrak{z}$ ist eine Lyndon-Zerlegung von w.*

(iii) *$\mathfrak{z}$ ist eine Zerlegung von w, und für jedes $j \in \underline{k}$ ist $z^{(j)}$ der lexikographisch kleinste Rechtsfaktor $\neq \imath$ von $z^{(1)} \cdots z^{(j)}$.*

Für jedes Wort über X *existiert* insbesondere eine – und zwar nach 1.21 eindeutig bestimmte – Lyndon-Zerlegung. Bevor wir uns dem Beweis zuwenden, illustrieren wir den Satz durch ein einfaches Beispiel: Sei $X = \{a, b, c\}$ und $a < b < c$, $w = bbabbacababc$. Dann ist

ababc der lexikographisch kleinste Rechtsfaktor $\neq \imath$ von w,
abbac der lexikographisch kleinste Rechtsfaktor $\neq \imath$ von *bbabbac*,
b der lexikographisch kleinste Rechtsfaktor $\neq \imath$ von *bb*.

Nach 1.22 ist also $(b, b, abbac, ababc)$ die Lyndon-Zerlegung von w, durch Beklammerung kenntlich gemacht in der Darstellung $w = (b)(b)(abbac)(ababc)$. Der Teil (iii) des Satzes macht es möglich, die Lyndon-Zerlegung eines Wortes auf einfache Weise – wie im Beispiel vorgeführt – zu bestimmen. Während die Aussagen 1.15, 1.17, 1.18 auf die „Innenwelt“ der Lyndon-Worte eingingen, beschreibt 1.22 deren Verhältnis zur „Außenwelt“, d.h. zu ganz X^*: Jedes Wort über X läßt sich in der dort beschriebenen Weise, und zwar nach 1.21 eindeutig, als Produkt von Lyndon-Worten schreiben; die Eindeutigkeit wird erzwungen durch die Minimalität der Faktorenzahl oder aber auch durch das monotone Fallen (im Sinne der lexikographischen Ordnung) der Faktoren. In diesem Sinne (einem ganz anderen als dem vermöge der Buchstaben) können die Lyndon-Worte als „Atome“ beim multiplikativen Aufbau beliebiger Worte angesehen werden. Dies erinnert an die Rolle der Primzahlen beim

multiplikativen Aufbau von $\mathbb{N}$. Definiert man als Primfaktorzerlegung einer natürlichen Zahl n ein Tupel von Primzahlen in monoton fallender Reihenfolge, deren Produkt n ergibt, so zeigt sich: Mit der „Übersetzung“

$$\begin{aligned} X^* &\to \mathbb{N}, \\ \text{Lyndon-Wort} &\to \text{Primzahl}, \\ \underset{lex}{\geq} &\to \geq, \\ \text{Rechtsfaktor} &\to \text{Teiler}, \\ \imath &\to 1 \end{aligned}$$

entsteht aus der Äquivalenz von (ii) und (iii) im wesentlichen der Satz von der Existenz und Eindeutigkeit der Primfaktorzerlegung. So wie der kleinste Teiler $\neq 1$ einer natürlichen Zahl $\neq 1$ stets eine Primzahl ist, gilt (als unmittelbare Folge von 1.15):

1.22.1 *Der lexikographisch kleinste Rechtsfaktor $\neq \imath$ eines beliebigen Wortes $\neq \imath$ ist stets ein Lyndon-Wort.* □

Beweis von 1.22: Wir zeigen (i)⇒(ii) durch Kontraposition: Ist $k \geq 2$ und $z^{(j)} \underset{lex}{<} z^{(j+1)}$ für ein $j \in \underline{k-1}$, so $z^{(j)}z^{(j+1)} \in \mathfrak{L}^X$ nach 1.18 und daher $(z^{(1)}, \ldots, z^{(j-1)}, z^{(j)}z^{(j+1)}, z^{(j+2)}, \ldots, z^{(k)})$ eine Zerlegung der Länge $k-1$ von w in Lyndon-Worte. Aus ¬(ii) folgt also ¬(i).

Da w trivialerweise eine Zerlegung in Lyndon-Worte besitzt (jedenfalls ja die in Buchstaben), gibt es auch eine solche mit minimaler Faktorenzahl. Jede solche ist aber, wie die eben bewiesene Implikation lehrt, eine Lyndon-Zerlegung von w. Letztere ist nach 1.21 eindeutig bestimmt. Es folgt die Äquivalenz von (i) und (ii). Insbesondere wissen wir damit, daß w genau eine Lyndon-Zerlegung besitzt. Zum Beweis der Äquivalenz mit (iii) genügt es nun zu zeigen:

(∗) Gilt (iii), so ist $\mathfrak{z}$ monoton fallend:

Denn vermöge 1.22.1 ist die in (iii) beschriebene Zerlegung dann notwendig die Lyndon-Zerlegung von w.

Beweis von (∗): Für $k \leq 1$ ist (∗) trivial. Sei $k \geq 2$. Falls $z^{(j-1)} \underset{\text{lex}}{<} z^{(j)}$ für ein $j \in \underline{k} \setminus \{1\}$, so $z^{(j-1)}z^{(j)} \in \mathfrak{L}^X$ nach 1.22.1 und 1.18, also $z^{(j-1)}z^{(j)} \underset{\text{lex}}{<} z^{(j)}$ nach 1.15, im Widerspruch zur Definition von $z^{(j)}$. Damit ist (∗) bewiesen und der Beweis von 1.22 komplett. □

Ohne Beweis sei ein Resultat angegeben, das allgemein klärt, welche Teilmengen von X^* in dem oben beschriebenen Sinn wie $\mathfrak{L}^X$ als Analogon für die Menge der Primzahlen in $\mathbb{N}$ in Frage kommen. Es zeigt, daß die in 1.14.2 beschriebene Eigenschaft der Lyndon-Worte kein Zufall ist. Es handelt sich um eine Folgerung aus einem Satz von Schützenberger (1965):

Satz. *Sei X eine Menge und H eine vollständig geordnete Teilmenge von X^*. Jedes Element von X^* besitze genau eine monoton fallende Zerlegung in Faktoren aus H. Dann ist H ein Repräsentantensystem für die Konjugiertenklassen primitiver Worte in X^+.*

Für einen Beweis siehe [Reu], 7.2, [Lot], 5.4.

Kapitel 2

Verschiedene freie Strukturen

Wir stellen in der nachfolgenden Definition einige Begriffe zusammen, die für verschiedenste algebraische Strukturen (Monoide, Gruppen, Algebren z.B.) eine wichtige Rolle spielen. Daher gehen wir von einer beliebigen Klasse $\mathfrak{K}$ algebraischer Strukturen aus, von der wir zunächst nur annehmen wollen, daß sie gegen verknüpfungstreue Abbildungen („$\mathfrak{K}$-Homomorphismen“) und gegen Durchschnittsbildungen abgeschlossen ist; unter letzterem verstehen wir, daß für alle $S \in \mathfrak{K}$ und für alle nichtleeren Mengen $\mathcal{M}$ von zu $\mathfrak{K}$ gehörigen Unterstrukturen von S gilt: $\bigcap_{T\in\mathcal{M}} T \in \mathfrak{K}$. Ist S die Trägermenge einer beliebigen algebraischen Struktur und sind bezüglich der auf S gegebenen Verknüpfungen keine Unklarheiten zu erwarten, so bedeute die Schreibweise $T \leq_{\mathfrak{K}} S$, daß T eine zu $\mathfrak{K}$ gehörige Teilstruktur von S ist. Es kommt häufig vor, daß auf S mehrere Verknüpfungen gegeben sind und S bezüglich einiger von ihnen einer Klasse $\mathfrak{K}$, bezüglich anderer einer Klasse $\mathfrak{K}'$ angehört. So gehört z.B. jeder Körper zum einen der Klasse der kommutativen Ringe, zum anderen bezüglich der Addition der Klasse der abelschen Gruppen, bezüglich der Multiplikation der Klasse der kommutativen Halbgruppen an. In vielen Kontexten genügt es bei mehrfacher Klassenzugehörigkeit dieser Art, einfach nur die jeweilige Klasse zu erwähnen (statt der jeweils zuständigen Verknüpfungen), in bezug auf welche die Trägermenge betrachtet wird. Häufig wird eine der Verknüpfungen der zu $\mathfrak{K}$ gehörigen Strukturen traditionell als „Multiplikation“ bezeichnet. Dann bezeichnen wir mit $\mathfrak{K}_1$ die Klasse der zu $\mathfrak{K}$ gehörigen Strukturen, die ein multiplikativ neutrales Element besitzen. Zum Beispiel bezeichnen wir mit

$\mathfrak{S}$ die Klasse der Halbgruppen (Semigruppen),
$\mathfrak{S}_1$ die Klasse der Monoide.

Die Schreibweise $T \leq_{\mathfrak{K}_1} S$ verwenden wir nur, wenn S und T übereinstimmende multiplikativ neutrale Elemente enthalten und $T \leq_{\mathfrak{K}} S$ gilt. Ähnlich

verstehen wir unter einem $\mathfrak{K}_1$-Homomorphismus von $T \in \mathfrak{K}_1$ in eine Struktur T' mit multiplikativ neutralem Element $1_{T'}$ einen $\mathfrak{K}$-Homomorphismus von T in T', der das multiplikativ neutrale Element von T auf $1_{T'}$ abbildet[3]. Zum Beispiel ist ein $\mathfrak{S}_1$-Homomorphismus ein Monoid-Homomorphismus im Sinne von Kapitel 1. Das sogenannte leere Produkt (in additivem Kontext: die leere Summe) ist in einer Struktur genau dann erklärt, wenn sie ein neutrales Element enthält, und ist dann definitionsgemäß dieses. Alle diese Verabredungen dienen allein einer bequemen *Darstellung* in der Folge, wollen pragmatisch und nicht etwa als Vorschlag einer systematischen Theorie verstanden werden. Eher handelt es sich um eine für unsere Zwecke ausreichende Vorform kategorieller Denkweisen, mit der wir uns begnügen, um den Aufwand an dieser Stelle minimal zu halten.

2.1 Definition Sei $\mathfrak{K}$ eine durchschnittsabgeschlossene Klasse algebraischer Strukturen, $S \in \mathfrak{K}$ und $X \subseteq S$. Wir nennen

$$\langle X \rangle_{\mathfrak{K}} := \bigcap_{X \subseteq T \leq_{\mathfrak{K}} S} T$$

das $\mathfrak{K}$-Erzeugnis von X. X heißt ein $\mathfrak{K}$-Erzeugendensystem von S, wenn gilt: $\langle X \rangle_{\mathfrak{K}} = S$. X heißt $\mathfrak{K}$-unabhängig, wenn sich jede Abbildung φ von X in eine Struktur $T \in \mathfrak{K}$ zu genau einem $\mathfrak{K}$-Homomorphismus von $\langle X \rangle_{\mathfrak{K}}$ in T fortsetzen läßt[4]. Eine $\mathfrak{K}$-Basis von S ist ein $\mathfrak{K}$-unabhängiges $\mathfrak{K}$-Erzeugendensystem von S. Ist X eine $\mathfrak{K}$-Basis von S, so heißt S von X frei $\mathfrak{K}$-erzeugt (kurz: frei über X). Eine zu $\mathfrak{K}$ gehörige Struktur heißt frei (bezüglich $\mathfrak{K}$), wenn sie eine $\mathfrak{K}$-Basis besitzt. In problemloser Verallgemeinerung von 1.4 erhalten wir

2.1.1 *Seien S, $S' \in \mathfrak{K}$ von Teilmengen X bzw. X' frei $\mathfrak{K}$-erzeugt. Es gebe eine Bijektion φ von X auf X'. Dann gibt es einen eindeutig bestimmten $\mathfrak{K}$-Isomorphismus $\bar{\varphi}$ von S auf S' mit $x\bar{\varphi} = x\varphi$ für alle $x \in X$.* □

(Im Falle $X = X'$, $\varphi = id$ nennt man einen Isomorphismus mit der letztgenannten Eigenschaft einen X-Isomorphismus.) Aufgrund von 2.1.1 spricht man, sofern die Existenz gesichert ist, von *der* von X frei $\mathfrak{K}$-erzeugten Struktur; denn eine solche ist bis auf X-Isomorphie eindeutig bestimmt.

[3] Ein solcher Homomorphismus heißt unital.

[4] Für zahlreiche wichtige Klassen $\mathfrak{K}$ gilt, daß sich das $\mathfrak{K}$-Erzeugnis als Abschluß von X gegenüber in S gegebenen Verknüpfungen beschreiben läßt, die von $\mathfrak{K}$-Homomorphismen respektiert werden müssen. Dies ist z. B. für die Klassen der Monoide, Gruppen, Vektorräume, Algebren der Fall. Dann gibt es trivialerweise stets *höchstens* eine Fortsetzung von φ zu einem $\mathfrak{K}$-Homomorphismus von $\langle X \rangle_{\mathfrak{K}}$ in T, so daß die $\mathfrak{K}$ -Unabhängigkeit von X auf die bloße *Existenz* einer Fortsetzung von φ hinausläuft.

Wir legen die folgenden Bezeichnungen für Klassen von Strukturen fest, die im folgenden eine Rolle spielen werden. Es sei

$\mathfrak{G}$ die Klasse der Gruppen,
$\mathfrak{M}$ die Klasse der abelschen Gruppen,
$\mathfrak{R}$ die Klasse der Ringe.

Zum Beispiel sind $\{1\}$, $\{-1\}$ $\mathfrak{G}$-Basen, auch $\mathfrak{M}$-Basen von $(\mathbb{Z},+)$. Es sind aber keine $\mathfrak{S}_1$-Basen, da sie zwar $\mathfrak{S}_1$-unabhängig, aber keine $\mathfrak{S}_1$-Erzeugendensysteme von $(\mathbb{Z},+)$ sind. Ein $\mathfrak{S}_1$-Erzeugendensystem von $(\mathbb{Z},+)$ wäre z.B. $\{1,-1\}$, aber dieses ist nicht $\mathfrak{S}_1$-unabhängig: Wählen wir etwa $\varphi : \{1,-1\} \to \mathbb{Z}$, $1\varphi = 1 = (-1)\varphi$, so hat φ keine Fortsetzung $\bar{\varphi}$ zu einem additiven Homomorphismus von $\mathbb{Z}$, denn sonst müßte gelten:

$$0 = 0\bar{\varphi} = (1+(-1))\bar{\varphi} = 1\bar{\varphi} + (-1)\bar{\varphi} = 2,$$

ein Widerspruch.– Die Klasse aller Körper enthält z.B. überhaupt kein freies Objekt. Bezüglich der Klasse $\mathfrak{S}_1$ ist X^* frei (siehe vor 1.5). Da jede Halbgruppe durch Hinzufügen eines einzigen Elementes auf triviale Weise zu einem Monoid erweitert werden kann, erhält man daraus leicht, daß die Halbgruppe X^+ bezüglich der Klasse $\mathfrak{S}$ frei ist.

Problem. Bei gegebener Klasse $\mathfrak{K}$ algebraischer Strukturen entscheide man, ob es zu jeder Menge X eine von X frei $\mathfrak{K}$-erzeugte Struktur gibt.

2.2 Definition Sei G eine Gruppe, $X \subseteq G$, $g \in G$. Ein Tupel

$$((x_1,\varepsilon_1),\dots,(x_k,\varepsilon_k)) \in \mathcal{T}(X \times \{1,-1\}) \text{ mit } x_1^{\varepsilon_1}\cdots x_k^{\varepsilon_k} = g$$

heißt eine X-Darstellung von g. Sie heißt gekürzt, wenn für alle $i \in \underline{k-1}$ mit $x_i = x_{i+1}$ gilt: $\varepsilon_i = \varepsilon_{i+1}$. Ob eine X-Darstellung gekürzt ist, entscheidet sich also allein durch die Betrachtung *aufeinanderfolgender Paare*. Ob man etwa in der Gleichung $x_1^{\varepsilon_1}\cdots x_k^{\varepsilon_k} = g$ noch andere (wie auch immer geartete) „Kürzungsmöglichkeiten“ (im intuitiven Sinne) erkennt, ist dagegen völlig ohne Belang.

2.2.1 *$\emptyset$ ist eine gekürzte X-Darstellung von 1_G.* □

2.2.2 *Ist $((x_1,\varepsilon_1),\dots,(x_k,\varepsilon_k))$ eine (gekürzte) X-Darstellung von g, so ist $((x_k,-\varepsilon_k),\dots,(x_1,-\varepsilon_1))$ eine (gekürzte) X-Darstellung von g^{-1}.* □

2.2.3 *Ist $((x_1,\varepsilon_1),\dots,(x_k,\varepsilon_k))$ eine X-Darstellung von g, so gibt es ein $r \in \underline{k} \cup \{0\}$ und Indizes $i_1 < \cdots < i_r$, so daß $((x_{i_1},\varepsilon_{i_1}),\dots,(x_{i_r},\varepsilon_{i_r}))$ eine gekürzte X-Darstellung von g ist,*

wie man durch triviale Induktion nach k beweist. $\square$

2.3 Proposition *Seien G eine Gruppe, $X \subseteq G$ und g, $g' \in G$. Es seien $((x_1, \varepsilon_1), \ldots, (x_k, \varepsilon_k))$ bzw. $((x'_1, \varepsilon'_1), \ldots, (x'_l, \varepsilon'_l))$ gekürzte X-Darstellungen von g bzw. g'. Dann gibt es ein $j \in \underline{k} \cup \{0\}$, so daß*

$$((x_1, \varepsilon_1), \ldots, (x_j, \varepsilon_j), (x'_{1+k-j}, \varepsilon'_{1+k-j}), \ldots, (x'_l, \varepsilon'_l))$$

eine gekürzte X-Darstellung von gg' ist und

$$x_k^{\varepsilon_k} {x'_1}^{\varepsilon'_1} = \cdots = x_{j+1}^{\varepsilon_{j+1}} {x'_{k-j}}^{\varepsilon'_{k-j}} = 1_G$$

gilt.

Verwendet man die beiden gegebenen gekürzten X-Darstellungen zu einer Produkt-Darstellung von gg', so läßt sich möglicherweise kürzen, jedoch höchstens dort, wo das Ende der Produkt-Darstellung von g mit dem Anfang der Produkt-Darstellung von g' zusammenstößt. Kürzt man von dort aus sukzessive so weit es geht, bleibt eine nicht weiter kürzbare Produkt-Darstellung von gg' übrig, bestehend aus einem Anfangsstück der Produkt-Darstellung von g und einem Endstück der Produkt-Darstellung von g':

$$gg' = x_1^{\varepsilon_1} \cdots x_j^{\varepsilon_j} \underbrace{\vdots\, x_{j+1}^{\varepsilon_{j+1}} \cdots \underbrace{x_k^{\varepsilon_k} {x'_1}^{\varepsilon'_1}}_{=1_G} \cdots {x'_{k-j}}^{\varepsilon'_{k-j}} \vdots {x'_{k-j+1}}^{\varepsilon'_{k-j+1}} \cdots {x'_l}^{\varepsilon'_l}}_{=1_G}.$$

Der formal korrekte Beweis besteht aus einer einfachen Induktion nach $\min\{k, l\}$: Ist $k = 0$ oder $l = 0$, so ist die Behauptung trivial (man setzt $j := 0$ bzw. $j := k$), ebenso falls $x_k \neq x'_1$ oder $x_k = x'_1$, $\varepsilon_k = \varepsilon'_1$ (man setzt $j := k$). Falls aber $x_k = x'_1$, $\varepsilon_k = -\varepsilon'_1$, so setzen wir $h := x_1^{\varepsilon_1} \cdots x_{k-1}^{\varepsilon_{k-1}}$, $h' := {x'_2}^{\varepsilon'_2} \cdots {x'_l}^{\varepsilon'_l}$ und wenden die Induktions-Voraussetzung auf die gekürzten X-Darstellungen $((x_1, \varepsilon_1), \ldots, (x_{k-1}, \varepsilon_{k-1}))$, $((x'_2, \varepsilon'_2), \ldots, (x'_l, \varepsilon'_l))$ von h bzw. h' an. Wegen $gg' = h x_k^{\varepsilon_k} {x'_1}^{\varepsilon'_1} h' = hh'$ folgt die Behauptung. $\square$

2.4 Proposition *Sei G eine Gruppe und $X \subseteq G$.*

(1) *Für alle $g \in G$ sind äquivalent:*

- (i) $g \in \langle X \rangle_{\mathfrak{G}}$,
- (ii) *g besitzt eine X-Darstellung,*
- (iii) *g besitzt eine gekürzte X-Darstellung*

(2) *Es sind äquivalent:*

(i) *X ist $\mathfrak{G}$-unabhängig,*

(ii) *$\emptyset$ ist die einzige gekürzte X-Darstellung von 1_G,*

(iii) *jedes $g \in G$ hat höchstens eine gekürzte X-Darstellung.*

Beweis. (1) Sei $g \in G$. Die Äquivalenz von (1ii) und (1iii) folgt aus 2.2.3.

(1ii)⇒(1i): Ist $H \leq_{\mathfrak{G}} G$ und $X \subseteq H$, so gilt für alle $x_1, \ldots, x_k \in X$, $\varepsilon_1, \ldots, \varepsilon_k \in \{1, -1\}$: $x_1^{\varepsilon_1} \cdots x_k^{\varepsilon_k} \in H$. Aus (1ii) folgt also: $g \in \bigcap_{X \subseteq H \leq G} H = \langle X \rangle_{\mathfrak{G}}$.

(1i)⇒(1ii): Sei G_0 die Menge aller Elemente $h \in G$, die eine X-Darstellung besitzen. Dann gilt $X \subseteq G_0$ und $G_0 \leq_{\mathfrak{G}} G$, denn G_0 enthält 1_G (nach 2.2.1) und ist (nach 2.3 und 2.2.2) gegen Produkt- und Inversenbildung abgeschlossen. Es folgt: $\bigcap_{X \subseteq H \leq G} H \subseteq G_0$, d.h. $\langle X \rangle_{\mathfrak{G}} \subseteq G_0$.

Der Beweis von (2) ist weniger trivial: Um zunächst (2ii)⇒ (2iii) zu zeigen, betrachten wir ein beliebiges Element $g \in G$ mit gekürzten X-Darstellungen $((x_1, \varepsilon_1), \ldots, (x_k, \varepsilon_k))$, $((x'_1, \varepsilon'_1), \ldots, (x'_l, \varepsilon'_l))$. Dann gilt:

$$x_1^{\varepsilon_1} \cdots x_k^{\varepsilon_k} {x'_l}^{-\varepsilon'_l} \cdots {x'_1}^{-\varepsilon'_1} = gg^{-1} = 1_G.$$

Nach 2.3 gibt es ein $j \in \underline{k} \cup \{0\}$, so daß

$$((x_1, \varepsilon_1), \ldots, (x_{k-j}, \varepsilon_{k-j}), (x'_{l-j}, -\varepsilon'_{l-j}), \ldots, (x'_1, -\varepsilon'_1))$$

eine gekürzte X-Darstellung von 1_G ist und

$$x_k^{\varepsilon_k} {x'_l}^{-\varepsilon'_l} = \cdots = x_{k-j+1}^{\varepsilon_{k-j+1}} {x'}_{l-j+1}^{-\varepsilon'_{l-j+1}} = 1_G.$$

Nach Voraussetzung folgt damit zunächst $l = j = k$, und ferner ist $((x_i, \varepsilon_i), (x'_i, -\varepsilon'_i))$ für jedes $i \in \underline{k}$ eine X-Darstellung von 1_G, mithin $x_i = x'_i$, $\varepsilon_i \neq -\varepsilon'_i$, d.h. $(x_i, \varepsilon_i) = (x'_i, \varepsilon'_i)$ für alle $i \in \underline{k}$, da ε_i, $\varepsilon'_i \in \{1, -1\}$.

(2iii)⇒(2i): Gilt (2iii), so ist insbesondere $\emptyset$ die einzige gekürzte X-Darstellung von 1_G; d.h. es gilt jedenfalls (2ii).– Sei φ eine Abbildung von X in eine Gruppe H. Zu jedem Element $g \in \langle X \rangle_{\mathfrak{G}}$ gibt es dann nach Voraussetzung und (1) genau eine gekürzte X-Darstellung $((x_1, \varepsilon_1), \ldots, (x_k, \varepsilon_k))$. Wir setzen $g\bar{\varphi} := (x_1\varphi)^{\varepsilon_1} \cdots (x_k\varphi)^{\varepsilon_k}$ und zeigen, daß $\bar{\varphi}$ ein Homomorphismus von $\langle X \rangle_{\mathfrak{G}}$ in H ist; daß *höchstens* diese als homomorphe Fortsetzung von φ auf $\langle X \rangle_{\mathfrak{G}}$ in Frage kommt, ist klar. Ist $g' \in \langle X \rangle_{\mathfrak{G}}$ und $((x'_1, \varepsilon'_1), \ldots, (x'_l, \varepsilon'_l))$ die gekürzte X-Darstellung von g', so gibt es nach 2.3 ein $j \in \underline{k} \cup \{0\}$, so daß

$$((x_1, \varepsilon_1), \ldots, (x_j, \varepsilon_j), (x'_{1+k-j}, \varepsilon'_{1+k-j}), \ldots, (x'_l, \varepsilon'_l))$$

die gekürzte X-Darstellung von gg' ist, und $x_k^{\varepsilon_k} {x'_1}^{\varepsilon'_1} = \cdots = x_{j+1}^{\varepsilon_{j+1}} {x'_{k-j}}^{\varepsilon'_{k-j}} = 1_G$ gilt, woraus

$$(x_k, \varepsilon_k) = (x'_1, -\varepsilon'_1), \ldots, (x_{j+1}, \varepsilon_{j+1}) = (x'_{k-j}, -\varepsilon'_{k-j})$$

folgt, da wir ja (2ii) anwenden dürfen. Damit gilt: $(x_k\varphi)^{\varepsilon_k} = (x'_1\varphi)^{-\varepsilon'_1}, \ldots,$ $(x_{j+1}\varphi)^{\varepsilon_{j+1}} = (x'_{k-j}\varphi)^{-\varepsilon'_{k-j}}$, und daher

$$\begin{aligned}(gg')\bar{\varphi} &= (x_1\varphi)^{\varepsilon_1} \cdots (x_j\varphi)^{\varepsilon_j}(x'_{1+k-j}\varphi)^{\varepsilon'_{1+k-j}} \cdots (x'_l\varphi)^{\varepsilon'_l} \\ &= (x_1\varphi)^{\varepsilon_1} \cdots (x_k\varphi)^{\varepsilon_k}(x'_1\varphi)^{\varepsilon'_1} \cdots (x'_l\varphi)^{\varepsilon'_l} \\ &= g\bar{\varphi} \cdot g'\bar{\varphi}.\end{aligned}$$

(2i)⇒(2ii) (nach Schreier): Wir werden Gebrauch von der folgenden trivialen Vorbemerkung machen:

(∗) Sei $m \in \mathbb{N}$. Jede injektive Funktion von einer Teilmenge von $\underline{m}$ in $\underline{m}$ läßt sich zu einer Permutation von $\underline{m}$ fortsetzen.

Sei $((x_1, \varepsilon_1), \ldots, (x_n, \varepsilon_n))$ eine gekürzte X-Darstellung von 1_G. Unser Ziel ist es, $n = 0$ zu beweisen. Die Idee dazu ist es, jedem x_i eine Permutation der Menge $\underline{n+1}$ so zuzuordnen, daß $x_i^{\varepsilon_i}$ bei dieser Zuordnung die Zahl i auf $i+1$ abbildet. Das Bild von 1 bei der Hintereinanderausführung ist dann $n+1$, andererseits aber 1; also gilt $n = 0$:

Für alle $x \in X$ sei

$$R_x := \{(i, i+1) | i \in \underline{n},\ (x_i, \varepsilon_i) = (x, 1)\} \cup \{(j+1, j) | j \in \underline{n},\ (x_j, \varepsilon_j) = (x, -1)\}$$

$(\subseteq \underline{n+1} \times \underline{n+1})$. Dann ist die Relation R_x links- und rechtseindeutig, also eine injektive Funktion:

Zur Rechtseindeutigkeit (d.h., Funktionseigenschaft): Gäbe es ein $i \in \underline{n}$ mit $(i, i+1), (i, i-1) \in R_x$, so wäre jedenfalls $i > 1$ und $(x_i, \varepsilon_i) = (x, 1)$, $(x_{i-1}, \varepsilon_{i-1}) = (x, -1)$, also $x_{i-1} = x_i$, $\varepsilon_{i-1} \neq \varepsilon_i$, ein Widerspruch.

Zur Linkseindeutigkeit (d.h., Injektivität): Gäbe es ein $i \in \underline{n-1}$ mit $(i, i+1)$, $(i+2, i+1) \in R_x$, so wäre $(x_i, \varepsilon_i) = (x, 1)$, $(x_{i+1}, \varepsilon_{i+1}) = (x, -1)$, also $x_i = x_{i+1}$, $\varepsilon_i \neq \varepsilon_{i+1}$, ein Widerspruch.

Nach (∗) gibt es eine Permutation π_x von $\underline{n+1}$, die R_x enthält. (Für alle $x \in X \setminus \{x_1, \ldots, x_n\}$ können wir $\pi_x = id$ setzen.) Es gilt:

(∗∗) $k\pi_{x_k}^{\varepsilon_k} = k+1$ für alle $k \in \underline{n}$.

Ist nämlich $\varepsilon_k = 1$, so $(k, k+1) \in R_{x_k} \subseteq \pi_{x_k} = \pi_{x_k}^{\varepsilon_k}$. Ist aber $\varepsilon_k = -1$, so $(k+1, k) \in R_{x_k} \subseteq \pi_{x_k}$, folglich $(k, k+1) \in \pi_{x_k}^{\varepsilon_k}$.

Sei nun $\varphi : X \to S_{\underline{n+1}}$, $x \mapsto \pi_x$. Nach Voraussetzung existiert ein $\mathfrak{G}$-Homomorphismus $\bar{\varphi}$ von $\langle X\rangle_{\mathfrak{G}}$ in $S_{\underline{n+1}}$ mit $\bar{\varphi}|_X = \varphi$. Es folgt:

$$id = 1_G\bar{\varphi} = (x_1^{\varepsilon_1} \cdots x_n^{\varepsilon_n})\bar{\varphi} = (x_1\varphi)^{\varepsilon_1} \cdots (x_n\varphi)^{\varepsilon_n} = \pi_{x_1}^{\varepsilon_1} \cdots \pi_{x_n}^{\varepsilon_n},$$

damit $1 = 1\,id = 1\pi_{x_1}^{\varepsilon_1} \cdots \pi_{x_n}^{\varepsilon_n} \underset{(**)}{=} n+1$, also $n = 0$. □

Entsprechend einigen Überlegungen zu freien Monoiden im vorigen Kapitel formulieren wir drei Fragen, die uns noch einige Zeit beschäftigen werden: (1) Gibt es zu jeder Menge X eine von X frei erzeugte Gruppe? [vgl. die Anwendung des Erweiterungsprinzips vor 1.5] (2) Welche Beziehungen bestehen zwischen zwei freien Erzeugendensystemen einer freien Gruppe? [vgl. 1.7(2)] (3) Sind Untergruppen freier Gruppen wieder frei? [vgl. 1.7.2] Die Klasse der Gruppen verhält sich, wie die späteren Antworten auf diese Fragen zeigen werden, in interessanter Weise anders als die der Monoide. $(\mathbb{Z}, +)$ ist eine freie Gruppe, die zwei freie Erzeugendensysteme hat, nämlich $\{1\}$ und $\{-1\}$: Ist H irgendeine Gruppe, $h \in H$, so ist die Abbildung $\{1\} \to H$, $1 \mapsto h$ zu genau einem $\mathfrak{G}$-Homomorphismus von $\mathbb{Z}$ in H fortsetzbar (dessen Bild die Untergruppe $\langle h\rangle$ von H ist); ebenso die Abbildung $\{-1\} \to H$, $-1 \mapsto h$. Leicht zu sehen ist außerdem, daß $\{1\}$ und $\{-1\}$ die beiden einzigen freien Erzeugendensysteme von $\mathbb{Z}$ sind. Jede freie Gruppe, die nicht nur aus dem neutralen Element besteht, „involviert“ notwendig $\mathbb{Z}$:

2.4.1 *Ist X eine Menge und F eine von X frei erzeugte Gruppe, so gilt $\langle x\rangle \cong \mathbb{Z}$ für alle $x \in X$.*

Hätte nämlich $x \in X$ eine endliche Ordnung n, so wäre $((x,1), \ldots, (x,1))$ aufgrund der Gleichung $x^n = 1_F$ eine gekürzte X-Darstellung von 1_F, ein Widerspruch zu 2.4(2). □

Später werden wir sehen, daß in einer freien Gruppe *jedes* Element mit Ausnahme des neutralen eine zu $\mathbb{Z}$ isomorphe Gruppe erzeugt.

2.4.2 *Ist $|X| = 1$, so gibt es eine von X frei erzeugte Gruppe. Sie ist isomorph zu $(\mathbb{Z}, +)$.*

Ist nämlich $X = \{x\}$, $\iota\colon X \to \mathbb{Z}$, $x \mapsto 1$, so gibt es nach dem Erweiterungsprinzip und dem Zusatz dazu eine Menge U mit $X \subseteq U$, eine Verknüpfung $\cdot$ auf U und einen Isomorphismus $\bar{\iota}\colon (U, \cdot) \to (\mathbb{Z}, +)$ mit $\bar{\iota}|_X = \iota$. Es folgt , daß U eine von X frei erzeugte Gruppe ist. □

So wie $\mathbb{Z}$ der „Großvater“ aller zyklischen Gruppen ist (jede zyklische Gruppe ist isomorph zu einer Faktorgruppe von $\mathbb{Z}$), ist allgemeiner eine freie Gruppe mit einem n-elementigen freien Erzeugendensystem „Großvater“ aller Gruppen, die mit n Elementen erzeugbar sind. Es gilt:

2.4.3 *Ist F eine von X frei erzeugte Gruppe und H irgendeine Gruppe, die von einer Teilmenge Y erzeugt wird mit $|Y| \leq |X|$, so ist H zu einer Faktorgruppe von F isomorph.*

Denn genauer läßt sich jede surjektive Abbildung φ von X auf Y zu einem $\mathfrak{G}$-Homomorphismus $\bar{\varphi}$ von F in H fortsetzen, dessen Bild folglich das Erzeugendensystem Y enthält und der damit surjektiv sein muß. Nach dem Homomorphiesatz gilt dann: $H \cong F/\mathrm{Kern}\,\bar{\varphi}$. □

Wir werden später einige interessante Einblicke in die innere Struktur freier Gruppen gewinnen. Der in 2.4.3 angesprochene Aspekt wird dabei keine wichtige Rolle spielen. Er stellt jedoch die (simple) Grundlage eines (ganz und gar nicht simplen) Zweigs der Gruppentheorie dar, in der beliebige Gruppen als Faktorgruppen freier Gruppen studiert werden. Eine von dessen Hauptfragen ist z.B., ob man Y und φ besonders „geschickt“ wählen kann, so daß die gegebene Gruppe H in besonders übersichtlicher Form als Faktorgruppe der freien Gruppe in Erscheinung tritt. Man nennt eine solche „Realisierung“ einer Gruppe H als Faktorgruppe einer freien Gruppe eine Präsentation von H.

Bevor wir mit dem Studium von Klassen algebraischer Strukturen einsetzen, die eine zentrale Rolle in der Algebra spielen, betrachten wir vorweg die in gewissem Sinne „allgemeinste“ Strukturklasse, in der nämlich für die Verknüpfung überhaupt keine Bedingungen verlangt werden:

2.5 Definition Ein Magma ist ein Paar $(S, \cdot)$, bei dem S eine Menge, $\cdot$ eine Verknüpfung auf S ist. Sei $\mathfrak{Ma}$ die Klasse aller Magmen. Ist S ein Magma, $X \subseteq S$, so setzen wir induktiv

$$X^{(1)} := X,$$
$$X^{(n)} := X^{(1)}X^{(n-1)} \cup X^{(2)}X^{(n-2)} \cup \cdots \cup X^{(n-1)}X^{(1)} \text{ für alle } n \in \mathbb{N} \setminus \{1\},$$

wobei $AB := \{\, ab \mid a \in A, b \in B\}$ für alle A, $B \subseteq S$. Im Falle, daß S ein neutrales Element 1_S enthält, setzen wir zusätzlich $X^{(0)} := \{1_S\}$.

2.5.1 *Sei X eine Teilmenge eines Magmas S.*

(a) $\langle X\rangle_{\mathfrak{Ma}} = \bigcup\limits_{n\in\mathbb{N}} X^{(n)}$,

(b) $\langle X\rangle_{\mathfrak{Ma}_1} = \bigcup\limits_{n\in\mathbb{N}_0} X^{(n)}$ *falls* $S \in \mathfrak{Ma}_1$,

denn ein Teilmagma S_0 von S mit $X \subseteq S_0$ enthält alle Mengen $X^{(n)}$ $(n \in \mathbb{N})$, also auch $\bigcup\limits_{n\in\mathbb{N}} X^{(n)}$. Da $\bigcup\limits_{n\in\mathbb{N}} X^{(n)}$ ein X enthaltendes Teilmagma von S ist, folgt (a). Unmittelbare Folge ist (b). □

2.6 Proposition *Zu jeder Menge X gibt es ein von X frei erzeugtes Magma.*

Beweis. Zunächst sei Y eine zu X gleichmächtige Menge, die kein Kuratowski-Paar (u, v) (also kein Element der Form $\{\{u\}, \{u, v\}\}$ enthält. [Dazu genügt es z.B., für Y die Menge der 3-Tupel (x, x, x) mit $x \in X$ zu betrachten, denn nach Definition besteht jedes 3-Tupel aus drei verschiedenen Kuratowski-Paaren, ist also 3-elementig, während Kuratowski-Paare höchstens 2-elementig sind.] Wir setzen induktiv

$$Y^{[1]} := Y,$$
$$Y^{[n]} := \bigcup_{i \in \underline{n-1}} Y^{[i]} \times Y^{[n-i]} \text{ für alle } n \in \mathbb{N} \smallsetminus \{1\},$$

und $\mathcal{M}(Y) := \bigcup_{n \in \mathbb{N}} Y^{[n]}$. Wir zeigen

$(*)$ Für alle $m, n \in \mathbb{N}$ gilt: $Y^{[m]} \cap Y^{[n]} \neq \emptyset \Rightarrow m = n$.

Wir beweisen $(*)$ durch Induktion nach $\min\{m, n\}$. Ist $\min\{m, n\} = 1$ und dabei o.B.d.A. $m = 1$, $n > 1$, so gilt $Y^{[m]} \cap Y^{[n]} = Y \cap Y^{[n]} = \emptyset$, da $Y^{[n]}$ nur Paare, Y aber kein einziges Paar enthält. Zum Induktionsschritt sei nun $\min\{m, n\} > 1$, etwa $m \leq n$ und $w \in Y^{[m]} \cap Y^{[n]}$. Dann existieren $i \in \underline{m-1}$, $j \in \underline{n-1}$ mit $w \in Y^{[i]} \times Y^{[m-i]}$ und $w \in Y^{[j]} \times Y^{[n-j]}$. Die erste Komponente des Paars w liegt daher sowohl in $Y^{[i]}$ als auch in $Y^{[j]}$, so daß also $Y^{[i]} \cap Y^{[j]} \neq \emptyset$, wegen $i < m$ nach Induktionsvoraussetzung folglich $i = j$ gilt. Ebenso erhält man durch Betrachtung der zweiten Komponenten (da auch $m - i < m$ gilt): $m - i = n - j$. Damit folgt: $m = (m-i)+i = (n-j)+j = n$, und $(*)$ ist bewiesen.

Wir definieren einen Verknüpfung $\cdot$ auf $\mathcal{M}(Y)$ durch:

$$v \cdot w := (v, w),$$

[das ist nichts anderes als die *Identität* auf $\mathcal{M}(Y) \times \mathcal{M}(Y)$!] Wir zeigen:

$(**)$ Y ist eine $\mathfrak{Ma}$-Basis von $\mathcal{M}(Y)$.

Aus der Definition von $Y^{[n]}$ folgt durch Induktion nach n: $Y^{[n]} = Y^{(n)}$ für alle $n \in \mathbb{N}$, also $\langle Y \rangle_{\mathfrak{Ma}} = \mathcal{M}(Y)$ nach 2.5.1. Es bleibt zu zeigen, daß Y $\mathfrak{Ma}$-unabhängig ist. Dazu sei $(T, \circ)$ ein Magma und φ eine Abbildung von Y in T. Induktiv definieren wir für jedes $n \in \mathbb{N}$ eine Abbildung $\overline{\varphi_n}$ von $\bigcup_{m \leq n} Y^{[m]}$ in T: Für alle $y \in Y$ sei $y\overline{\varphi_1} := y\varphi$. Sei nun $n > 1$ und $\overline{\varphi_m}$ bereits für alle $m < n$ definiert. Ist $w \in Y^{[n]}$, so gibt es (wegen $n > 1$) ein $i \in \underline{n-1}$ mit $w \in Y^{[i]} \times Y^{[n-i]}$, und nach $(*)$ gilt: $w \notin \bigcup_{m<n} Y^{[m]}$. Seien $u \in Y^{[i]}$,

$v \in Y^{[n-i)}$ mit $w = (u, v)$. Dann sind u, v als Komponenten des Paars w und die Zahl i nach $(*)$ eindeutig bestimmt, so daß die Setzung $w\overline{\varphi_n} := u\overline{\varphi_i} \circ v\overline{\varphi_{n-i}}$ wohldefiniert ist. Für alle $w \in Y^{[m)}$ mit $m < n$ sei $w\overline{\varphi_n} := w\overline{\varphi_m}$. Dann ist $\overline{\varphi_n}$ eine Fortsetzung von $\overline{\varphi_{n-1}}$, für alle $n \in \mathbb{N}$. Nun sei $\overline{\varphi} := \bigcup_{n\in\mathbb{N}} \overline{\varphi_n}$. Dann ist $\overline{\varphi}$ die Abbildung von $\mathcal{M}(Y)$ in T, die jedes Element w auf $w\overline{\varphi_n}$ abbildet, wenn n der Index ist mit $w \in Y^{[n)}$. Offenbar gilt $\overline{\varphi}|_Y = \varphi$, und für alle u, $v \in \mathcal{M}(Y)$ gilt

$$(u \cdot v)\overline{\varphi} = (u, v)\overline{\varphi} = u\overline{\varphi_i} \circ v\overline{\varphi_j} = u\overline{\varphi} \circ v\overline{\varphi},$$

wobei i, j die natürlichen Zahlen sind mit $u \in Y^{[i)}$, $v \in Y^{[j)}$. Also gilt $(**)$.

Damit sind die Voraussetzungen für den routinemäßigen Einsatz des Erweiterungsprinzips gegeben: Sei ι eine Bijektion von X auf Y. Dann gibt es ein X enthaltendes, zu $\mathcal{M}(Y)$ isomorphes Magma M und eine Fortsetzung von ι zu einem Isomorphismus $\overline{\iota}$ von M auf $\mathcal{M}(Y)$. Wegen $X\overline{\iota} = X\iota = Y$ ist daher M ein von X frei erzeugtes Magma. □

Das nach 2.1.1 bis auf X-Isomorphie eindeutig bestimmte freie Magma über X bezeichnen wir mit $X^{(+)}$. Es gilt: $X^{(+)} = \dot{\bigcup}_{n\in\mathbb{N}} X^{(n)}$, und das für jedes $w \in X^{(+)}$ eindeutig bestimmte $n \in \mathbb{N}$ mit $w \in X^{(n)}$ nennen wir den Grad von w. Das Produkt eines Elements u mit einem Element v von $X^{(+)}$ schreiben wir in der Form (uv). Für kleine Grade ist es leicht, eine komplette Elementliste aufzustellen:

Grad 1: x
Grad 2: (x_1x_2)
Grad 3: $((x_1x_2)x_3)$, $(x_1(x_2x_3))$
Grad 4: $(((x_1x_2)x_3)x_4)$, $(x_1((x_2x_3)x_4))$, $((x_1x_2)(x_3x_4))$, $((x_1(x_2x_3))x_4)$, $(x_1(x_2(x_3x_4)))$

(wobei x, x_1, x_2, x_3, $x_4 \in X$). Aus ein und derselben n-stelligen Aufeinanderfolge von Elementen von X lassen sich so viele verschiedene Elemente von $X^{(+)}$ bilden, wie es sinnvolle vollständige Beklammerungen von ihr gibt.

In dem im Beweis von 2.6 konstruierten Magma $\mathcal{M}(Y)$ dient die Kuratowski-Paar-Bildung als Verknüpfung. Da Komponenten von Kuratowski-Paaren stets eindeutig bestimmt sind, folgt damit für das freie Magma $X^{(+)}$ über X:

2.6.1 *Zu jedem $t \in X^{(+)} \setminus X$ gibt es eindeutig bestimmte Elemente r, $s \in X^{(+)}$ mit $t = (rs)$. Insbesondere ist X die einzige $\mathfrak{Ma}$-Basis von $X^{(+)}$.* □

Ist T ein Teilmagma von $X^{(+)}$ und $U(T) := T \setminus T^2$, so folgt $T = \langle U(T)\rangle_{\mathfrak{Ma}} \cong U(T)^{(+)}$; letzteres nach 2.6.1. Es folgt:

2.6.2 *Jedes Teilmagma eines freien Magmas ist frei.* □

Durch Hinzufügen genau eines Elements zu $X^{(+)}$, das als neutral definiert wird, erhält man ein unitäres Magma, welches offensichtlich bezüglich der Klasse aller unitären Magmen von X frei erzeugt wird. Wir bezeichnen dieses von X frei erzeugte unitäre Magma mit $X^{(*)}$. Wir haben einen kanonischen Epimorphismus von $X^{(*)}$ auf das freie Monoid X^*, der bei den hier eingeführten üblichen Schreibweisen einfach durch das „Weglassen der Klammern“ gegeben ist, ebenso von $X^{(+)}$ auf X^+. Der Homomorphiesatz ergibt daher:

2.6.3 *Für jede Menge X ist das freie Monoid X^* zu einer Faktorstruktur des freien unitären Magmas $X^{(*)}$, die freie Halbgruppe X^+ zu einer Faktorstruktur des freien Magmas $X^{(+)}$ isomorph.* □

Dem Freiheitsbegriff begegnet man im Mathematik-Studium schon früh, nämlich beim Studium von Vektorräumen über Körpern, wenngleich jener in der Linearen Algebra in der Regel nicht thematisiert wird und auch der Name nicht fällt: Bezüglich der Klasse ${}^K\mathfrak{M}$ der (Links-)Vektorräume über einem Körper K sehen wir die K-linearen Abbildungen als die ${}^K\mathfrak{M}$-Homomorphismen an. Ist $V \in {}^K\mathfrak{M}$ und B eine K-Basis von V, so läßt sich jede Abbildung von B in einen beliebigen K-Vektorraum zu genau einer K-linearen Abbildung fortsetzen. Es ist also *jeder* K-Vektorraum frei bezüglich ${}^K\mathfrak{M}$, und die K-Basen (im Sinne der Linearen Algebra) sind genau die ${}^K\mathfrak{M}$-Basen (im Sinne von 2.1). Genauer sind die K-linearen Erzeugendensysteme bzw. die K-linear unabhängigen Teilmengen eines K-Vektorraums genau seine ${}^K\mathfrak{M}$-Erzeugendensysteme bzw. seine ${}^K\mathfrak{M}$-unabhängigen Teilmengen. Wir wollen diesen einfachen Zusammenhang in etwas allgemeinerer Form festhalten, indem wir statt des Körpers K nur einen unitären Ring voraussetzen. Da sich dann allerdings zahlreiche der für Vektorräume gewohnten Aussagen nicht mehr in die dann größere Allgemeinheit übertragen lassen, bedarf es für letztere eines besonderen Begriffs, um Irrtümer zu vermeiden.

2.7 Definition Sei $R \in \mathfrak{R}_1$, $M \in \mathfrak{M}$ und eine (im folgenden, wie üblich, als „Produkt zwischen R und M“ geschriebene) Abbildung von $R \times M$ in M mit

$$\begin{aligned} r(m+m') &= rm + rm' \\ (r+r')m &= rm + r'm \\ (rr')m &= r(r'm) \\ 1_R m &= m \end{aligned}$$

für alle $r, r' \in R$, $m, m' \in M$. Dann heißt M ein R-Links-Modul und das Produkt zwischen R und M eine Links-Aktion von R auf M. R-Rechts-Moduln und Rechts-Aktionen werden analog mit Hilfe einer Abbildung von $M \times R$ in M definiert. Die folgenden Ausführungen gehen auf Links-Moduln ein und lassen sich für Rechts-Moduln ohne weiteres direkt übertragen. Ein nennenswerter Unterschied ist der, daß die Abbildung von R in den Endomorphismenring von M, die jedem $r \in R$ das Produktbilden mit r zuordnet, im Falle einer Rechts-Aktion ein $\mathfrak{R}_1$-*Homomorphismus*, im Falle einer Links-Aktion dagegen ein $\mathfrak{R}_1$-*Antihomomorphismus* ist. Wir bezeichnen mit

${}^{R}\mathfrak{M}$ die Klasse der R-Links-Moduln.

Jeder Vektorraum über einem Körper K ist offensichtlich ein K-Modul. Für beliebiges $R \in \mathfrak{R}_1$ ist die additive Gruppe von R in natürlicher Weise selbst ein R-Links- und ein R-Rechts-Modul, wobei die Links- bzw. Rechts-Aktion durch die Multiplikation in R gegeben ist; allgemeiner gilt dies für die abelsche Gruppe R^n (bei beliebigem $n \in \mathbb{N}$) und komponentenweise Aktion von R. Ist $M \in {}^{R}\mathfrak{M}$, so heißt eine Untergruppe N von M mit $RN \subseteq N$ ein R-Teilmodul von M. Schreibweise: $N \underset{R}{\leq} M$. Die Linksideale von R sind genau die unter den Linksmultiplikationen mit beliebigen Elementen von R invarianten Untergruppen, entsprechend die Rechtsideale von R die unter allen Rechtsmultiplikationen invarianten Untergruppen der additiven Gruppe von R, illustrieren also den Teilmodulbegriff.

2.7.1 *Ist $M \in {}^{R}\mathfrak{M}$ und $N \underset{R}{\leq} M$, so folgt: $N \in {}^{R}\mathfrak{M}$, also $N \leq_{{}^{R}\mathfrak{M}} M$.* □

2.7.2 *${}^{R}\mathfrak{M}$ ist gegen Durchschnittsbildung und gegen R-lineare Abbildungen abgeschlossen.* □

Ein wichtiger Spezialfall tritt durch die Wahl $R = \mathbb{Z}$ auf:

2.7.3 *Jede abelsche Gruppe M ist ein $\mathbb{Z}$-Modul vermöge der Produktbildung*

$$zm := \begin{cases} \underbrace{m + \cdots + m}_{z} & \text{falls } z \geq 0, \\ \underbrace{-m - \cdots - m}_{|z|} & \text{falls } z < 0. \end{cases}$$

□

Das Weglassen des oberen Index ^{R} bedeutet also, daß $R = \mathbb{Z}$ gewählt wird; es bedeutet den Übergang von der Betrachtung von Moduln zu der von abelschen Gruppen.

Sei X eine Teilmenge eines R-Links-Moduls M. Ein Element $m \in M$ heißt eine R-Linearkombination über X, wenn es ein $k \in \mathbb{N}_0$ und k-Tupel $(x_1, \ldots, x_k) \in X^k$, $(r_1, \ldots, r_k) \in R^k$ gibt mit $m = \sum_{j \in \underline{k}} r_j x_j$. Die Menge der R-Linearkombinationen über X ist ein R-Teilmodul von M, der das R-Erzeugnis von X genannt wird; Schreibweise: $\langle X \rangle_R$. Gilt $\langle X \rangle_R = M$, so heißt X ein R-lineares Erzeugendensystem von M. Eine beliebige Teilmenge X von M heißt R-linear unabhängig, wenn für je endlich viele paarweise verschiedene Elemente $x_1, \ldots, x_k \in X$ gilt: Sind $r_1, \ldots, r_k \in R$ mit $r_1 x_1 + \cdots + r_k x_k = 0_M$, so gilt $r_1 = \cdots = r_k = 0_R$. Äquivalent dazu ist die Eindeutigkeit der Darstellbarkeit eines Elementes als R-Linearkombination über X. Ein R-linear unabhängiges R-lineares Erzeugendensystem von M heißt eine R-Basis von M. Es gilt also:

2.7.4 *X ist genau dann eine R-Basis von M, wenn es zu jedem $m \in M$ eine eindeutig bestimmte fast überall verschwindende Funktion $X \to R$, $x \mapsto r_x$ gibt mit $m = \sum_{x \in X, r_x \neq 0_R} r_x x$.* □

Diese naheliegenden Verallgemeinerungen aus der Vektorraum-Theorie bekannter Begriffe fügen sich ohne weiteres in den Kontext von 2.1 ein:

2.7.5 *Es gilt:*

(a) *X R-lineares Erzeugendensystem von $M \Leftrightarrow X$ ${}^R\mathfrak{M}$-Erzeugendensystem von M,*

(b) *X R-linear unabhängig $\Leftrightarrow X$ ${}^R\mathfrak{M}$-unabhängig,*

(c) *X R-Basis von $M \Leftrightarrow X$ ${}^R\mathfrak{M}$-Basis von M.* □

Enthält R Elemente s, $s' \neq 0_R$ mit $ss' = 0_R$, so enthält jeder R-Modul $M \neq \{0_M\}$ ein Element $m \neq 0_M$, zu dem es ein $r \in R \setminus \{0_R\}$ gibt mit $rm = 0_M$: Ist nämlich $m \in M \setminus \{0_M\}$, so gilt $s(s'm) = (ss')m = 0_M$; wir brauchen also (im Falle $s'm = 0_M$) nur $r := s'$ oder (im Falle $s'm \neq 0_M$) $r := s$ zu setzen. Ist R nullteilerfrei, so kann es natürlich dennoch vorkommen, daß das Produkt eines Elements $\neq 0_R$ mit einem Modulelement $\neq 0_M$ das Nullelement des Moduls ergibt. Im Falle eines *freien* R-Moduls ist dies jedoch nicht so:

2.7.6 *Ist M ein freier R-Modul und R nullteilerfrei, so gilt für alle $r \in R$, $m \in M$: $rm = 0_M \Rightarrow r = 0_R$ oder $m = 0_M$.*

Ist nämlich X eine R-Basis von M und $rm = 0_M$, so gibt es ein $k \in \mathbb{N}_0$ und $x_1, \ldots, x_k \in X$, $r_1, \ldots, r_k \in R \setminus \{0_R\}$ mit $m = r_1 x_1 + \cdots r_k x_k$. Es folgt:

$(rr_1)x_1 + \cdots + (rr_k)x_k = rm = 0_M$, also $rr_j = 0_R$ für alle $j \in \underline{k}$. Wegen der Nullteilerfreiheit von R ist $r = 0_R$ oder $r_j = 0_R$ für alle $j \in \underline{k}$, d.h. $m = 0_M$.□

2.8 Proposition *Sei J ein Linksideal eines unitären Ringes R, $Y \subseteq J$. Für alle $k \in \mathbb{N}$ sei $Y(k) := \{y_1 \cdots y_k | y_i \in Y\}$.*

(1) *Gilt $\langle Y \rangle_{R\mathfrak{M}} = J$, so $\langle Y(k) \rangle_{R\mathfrak{M}} = J^k$ [5] für alle $k \in \mathbb{N}_0$.*

(2) *Ist Y ${}^R\mathfrak{M}$-unabhängig, so für alle $k \in \mathbb{N}$ auch $Y(k)$, und die Abbildung $\pi_k : Y \times \overset{k}{\cdots} \times Y \to Y(k)$, $(y_1, \ldots, y_k) \mapsto y_1 \cdots y_k$, ist eine Bijektion.*

Beweis. (1) Sei $\langle Y \rangle_{R\mathfrak{M}} = J$. Für $k = 0$ gilt die Behauptung, da $Y(0) = \{1_R\}$. Ist $k > 0$ und gilt $\langle Y(k-1) \rangle_{R\mathfrak{M}} = J^{k-1}$, so folgt

$$J^k = J^{k-1}J = \langle Y(k-1) \rangle_{R\mathfrak{M}} \langle Y \rangle_{R\mathfrak{M}} = \langle Y(k) \rangle_{R\mathfrak{M}}.$$

(2) Sei Y ${}^R\mathfrak{M}$-unabhängig. Für $k = 1$ ist nichts zu zeigen. Sei $k > 1$, und es gelte die Behauptung für $k-1$. Sei T eine endliche Teilmenge von $Y \times \overset{k}{\cdots} \times Y$, und sei $\sum_{(y_1,\ldots,y_k) \in T} r_{y_1,\ldots,y_k} y_1 \cdots y_k = 0_R$ mit $r_{y_1,\ldots,y_k} \in R$. Sei Z die Menge der letzten Komponenten der Elemente von T, und für jedes $y \in Z$ sei $T(y) := \{(y_1, \ldots, y_{k-1}) | y_1, \ldots, y_{k-1} \in Y,\ (y_1, \ldots, y_{k-1}, y) \in T\}$. Dann gilt: $\sum_{y \in Z} \big(\sum_{(y_1,\ldots,y_{k-1}) \in T(y)} r_{y_1,\ldots,y_{k-1},y} y_1 \cdots y_{k-1} \big) y = 0_R$. Da Z ${}^R\mathfrak{M}$-unabhängig ist, folgt $\sum_{(y_1,\ldots,y_{k-1}) \in T(y)} r_{y_1,\ldots,y_{k-1},y} y_1 \cdots y_{k-1} = 0_R$ für alle $y \in Z$, wegen der Bijektivität von π_{k-1} also für alle $y \in Z$, $(y_1, \ldots, y_{k-1}) \in T(y)$ nach Induktionsvoraussetzung $r_{y_1,\ldots,y_{k-1},y} = 0_R$. Insbesondere ist π_k bijektiv. □

Ähnlich wie im Falle der Klassen $\mathfrak{S}_1$ und $\mathfrak{Ma}$ gilt für ${}^R\mathfrak{M}$ die folgende Aussage:

2.9 Proposition *Sei $R \in \mathfrak{R}_1$. Zu jeder Menge X gibt es einen von X frei erzeugten R-Links-Modul.*

Spezialfall: *Ist X endlich und $n := |X|$, so ist jeder von X frei erzeugte R-Modul ${}^R\mathfrak{M}$-isomorph zu R^n.*

Beweis. Setzen wir für jedes $r \in R$, $f \in R^X$

$$rf : X \to R,\ x \mapsto r(xf),$$

so erhalten wir eine Links-Aktion von R auf R^X; damit wird R^X ein R-Links-Modul. Für alle $x \in X$ sei $f_x : X \to R,\ y \mapsto \begin{cases} 1_R & \text{falls } y = x \\ 0_R & \text{sonst.} \end{cases}$. Die

[5] J^k ist der additive Abschluß von $J(k)$ und offensichtlich ein Linksideal von R.

Abbildung $\iota : X \to R^X$, $x \mapsto f_x$ ist injektiv. Aus dem Erweiterungsprinzip folgt, daß es einen X enthaltenden R-Links-Modul M gibt, der vermöge einer Fortsetzung $\overline{\iota}$ von ι zu R^X ${}^R\mathfrak{M}$-isomorph ist. Wir wollen zeigen, daß X eine R-linear unabhängige Teilmenge von M ist. Es genügt dazu einzusehen, daß $X\iota$ eine R-linear unabhängige Teilmenge von R^X ist. Sind $x_1, \ldots, x_k \in X$, paarweise verschieden, $r_1, \ldots, r_k \in R$ mit $r_1 f_{x_1} + \cdots + r_k f_{x_k} = 0_{R^X}$, so gilt für alle $j \in \underline{k}$: $0_R = x_j 0_{R^X} = r_1(x_j f_{x_1}) + \cdots + r_k(x_j f_{x_k}) = r_j(x_j f_{x_j}) = r_j$. Aus 2.7.5(b) folgt, daß $X\iota$ ${}^R\mathfrak{M}$-unabhängig ist. Also ist $\langle X\rangle_{R_{\mathfrak{M}}}$ ein von X ${}^R\mathfrak{M}$-frei erzeugter R-Links-Modul.

Ist X endlich, $n := |X|$, so folgt: $\langle X\rangle_{R_{\mathfrak{M}}} = M = R^X \underset{R}{\cong} R^n$. Damit folgt auch die Aussage des Spezialfalls. □

Ist X unendlich, so gilt $\langle X\rangle_{R_{\mathfrak{M}}} \neq M$, da dann z.B. jede Abbildung von X in $R \setminus \{0_R\}$ in $R^X \setminus \langle X\iota\rangle_{R_{\mathfrak{M}}}$ liegt. Es ist dann also der in 2.9 als existent behauptete R-Modul ein *echter* R-Teilmodul des Moduls M aus dem Beweis von 2.9. Letzteren (bis auf R-Modul-Isomorphie eindeutig bestimmten) *gesamten* R-Modul M nennen wir den Modul der formalen unendlichen Linearkombinationen und bezeichnen ihn mit $\overline{RX}$; das R-Erzeugnis von X dagegen bezeichnen wir mit RX und nennen es den Modul der formalen Linearkombinationen über X. Ist $f \in R^X$ und $xf =: r_x$ für alle $x \in X$, so schreiben wir $\sum_{x\in X} r_x x$ für das Element von $\overline{RX}$, dessen Bild unter $\overline{\iota}$ gerade f ist. Genau dann gilt $\sum_{x\in X} r_x x \in RX$, wenn $r_x = 0_R$ für fast alle $x \in X$ gilt. In dem Fall hat die oben für *alle* Elemente von $\overline{RX}$ eingeführte Schreibweise $\sum_{x\in X} r_x x$ die gewöhnliche „Summenbedeutung“, wenn man jeden Summanden, der gleich $0_{\overline{RX}}$ ist, fortläßt.

Ist B eine Basis, C ein Erzeugendensystem eines *Vektorraums*, so gilt bekanntlich: $|B| \leq |C|$; je zwei Basen sind gleichmächtig; ist B endlich und $|C| = |B|$, so ist auch C eine Basis. Das nächste Ziel wird sein, diese Aussagen auf R-Moduln zu verallgemeinern. Dazu muß R kommutativ sein, damit es gelingt, die gewünschte Aussage auf den Fall eines Vektorraums zurückzuführen. Einige vorbereitende Aussagen gelten jedoch wesentlich allgemeiner; wir benötigen für die folgende ebenso grundlegende wie triviale Proposition zunächst nur, daß R irgendein unitärer Ring ist:

2.10 Proposition *Sei $R \in \mathfrak{R}_1$, $M \in {}^R\mathfrak{M}$.*

(1) *Sei $N \underset{R}{\leq} M$. Sind $m, m' \in M$ mit $N + m = N + m'$, so gilt für alle $r \in R$: $N + rm = N + rm'$.*

(2) *Sei* $J \trianglelefteq R$ *mit* $JM = \{0_M\}$. *Dann gilt für alle* $r \in R$, $m \in M$: $(J+r)m = \{rm\}$.

Folgerung

(1') *Ist* $N \underset{R}{\leq} M$, *so ist* M/N *ein* R*-Modul vermöge der Operation*

$$r \bullet (N+m) = N + rm \quad \text{für alle } m \in M,\ r \in R.$$

(2') *Ist* $J \trianglelefteq R$ *mit* $JM = \{0_M\}$, *so ist* M *ein* R/J*-Modul vermöge der Operation*

$$(J+r) \cdot m = rm \quad \text{für alle } m \in M,\ r \in R.$$

Beweis. Die beiden Teile der Proposition sind trivial; sie dienen dazu, den einzigen springenden Punkt der Folgerung vorzubereiten: daß nämlich $\bullet$ und $\cdot$ tatsächlich *Abbildungen* sind. Sind m, $m' \in M$ mit $N + m = N + m'$, so folgt $m - m' \in N$, also auch $r(m - m') \in N$, mithin $N + rm = N + rm'$ für alle $r \in R$. Daher ist $\{((r, N+m), N+rm) | r \in R,\ m \in M\}$ eine Abbildung. Für alle $s \in J$, $r \in R$, $m \in M$ gilt $(s+r)m = sm + rm = rm$, also folgt $(J+r)m = \{rm\}$, und $\{((J+r, m), rm) | r \in R,\ m \in M\}$ ist eine Abbildung. Die Nachweise der Modulgesetze sind trivial. □

2.11 Proposition *Sei* $R \in \mathfrak{R}_1$, $J \trianglelefteq R$, $M \in {}^R\mathfrak{M}$, $JM := \langle sm | s \in J,\ m \in M \rangle_{\mathfrak{M}}$. *Dann gilt:*

(1) $JM \underset{R}{\leq} M$.

(2) M/JM *ist ein* R/J*-Links-Modul vermöge der Operation*

$$(J+r) \bullet (JM+m) = JM + rm \quad \text{für alle } r \in R,\ m \in M.$$

(3) *Ist* C *ein* ${}^R\mathfrak{M}$*-Erzeugendensystem von* M, *so ist* $\{JM + c | c \in C\}$ *ein* ${}^{R/J}\mathfrak{M}$*-Erzeugendensystem von* M/JM, *und für alle* $z \in JM$ *existieren* $k \in \mathbb{N}_0$, $c_1, \ldots, c_k \in C$, $s_1, \ldots, s_k \in J$ *mit* $z = s_1c_1 + \cdots + s_kc_k$.

(4) *Ist* B *eine* ${}^R\mathfrak{M}$*-Basis von* M *und* $J \neq R$, *so ist* $\{JM + b | b \in B\}$ *eine* ${}^{R/J}\mathfrak{M}$*-Basis von* M/JM *und gleichmächtig zu* B.

Beweis. (1) ist trivial.

(2): Nach (1) gilt $JM \underset{R}{\leq} M$. Also gilt $M/JM \in {}^R\mathfrak{M}$ gemäß Folgerung (1') aus

2.10. Es gilt $J(M/JM) = \{0_{M/JM}\}$, also $M/JM \in {}^{R/J}\mathfrak{M}$ gemäß Folgerung (2') aus 2.10, wie behauptet.

Nur der zweite Teil von (3) ist nicht völlig trivial: Ist $z \in JM$, so existieren $l \in \mathbb{N}_0$, $t_1, \ldots, t_l \in J$, $m_1, \ldots, m_l \in M$ mit $z = t_1m_1 + \cdots + t_lm_l$. Daher genügt es, für beliebige $t \in J$, $m \in M$ zu zeigen:

$$(*) \qquad \exists k \in \mathbb{N}_0\ \exists s_1, \ldots, s_k \in J\ \exists c_1, \ldots, c_k \in C \quad tm = s_1c_1 + \cdots + s_kc_k.$$

Da C ein ${}^R\mathfrak{M}$-Erzeugendensystem von M ist, gibt es $k \in \mathbb{N}_0$, $r_1, \ldots, r_k \in R$, $c_1, \ldots, c_k \in C$ mit $m = r_1c_1 + \cdots + r_kc_k$. Es folgt: $tm = (tr_1)c_1 + \cdots + (tr_k)c_k$ und $tr_i \in J$ für alle $i \in \underline{k}$, da $J \trianglelefteq R$. Also gilt $(*)$.

Als Vorbemerkung zum Beweis von (4) halten wir zunächst fest, daß es zu jedem $z \in JM$ nach dem 2. Teil von (3) $k \in \mathbb{N}_0$, $b_1, \ldots, b_k \in B$ und $s_1, \ldots, s_k \in J$ gibt mit $z = s_1b_1 + \cdots + s_kb_k$. O.B.d.A. dürfen wir dabei annehmen, daß die Elemente b_i paarweise und die Elemente s_i von 0_R verschieden sind. Da B eine R-Basis von M ist, ist dies zugleich die einzige Möglichkeit, z als R-Linearkombination über B mit sämtlich nicht-verschwindenden Koeffizienten zu schreiben. Nach (3) bleibt nur zu zeigen, daß $\{JM+b|b \in B\}$ R/J-linear unabhängig und die Abbildung $B \to M/JM$, $b \mapsto JM+b$, injektiv ist. Seien $k \in \mathbb{N}_0$, $b_1, \ldots, b_k \in B$ paarweise verschieden und $r_1, \ldots, r_k \in R \setminus \{0_R\}$ mit $\sum_{i \in \underline{k}} (J + r_i) \bullet (JM + b_i) = JM$. Nach (2) ist die linke Seite gleich $JM + \sum_{i \in \underline{k}} r_ib_i$, so daß folgt: $\sum_{i \in \underline{k}} r_ib_i \in JM$. Aus der Vorbemerkung erhalten wir nun: $r_1, \ldots, r_k \in J$. Speziell ist damit auch die Abbildung $B \to M/JM$, $b \mapsto JM + b$ injektiv. Damit ist alles bewiesen. □

2.12 Korollar *Sei $R \in \mathfrak{R}_1$ kommutativ, M ein freier R-Links-Modul mit R-Basis B, C ein R-Erzeugendensystem von M. Dann gilt:*

(1) $|B| \leq |C|$,

(2) *Je zwei R-Basen von M sind gleichmächtig.*

(3) *Ist R ein Integritätsbereich, B endlich und $|C| \leq |B|$, so ist C eine R-Basis von M.*

Beweis. Nach dem Zorn'schen Lemma besitzt R ein maximales Ideal J. Es ist R/J ein Körper. Nach 2.11(3) ist $\{JM+c|c \in C\}$ ein ${}^{R/J}\mathfrak{M}$-Erzeugendensystem von M/JM, und nach 2.11(4) ist $\{JM + b|b \in B\}$ eine zu B gleichmächtige ${}^{R/J}\mathfrak{M}$-Basis von M/JM. Da die Behauptung für Vektorräume über Körpern wahr ist, gilt

$$|B| = |\{JM + b|b \in B\}| \leq |\{JM + c|c \in C\}| \leq |C|,$$

also (1).

(2) Ist auch C eine R-Basis von M, so gilt auch $|C| \leq |B|$, also $|B| = |C|$.

(3) Sei K Quotientenkörper von R. Der K-Vektorraum KB enthält den R-Teilmodul RB, der zu M ${}^R\mathfrak{M}$-isomorph ist. Sei C' das Bild von C unter einem ${}^R\mathfrak{M}$-Isomorphismus von M auf $RB \subseteq KB$. Dann gilt: $B \subseteq \langle C' \rangle_{R\mathfrak{M}}$, also ist C' ein ${}^K\mathfrak{M}$-Erzeugendensystem von KB. Wegen $|C'| \leq |B|$ ist daher C' eine K-Basis von KB und insbesondere R-linear unabhängig. Also ist auch C R-linear unabhängig und folglich eine R-Basis von M. □

Der Beweis zeigt, daß 2.12(1),(2) allgemeiner für unitäre Ringe R gelten, die einen echten kommutativen Faktorring besitzen. Ohne Beweis sei vermerkt, daß ohne die Voraussetzung der Kommutativität von R die Aussagen aus 2.12(1),(2) noch im Falle *unendlicher* Basen gültig bleiben, aber bei endlichen Basen falsch werden können ([Bo1], II. §1.12 Cor. 2 zu Prop. 23).

Sind je zwei Basen eines freien R-Moduls gleichmächtig, so heißt die Mächtigkeit einer Basis der Rang von M; Schreibweise: $rk_R(M)$. Der Rangbegriff spezialisiert sich im Falle eines *Körpers* R zum Begriff der Vektorraum-Dimension. Im Falle $R = \mathbb{Z}$ schreiben wir kurz $rk(M)$ statt $rk_Z(M)$. Die Spezialisierung $R = \mathbb{Z}$ ergibt:

2.13 Korollar *Zu jeder Menge X gibt es eine von X frei erzeugte abelsche Gruppe A. Ist Y ein $\mathfrak{M}$-Erzeugendensystem von A, so ist $|X| \leq |Y|$. Je zwei $\mathfrak{M}$-Basen von A sind gleichmächtig. Ist X endlich und $n := |X|$, so gilt $A \cong \mathbb{Z} \oplus \overset{n}{\cdots} \oplus \mathbb{Z}$. Ist Y ein $\mathfrak{M}$-Erzeugendensystem von A mit $|Y| \leq n$, so ist Y eine $\mathbb{Z}$-Basis von A.* □

Ohne Beweis vermerken wir:

Satz *Jede Untergruppe einer freien abelschen Gruppe ist frei.*

(Siehe etwa [Kur], Kap. VI, §19.) Im Falle einer *endlich erzeugten* freien abelschen Gruppe beweisen wir diese Aussage und erhalten noch weit schärfere Zusätze:

2.14 Satz *Sei $n \in \mathbb{N}$, A eine freie abelsche Gruppe vom Rang n, $B \leq_{\mathfrak{M}} A$. Dann gilt:*

(1) *Es gibt eine $\mathfrak{M}$-Basis Y von A, so daß es zu allen $y \in Y$ Zahlen $n_y \in \mathbb{N}_0$ gibt mit:*

- *$\{n_y y | y \in Y,\ n_y \neq 0\}$ ist eine $\mathfrak{M}$-Basis von B,*
- *$A/B \cong \bigoplus_{y \in Y} \mathbb{Z}/\mathbb{Z}n_y$.*

Insbesondere ist B frei und $rk(B) \leq n$; gilt $A/B \cong A$, so ist $B = \{0_A\}$.

(2) *$rk(B) = n \Leftrightarrow A/B$ endlich.*

Beweis. Im Falle $n = 1$ sind die Behauptungen wohlbekannte Aussagen über die Gruppe $(\mathbb{Z}, +)$. Wir führen den Beweis von (1) durch Induktion nach n. Für den Induktionsschritt sei $n > 1$, A eine freie abelsche Gruppe vom Rang n und $B \leq A$. Wir nehmen die Behauptungen als wahr an für freie abelsche Gruppen vom Rang $< n$. Ist $B = \{0_A\}$, so ist alles Behauptete trivial. Sei also $B \neq \{0_A\}$. Für jede $\mathfrak{M}$-Basis X von A sei

$$m(X) := \min\{k | k \in \mathbb{N}, \exists x_0 \in X \quad kx_0 \in B + \langle X \setminus \{x_0\}\rangle_{\mathfrak{M}}\},$$

d.h. $m(X)$ ist der kleinste natürlichzahlige Koeffizient, der bei einer in B liegenden $\mathbb{Z}$-Linearkombination über X auftritt. Sei ferner

$$\begin{aligned} m &:= \min\{m(X) \mid X \text{ } \mathfrak{M}\text{-Basis von } A\} \\ &= \min\{n \mid n \in \mathbb{N}, \text{ es gibt eine } \mathfrak{M}\text{-Basis } X \text{ von } A \text{ und } x, x_1, \ldots, x_k \in X, \\ &\qquad\qquad l_1, \ldots, l_k \in \mathbb{Z} \text{ mit } nx + l_1x_1 + \cdots + l_kx_k \in B\} \end{aligned}$$

und eine $\mathfrak{M}$-Basis X von A so gewählt, daß $m(X) = m$, also minimal im Vergleich mit allen $\mathfrak{M}$-Basen von A ist. Seien dann $b \in B$, $x_0 \in X$, $l_x \in \mathbb{Z}$ für alle $x \in X \setminus \{x_0\}$, $l_x = 0$ für fast alle $x \in X$, so daß gilt:

$$b = mx_0 + \sum_{x \in X \setminus \{x_0\}, l_x \neq 0} l_x x.$$

Für jedes $x \in X \setminus \{x_0\}$ mit $l_x \neq 0$ seien q_x, $r_x \in \mathbb{Z}$ mit $l_x = mq_x + r_x$, $0 \leq r_x < m$. Wir setzen

$$\begin{aligned} \overline{x_0} &:= x_0 + \sum_{x \in X \setminus \{x_0\}, l_x \neq 0} q_x x, \\ \overline{X} &:= \{\overline{x_0}\} \cup (X \setminus \{x_0\}). \end{aligned}$$

Dann ist $\overline{X}$ eine $\mathfrak{M}$-Basis von A (nach der letzten Aussage in 2.13, denn $x_0 \in \langle \overline{X} \rangle_{\mathfrak{M}}$) und

$$b = mx_0 + \sum_{x \in X \setminus \{x_0\}, l_x \neq 0} (mq_x + r_x)x = m\overline{x_0} + \sum_{x \in X \setminus \{x_0\}, l_x \neq 0} r_x x.$$

Wäre $r_x \neq 0$ für ein $x \in X \setminus \{x_0\}$ mit $l_x \neq 0$, so $m(\overline{X}) \leq r_x < m$, ein Widerspruch; also ist $r_x = 0$ für alle $x \in X \setminus \{x_0\}$ mit $l_x \neq 0$, d.h.: $b = m\overline{x_0}$.

Sei nun $A' := \langle X \setminus \{x_0\}\rangle_{\mathfrak{M}}$ und $B' := B \cap A'$. Es gilt:

$$\langle b\rangle_{\mathfrak{M}} \cap B' \subseteq \langle \overline{x_0}\rangle_{\mathfrak{M}} \cap A' = \langle \overline{x_0}\rangle_{\mathfrak{M}} \cap \langle \overline{X} \setminus \{\overline{x_0}\}\rangle_{\mathfrak{M}} = \{0_A\}.$$

Wir behaupten:

$$(*) \qquad \langle b\rangle_{\mathfrak{M}} + B' = B.$$

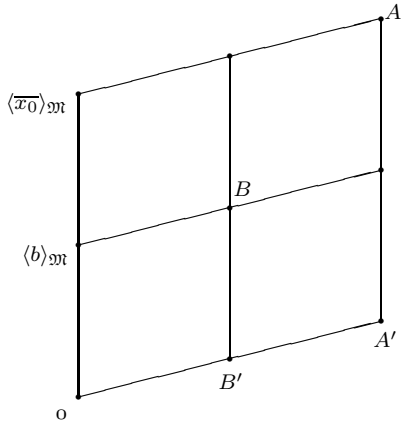
Die Inklusion „$\subseteq$“ ist trivial. Zum Beweis der anderen Inklusion seien $c \in B$ und $k,\ j_x \in \mathbb{Z}$ für alle $x \in X \setminus \{x_0\}$, $j_x = 0$ für fast alle $x \in X$ mit $c = k\overline{x_0} + \sum\limits_{x \in X \setminus \{x_0\}} j_x x$. Seien $q,\ r \in \mathbb{Z}$ mit $k = mq + r$, $0 \leq r < m$. Dann gilt:

$$\underbrace{c - qb}_{\in B} = c - mq\overline{x_0} = r\overline{x_0} + \sum_{x \in X \setminus \{x_0\}} j_x x.$$

Wäre $r \neq 0$, so $m(\overline{X}) \leq r < m$, ein Widerspruch; also ist $r = 0$, $c - qb = \sum\limits_{x \in X \setminus \{x_0\}} j_x x \in B'$ und folglich $c \in \langle b\rangle_{\mathfrak{M}} + B'$. Also gilt $(*)$.

Es folgt:

$$(**) \qquad \begin{aligned} B &= \langle b\rangle_{\mathfrak{M}} \oplus B', \\ A/B &\cong \langle \overline{x_0}\rangle_{\mathfrak{M}}/\langle b\rangle_{\mathfrak{M}} \oplus A'/B' \cong \mathbb{Z}/\mathbb{Z}m \oplus A'/B'. \end{aligned}$$



A' ist eine freie abelsche Gruppe vom Rang $n-1$, $B' \leq_{\mathfrak{M}} A'$. Nach Induktions-Voraussetzung gibt es also eine $\mathfrak{M}$-Basis Y' von A' und zu jedem $y \in Y'$ ein $n_y \in \mathbb{N}_0$, so daß $\{n_y y | y \in Y', n_y \neq 0\}$ eine $\mathfrak{M}$-Basis von B' ist, und $A'/B' \cong \bigoplus\limits_{y \in Y'} \mathbb{Z}/\mathbb{Z}n_y$. Mit $Y := \{\overline{x_0}\} \cup Y'$ und $n_{\overline{x_0}} := m$ ist nun die Behauptung erfüllt, da $(**)$ gilt. Also gilt (1).

(2) Es gilt:

$$\begin{aligned} rk(B) = n &\underset{(1)}{\Leftrightarrow} n_y \neq 0 \text{ für alle } y \in Y \\ &\Leftrightarrow \mathbb{Z}/\mathbb{Z}n_y \text{ endlich für alle } y \in Y \\ &\underset{(1)}{\Leftrightarrow} A/B \text{ endlich.} \end{aligned}$$

□

Wir führen nun die bislang getrennt betrachteten Magmen (insbesondere also auch Monoide und Gruppen als deren besonders wichtige Spezialfälle) und R-Moduln (für einen gegebenen unitären Ring R) in einem weiteren wichtigen Strukturbegriff zusammen. Dazu betrachten wir zunächst eine spezielle, im weiteren aber besonders wichtige Situation:

Sei $R \in \mathfrak{R}_1$. Wir betrachten den R-Links-Modul RY, wobei jetzt Y nicht mehr eine beliebige Menge ist, sondern selbst schon eine („Multiplikation“ genannte) Verknüpfung trägt, also ein Magma ist. Dann definieren wir durch distributive Fortsetzung eine („Multiplikation“ genannte) Verknüpfung auf RY:

$$\left(\sum_{y\in Y} r_y y\right)\left(\sum_{y\in Y} s_y y\right) := \sum_{y,z\in Y} r_y s_z yz = \sum_{y\in Y}\left(\sum_{\substack{u,v\in Y,\\ uv=y}} r_u s_v\right) y.$$

(Man beachte, daß nur endlich viele der Elemente r_y, s_y von 0_R verschieden sind!) Ist R kommutativ, so gilt

$$\left(\sum_{y\in Y} r_y y\right)\left(r\sum_{y\in Y} s_y y\right) = \left(r\sum_{y\in Y} r_y y\right)\left(\sum_{y\in Y} s_y y\right) = r\left(\sum_{y\in Y} r_y y \sum_{y\in Y} s_y y\right).$$

2.15 Definition Sei $K \in \mathfrak{R}_1$, kommutativ. Sei $M \in {}^K\mathfrak{M}$ und $\cdot$ eine Verknüpfung auf M, in bezug auf die beide Distributivgesetze erfüllt sind. Gilt für alle m, $m' \in M$, $r \in K$ dann

$$m \cdot (rm') = (rm) \cdot m' = r(m \cdot m'),$$

so heißt M eine K-Algebra. (Genauer könnte man A als „K-Links-Algebra“ bezeichnen, um von „K-Rechts-Algebren“ zu unterscheiden, von denen man entsprechend im Fall eines K-*Rechts*-Moduls M spräche; letztere spielen aber für uns im weiteren keine Rolle.) Es sei

${}^K\mathfrak{A}$ die Klasse der K-Algebren,
${}^K\mathfrak{R}$ die Klasse der assoziativen K-Algebren.

2.15.1 *Die Ringe sind genau die assoziativen $\mathbb{Z}$-Algebren.* □

Wie schon im Falle der Moduln lassen wir auch bei Algebrenklassen den oberen Index ^{K} fort, wenn $K = \mathbb{Z}$ gilt. Sind M, $N \in {}^K\mathfrak{A}$ (bzw. ${}^K\mathfrak{R}$), so heißt eine Abbildung von M in N ein Algebren-Homomorphismus (${}^K\mathfrak{A}$-Homomorphismus bzw. ${}^K\mathfrak{R}$-Homomorphismus), wenn sie zugleich ein multiplikativer und ein ${}^K\mathfrak{M}$-Homomorphismus ist. Die Klassen ${}^K\mathfrak{A}$ und ${}^K\mathfrak{R}$ sind offensichtlich gegen Durchschnittsbildungen und gegen Algebren-Homomorphismen abgeschlossen. Für jedes Magma Y ist KY eine K-Algebra, wie oben ausgeführt. Ist $Y \in \mathfrak{G}$ (bzw. $\mathfrak{S}$), so heißt KY der Gruppenring (bzw. Halbgruppenring) von Y über K.

2.15.2 *Sei $Y \in \mathfrak{Ma}$. Dann gilt: $KY \in {}^K\mathfrak{R} \Leftrightarrow Y \in \mathfrak{S}$.* □

Weitere Beispiele für Algebren sind Matrixringe über K, Polynomringe (in beliebig vielen Veränderlichen) über K, allgemeiner: kommutative Oberringe von K.

2.15.3 *Ist $M \in {}^K\mathfrak{A}$ und $X \subseteq M$, so gilt: $\langle X\rangle_{K_{\mathfrak{A}}} = \langle\langle X\rangle_{\mathfrak{Ma}}\rangle_{K_{\mathfrak{M}}}$,* □

d.h. die von X erzeugte K-Teilalgebra von M besteht aus den K-Linearkombinationen des multiplikativen Abschlusses von X.

2.16 Proposition *Sei X eine Menge, $K \in \mathfrak{R}_1$, kommutativ. Dann gilt:*

(1) *$KX^{(+)}$ (bzw. $KX^{(*)}$) ist eine von X ${}^K\mathfrak{A}$-frei (bzw. ${}^K\mathfrak{A}_1$-frei) erzeugte K-Algebra.*

(2) *KX^{+} (bzw. KX^{*}) ist eine von X ${}^K\mathfrak{R}$-frei (bzw. ${}^K\mathfrak{R}_1$-frei) erzeugte K-Algebra.*

Einige Bemerkungen bezüglich der freien Algebren KX^* und $KX^{(*)}$ seien eingefügt, bevor wir den einfachen Beweis führen:

2.16.1 *Für alle $n \in \mathbb{N}_0$ gilt: $KX^n \leq_{K_{\mathfrak{M}}} KX^*$ (bzw. $KX^{(n)} \leq_{K_{\mathfrak{M}}} KX^{(*)}$).* $KX^* = KX^0 \dot{\oplus} KX^+$, $KX^{(*)} = KX^{(0)} \dot{\oplus} KX^{(+)}$, □

KX^n (bzw. $KX^{(n)}$) heißt der Teilmodul der homogenen Elemente vom Grad n von KX^* (bzw. $KX^{(*)}$).

Die Teilmoduln der homogenen Elemente bilden eine direkte Zerlegung von KX^* (bzw. $KX^{(*)}$), d.h.: Zu jedem $w \in KX^*$ (bzw. $KX^{(*)}$) gibt es eindeutig bestimmte, vom Null-Element verschiedene homogene Elemente $w_1, \ldots, w_k$ von paarweise verschiedenen Graden $n_1, \ldots, n_k$, so daß gilt: $w = w_1 + \cdots + w_k$.

Die Menge $\{w_1, \ldots, w_k\}$ heißt die Zerlegung von w in seine homogenen Komponenten; für jedes $j \in \underline{k}$ heißt w_j der homogene Bestandteil (oder die homogene Komponente) vom Grad n_j von w.

Beweis zu 2.16: Offenbar ist X ein ${}^K\mathfrak{A}$- (${}^K\mathfrak{A}_1$-, ${}^K\mathfrak{R}$- bzw. ${}^K\mathfrak{R}_1$-) Erzeugendensystem von $KX^{(+)}$ ($KX^{(*)}$, KX^+ bzw. KX^*). Wir müssen also nur noch jeweils die Unabhängigkeit der Teilmenge X zeigen.

Sei φ eine Abbildung von X in eine K-Algebra M. Aufgrund der $\mathfrak{Ma}$-Freiheit von $X^{(+)}$ läßt sich φ eindeutig zu einem $\mathfrak{Ma}$-Homomorphismus $\overline{\varphi}$ von $X^{(+)}$ in das Magma $(M, \cdot)$ fortsetzen. Aufgrund der ${}^K\mathfrak{M}$-Freiheit von $KX^{(+)}$ läßt sich $\overline{\varphi}$ eindeutig zu einem ${}^K\mathfrak{M}$-Homomorphismus $\overline{\overline{\varphi}}$ von $KX^{(+)}$ in den K-Modul $(M, +)$ fortsetzen. Wir zeigen

$(*)$ $\qquad\qquad$ $\overline{\overline{\varphi}}$ ist ein ${}^K\mathfrak{A}$-Homomorphismus:

Seien $v, w \in KX^{(+)}$ und $v_1, \ldots, v_k, w_1, \ldots, w_l \in X^{(+)}$, $r_1, \ldots, r_k, s_1, \ldots, s_l \in K$ mit $v = \sum\limits_{i \in \underline{k}} r_i v_i$, $w = \sum\limits_{i \in \underline{l}} s_i w_i$. Da $\overline{\overline{\varphi}}$ ein ${}^K\mathfrak{M}$-Homomorphismus, $\overline{\varphi}$ ein $\mathfrak{Ma}$-Homomorphismus und $\overline{\overline{\varphi}}$ eine Fortsetzung von $\overline{\varphi}$ ist, gilt:

$$\begin{aligned}(vw)\overline{\overline{\varphi}} &= \Big(\sum_{i \in \underline{k}, j \in \underline{l}} r_i s_j v_i w_j\Big)\overline{\overline{\varphi}} = \sum_{i \in \underline{k}, j \in \underline{l}} r_i s_j (v_i w_j)\overline{\varphi} = \sum_{i \in \underline{k}, j \in \underline{l}} r_i s_j (v_i\overline{\varphi})(w_j\overline{\varphi}) \\ &= \sum_{i \in \underline{k}} r_i (v_i\overline{\varphi}) \sum_{j \in \underline{l}} s_j (w_j\overline{\varphi}) = \Big(\sum_{i \in \underline{k}} r_i v_i\Big)\overline{\overline{\varphi}}\Big(\sum_{j \in \underline{l}} s_j w_j\Big)\overline{\overline{\varphi}} = v\overline{\overline{\varphi}}\, w\overline{\overline{\varphi}}.\end{aligned}$$

Also gilt $(*)$. Ist weiter $\tilde{\varphi}$ irgendeine Forsetzung von φ zu einem ${}^K\mathfrak{A}$-Homomorphismus von $KX^{(+)}$ in M, so ist $\tilde{\varphi}|_{X^{(+)}}$ ein $\mathfrak{Ma}$-Homomorphismus mit $\tilde{\varphi}|_X = \varphi$, also $\tilde{\varphi}|_{X^{(+)}} = \overline{\varphi}$. Da $\tilde{\varphi}$ auch ein ${}^K\mathfrak{M}$-Homomorphismus ist, folgt daraus $\tilde{\varphi} = \overline{\overline{\varphi}}$. Also ist X eine ${}^K\mathfrak{A}$-unabhängige Teilmenge von $KX^{(+)}$.

Ist M unitär, so setzen wir $\overline{\overline{\varphi}}$ K-linear auf $KX^{(*)}$ so fort, daß das Element von $X^{(0)}$ auf 1_M abgebildet wird. Diese Fortsetzung ist, wie aufgrund von 2.16.1 leicht zu sehen, ein ${}^K\mathfrak{A}_1$-Homomorphismus von $KX^{(+)}$ in M und auch die einzig mögliche Fortsetzung zu einem solchen. Ist andererseits M assoziativ, so ist $\overline{\overline{\varphi}}$ kompatibel mit der Partition von $X^{(+)}$, die durch die Bildgleichheit unter dem kanonischen Epimorphismus $X^{(+)} \to X^+$ gegeben ist, und induziert folglich einen φ fortsetzenden ${}^K\mathfrak{R}$-Homomorphismus von KX^+ in M. Dessen eindeutige Bestimmtheit folgt wieder daraus, daß X ein ${}^K\mathfrak{R}$-Erzeugendensystem von KX^+ ist. Ähnlich wie oben auf $KX^{(*)}$ (und damit wieder auf die einzig mögliche Weise) wird schließlich φ auf KX^* fortgesetzt, wenn M sowohl unitär als auch assoziativ ist. □

2.17 Definition Sei $K \in \mathfrak{R}_1$, kommutativ, X eine Menge. Ist $u \in X^*$ mit $l(u) = n$, so gibt es genau $n+1$ Paare $(v, w) \in X^* \times X^*$ mit $vw = u$. Daher können wir eine Multiplikation auf $\overline{KX^*}$ definieren durch

$$\sum_{v \in X^*} r_v v \sum_{w \in X^*} s_w w := \sum_{u \in X^*} \Big(\sum_{v,w \in X^*, vw=u} r_v s_w \Big) u.$$

Damit wird die vor 2.15 auf KX^* eingeführte Produktbildung auf $\overline{KX^*}$ fortgesetzt und letzteres zu einer K-Links-Algebra gemacht. Durch eine routinemäßige Anwendung des Erweiterungsprinzips können wir $\overline{KX^*}$ durch eine zu ihr isomorphe, aber K *enthaltende* Algebra ersetzen; dazu nehmen wir im folgenden an: $K \cap X^+ = \emptyset$ (was man durch Anwendung des Entgiftungssatzes stets erreichen kann). Die Abbildung

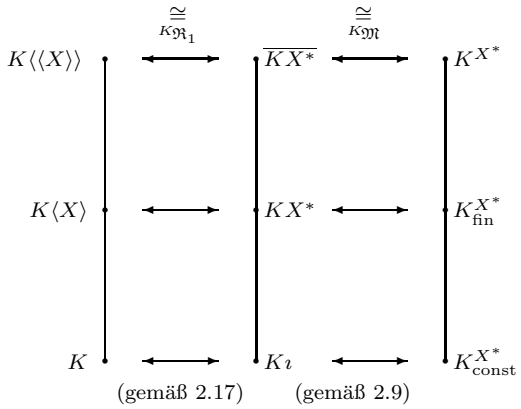
$$\iota : K \cup X^+ \to \overline{KX^*}, \quad \begin{cases} r & \mapsto r\iota \text{ für } r \in K \\ v & \mapsto v \text{ für } v \in X^+ \end{cases}$$

ist injektiv und bildet 1_K auf ι ab. Nach dem Erweiterungsprinzip läßt sich ι zu einem ${}^K\mathfrak{R}_1$-Isomorphismus $\overline{\iota}$ einer $K \cup X^+$ enthaltenden K-Algebra $K\langle\langle X \rangle\rangle$ (die o.B.d.A. das Element ι nicht enthält) auf $\overline{KX^*}$ fortsetzen. Die assoziative unitäre K-Algebra $K\langle\langle X \rangle\rangle$ ist bis auf $(K \cup X^+)$-Isomorphie eindeutig bestimmt und heißt die Algebra der Potenzreihen in der nichtkommutierenden Variablenmenge X über K (oder: nichtkommutative Potenzreihenalgebra in X über K). Ihre Elemente heißen Potenzreihen in der nichtkommutierenden Variablenmenge X über K. Das Urbild von KX^* in $K\langle\langle X \rangle\rangle$ heißt die Algebra der Polynome in der nichtkommutierenden Variablenmenge X über K (oder: nichtkommutative Polynomalgebra in X über K); Schreibweise: $K\langle X \rangle$. Ihre Elemente heißen Polynome in der nichtkommutierenden Variablenmenge X über K. Ein Element der Form rv mit $r \in K \setminus \{0_K\}$, $v \in X^*$ heißt ein Monom.

Das nachstehende Diagramm skizziert die in 2.17 bzw. 2.9 eingeführten Strukturen und ihre Beziehungen zueinander, wobei $K^{X^*}_{\text{fin}}$ die Menge der fast überall verschwindenden, $K^{X^*}_{\text{const}}$ die der konstanten Abbildungen von X^* in K ist. (Man beachte, daß die Rolle der Menge X in 2.9 hier von X^* übernommen wird!) Wir erlauben uns, die Elemente von $K\langle\langle X \rangle\rangle$ genauso zu schreiben wie die Elemente von $\overline{KX^*}$, nämlich in der Form $\sum_{v \in X^*} r_v v$, wobei der wegen $\iota \notin K\langle\langle X \rangle\rangle$ in $K\langle\langle X \rangle\rangle$ undefinierte Term $r_\iota \iota$ zu r_ι definiert wird; er wird das absolute Glied der Potenzreihe genannt. Gilt $w = \sum_{v \in X^*} r_v v \neq 0_K$, so wird die Zahl

$$w\underline{\delta} := \min\{n | n \in \mathbb{N}_0, \text{ es gibt ein } v \in X^n \text{ mit } r_v \neq 0_K\}$$

der Minimalgrad von w genannt. Das Element $\sum_{v\in X^n} r_v v$ heißt der homogene Bestandteil vom Grad n von w (vgl.2.16.1).



Für die *kommutative* Potenzreihen-Algebra in X über K schreiben wir hingegen $K[[X]]$ und $K[X]$, wie üblich, für den *kommutativen* Polynomring in X über K. Ist X einelementig, so verwenden wir diese Schreibweise mit dem Element von X anstelle von X. Offensichtlich gilt $K\langle\langle x\rangle\rangle = K[[x]]$, $K\langle x\rangle = K[x]$.

2.17.1 *Die Abbildung $\alpha : K\langle\langle X\rangle\rangle \to K$, $\sum_{v\in X^*} r_v v \mapsto r_\imath$ ist ein ${}^K\mathfrak{R}_1$-Homomorphismus. Der Kern von α besteht aus den Potenzreihen, deren absolutes Glied 0_K ist.* □

2.17.2 *Für alle $w, w' \in K\langle\langle X\rangle\rangle$ mit $ww' \neq 0_K$ gilt: $(ww')\underline{\delta} \geq w\underline{\delta} + w'\underline{\delta}$. Insbesondere gilt für alle $j \in \mathbb{N}_0$ mit $w^j \neq 0_K$: $w^j\underline{\delta} \geq j \cdot w\underline{\delta}$. Sind v, v' die homogenen Bestandteile von w, w' vom Grad $w\underline{\delta}$, $w'\underline{\delta}$ und gilt $vv' \neq 0_K$, so ist vv' der homogene Bestandteil minimalen Grades von ww', und es folgt $(ww')\underline{\delta} = w\underline{\delta} + w'\underline{\delta}$.* □

2.17.3 *Für alle $w \in K\langle\langle X\rangle\rangle$ mit $w\underline{\delta} \geq 1$ ist $1_K + w$ ein Element der Einheitengruppe von $K\langle\langle X\rangle\rangle$, und zwar gilt: $(1_K + w)\sum_{j\in\mathbb{N}_0}(-w)^j = 1_K$.*

Beweis: Sei $w \in K\langle\langle X\rangle\rangle$ mit $w\underline{\delta} \geq 1$. Für alle $n \in \mathbb{N}$ hat

$$\begin{aligned}(1_K + w)\sum_{j\in\mathbb{N}_0}(-w)^j &= (1_K + w)\sum_{j=0}^{n-1}(-w)^j + (1_K + w)\sum_{j\geq n}(-w)^j \\ &= 1_K + (-1)^n w^{n+1} + (1_K + w)\sum_{j>n}(-w)^j\end{aligned}$$

0_K als homogenen Bestandteil vom Grad n (2.17.2). □

Folgerung *Sei α wie in* 2.17.1 *und $J := \operatorname{Kern}\alpha$. Dann ist J eine Radikalalgebra (d.h. eine Algebra, die gleich ihrem eigenen Jacobson-Radikal* [6] *ist). Genau dann ist J das Jacobson-Radikal von $K\langle\langle X\rangle\rangle$, wenn das Jacobson-Radikal von K gleich $\{0_K\}$ ist.*

Denn aus 2.17.3 folgt, daß jedes Element von J im Jacobson-Radikal von J liegt. Es folgt die erste Aussage. Nach 2.17.3 und dem Homomorphiesatz gilt $K\langle\langle X\rangle\rangle/J \underset{K_{\mathfrak{R}_1}}{\cong} K$, woraus die zweite Aussage folgt. □

Die folgende Bemerkung wird uns anschließend im Spezialfall $K = \mathbb{Z}$ einen wichtigen Dienst leisten:

2.17.4 *Sei $w \in K\langle\langle X\rangle\rangle$ mit $w\underline{\delta} \geq 1$. Dann gibt es zu jedem $z \in \mathbb{Z}$ eine Potenzreihe $u \in K[[w]]$ mit $(1_K + w)^z = 1_K + zw + w^2u$.*

Beweis: Für jedes Element r eines unitären Ringes R gilt für beliebiges $n \in \mathbb{N}$:

$$(1_R + r)^n = 1_R + nr + \binom{n}{2}r^2 + \cdots + \binom{n}{n}r^n.$$

Daraus folgt zunächst unsere Behauptung im Falle $z \in \mathbb{N}$, und darüber hinaus für alle $n \in \mathbb{N}$ nach 2.17.3:

$$\begin{aligned}(1_K + w)^{-n} &= \big(1_K + \sum_{j\in\mathbb{N}}(-w)^j\big)^n \\ &= 1_K + n\sum_{j\in\mathbb{N}}(-w)^j + \binom{n}{2}\big(\sum_{j\in\mathbb{N}}(-w)^j\big)^2 + \cdots + \binom{n}{n}\big(\sum_{j\in\mathbb{N}}(-w)^j\big)^n \\ &= 1_K - nw + nw^2\sum_{j\in\mathbb{N}_{>1}}(-1)^{j-1}w^{j-2} + \sum_{k=2}^{n}\binom{n}{k}\big(w\sum_{j\in\mathbb{N}}(-1)^j w^{j-1}\big)^k \\ &= 1_K - nw + w^2u\end{aligned}$$

für ein $u \in K[[w]]$, wie behauptet. □

2.18 Lemma (Magnus 1935) *Für jede Menge X ist $1 + X$ eine $\mathfrak{G}$-unabhängige Teilmenge der Einheitengruppe von $\mathbb{Z}\langle\langle X\rangle\rangle$.*

Beweis. Seien $l > 0$, $y_1, \ldots, y_l \in X$, $\varepsilon_1, \ldots, \varepsilon_l \in \{1, -1\}$, so daß für alle $i \in \underline{l-1}$ gilt: $y_i = y_{i+1} \Rightarrow \varepsilon_i = \varepsilon_{i+1}$. Wir zeigen: $(1 + y_1)^{\varepsilon_1} \cdots (1 + y_l)^{\varepsilon_l} \neq 1$.

[6] Das Jacobson-Radikal einer assoziativen Algebra ist das größte Ideal, das bezüglich $*$ eine Gruppe ist, wo $u * v = u + v + uv$.

Durch maximales Zusammenfassen im Produkt benachbarter Faktoren $(1+y_i)^{\varepsilon_i}$, $(1+y_{i+1})^{\varepsilon_{i+1}}, \cdots$ mit $y_i = y_{i+1} = \cdots$ zu einer einzigen Potenz erhalten wir: Es gibt ein $k \in \underline{l}$ und $x_1, \ldots, x_k \in X$, $z_1, \ldots, z_k \in \mathbb{Z} \setminus \{0\}$ mit $x_i \neq x_{i+1}$ für alle $i \in \underline{k-1}$, so daß unter Verwendung von 2.17.4 gilt:

$$\begin{aligned}(1+y_1)^{\varepsilon_1} \ldots (1+y_l)^{\varepsilon_l} &= (1+x_1)^{z_1} \cdots (1+x_k)^{z_k} \\ &= (1+z_1x_1+x_1^2u_1) \cdots (1+z_kx_k+x_k^2u_k)\end{aligned}$$

für geeignete Elemente $u_1, \ldots, u_k \in K\langle\langle X\rangle\rangle$. Alle Monome, die in $x_j^2u_j$ auftreten, haben den quadratischen Linksfaktor x_j^2. Da in dem Monom $x_1 \cdots x_k$ kein Quadrat eines Buchstabens als Teilsilbe vorkommt, hat in dem Produkt $(1+z_1x_1+x_1^2u_1) \cdots (1+z_kx_k+x_k^2u_k)$ das Monom $x_1 \cdots x_k$ denselben Koeffizienten wie in $(1+z_1x_1) \cdots (1+z_kx_k)$, m.a.W.: dieser ist gleich $z_1 \cdots z_k$ und damit $\neq 0$. Es folgt: $(1+y_1)^{\varepsilon_1} \ldots (1+y_l)^{\varepsilon_l} \neq 1$. Aus 2.4(2) folgt die Behauptung. □

2.19 Satz *Für jede Menge X gilt:*

(1) *Es gibt eine von X $\mathfrak{G}$-frei erzeugte Gruppe.*
(Im folgenden bezeichne F „die“ (siehe 2.1.1) freie Gruppe über X.)

(2) *F/F' ist eine von $\{F'x | x \in X\}$ $\mathfrak{M}$-frei erzeugte abelsche Gruppe vom Rang $|X|$; die Zuordnung $X \to F/F'$, $x \mapsto F'x$ ist injektiv.*

(3) *Ist Y ein $\mathfrak{G}$-Erzeugendensystem von F, so gilt $|X| \leq |Y|$. Je zwei $\mathfrak{G}$-Basen von F sind gleichmächtig.*

(4) **(Nielsen 1918)** *Ist X endlich und Y ein $\mathfrak{G}$-Erzeugendensystem von F mit $|Y| \leq |X|$, so ist $|Y| = |X|$ und Y eine $\mathfrak{G}$-Basis von F.*

Beweis. (1) Sei $\iota : X \to \mathbb{Z}\langle\langle X\rangle\rangle$, $x \mapsto 1+x$. Nach 2.18 ist $\langle X\iota\rangle_{\mathfrak{G}}$ eine von $X\iota$ frei erzeugte Untergruppe der Einheitengruppe von $\mathbb{Z}\langle\langle X\rangle\rangle$. Trivialerweise ist ι injektiv. Nach dem Erweiterungsprinzip gibt es also eine von X $\mathfrak{G}$-frei erzeugte Gruppe.[7]

(2) Sei $\overline{X} := \{F'x | x \in X\}$. Offensichtlich gilt: $\langle\overline{X}\rangle_{\mathfrak{M}} = F/F'$. Nun sei ψ eine Abbildung von $\overline{X}$ in eine abelsche Gruppe A. Für jedes $x \in X$ setzen wir $x\varphi := (F'x)\psi$. Da F von X frei erzeugt wird, gibt es einen $\mathfrak{G}$-Homomorphismus $\overline{\varphi}$ von F in A mit $\overline{\varphi}|_X = \varphi$. Da A abelsch ist, folgt:

[7] Unter Verwendung der Folgerung zu 2.17.3 genügt es festzustellen, daß X nach 2.18 eine $\mathfrak{G}$-unabhängige Teilmenge der Gruppe $(J, *)$ ist, wobei J das Jacobson-Radikal von $\mathbb{Z}\langle\langle X\rangle\rangle$ ist.

$F' \subseteq \operatorname{Kern}\overline{\varphi}$, also $(F'g)\overline{\varphi} = \{g\overline{\varphi}\}$ für alle $g \in F$. Daher ist die Zuordnung $\overline{\psi} : F'g \mapsto g\overline{\varphi}$ ein $\mathfrak{M}$-Homomorphismus von F/F' in A mit

$$(F'x)\overline{\psi} = x\overline{\varphi} = x\varphi = (F'x)\psi \quad \text{für alle } x \in X.$$

Sind x_1, $x_2 \in X$ mit $F'x_1 = F'x_2$, so liegt $x_1x_2^{-1}$ in F' und damit im Kern eines jeden Homomorphismus von F in eine abelsche Gruppe. Sei

$$\varphi : X \to \mathbb{Z},\ x \mapsto \begin{cases} 1 & \text{falls } x = x_1, \\ 0 & \text{sonst.} \end{cases}$$

Da F von X frei erzeugt wird, gibt es einen Homomorphismus $\overline{\varphi}$ von F in $\mathbb{Z}$ mit $\overline{\varphi}|_X = \varphi$. Es folgt: $0 = (x_1x_2^{-1})\overline{\varphi} = x_1\overline{\varphi} - x_2\overline{\varphi} = 1 - x_2\varphi$, also $x_2\varphi = 1$ und damit $x_2 = x_1$.

(3) Ist Y ein $\mathfrak{G}$-Erzeugendensystem von F, so ist $\{F'y \mid y \in Y\}$ ein $\mathfrak{M}$-Erzeugendensystem von F/F'. Nach (2) und 2.13 gilt also:

$$|Y| \geq \big|\{F'y \mid y \in Y\}\big| \geq \big|\{F'x \mid x \in X\}\big| = |X|.$$

Die letzte Behauptung folgt nun aus dem Satz von Schröder-Bernstein.

(4) Sei $\overline{Y} := \{F'y \mid y \in Y\}$. Dann ist $\overline{Y}$ ein $\mathfrak{M}$-Erzeugendensystem von F/F', und

$$|\overline{Y}| \leq |Y| \leq |X| = \big|\{F'x \mid x \in X\}\big| \quad \text{nach (2).}$$

Nach dem letzten Teil von 2.13 gilt nun $|\overline{Y}| = \big|\{F'x \mid x \in X\}\big|$ und folglich $|Y| = |X|$. Weiter ist wegen $|\overline{Y}| = |Y|$ die Abbildung $Y \to \overline{Y}$, $y \mapsto F'y$ bijektiv. Da $\overline{Y}$ eine $\mathfrak{M}$-Basis von F/F' ist, gibt es keine nichttriviale $\overline{Y}$-Darstellung von $1_{F/F'}$. Daher gibt es erst recht keine nichttriviale Y-Darstellung von 1_F. Nach 2.4(2) ist Y also $\mathfrak{G}$-unabhängig. □

Ist F eine freie Gruppe, so heißt die nach 2.19(3) eindeutig bestimmte Mächtigkeit einer $\mathfrak{G}$-Basis von F der Rang von F. Aus 2.19(3) erhält man speziell:

2.19.1 *Eine endlich erzeugte freie Gruppe ist von endlichem Rang.* □

Kapitel 3

Freie Gruppen

Ist G eine Gruppe und $X \subseteq G$, so setzen wir $X^{-1} := \{x^{-1} | x \in X\}$ und $\widetilde{X} := X \cup X^{-1}$. Offenbar gilt dann:

3.0.1 $\langle X \rangle_{\mathfrak{G}} = \langle \widetilde{X} \rangle_{\mathfrak{S}_1}$. □

Da auf Gruppen als spezielle Monoide die Begriffsbildungen aus Kapitel 1 anwendbar sind, können wir insbesondere den Begriff der $\widetilde{X}$-Länge (siehe 1.6) eines Elements $g \in \langle X \rangle_{\mathfrak{G}}$ betrachten. Es gilt:

3.0.2 $\forall g, g' \in \langle X \rangle_{\mathfrak{G}} \quad l_{\widetilde{X}}(gg') \leq l_{\widetilde{X}}(g) + l_{\widetilde{X}}(g')$. □

3.0.3 *Für alle $g \in \langle X \rangle_{\mathfrak{G}}$ gilt: Ist $((x_1, \varepsilon_1), \ldots, (x_k, \varepsilon_k))$ eine X-Darstellung von g, so ist $l_{\widetilde{X}}(g) \leq k$.* □

Der Konjugiertheitsbegriff in der Gruppentheorie entspricht dem für das freie Monoid definierten (siehe 1.8 und 1.9), denn es gilt:

3.0.4 *Für je zwei Elemente g, g' einer Gruppe G sind äquivalent:*
$\exists h \in G \quad h^{-1}gh = g', \quad \exists h \in G \quad gh = hg', \quad \exists u, v \in G \quad g = uv,\ g' = vu,$

Gilt nämlich $gh = hg'$ für ein $h \in G$, so setzen wir $u := h$, $v := g'h^{-1}$ und erhalten die Gleichungen $g = uv$, $g' = vu$. Alles Übrige ist trivial. □

Ist X ein $\mathfrak{G}$-Erzeugendensystem der Gruppe G, also (nach 3.0.1) $\langle \widetilde{X} \rangle_{\mathfrak{S}_1} = G$, so gibt es aufgrund der Freiheit des Monoids $\widetilde{X}^*$ über $\widetilde{X}$ einen $\mathfrak{S}_1$-Epimorphismus ψ von $\widetilde{X}^*$ auf G mit $\psi|_{\widetilde{X}} = id_{\widetilde{X}}$. Wir schreiben (zur Unterscheidung von Produktbildungen in G) „.“ für die Multiplikation in $\widetilde{X}^*$. Es gilt:

$$\iota\psi = 1_G = yy^{-1} = (y\psi)(y^{-1}\psi) = (y.y^{-1})\psi \quad \text{für alle } y \in \widetilde{X}.$$

Die Relation $\underset{\psi}{\sim}$ auf $\widetilde{X}^*$ mit

$$u \underset{\psi}{\sim} v \Leftrightarrow u\psi = v\psi \quad \text{für alle } u,\ v \in \widetilde{X}$$

ist daher eine verknüpfungsverträgliche Äquivalenzrelation (Kongruenzrelation), die alle Paare $(y.y^{-1}, \imath)$ mit $y \in \widetilde{X}$ enthält. Die folgende Proposition beschreibt den Zusammenhang zwischen der freien Gruppe über X und dem freien Monoid über $\widetilde{X}$:

3.1 Proposition *Sei X eine Menge, F die freie Gruppe über X, $\widetilde{X} := X \cup X^{-1}$ und $\sim$ die kleinste Kongruenzrelation auf $\widetilde{X}^*$, die alle Paare $(y.y^{-1}, \imath)$ mit $y \in \widetilde{X}$ enthält. Dann gilt:*

$$\sim = \underset{\psi}{\sim}, \quad \widetilde{X}^*/_{\sim} \cong F.$$

Beweis. Für alle $w \in \widetilde{X}^*$ bezeichne $[w]$ die Äquivalenzklasse bezüglich $\sim$, die w enthält. Wir wollen zunächst feststellen, daß das Monoid $\widetilde{X}^*/_{\sim}$ eine Gruppe ist und müssen dazu nur einsehen, daß das $\mathfrak{S}_1$-Erzeugendensystem $\{[y] | y \in \widetilde{X}\}$ aus invertierbaren Elementen besteht. Dies ergibt sich aus der für alle $x \in X$ aufgrund der Definition von $\sim$ geltenden Gleichungskette

$$[x][x^{-1}] = [x.x^{-1}] = [\imath] = [x^{-1}.x] = [x^{-1}][x].$$

Da damit für alle $x \in X$ gilt: $[x]^{-1} = [x^{-1}]$, ist $\{[x] | x \in X\}$ ein $\mathfrak{G}$-Erzeugendensystem von $\widetilde{X}^*/_{\sim}$. Aufgrund der Freiheit von F über X gibt es nun einen $\mathfrak{G}$-Epimorphismus σ von F auf $\widetilde{X}^*/_{\sim}$ mit $x\sigma = [x]$ für alle $x \in X$.

Nach der Definition von $\sim$ gilt offensichtlich $\sim \subseteq \underset{\psi}{\sim}$. Für alle $w \in \widetilde{X}^*$ bezeichne $[w]_{\underset{\psi}{\sim}}$ die w enthaltende Kongruenzklasse bezüglich $\underset{\psi}{\sim}$ in $\widetilde{X}^*$. Die Inklusions-Abbildung

$$\chi : \ \widetilde{X}^*/_{\sim} \to \widetilde{X}^*/_{\underset{\psi}{\sim}}, \quad [w] \mapsto [w]_{\underset{\psi}{\sim}}$$

ist ein Epimorphismus. Bezeichnet Ψ den von ψ induzierten Isomorphismus von $\widetilde{X}^*/_{\underset{\psi}{\sim}}$ auf F (also: $[w]_{\underset{\psi}{\sim}}\Psi = w\psi$ für alle $w \in \widetilde{X}^*$), so gilt $x\sigma\chi\Psi = x\psi = x$ für jedes $x \in X$, d.h. es gilt: $\sigma(\chi\Psi) = id_F$. Also sind σ und χ injektiv. Es folgt: $\sim = \underset{\psi}{\sim}$, $\widetilde{X}^*/_{\sim} \cong F$. □

3.2 Proposition *Sei X eine Menge und F die freie Gruppe über X, $\widetilde{X} := X \cup X^{-1}$, $g \in F$, $((x_1, \varepsilon_1), \ldots, (x_k, \varepsilon_k))$ eine X-Darstellung von g. Es sind äquivalent:*

(i) $l_{\widetilde{X}}(g) = k$,

(ii) $x_i^{\varepsilon_i} x_{i+1}^{\varepsilon_{i+1}} \neq 1_F$ *für alle* $i \in \underline{k-1}$,

(iii) $((x_1, \varepsilon_1), \ldots, (x_k, \varepsilon_k))$ *ist gekürzt.*

Beweis. (i)⇒(ii): Wäre $x_i^{\varepsilon_i} x_{i+1}^{\varepsilon_{i+1}} = 1_F$ für ein $i \in \underline{k-1}$, so $l_{\widetilde{X}}(g) < k$ wegen $g = x_1^{\varepsilon_1} \cdots x_{i-1}^{\varepsilon_{i-1}} x_{i+2}^{\varepsilon_{i+2}} \cdots x_k^{\varepsilon_k}$.

(ii)⇒(iii): Sei $i \in \underline{k-1}$ mit $x_i = x_{i+1}$. Wäre $\varepsilon_i \neq \varepsilon_{i+1}$, so $\varepsilon_i = -\varepsilon_{i+1}$ und $x_i^{\varepsilon_i} x_{i+1}^{\varepsilon_{i+1}} = x_i^{\varepsilon_i} x_i^{-\varepsilon_i} = 1_F$, im Widerspruch zu (ii).

(iii)⇒(i): Nach 3.0.3 gilt $l_{\widetilde{X}}(g) \leq k$. Gälte $l_{\widetilde{X}}(g) < k$, so wäre nach 2.2.3 die (nach 2.4(2) eindeutig bestimmte) gekürzte X-Darstellung von g von einer Länge kleiner als k, im Widerspruch zu (iii). □

3.3 Definition Sei X eine Menge und F die freie Gruppe über X. Für Elemente $g, g_1, \ldots, g_n \in F$ bedeute die Schreibweise

$$g = g_1 | g_2 | \cdots | g_n : \quad g = g_1 g_2 \cdots g_n \text{ und } l_{\widetilde{X}}(g) = l_{\widetilde{X}}(g_1) + l_{\widetilde{X}}(g_2) + \cdots + l_{\widetilde{X}}(g_n).$$

Sie drückt aus, daß die gekürzte X-Darstellung von g durch „Aneinanderhängen“ der gekürzten X-Darstellungen der g_i entsteht. Gilt $g = g_1 | g_2$, so heißt g_1 ein Linksfaktor (Schreibweise in Anlehnung an die bei der entsprechenden Begriffsbildung bei freien Monoiden, s. 1.8: $g_1 \restriction g$), g_2 ein Rechtsfaktor ($g_2 \upharpoonright g$) von g; ein Links- oder Rechtsfaktor g_i von g wird echt genannt, wenn er von von g verschieden ist. Bei der Verwendung der eingeführten Schreibweise mit mehr als zwei Faktoren gilt es zu beachten, daß sie nur im Falle $g_2, \ldots, g_{n-1} \neq 1_F$ dasselbe aussagt wie „$g_i g_{i+1} = g_i | g_{i+1}$ für alle $i \in \underline{n-1}$“:

3.3.1 *Für alle* g_1, g_2, $g_3 \in F$ *gilt:*

$$g = g_1 | g_2 | g_3 \Leftrightarrow g_1 g_2 = g_1 | g_2, g_2 g_3 = g_2 | g_3 \text{ und: } g_1 g_3 = g_1 | g_3 \text{ im Falle } g_2 = 1_F.$$

□

Jedes Element $g \in F \setminus \{1_F\}$ hat einen eindeutig bestimmten echten Linksfaktor g' maximaler $\widetilde{X}$-Länge. Dies ist der Linksfaktor, zu dem es ein (dann ebenfalls eindeutig bestimmtes) Paar $(x_g, \varepsilon_g) \in X \times \{1, -1\}$ gibt mit $g = g' \restriction x_g^{\varepsilon_g}$. Wir nennen g positiv, wenn $\varepsilon_g = 1$ gilt, andernfalls negativ, und bemerken:

3.3.2 $f : F \setminus \{1_F\} \to F \times X$, $g \mapsto \begin{cases} (g', x_g) \text{ falls } g \text{ positiv} \\ (g, x_g) \text{ falls } g \text{ negativ} \end{cases}$ *ist injektiv.*

Haben nämlich $g, h \in F$ unter f dasselbe Bild, so gilt $x_g = x_h$. Wäre nun etwa g positiv, h negativ, so $g' = h$ aufgrund der Gleichheit der ersten Komponenten von gf und hf, mit dem Widerspruch $g = g' \shortmid x_g = h \shortmid x_g,\ x_g^{-1} \shortmid h$. Wir erschließen damit $\varepsilon_g = \varepsilon_h$ und weiter $(g', x_g) = (h', x_h)$, falls g, h positiv, bzw. $(g, x_g) = (h, x_h)$ falls g, h negativ. In jedem Fall folgt $g = h$. □

Ein Element $g \in F$ heißt zyklisch (X-)gekürzt, wenn $g = 1_F$ oder für die gekürzte X-Darstellung $((x_1, \varepsilon_1), \ldots, (x_k, \varepsilon_k))$ von g gilt: $x_k^{\varepsilon_k} x_1^{\varepsilon_1} \neq 1_F$, d.h. wenn über das Produkt je zweier benachbarter Faktoren $x_i^{\varepsilon_i} x_{i+1}^{\varepsilon_{i+1}}$ hinaus auch das Produkt des letzten Faktors mit dem ersten $\neq 1_F$ ist. Besteht (wie hier) bezüglich der Menge X als unabhängiger Erzeugermenge von F keine Gefahr einer Verwechslung, so lassen wir das Präfix „(X-)“ fort. Die folgende Bemerkung zeigt insbesondere, daß jedes Element von F zu einem zyklisch gekürzten Element konjugiert ist:

3.3.3 *Für jedes $g \in F$ gibt es ein zyklisch gekürztes Element $\underline{g} \in F$ und ein $h \in F$ mit $g = h^{-1} \shortmid \underline{g} \shortmid h$.*

Dies ist für $g = 1_F$ trivial. Sei $g \in F \setminus \{1_F\}$ und $((x_1, \varepsilon_1), \ldots, (x_k, \varepsilon_k))$ die gekürzte X-Darstellung von g. Dann gibt es ein $i \in \underline{k}$ mit $x_i^{\varepsilon_i} x_{k+1-i}^{\varepsilon_{k+1-i}} \neq 1_F$, und wir nehmen hierbei i als minimal an. Das Element $\underline{g} := x_i^{\varepsilon_i} x_{i+1}^{\varepsilon_{i+1}} \cdots x_{k+1-i}^{\varepsilon_{k+1-i}}$ ist dann zyklisch gekürzt, und für $h := x_{k-i+2}^{\varepsilon_{k-i+2}} \cdots x_k^{\varepsilon_k}$ gilt:

$$\begin{aligned} h^{-1}\underline{g}h &= (x_{k-i+2}^{\varepsilon_{k-i+2}} \cdots x_k^{\varepsilon_k})^{-1} \underline{g} (x_{k-i+2}^{\varepsilon_{k-i+2}} \cdots x_k^{\varepsilon_k}) = x_k^{-\varepsilon_k} \cdots x_{k-i+2}^{-\varepsilon_{k-i+2}} \underline{g} x_{k-i+2}^{\varepsilon_{k-i+2}} \cdots x_k^{\varepsilon_k} \\ &= x_1^{\varepsilon_1} \cdots x_{i-1}^{\varepsilon_{i-1}} \shortmid \underline{g} \shortmid x_{k-i+2}^{\varepsilon_{k-i+2}} \cdots x_k^{\varepsilon_k} = g, \end{aligned}$$

nach Wahl von i. Also gilt die Behauptung. □

3.4 Korollar *Jede freie Gruppe ist torsionsfrei (d.h. für jedes nicht-neutrale Element g einer freien Gruppe gilt: $\langle g \rangle_{\mathfrak{G}} \cong \mathbb{Z}$.)*

Beweis. Sei F eine freie Gruppe, X eine $\mathfrak{G}$-Basis von F, $g \in F \setminus \{1_F\}$. Nach 3.3.3 gibt es ein zyklisch gekürztes Element $\underline{g} \neq 1_F$ und ein $h \in F$ mit $g = h^{-1} \shortmid \underline{g} \shortmid h$. Wir zeigen:

$$(*) \qquad l_{\widetilde{X}}(g^n) \geq n l_{\widetilde{X}}(\underline{g}) \quad \text{für alle } n \in \mathbb{N}:$$

Es gilt:

$$g^n = (h^{-1} \shortmid \underline{g} \shortmid h)(h^{-1} \shortmid \underline{g} \shortmid h) \cdots = h^{-1} \shortmid (\underline{g})^n \shortmid h = h^{-1} \shortmid \underbrace{\underline{g} \shortmid \cdots \shortmid \underline{g}}_{n} \shortmid h,$$

da $\underline{g}$ zyklisch gekürzt ist. Daraus folgt: $l_{\widetilde{X}}(g^n) = 2 l_{\widetilde{X}}(h) + n l_{\widetilde{X}}(\underline{g})$, damit $(*)$. Aus $(*)$ folgt $g^n \neq 1_F$ für alle $n \in \mathbb{N}$. □

Die folgende Proposition klärt, wann zwei zyklisch gekürzte Elemente konjugiert sind:

3.5 Proposition *Sei X eine Menge und F die freie Gruppe über X. Seien g, g' zyklisch gekürzte Elemente von F. Genau dann sind g, g' in F konjugiert, wenn es f, $h \in F$ gibt mit $g = f|h$, $g' = h|f$. D.h.: g, g' sind genau dann in der freien Gruppe F konjugiert, wenn die gekürzten X-Darstellungen von g und g' im freien Monoid $\mathcal{T}(X \times \{1, -1\})$ konjugiert (im Sinne von* 1.8*) sind.*

Beweis. Daß die angegebene Bedingung hinreichend für die Konjugiertheit von g und g' ist, folgt aus 3.0.4. Zum Beweis der Notwendigkeit sei o. B. d. A. $l_{\widetilde{X}}(g') \leq l_{\widetilde{X}}(g)$ und $a \in F$ mit $g' = a^{-1}ga$. Wir zeigen die Behauptung durch Induktion nach $l_{\widetilde{X}}(a)$. Im Falle $l_{\widetilde{X}}(a) = 0$, also $a = 1$, gilt $g' = g$, und die Behauptung ist trivial. Sei zum Induktionsschritt nun $l_{\widetilde{X}}(a) > 0$, $((x_1, \varepsilon_1), \ldots, (x_k, \varepsilon_k))$ die gekürzte X-Darstellung von g, $((y_1, \delta_1), \ldots, (y_l, \delta_l))$ die von a, z die von g' und z^- die von g'^{-1}. Es gilt:

$$g' = a^{-1}ga = y_l^{-\delta_l} \cdots y_1^{-\delta_1} x_1^{\varepsilon_1} x_2^{\varepsilon_2} \cdots x_k^{\varepsilon_k} y_1^{\delta_1} \cdots y_l^{\delta_l}.$$

1. Fall: $y_1^{-\delta_1} x_1^{\varepsilon_1} = 1_F$. Da g zyklisch gekürzt ist, folgt $x_k^{\varepsilon_k} y_1^{\delta_1} \neq 1_F$ und

$$g' = y_l^{-\delta_l} \cdots y_2^{-\delta_2} x_2^{\varepsilon_2} \cdots x_k^{\varepsilon_k} x_1^{\varepsilon_1} y_2^{\delta_2} \cdots y_l^{\delta_l}.$$

Induktiv folgt: $z \sim ((x_2, \varepsilon_2), \ldots, (x_k, \varepsilon_k), (x_1, \varepsilon_1)) \sim ((x_1, \varepsilon_1), \ldots, (x_k, \varepsilon_k))$.

2. Fall: $y_1^{-\delta_1} x_1^{\varepsilon_1} \neq 1_F$. Da g' zyklisch gekürzt ist, folgt dann $x_k^{\varepsilon_k} y_1^{\delta_1} = 1_F$, da andernfalls der Widerspruch $g' = a^{-1}|g|a$, also $l_{\widetilde{X}}(g') > l_{\widetilde{X}}(g)$ einträte. Wie die Gleichung

$$g'^{-1} = a^{-1}g^{-1}a = y_l^{-\delta_l} \cdots y_1^{-\delta_1} x_k^{-\varepsilon_k} \cdots x_1^{-\varepsilon_1} y_1^{\delta_1} \cdots y_l^{\delta_l}$$

zeigt, sind dann bezüglich g^{-1} statt g und g'^{-1} statt g' die Voraussetzungen des bereits erledigten 1. Falles gegeben, woraus $z^- \sim ((x_k, -\varepsilon_k), \ldots, (x_1, -\varepsilon_1))$ folgt, also $z \sim ((x_1, \varepsilon_1), \ldots, (x_k, \varepsilon_k))$. □

Nach den Konjugiertenklassen in freien Gruppen werden wir nun Untergruppen freier Gruppen betrachten. Eine alternative Möglichkeit, den Inhalt von 3.4 auszudrücken, ist die folgende:

3.4' *Jede zyklische Untergruppe einer freien Gruppe ist frei.*

Diese Einsicht soll durch unser nächstes Ziel, den wichtigen Satz von Nielsen und Schreier, in ihre volle Allgemeinheit gehoben werden: Nicht nur alle *zyklischen*, sondern *alle* Untergruppen einer freien Gruppe sind frei. Schreier bewies den Satz 1927, nachdem dieser zuvor von Nielsen schon für endlich erzeugte Untergruppen eingesehen worden war. Jahrzehnte unbemerkt blieb die Tatsache, daß das Resultat sich bereits in einer Arbeit von P. Hoyer aus dem Jahr 1902 findet. Die grobe Beweisskizze enthält nur zwei Schritte:

1) Zu einer beliebigen Gruppe G mit Erzeugendensystem X, einer Untergruppe H von G und einem Rechts-Repräsentantensystem R für H in G wird eine kanonische Konstruktion eines Erzeugendensystems $X_R^{(H)}$ von H angegeben. (Die Methode, nach der das geschieht, wird nach K. Reidemeister benannt: Ein in H liegendes Produkt über $\widetilde{X}$ wird als Produkt über $\widetilde{X_R^{(H)}}$ „umgeschrieben“ (3.6).) 2) Im Falle einer *freien* Gruppe G mit $\mathfrak{G}$-Basis X kann man R so wählen, daß $X_R^{(H)}$ $\mathfrak{G}$-unabhängig, also eine $\mathfrak{G}$-Basis von H ist. (Hinreichend dazu ist es, für R ein sog. „Schreiersystem“ zu wählen, das in 3.8 definiert wird.)

Schritt 1) ist der wesentlich einfachere, führt aber bereits zu Konsequenzen, die unabhängiges Interesse verdienen; siehe 3.7. Wir beginnen mit einigen Vorbemerkungen, in denen eine beliebige Gruppe G gegeben sein möge:

3.5.1 *Ist* $\overline{\ }$ *eine beliebige Abbildung von* G *in* G *mit* $\overline{1_G} = 1_G$*, so gilt für alle* $g_1, \ldots, g_n \in G$*:*

$$g_1 \cdots g_n = \Big(\prod_{j=1}^{n} \overline{g_1 \cdots g_{j-1}}\, g_j\, \overline{g_1 \cdots g_j}^{\,-1}\Big)\overline{g_1 \cdots g_n},$$

wobei das Produkt natürlich so zu verstehen ist, daß die Reihenfolge der Faktoren die des ansteigenden Laufindex j ist. □

Wir wählen für $\overline{\ }$ die Repräsentanten-Zuordnung bezüglich eines Rechts-Repräsentantensystems R einer Untergruppe H von G mit $1_G \in R$: Es sei

$$\begin{aligned} \overline{\ } : G &\to R \\ g &\mapsto r \text{ mit } Hg = Hr. \end{aligned}$$

3.5.2 $\forall g, g' \in G \quad \overline{gg'} = \overline{\overline{g}g'}, \quad \forall g \in G, r \in R \quad r = \overline{\overline{rg^{-1}}g},$

denn aus $Hg = H\overline{g}$ folgt: $Hgg' = H\overline{g}g'$, also gilt die erste Aussage. Die zweite ergibt sich durch Anwendung der ersten: $r = \overline{r} = \overline{rg^{-1}g} = \overline{\overline{rg^{-1}}g}$. □

Weiter setzen wir für alle $r \in R$, $g \in G$:

$$q(r,g) := rg\,\overline{rg}^{\,-1}.$$

Der „Quotient“ $q(r,g)$ ist der (in H liegende) Faktor, der das Element rg von dem in R enthaltenen Repräsentanten der Restklasse Hrg unterscheidet: $rg = q(r,g) \cdot \overline{rg}$. Wir bemerken für alle $g \in G$, $r \in R$:

3.5.3 $q(r, g^{-1}) = q(\overline{rg^{-1}}, g)^{-1}$,

denn $q(r, g^{-1})q(\overline{rg^{-1}}, g) = rg^{-1}\overline{rg^{-1}}^{-1}\overline{rg^{-1}}\, g\, \overline{\overline{rg^{-1}}g}^{-1} \underset{3.5.2}{=} rr^{-1} = 1_G.$ □

„Rollentausch von g und g^{-1}“ erweist die Äquivalenz von 3.5.3 mit

3.5.4 $q(r, g)^{-1} = q(\overline{rg}, g^{-1})$. □

3.5.5 $rg \in R \Leftrightarrow q(r, g) = 1_G \Leftrightarrow q(\overline{rg}, g^{-1}) = 1_G \Leftrightarrow \overline{rg}g^{-1} \in R$. □

3.6 Lemma *Sei G eine Gruppe, $H \leq_{\mathfrak{G}} G$, R ein Rechts-Repräsentantensystem für H in G mit $1_G \in R$, $\overline{\ }$ die zugehörige Repräsentanten-Zuordnung. Sei X ein $\mathfrak{G}$-Erzeugendensystem von G und*

$$X_R^{(H)} := \{q(r, x) \mid r \in R,\, x \in X\} \smallsetminus \{1_G\}.$$

Dann ist $X_R^{(H)}$ ein $\mathfrak{G}$-Erzeugendensystem von H und $l_{\widetilde{X_R^{(H)}}}(h) \leq l_{\widetilde{X}}(h)$ für alle $h \in H$.

Beweis. Trivialerweise gilt $\langle X_R^{(H)}\rangle_{\mathfrak{G}} \subseteq H$. Zum Beweis der umgekehrten Inklusion sei $h \in H$. Dann gilt $\overline{h} = 1_G$, da 1_G nach Voraussetzung der in R liegende Repräsentant von H ist. Wegen $\langle X\rangle_{\mathfrak{G}} = G$ gibt es $y_1, \ldots, y_n \in \widetilde{X}$ mit

$$\begin{aligned} h = y_1 \cdots y_n &= \Big(\prod_{j=1}^{n} \overline{y_1 \cdots y_{j-1}}\, y_j\, \overline{y_1 \cdots y_j}^{-1}\Big) \underbrace{\overline{y_1 \cdots y_n}}_{=\overline{h}=1_G} \qquad \text{nach 3.5.1} \\ &= \prod_{j=1}^{n} \overline{y_1 \cdots y_{j-1}}\, y_j\, \overline{\overline{y_1 \cdots y_{j-1}} y_j}^{-1} \qquad \text{nach 3.5.2} \\ &= \prod_{j=1}^{n} q(\overline{y_1 \cdots y_{j-1}}, y_j) \in \langle X_R^{(H)}\rangle_{\mathfrak{G}}, \end{aligned}$$

denn im Falle $y_j \in X^{-1}$ gilt $q(\overline{y_1 \cdots y_{j-1}}, y_j) \in \big(X_R^{(H)}\big)^{-1}$ nach 3.5.3. □

3.7 Korollar *Ist G eine endlich erzeugte Gruppe, $H \leq_{\mathfrak{G}} G$ und der Index $|G : H|$ endlich, so ist auch H endlich erzeugt.*
Genauer: Bezeichnet allgemein $d(B)$ die minimale Erzeugendenzahl einer endlich erzeugten Gruppe B, so gilt:

$$d(H) \leq |G : H| \cdot d(G).$$

Beweis. Sei X ein Erzeugendensystem von G mit $|X| = d(G)$. Nach 3.6 gilt

$$d(H) \leq |X_R^{(H)}| \leq |R \times X| = |R| \cdot |X| = |G : H| \cdot d(G).$$

□

Zur Illustration seien zwei Beispiele angegeben, die zugehörigen Nachweise aber dem Leser überlassen: Sei F eine freie Gruppe vom Rang 2, $F = \langle x, y \rangle_{\mathfrak{G}}$.
a) Sei $Y := \{x^z y x^{-z} | z \in \mathbb{Z}\}$.
b) Sei $Y := \{x^z y^u x y^{-u} x^{-(z+1)} | z, u \in \mathbb{Z}, u \neq 0\}$.
In beiden Fällen ist $\langle Y \rangle_{\mathfrak{G}}$ ein Normalteiler von F, der von Y frei erzeugt, insbesondere also von *unendlichem* Rang ist. Später (siehe 3.11.1) beschreiben wir auch die jeweilige Faktorgruppe $F/\langle Y \rangle_{\mathfrak{G}}$.

Wir wenden uns nun dem komplizierteren Schritt 2) der obigen Beweisskizze zu. Dieser setzt sich erneut aus zwei Teilschritten zusammen: Zum einen wird gezeigt, daß eine Untergruppe H einer freien Gruppe F von $X_R^{(H)}$ *frei* erzeugt wird, wenn das Rechts-Repräsentantensystem R die Eigenschaft hat, *mit einem Element stets auch alle seine Linksfaktoren zu enthalten.* Zum anderen ist dann nur noch einzusehen, daß tatsächlich stets ein solches Rechts-Repräsentantensystem *existiert.* Letztere Einsicht ist unsere nächstes Ziel (3.9). Bemerkenswert ist, daß dieser Existenzbeweis ohne einschränkende Voraussetzungen über die gegebene Gruppe, die Untergruppe H und das Erzeugendensystem X auskommt. Im Kern stammt das Resultat von M. Hall (1949).

3.8 Definition Sei G eine Gruppe und X ein $\mathfrak{G}$-Erzeugendensystem von G. Eine Teilmenge R von G heißt eine X-Schreiermenge, wenn es zu jedem $g \in R$ eine gekürzte X-Darstellung $((x_1, \varepsilon_1), \ldots, (x_k, \varepsilon_k))$ gibt mit $x_1^{\varepsilon_1} \cdots x_j^{\varepsilon_j} \in R$ für alle $j \in \underline{k} \cup \{0\}$. Trivial sind die Feststellungen:

3.8.1 *Ist $M \subseteq G$ und zu $m \in M$ stets z_m eine gekürzte X-Darstellung von m, so ist $\{x_1^{\varepsilon_1} \cdots x_j^{\varepsilon_j} \mid$ Es gibt ein $m \in M$, für das $((x_1, \varepsilon_1), \ldots, (x_j, \varepsilon_j))$ ein Linksfaktor von z_m in $\mathcal{T}(X, \{1, -1\})$ ist$\}$ eine X-Schreiermenge, die M enthält.* □

3.8.2 *Vereinigungen von X-Schreiermengen sind X-Schreiermengen.* □

Sei $H \leq_{\mathfrak{G}} G$. Ein X-Schreiersystem für H ist ein Rechts-Repräsentantensystem für H in G, das eine X-Schreiermenge ist. Insbesondere muß jedes X-Schreiersystem das Element 1_G enthalten. Wir betrachten ein paar sehr einfache konkrete Fälle:

Sei F eine freie Gruppe vom Rang 2, $F = \langle x, y \rangle_{\mathfrak{G}}$, V eine nicht-zyklische

Gruppe der Ordnung 4. z.B. $V = \{(1,1),(1,-1),(-1,1),(-1,-1)\}$ (mit komponentenweiser Multiplikation als Verknüpfung). Sei dann ψ der (eindeutig bestimmte) Epimorphismus von F auf V mit $x\psi = (1,-1)$, $y\psi = (-1,1)$ und $H := \text{Kern}\psi$. Wegen $F/H \cong V$ hat insbesondere jedes Rechts-Repräsentantensystem für H in F genau 4 Elemente. Daß die nachstehend angegebenen Teilmengen von F solche sind, erhält man unschwer daraus, daß ihre Bilder unter ψ jeweils gleich V sind. Es gilt, wie unmittelbar aus der Definition ablesbar:

a) $\{1_F, x, y, xy\}$ ist ein X-Schreiersystem für H, (vgl. auch 3.11.1)
b) $\{1_F, x, y, xy^{-1}\}$ ist ein X-Schreiersystem für H,
c) $\{1_F, x, y, x^{-1}y\}$ ist kein X-Schreiersystem für H,
d) $\{1_F, x, xy, xyx\}$ ist ein X-Schreiersystem für H.

Wir verallgemeinern nun in naheliegender Weise den Begriff der $\widetilde{X}$-Länge eines Elements auf beliebige nichtleere Teilmengen von $G(= \langle X\rangle_{\mathfrak{G}})$. Für jede Teilmenge $T \neq \emptyset$ von G sei

$$l_{\widetilde{X}}(T) := \min\{l_{\widetilde{X}}(g) | g \in T\}.$$

Offenbar gilt:

3.8.3 *Ist $g \in T$ und $((x_1,\varepsilon_1),\ldots,(x_k,\varepsilon_k))$ eine X-Darstellung von g, so gilt $l_{\widetilde{X}}(T) \leq l_{\widetilde{X}}(g) \leq k$.* □

3.9 Lemma *Sei G eine Gruppe, X ein $\mathfrak{G}$-Erzeugendensystem von G und $H \leq_{\mathfrak{G}} G$. Dann gibt es ein X-Schreiersystem R für H mit der Eigenschaft:*

$$\forall g \in R \quad l_{\widetilde{X}}(g) = l_{\widetilde{X}}(Hg)$$

Beweis (unter Verwendung des Auswahlaxioms): Wir definieren induktiv eine Kette von Schreiermengen

$$\{1_G\} = R_0 \subseteq R_1 \subseteq R_2 \subseteq \ldots,$$

so daß für alle $n \in \mathbb{N}_0$ gilt:

(∗) R_n enthält aus jeder Rechtsrestklasse T von H in G mit $l_{\widetilde{X}}(T) \leq n$ genau ein Element g_T, und für dieses gilt: $l_{\widetilde{X}}(g_T) = l_{\widetilde{X}}(Hg_T)$.

Unter Beachtung von 3.8.2 genügt es dann, $R := \bigcup_{n\in\mathbb{N}_0} R_n$ zu setzen, um die Behauptung des Lemmas zu erhalten.

Für $n = 0$ ist (∗) erfüllt mit $R_0 := \{1_G\}$ (also: $g_H = 1_G$). Sei nun $n \in \mathbb{N}_0$ und

R_n eine X-Schreiermenge mit $(*)$. Wir zeigen:

$(**)$ Zu jeder Rechtsrestklasse S von H in G mit $l_{\widetilde{X}}(S) = n+1$ gibt es ein $g_S \in S$ mit $l_{\widetilde{X}}(g_S) = n+1$, so daß $R_n \cup \{g_S\}$ eine X-Schreiermenge ist.

Beweis: Sei $s \in S$ mit $l_{\widetilde{X}}(s) = n+1$, und seien $y_1, \ldots, y_{n+1} \in \widetilde{X}$ mit $s = y_1 \cdots y_{n+1}$. Sei $t := y_1 \cdots y_n$ und $T := Ht$. Dann folgt: $l_{\widetilde{X}}(T) \leq l_{\widetilde{X}}(t) \leq n$ (3.8.3). Für das nach Induktions-Voraussetzung gemäß $(*)$ gegebene Element g_T gilt dann: $Hg_T = T = Ht$. Es folgt: $Hg_T y_{n+1} = Hty_{n+1} = Hs = S$, also

$$n+1 \leq l_{\widetilde{X}}(S) \leq l_{\widetilde{X}}(g_T y_{n+1}) \leq l_{\widetilde{X}}(g_T)+1 = l_{\widetilde{X}}(Hg_T)+1 \leq l_{\widetilde{X}}(t)+1 = n+1.$$

Setzen wir nun $g_S := g_T y_{n+1}$, so folgt: $g_S \in S$, $l_{\widetilde{X}}(g_S) = n+1$, und $R_n \cup \{g_S\}$ ist eine X-Schreiermenge. Es gilt also $(**)$.

Wir setzen (unter Verwendung des Auswahlaxioms)

$$R_{n+1} := R_n \cup \{g_S \mid S \text{ Rechtsrestklasse von } H \text{ in } G,\ l_{\widetilde{X}}(S) = n+1\}.$$

Nach 3.8.2 ist R_{n+1} eine X-Schreiermenge. Es gilt $(*)$ mit $n+1$ an Stelle von n, womit die induktive Definition der Kette abgeschlossen ist. □

Wie d) im Beispiel vor 3.8.3 zeigt, kann es durchaus X-Schreiersysteme geben, die die Längenbedingung in 3.9 nicht erfüllen.

Wir betrachten nun Schreiersysteme in freien Gruppen: Die folgende Aussage stellt den Kern des Beweises des Satzes von Nielsen und Schreier dar, denn seine beiden letzten Teile beschreiben die Multiplikation zwischen zwei Elementen von $\widetilde{X_R^{(H)}}$, wenn R ein X-Schreiersystem ist.

3.10 Lemma *Seien X eine Menge, F die freie Gruppe über X, $H \leq_{\mathfrak{G}} F$, R ein X-Schreiersystem für H, r, $s \in R$, y, $z \in \widetilde{X}$ mit $q(r,y)$, $q(s,z) \neq 1_F$.*

(1) $ry = r|y$, $y\,\overline{ry}^{-1} = y|\overline{ry}^{-1}$, $q(r,y) = r|y|\overline{ry}^{-1}$,

(2) $q(r,y) = q(s,z) \Leftrightarrow r = s,\ y = z$; $\quad q(r,y) = q(s,z)^{-1} \Leftrightarrow \overline{ry} = s,\ y = z^{-1}$,

(3) *Gilt $q(r,y) \neq q(s,z)^{-1}$, so folgt:* $\quad q(r,y)q(s,z) = r|y|\overline{ry}^{-1}s|z|\overline{sz}^{-1}$,
$q(r,y)^{-1}q(s,z)^{-1} = \overline{ry}|y^{-1}|r^{-1}\overline{sz}|z^{-1}|s^{-1}$,

(4) *Gilt $q(r,y) \neq q(s,z)$, so folgt:* $\quad q(r,y)q(s,z)^{-1} = r|y|\overline{ry}^{-1}\overline{sz}|z^{-1}|s^{-1}$,
$q(r,y)^{-1}q(s,z) = \overline{ry}|y^{-1}|r^{-1}s|z|\overline{sz}^{-1}$.

Beweis. (1) Gälte $y^{-1} \restriction r$, so $ry \upharpoonright r \in R$ und folglich $ry \in R$, da R X-Schreiermenge ist, nach 3.5.5 ein Widerspruch. Damit gilt die erste Behauptung, und durch Anwendung derselben mit $\overline{ry}$ statt r, y^{-1} statt y, auch die zweite, denn nach 3.5.5 gilt $q(\overline{ry}, y^{-1}) \neq 1_G$. Da $y \neq 1_F$, folgt $q(r,y) = r \shortmid y \shortmid \overline{ry}^{-1}$ mit 3.3.1.

(2) 1. Teil: Zum Beweis der nichttrivialen Implikation („$\Rightarrow$“) machen wir zunächst die Annahme, es gälte $l_{\widetilde{X}}(r) < l_{\widetilde{X}}(s)$. Mit unserer Voraussetzung und (1) erhalten wir die Gleichung $r \shortmid y \shortmid \overline{ry}^{-1} = s \shortmid z \shortmid \overline{sz}^{-1}$, somit: $ry \upharpoonright s$. Da R X-Schreiermenge ist und $s \in R$ gilt, folgt $ry \in R$, d.h. $ry = \overline{ry}$ und damit der Widerspruch $q(r,y) = ry\overline{ry}^{-1} = 1_F$. Also gilt $l_{\widetilde{X}}(r) \geq l_{\widetilde{X}}(s)$, ebenso $l_{\widetilde{X}}(s) \geq l_{\widetilde{X}}(r)$ und damit $l_{\widetilde{X}}(r) = l_{\widetilde{X}}(s)$. Aus $r \shortmid y \shortmid \overline{ry}^{-1} = s \shortmid z \shortmid \overline{sz}^{-1}$ folgt nun $r = s$, $y = z$. – Vermöge 3.5.4 erhält man aus dem 1. Teil unmittelbar auch den 2. Teil.

Wir zeigen nun zunächst folgende Hilfsaussage::

$$(*) \qquad \forall t \in R \quad z^{-1}s^{-1}t = z^{-1} \shortmid s^{-1}t$$

Gälte nämlich $z \upharpoonright s^{-1}t$, so gäbe es ein $a \in F$ mit $z \shortmid a = s^{-1}t$, und nach (1) sowie 3.3.1 folgte $s \shortmid z \shortmid a = t \in R$. Da R X-Schreiermenge ist, folgte $sz \in R$, mit 3.5.5 ein Widerspruch. Als Spezialfall von $(*)$ erhalten wir zunächst $z^{-1}s^{-1}\overline{ry} = z^{-1} \shortmid s^{-1}\overline{ry}$, somit $\overline{ry}^{-1}sz = \overline{ry}^{-1}s \shortmid z$. Als weiterer Spezialfall von $(*)$ ergibt sich $y\overline{ry}^{-1}s = y \shortmid \overline{ry}^{-1}s$.

Die Voraussetzung von (3) besagt nach (2), daß im Falle $\overline{ry}^{-1}s = 1_F$ gilt: $yz \neq 1_F$. Nun folgt die erste Behauptung in (3) mit 3.3.1. Die zweite Behauptung in (3) ergibt sich aus der ersten durch Invertieren beider Seiten der Gleichung und Rollentausch von (r,y) und (s,z).

(4) Nach Voraussetzung und 3.5.4 gilt: $q(\overline{ry}, y^{-1})q(s,z) = q(r,y)^{-1}q(s,z) \neq 1_F \neq q(r,y)q(s,z)^{-1} = q(r,y)q(\overline{sz}, z^{-1})$. Wendet man (3) auf das erste sowie auf das letzte in dieser Kette genannte Produkt an, so erhält man beide Behauptungen aus (4). □

Wählt man unter den Voraussetzungen von 3.10 $y, z \in X$, so zeigt (2), daß die Voraussetzung von (3) erfüllt ist. Als Konsequenzen der letzten beide Teile von 3.10 erhalten wir daher die folgenden Aussagen, die uns einen kurzen Beweis von 3.11 ermöglichen werden:

$$y\overline{ry}^{-1}q(s,z) = y\overline{ry}^{-1}s \shortmid z \shortmid \overline{sz}^{-1}, \quad y^{-1}r^{-1}q(s,z)^{-1} = y^{-1} \shortmid r^{-1}\overline{sz} \shortmid z^{-1} \shortmid s^{-1},$$

und falls $q(r,y) \neq q(s,z)$:

$$y\overline{ry}^{-1}q(s,z)^{-1} = y \shortmid \overline{ry}^{-1}\overline{sz} \shortmid z^{-1} \shortmid s^{-1}, \quad y^{-1}r^{-1}q(s,z) = y^{-1} \shortmid r^{-1}s \shortmid z \shortmid \overline{sz}^{-1}.$$

3.11 Satz (Nielsen, Schreier 1927) *Sei F eine freie Gruppe mit $\mathfrak{G}$-Basis X, $H \leq_{\mathfrak{G}} F$, R ein X-Schreiersystem für H in F und $X_R^{(H)}$ wie in 3.6. Dann ist $X_R^{(H)}$ eine $\mathfrak{G}$-Basis von H. Insbesondere ist jede Untergruppe einer freien Gruppe frei.*

Beweis. Nach 3.6 genügt es zu zeigen, daß $X_R^{(H)}$ $\mathfrak{G}$-unabhängig ist. Da die Existenz eines X-Schreiersystems für H in F nach 3.9 gesichert ist, folgt dann auch die Schlußbehauptung. Für alle $k \in \mathbb{N}$ zeigen wir die folgende

Behauptung: Sind $x_1, \ldots, x_k \in X$, $r_1, \ldots, r_k \in R$, $\varepsilon_1, \ldots, \varepsilon_k \in \{1, -1\}$ mit $q(r_i, x_i) \neq 1_F$ für alle $i \in \underline{k}$ und gilt

$$q(r_i, x_i) = q(r_{i+1}, x_{i+1}) \Rightarrow \varepsilon_i = \varepsilon_{i+1}$$

für alle $i \in \underline{k-1}$, so folgt:

$$(*) \qquad q(r_1, x_1)^{\varepsilon_1} \cdots q(r_k, x_k)^{\varepsilon_k} = u_1 | x_1^{\varepsilon_1} | u_2 | x_2^{\varepsilon_2} | \cdots | u_k | x_k^{\varepsilon_k} | \begin{cases} \overline{r_k x_k}^{-1} & \text{f. } \varepsilon_k = 1 \\ r_k^{-1} & \text{f. } \varepsilon_k = -1 \end{cases}$$

für geeignete $u_1, \ldots, u_k \in R^{-1}R$.

Haben wir dies nämlich bewiesen, so folgt insbesondere:

$$l_{\widetilde{X}}\big(q(r_1, x_1)^{\varepsilon_1} \cdots q(r_k, x_k)^{\varepsilon_k}\big) \geq k > 0,$$

also $q(r_1, x_1)^{\varepsilon_1} \cdots q(r_r, x_k)^{\varepsilon_k} \neq 1_F$, und damit nach 2.4(2) die $\mathfrak{G}$-Unabhängigkeit von $X_R^{(H)}$.

Den Beweis der Behauptung führen wir durch Induktion nach k. Der Induktionsanfang ist bereits durch 3.10(1) erledigt, denn es gilt $1_F \in R$. Für den Induktionsschritt seien $x_1, \ldots, x_{k+1} \in X$, $r_1, \ldots, r_{k+1} \in R$, $\varepsilon_1, \ldots, \varepsilon_{k+1} \in \{1, -1\}$ mit $q(r_i, x_i) \neq 1_F$ für alle $i \in \underline{k+1}$, und es gelte:

$$\forall i \in \underline{k} \quad q(r_i, x_i) = q(r_{i+1}, x_{i+1}) \Rightarrow \varepsilon_i = \varepsilon_{i+1}.$$

Es gelte $(*)$. Setzen wir nun $a := \prod_{i=1}^{k-1} u_i {x_i}^{\varepsilon_i}$, so ergibt sich:

$$\prod_{i=1}^{k+1} q(r_i, x_i)^{\varepsilon_i} = \begin{cases} a | u_k | x_k | \overline{r_k x_k}^{-1} q(r_{k+1}, x_{k+1})^{\varepsilon_{k+1}} & \text{falls } \varepsilon_k = 1 \\ a | u_k | x_k^{-1} | r_k^{-1} q(r_{k+1}, x_{k+1})^{\varepsilon_{k+1}} & \text{falls } \varepsilon_k = -1 \end{cases}$$

$$= \begin{cases} a | u_k | x_k | \overline{r_k x_k}^{-1} r_{k+1} | x_{k+1} | \overline{r_{k+1} x_{k+1}}^{-1} & \text{falls } \varepsilon_k = 1 = \varepsilon_{k+1}, \\ a | u_k | x_k | \overline{r_k x_k}^{-1} \overline{r_{k+1} x_{k+1}} | x_{k+1}^{-1} | r_{k+1}^{-1} & \text{falls } \varepsilon_k = 1 \neq \varepsilon_{k+1}, \\ a | u_k | x_k^{-1} | r_k^{-1} \overline{r_{k+1} x_{k+1}} | x_{k+1}^{-1} | r_{k+1}^{-1} & \text{falls } \varepsilon_k = -1 = \varepsilon_{k+1}, \\ a | u_k | x_k^{-1} | r_k^{-1} r_{k+1} | x_{k+1} | \overline{r_{k+1} x_{k+1}}^{-1} & \text{falls } \varepsilon_k = -1 \neq \varepsilon_{k+1}, \end{cases}$$

unter Anwendung der aus 3.10 erschlossenen Gleichungen, denn im Falle $\varepsilon_k \neq \varepsilon_{k+1}$ gilt $q(r_k, x_k) \neq q(r_{k+1}, x_{k+1})$. Die Behauptung folgt. $\square$

3.11.1 Beispiele Sei $X = \{x, y\}$ mit $x \neq y$ und F die freie Gruppe über X. Wir beschreiben einige $\mathfrak{G}$-Basen gewisser Normalteiler von F:

(1) Sei $\varphi : X \to \mathbb{Z}$, $x \mapsto 1$, $y \mapsto 0$ und $\overline{\varphi} : F \to \mathbb{Z}$ die Fortsetzung von φ zu einem $\mathfrak{G}$-Homomorphismus, $H := \operatorname{Kern} \overline{\varphi}$. Dann gilt: $x^z\overline{\varphi} = z$ für alle $z \in \mathbb{Z}$, also $F = \dot{\bigcup}_{z\in\mathbb{Z}} Hx^z$. Sei $R := \{x^z \mid z \in \mathbb{Z}\}$. Dann ist R offensichtlich gegen Linksfaktoren abgeschlossen, also ein X-Schreiersystem für H in F, und es gilt:

$$q(x^z, x) = x^z x\overline{x^z x}^{-1} = x^{z+1}(x^{z+1})^{-1} = 1_F$$
$$q(x^z, y) = x^z y\overline{x^z y}^{-1} = x^z y x^{-z}, \text{ da } H(x^z y)^{-1} = Hx^{-z}.$$

Also ist $X_R^{(H)} = \{x^z y x^{-z} | z \in \mathbb{Z}\}$ nach 3.11 eine $\mathfrak{G}$-Basis von H. Es folgt: $F/\langle X_R^{(H)}\rangle_{\mathfrak{G}} = F/H \cong \mathbb{Z}$.

(2) Sei $\varphi : X \to \{1, -1\}$, $x \mapsto -1$, $y \mapsto 1$ und $\overline{\varphi} : F \to \{1, -1\}$ die Fortsetzung von φ zu einem $\mathfrak{G}$-Homomorphismus, $H := \operatorname{Kern} \overline{\varphi}$. Dann ist F/H von der Ordnung 2, $F = H\dot{\cup}Hx$ und $R := \{1, x\}$ ein X-Schreiersystem für H in F. Es gilt:

$$q(1_F, x) = x\overline{x}^{-1} = 1_F, \quad q(x, x) = xx\overline{xx}^{-1} = x^2,$$
$$q(1_F, y) = y\overline{y}^{-1} = y, \qquad q(x, y) = xy\overline{xy}^{-1} = xyx^{-1}$$

Also ist $X_R^{(H)} = \{x^2, y, xyx^{-1}\}$ nach 3.11 eine $\mathfrak{G}$-Basis von H.

(3) Sei $\varphi : X \to \{(1, 1), (1, -1), (-1, 1), (-1, -1)\}$, $x \mapsto (1, -1)$, $y \mapsto (-1, 1)$ und $\overline{\varphi}$ die Fortsetzung von φ zu einem $\mathfrak{G}$-Epimorphismus von F, $H := \operatorname{Kern} \overline{\varphi}$. Dann ist F/H nicht-zyklisch von der Ordnung 4 und $R := \{1_F, x, y, xy\}$ ein X-Schreiersystem für H in F. Es gilt:

$$q(1_F, x) = q(1_F, y) = q(x, y) = 1_F, \quad q(x, x) = x^2, \quad q(y, y) = y^2,$$
$$q(y, x) = yx\overline{yx}^{-1} = yxy^{-1}x^{-1}, \qquad q(xy, x) = xyx\overline{xyx}^{-1} = xyxy^{-1},$$
$$q(xy, y) = xy^2\overline{xy^2}^{-1} = xy^2x^{-1}.$$

Also ist $X_R^{(H)} = \{x^2, y^2, yxy^{-1}x^{-1}, xyxy^{-1}, xy^2x^{-1}\}$ nach 3.11 eine $\mathfrak{G}$-Basis von H.

(4) Sei $H = F'$. Nach 2.19(2) ist F/F' eine freie abelsche Gruppe vom Rang 2 und $\{F'x, F'y\}$ eine $\mathfrak{M}$-Basis von F/F'. Also ist $R := \{x^z y^u | z, u \in \mathbb{Z}\}$

ein Repräsentantensystem für F' in F. Offensichtlich ist R auch eine X-Schreiermenge. Es gilt:

$$q(x^z y^u, x) = x^z y^u x \overline{x^z y^u x}^{-1} = x^z y^u x y^{-u} x^{-(z+1)} (= 1_F \Leftrightarrow u = 0),$$
$$q(x^z y^u, y) = x^z y^u y \overline{x^z y^{u+1}}^{-1} = 1_F.$$

Also ist $X_R^{(F')} = \{x^z y^u x y^{-u} x^{-(z+1)} | z, u \in \mathbb{Z}, u \neq 0\}$ nach 3.11 eine $\mathfrak{G}$-Basis von F'. Diese besteht sogar aus Kommutatoren, denn es gilt: $x^z y^u x y^{-u} x^{-(z+1)} = [(x^z y^u)^{-1}, x^{-1}]$ für alle z, $u \in \mathbb{Z}$.

Wir ordnen in dem nachstehenden Diagramm unsere bislang betrachteten Beispiele von Untergruppen H einer freien Gruppe F vom Rang > 1 hinsichtlich des Aspekts von Endlichkeit/Unendlichkeit des Ranges bzw. des Index. Es sei $x \in F \setminus \{1_F\}$; die Zahlen beziehen sich auf die Numerierung in 3.11.1:

	$\lvert F : H \rvert$ endlich	$\lvert F : H \rvert$ unendlich
$rk(H)$ endlich	(2), (3)	$\langle x \rangle_{\mathfrak{G}}$
$rk(H)$ unendlich	(siehe Text)	(1), (4)

Daß der Index einer Untergruppe von unendlichem Rang ebenfalls unendlich sein muß, wissen wir – jedenfalls im Falle einer *endlich erzeugten* Gruppe F – schon aus 3.7; es gibt also kein Beispiel für das Feld links unten im Diagramm, wenn F endlich erzeugt ist. Hat F dagegen unendlichen Rang und ist H eine beliebige Untergruppe von endlichem Index, so ist H *stets* von unendlichem Rang, da ja die Vereinigung eines $\mathfrak{G}$-Erzeugendensystems von H mit einem Rechts-Repräsentantensystem für H in F ein $\mathfrak{G}$-Erzeugendensystem von F ist, also nicht endlich sein kann. Bei den Beispielen in 3.11.1 ist H sogar ein Normalteiler von F, bei dem für das Feld rechts oben im Diagramm angegebenen Beispiel jedoch nicht. Aus einer noch zu beweisenden allgemeineren Bemerkung (siehe 3.12.2) folgt, daß es hier kein Beispiel geben kann, bei dem H Normalteiler von F wäre.

3.12 Proposition *Sei F eine freie Gruppe mit $\mathfrak{G}$-Basis X, $H \leq_{\mathfrak{G}} F$, R ein X-Schreiersystem für H, $P := \{(r, x) | r \in R, x \in X, rx \in R\}$. Dann gilt:* $|P| = |R \setminus \{1_F\}|$

Folgerung *Ist $|F : H|$ endlich und $rk(F) = n \in \mathbb{N}$, so folgt:*

$$rk(H) = |F : H|(n-1) + 1.$$

Beweis. Sei f wie in 3.3.2. Wir zeigen: $(R \setminus \{1_F\})f = P$. Da f nach 3.3.2 injektiv ist, beweist das die Proposition. Da R eine X-Schreiermenge ist, gilt

jedenfalls $(R \setminus \{1_F\})f \subseteq P$. Sei nun $(r,x) \in P$, also $r \in R$, $x \in X$, $rx \in R$. Gilt $rx = r \cdot x$, so $rx \neq 1_F$ und $(rx)f = (r,x)$. Andernfalls gilt $r = r' \cdot x^{-1}$ und $rf = (r,x)$.

Beweis der Folgerung: Nach 3.11 ist $X_R^{(H)}$ eine $\mathfrak{G}$-Basis von H. Nach 3.5.5 und 3.10(2) ist die Zuordnung $(r,x) \mapsto q(r,x)$ eine Bijektion von $(R \times X) \setminus P$ auf $X_R^{(H)}$. Es folgt:

$rk(H) = |X_R^{(H)}| = |R \times X| - |P| = |R| \cdot |X| - (|R| - 1) = |F : H|(n-1) + 1$. $\square$

In der Linearen Algebra spielt der sogenannte „Basis-Ergänzungssatz“ eine wichtige Rolle, der besagt, daß eine Teilraumbasis stets zu einer Basis des ganzen Raumes ergänzt werden kann. Das Analogon für freie Gruppen kann nicht erwartet werden; ganz im Gegenteil folgt aus 3.12 bereits

3.12.1 *Ist H eine echte Untergruppe einer freien Gruppe F endlichen Ranges > 1 und von endlichem Index, so gilt $rk(H) > rk(F)$,*

denn $|F : H|(n-1) + 1 \geq 2(n-1) + 1 = 2n - 1 > n$ für jedes $n \in \mathbb{N}_{>1}$. $\square$

Insbesondere liegt unter den Voraussetzungen von 3.12.1 keine $\mathfrak{G}$-Basis von H in einer $\mathfrak{G}$-Basis von F. In dieselbe Richtung zielt die folgende einfache Aussage:

3.12.2 *Ist H eine echte Untergruppe einer freien Gruppe F, die einen von $\{1_F\}$ verschiedenen Normalteiler von F enthält, so gibt es keine $\mathfrak{G}$-Basen X von F und Y von H mit $Y \subseteq X$.*

Ist nämlich $N \trianglelefteq F$ mit $N \leq H$, $x \in X \setminus Y$ und $1_F \neq z \in N$, so gibt es eine gekürzte Y-Darstellung von z, und diese ist zugleich eine, mithin *die* gekürzte X-Darstellung von z (2.4(2)). H kann nicht x enthalten, enthält aber das Element $x^{-1}zx = x^{-1} \cdot z \cdot x$, das daher eine gekürzte Y-Darstellung besitzen muß. Wieder ist diese zugleich die gekürzte X-Darstellung, so daß letztere nicht $(x,1)$ als letzte Komponente haben kann, ein Widerspruch. $\square$

Andererseits gilt der folgende wichtige „schwache Basis-Ergänzungssatz“, der auf M. Hall zurückgeht:

3.13 Satz *Sei F eine freie Gruppe, X eine $\mathfrak{G}$-Basis, H eine Untergruppe endlichen Ranges von F und R ein X-Schreiersystem für H. Dann gibt es eine $\mathfrak{G}$-unabhängige Teilmenge Z von F, so daß gilt:*

(i) *$|F : \langle Z \rangle_{\mathfrak{G}}|$ ist endlich,*

(ii) *$X_R^{(H)} \subseteq Z$.*

Beweis. Sei $R' := \{s | s \in R, \exists x \in \widetilde{X} \quad q(s,x) \neq 1_F\}$ und $L(R')$ die Menge aller Linksfaktoren der Elemente von R' (siehe 3.8.1). Dann gilt: $L(R') \subseteq R$, weil R eine Schreiermenge ist. Da H endlichen Rang hat, ist nach 3.11 die $\mathfrak{G}$-Basis $X_R^{(H)}$ von H endlich. Aus 3.10(2) folgt nun die Endlichkeit von R'. Also ist $L(R')$ eine in R enthaltene endliche Schreiermenge.

Wir werden zeigen, daß es eine Untergruppe $\widehat{H}$ von F mit den folgenden beiden Eigenschaften gibt:

(i') $L(R')$ ist ein Rechts-Repräsentantensystem für $\widehat{H}$ in F,

(ii') $X_R^{(H)} \subseteq X_{L(R')}^{(\widehat{H})}$.

Nach 3.11 ist, wenn (i') gilt, $X_{L(R')}^{(\widehat{H})}$ eine $\mathfrak{G}$-Basis von $\widehat{H}$. Mit (ii') genügt es also, $Z := X_{L(R')}^{(\widehat{H})}$ zu setzen, um die Behauptung des Satzes zu erhalten.

Die Konstruktion einer solchen Untergruppe $\widehat{H}$ nimmt einen Gedanken auf, der schon im Beweis von 2.4(2) eine entscheidende Rolle gespielt hat: Basierend auf Fortsetzungen gewisser injektiver Abbildungen zu Permutationen nämlich erhält man einen Homomorphismus von F in die (endliche) symmetrische Gruppe $\mathcal{S}_{L(R')}$. Wir werden einsehen, daß $\widehat{H}$ als der Stabilisator von 1_F bezüglich einer solchen Permutationsdarstellung gewählt werden kann.

Für jedes $x \in X$ sei $T(x) := \{\, t \,|\, t \in L(R'), \overline{tx} \in L(R')\}$. Dann ist die Abbildung

$$T(x) \to L(R'),\ t \mapsto \overline{tx}$$

injektiv: Sind $t, t^* \in T(x)$ mit $\overline{tx} = \overline{t^*x}$, so folgt $Htx = Ht^*x$, also $Ht = Ht^*$ und damit $t = t^*$, denn $t, t^* \in L(R') \subseteq R$. Daher gibt es eine Permutation π_x von $L(R')$ mit $t\pi_x = \overline{tx}$ für alle $t \in T(x)$. Sei [8]

$$\varphi : X \to \mathcal{S}_{L(R')},\ x \mapsto \pi_x.$$

Da F frei über X ist, gibt es (genau) einen Homomorphismus ψ von F in $S_{L(R')}$ mit $\psi|_X = \varphi$. Der Beweis gliedert sich nun in drei Schritte:

1. Ist $s \in L(R')$ und $t \in F$, $y \in \widetilde{X}$ mit $s = t\iota y$, so gilt: $t(y\psi) = s$:

Aus den Voraussetzungen folgt zunächst: $t \in L(R')$, da $L(R')$ eine Schreiermenge ist. Seien $x \in X$, $\varepsilon \in \{1, -1\}$ mit $y = x^\varepsilon$.

[8] Diese Setzung benötigt nicht etwa das Auswahlaxiom: Man braucht nur z.B. die endliche Menge $L(R')$ vollständig zu ordnen und kann dann zu jedem $x \in X$ die Fortsetzung π_x *kanonisch* wählen, z.B. durch Zuordnen der noch nicht zugeordneten Elemente in der monoton wachsenden Reihenfolge.

1. Fall: $\varepsilon = 1$. Dann gilt: $t(y\psi) = t(x\psi) = t\pi_x = \overline{tx} = s$, denn $tx = s \in R$.
2. Fall: $\varepsilon = -1$. Dann gilt: $\overline{sx} = \overline{t} = t$, also $s \in T(x)$ und $s(y\psi)^{-1} = s(x\psi) = s\pi_x = \overline{sx} = t$. Es folgt: $t(y\psi) = s$.

2. Für alle $u \in L(R')$ gilt: $1_F(u\psi) = u$:

Ist nämlich $u \in L(R')$ und $((x_1, \varepsilon_1), \ldots, (x_n, \varepsilon_n))$ die gekürzte X-Darstellung von u, so gilt $x_1^{\varepsilon_1} \cdots x_k^{\varepsilon_k} \in L(R')$ für alle $k \in \underline{n} \cup \{0\}$. Es folgt: $1_F(u\psi) = 1_F((x_1^{\varepsilon_1} \cdots x_n^{\varepsilon_n})\psi) = 1_F(x_1^{\varepsilon_1}\psi) \cdots (x_n^{\varepsilon_n}\psi) = x_1^{\varepsilon_1} \cdots x_n^{\varepsilon_n} = u$ vermöge iterierter Anwendung von Schritt 1.

Sei nun $\widehat{H} := \{\, g \,|\, g \in F,\ 1_F(g\psi) = 1_F\}$, der Stabilisator von 1_F in F bezüglich der Operation ψ. Dann ist $\widehat{H}$ eine Untergruppe von F, bezüglich der wir nun die Aussage (i') beweisen können: Ist nämlich $g \in F$ und $u := 1_F(g\psi)$, so folgt $u \in L(R')$, nach Schritt 2 also $1_F(u\psi) = 1_F(g\psi)$ und damit $gu^{-1} \in \widehat{H}$, d.h. $\widehat{H}g = \widehat{H}u$. Sind u, $u' \in L(R')$ mit $\widehat{H}u = \widehat{H}u'$, so gibt es ein $g \in \widehat{H}$ mit $u' = gu$. Mit Schritt 2 folgt : $u' = 1_F(u'\psi) = (1_F(g\psi))(u\psi) = 1_F(u\psi) = u$.

Sei nun $\overline{\overline{\ }}$ die Repräsentantenzuordnung bezüglich des Rechts-Repräsentantensystems $L(R')$ für $\widehat{H}$ in F.

3. Für alle $f \in F$ gilt: $1_F(f\psi) = \overline{\overline{f}}$:

Sei dazu $g \in \widehat{H}$ mit $f = g\overline{\overline{f}}$. Es gilt: $1_F(f\psi) = (1_F(g\psi))(\overline{\overline{f}}\psi) = 1_F(\overline{\overline{f}}\psi) = \overline{\overline{f}}$, nach Schritt 2.

Jetzt können wir auch die Aussage (ii') beweisen: Seien $r \in R$, $x \in X$ mit $q(r, x) = rx\overline{rx}^{-1} \neq 1_F$. Es gilt $r \in R' \subseteq L(R')$ und wegen $q(\overline{rx}, x^{-1}) = q(r, x)^{-1} \neq 1_F$ (siehe 3.5.4) auch $\overline{rx} \in R' \subseteq L(R')$, also $r \in T(x)$. Die Schritte 2 und 3 ergeben nun:

$$\overline{\overline{rx}} = 1_F((rx)\psi) = (1_F(r\psi))(x\psi) = r(x\psi) = \overline{rx}$$

und damit $q(r, x) = rx\overline{\overline{rx}}^{-1} \in X^{(\widehat{H})}_{L(R')}$. □

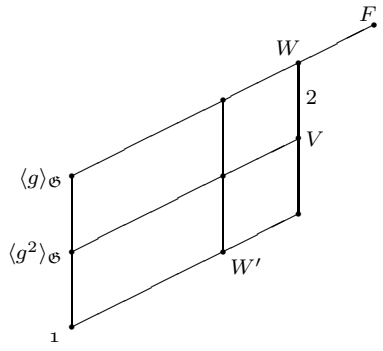
3.14 Korollar (Greenberg 1960) *Sei F eine freie Gruppe und H eine Untergruppe endlichen Ranges von F. Enthält H einen Normalteiler $\neq \{1_F\}$ von F, so ist $|F : H|$ endlich. Insbesondere ist auch $rk(F)$ endlich.*

Beweis. Nach 3.13 (und 3.9) gibt es eine Untergruppe $\widehat{H}$ von F von endlichem Index mit einer $\mathfrak{G}$-Basis, die eine $\mathfrak{G}$-Basis von H enthält. Aus 3.12.2 folgt: $H = \widehat{H}$. Insbesondere ist F endlich erzeugt, also nach 2.19.1 von endlichem Rang. □

Eine unmittelbare Folgerung ist zum Beispiel, daß eine freie Gruppe F unendlichen Ranges keinen Normalteiler $\neq \{1_F\}$ endlichen Ranges besitzen kann. Aus 3.14 und 3.12 ergibt sich ferner unmittelbar

3.14.1 *Sei F eine freie Gruppe von endlichem Rang. Dann hat ein Normalteiler $\neq \{1_F\}$ von F genau dann endlichen Rang, wenn sein Index in F endlich ist.* □

3.15 Korollar (Levi 1930) *Für jede freie Gruppe F gilt:* $\bigcap_{\substack{H \leq F \\ |F:H| \text{ endlich}}} H = \{1_F\}$



Beweis. Sei $g \in F \setminus \{1_F\}$. Wir zeigen: Es gibt eine Untergruppe von F, deren Index endlich ist und die g nicht enthält.

Die einzigen $\mathfrak{G}$-Basen der Untergruppe $\langle g \rangle_{\mathfrak{G}}$ von F sind $\{g\}$ und $\{g^{-1}\}$. Daher gibt es nach 3.13 eine $\mathfrak{G}$-unabhängige Teilmenge Z von F mit $g \in Z$ oder $g^{-1} \in Z$, so daß $W := \langle Z \rangle_{\mathfrak{G}}$ endlichen Index in F hat. Nach 2.19(2) ist W/W' eine von $\{W'z | z \in Z\}$ frei erzeugte abelsche Gruppe. Sei V die Untergruppe von W mit

$$W' \leq V < W \text{ und } V/W' = \langle\, W'z \,|\, z \in Z \setminus \{g, g^{-1}\} \text{ oder } z = g^2 \rangle_{\mathfrak{M}}.$$

Dann gilt: $g \notin V$, $|F : V| = |F : W||W : V| = 2|F : W|$. □

3.16 Lemma *Sei G eine Gruppe, M eine Menge und $\mathfrak{Y} := \{\, H \,|\, H \leq_{\mathfrak{G}} G, |G : H| = |M|\}$. Dann gilt:*

$$|\mathfrak{Y}| \leq |\mathrm{Hom}(G, S_M)|,$$

wobei $\mathrm{Hom}(G, S_M)$ *die Menge der Homomorphismen von G in S_M ist.*

Folgerung: *Ist G endlich erzeugt, $d(G)$ die minimale Erzeugendenzahl von G und $k \in \mathbb{N}$, so gibt es höchstens $k!^{d(G)}$ Untergruppen H von G mit $|G : H| = k$. Insbesondere gibt es dann zu jeder endlichen Gruppe $\widetilde{G}$ nur endlich viele Normalteiler N von G mit $G/N \cong \widetilde{G}$.*

Im Beweis spielt die folgende einfache Bemerkung eine Rolle:

3.16.1 *Ist β eine Bijektion von einer Menge N auf eine Menge M, so ist $\overline{\beta} : S_N \to S_M$, $\rho \mapsto \beta^{-1}\rho\beta$ ein Isomorphismus von S_N auf S_M.* □

Beweis von 3.16. Wir konstruieren eine injektive Abbildung von $\mathfrak{Y}$ nach $\mathrm{Hom}(G, S_M)$: Sei $m \in M$. Für alle $H \in \mathfrak{Y}$ sei $R(H)$ die Menge der Rechts-Restklassen von H in G und β_H eine Bijektion von $R(H)$ auf M mit $H\beta_H = m$. Für alle $g \in G$ sei

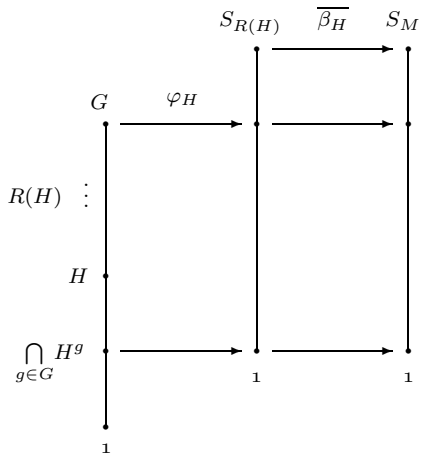
$$\rho_g : \ R(H) \to R(H), \ Hx \mapsto Hxg$$

und

$$\varphi_H : \ G \to S_{R(H)}, \ g \mapsto \rho_g.$$

Es gilt: $\varphi_H\overline{\beta_H} \in \ \mathrm{Hom}(G, S_M)$ und $Hg\beta_H = H\beta_H{\beta_H}^{-1}\rho_g\beta_H = m(g\varphi_H\overline{\beta_H})$, also

$$(*) \qquad g \in H \Leftrightarrow Hg = H \Leftrightarrow Hg\beta_H = m \Leftrightarrow m(g\varphi_H\overline{\beta_H}) = m.$$



Es genügt, die Injektivität der Abbildung

$$\alpha : \mathfrak{Y} \to \mathrm{Hom}(G, S_M), \ H \mapsto \varphi_H\overline{\beta_H}$$

nachzuweisen. Sind H, $H^* \in \mathfrak{Y}$ mit $H\alpha = H^*\alpha$, so gilt $\varphi_H\overline{\beta_H} = \varphi_{H^*}\overline{\beta_{H^*}}$, also unter Verwendung von $(*)$:

$$\begin{aligned} g \in H &\Leftrightarrow m(g\varphi_H\overline{\beta_H}) = m \\ &\Leftrightarrow m(g\varphi_{H^*}\overline{\beta_{H^*}}) = m \\ &\Leftrightarrow g \in H^* \end{aligned}$$

für alle $g \in G$.

Zum Beweis der Folgerung sei X ein $\mathfrak{G}$-Erzeugendensystem von G mit $|X| = d(G)$, $M := \underline{k}$. Es gibt höchstens so viele Homomorphismen von G in S_k wie es Abbildungen von X in S_k gibt. Damit folgt:

$$|\mathfrak{Y}| \leq |\mathrm{Hom}(G, S_k)| \leq |{S_k}^X| = k!^{d(G)}.$$

□

3.17 Satz (Mal'cev 1940) *Sei G eine endlich erzeugte Gruppe und $M \trianglelefteq G$ mit $G/M \cong G$. Dann gilt: $M \subseteq \bigcap_{\substack{H \leq G \\ |G:H| \text{ endlich}}} H$.*

Beweis. Sei $k \in \mathbb{N}$, $\mathfrak{Y}_G$ bzw. $\mathfrak{Y}_{G/M}$ die Menge der Untergruppen vom Index k von G bzw. G/M. Die Umkehrung des kanonischen Epimorphismus von G auf

G/M induziert eine injektive Abbildung α von $\mathfrak{Y}_{G/M}$ in $\mathfrak{Y}_G$. Die Isomorphie von G/M und G impliziert die Gleichmächtigkeit von $\mathfrak{Y}_{G/M}$ und $\mathfrak{Y}_G$, und nach 3.16 ist $\mathfrak{Y}_G$ endlich. Also ist α surjektiv, d.h. für alle $H \in \mathfrak{Y}_G$ gilt $M \subseteq H$. Es folgt:

$$M \subseteq \bigcap_{\substack{H \leq G \\ |G:H|=k}} H.$$

Da dies für jedes $k \in \mathbb{N}$ gilt, folgt die Behauptung. □

3.18 Korollar (Magnus 1935) *Ist F eine endlich erzeugte freie Gruppe und $\{1_F\} \neq M \trianglelefteq F$, so gilt $F/M \ncong F$.*

Denn sonst gälte nach 3.17 und 3.15 $\{1_F\} < M \subseteq \bigcap_{\substack{H \leq G \\ |G:H| \text{ endlich}}} H = \{1_F\}$, ein Widerspruch. □

Eine Gruppe mit der soeben für F nachgewiesenen Eigenschaft, zu keiner ihrer echten Faktorgruppen isomorph zu sein, heißt eine Hopf-Gruppe. Die Tatsache, daß jede endlich erzeugte freie Gruppe eine Hopf-Gruppe ist, ist schon implizit in Nielsens Resultaten aus dem Jahr 1921 enthalten, was jedoch erst fast drei Jahrzehnte später entdeckt wurde. Bereits der Begriff „Hopf-Gruppe“ ist jünger als Nielsens Arbeit, denn er entstand mit Bezug auf Arbeiten von H. Hopf aus den Jahren 1930 und 1931. Die Hopf-Eigenschaft überträgt sich nicht auf Faktorgruppen: Sind F eine freie Gruppe vom Rang 2 mit $\mathfrak{G}$-Basis $\{x, y\}$, k, l zueinander teilerfremde natürliche Zahlen und N der kleinste Normalteiler von F, der das Element $x^{-1}y^k xy^{-l}$ enthält, so ist F/N keine Hopf-Gruppe, wie von Baumslag und Solitar (1962) gezeigt wurde. Besonders bemerkenswert an diesem Beispiel ist, daß es sich um eine sog. one-relator group handelt, d.h. daß der Normalteiler N, nach dem F faktorisiert wird, von einem einzigen Element von F (als Normalteiler) erzeugt wird. Beispiele ohne diese Besonderheit wurden bereits über 10 Jahre zuvor von B. H. Neumann und von G. Higman entdeckt.

3.18.1 *Sei $G \in \mathfrak{G}$, $V \leq_{\mathfrak{G}} G$, R ein Rechts-Repräsentantensystem für V in G. Dann ist $\bigcap_{r \in R} V^r$ der größte in V enthaltene Normalteiler von G. Ist $|G:V|$ endlich, so auch $|G : \bigcap_{r \in R} V^r|$.*

Beweis. Jeder in V enthaltene Normalteiler von G liegt für jedes $g \in G$ in V^g, also insbesondere in $\bigcap_{r \in R} V^r$. Andererseits gibt es zu jedem $g \in G$ und $r \in R$ ein $s \in R$ mit $rg \in Vs$. Es folgt: $\big(\bigcap_{r \in R} V^r\big)^g = \bigcap_{r \in R} V^{rg} \subseteq \bigcap_{s \in R} V^s$. Ist $|G:V| = |R|$ endlich, so gilt $|G : \bigcap_{r \in R} V^r| \leq \prod_{r \in R} |G:V^r| = |G:V|^{|G:V|}$. □

Folgerung: *Für jede Gruppe G gilt:* $\bigcap_{\substack{H \leq G \\ |G:H| \text{ endlich}}} H = \bigcap_{\substack{N \trianglelefteq G \\ G/N \text{ endlich}}} N.$

Denn nach 3.18.1 enthält jede Untergruppe von endlichem Index einen Normalteiler von endlichem Index. □

Eine Gruppe, bei der der Durchschnitt aller Normalteiler von endlichem Index trivial ist, wird residuell endlich genannt. Nach der eben bemerkten Folgerung besagt also 3.15:

3.15' *Jede freie Gruppe ist residuell endlich.*

Wir beschließen das Kapitel mit einem alternativen Beweis des Satzes von Nielsen (2.19(4)):

Dazu sei X eine $\mathfrak{G}$-Basis einer freien Gruppe F von endlichem Rang und Y ein $\mathfrak{G}$-Erzeugendensystem von F mit $|Y| \leq |X|$. Wir betrachten eine surjektive Abbildung φ von X auf Y und setzen diese vermöge der Freiheit von F zu einem Endomorphismus $\overline{\varphi}$ von F fort. Wegen $X\overline{\varphi} = X\varphi = Y$ und $\langle Y \rangle_{\mathfrak{G}} = F$ ist $\overline{\varphi}$ surjektiv, also $F/\text{Kern}\,\overline{\varphi} \cong F$, nach 3.18 damit Kern $\overline{\varphi} = \{1_F\}$ und $\overline{\varphi}$ ein Automorphismus von F. Als Bild der $\mathfrak{G}$-Basis X unter $\overline{\varphi}$ ist dann auch Y eine $\mathfrak{G}$-Basis von F und zu X gleichmächtig. □

Kapitel 4

Freie Lie-Algebren

In Gruppen drückt der Kommutator $[a,b] = a^{-1}b^{-1}ab$ zweier Elemente a, b den „Unterschied“ zwischen ab und ba aus, denn es gilt: $ab = ba[a,b]$. In Algebren gibt es in der Regel zwar nicht diese auf multiplikative Invertierbarkeit angewiesene Bildung, jedoch statt dessen die Möglichkeit, den „additiven Unterschied“ zwischen ab und ba, nämlich $ab-ba$ zu betrachten; man bezeichnet dieses Element in der Algebrentheorie daher ebenfalls als Kommutator von a und b und schreibt es – solange keine Verwechslungen zu befürchten sind – ebenso wie in der Gruppentheorie, also in der Form $[a,b]$. Jedoch treten bisweilen in einem Kontext beide Typen von Kommutatoren auf, so daß es dann unterscheidender Schreibweisen bedarf (s. S. 100). Im Falle von Kommutatoren von Elementen der Einheitengruppe eines Ringes hat man folgende triviale Beziehung zwischen beiden Typen:

4.0.1 *Sind a, b Einheiten eines unitären Ringes R, so gilt*

$$a^{-1}b^{-1}ab = 1_R + a^{-1}b^{-1}(ab - ba)$$

□

Die Kommutatorbildung [.,.] stellt eine neue Verknüpfung auf der Ausgangs-Trägermenge dar. In einer *assoziativen* Algebra führt sie auf einen besonders wichtigen Typus von algebraischen Strukturen. Sei im folgenden K ein beliebiger kommutativer unitärer Ring.

4.0.2 *Gilt $((A,+),\cdot) \in {}^{K}\mathfrak{A}$, so auch $((A,+),[.,.]) \in {}^{K}\mathfrak{A}$, und $[a,a] = 0_A$ für alle $a \in A$,*

denn für alle a, b, $c \in A$, $r \in K$ gilt:

$$[a, b+c] = a(b+c) - (b+c)a = ab - ba + ac - ca = [a,b] + [a,c],$$

ebenso $[b+c,a] = [b,a] + [c,a]$, und

$$[a,rb] = a(rb) - (rb)a = (ra)b - b(ra) = [ra,b] = r(ab) - r(ba) = r[a,b],$$

ferner $[a,a] = aa - aa = 0_A$. □

4.0.3 *Gilt $A \in {}^K\mathfrak{R}$, so folgt für alle $a,b,c \in A$:*

$$[[a,b],c] + [[b,c],a] + [[c,a],b] = 0_A,$$

denn

$$\begin{aligned}&[[a,b],c] + [[b,c],a] + [[c,a],b]\\ &= (ab-ba)c - c(ab-ba) + (bc-cb)a - a(bc-cb) + (ca-ac)b - b(ca-ac)\\ &= 0_A.\end{aligned}$$

□

Die Assoziativität überträgt sich keineswegs von A auf die vermöge Kommutatorbildung gemäß 4.0.2 neu entstehende Algebra. Vielmehr tritt das in 4.0.3 aufgeführte Gesetz sozusagen an ihre Stelle, insofern es – wie die Assoziativität, nur in anderer Weise – eine harmonische Regel über das Produkt dreier Elemente ausdrückt.

4.1 Definition Sei $K \in \mathfrak{R}_1$ kommutativ, $((A,+),\circ) \in {}^K\mathfrak{A}$. A heißt eine K-Lie-Algebra (oder: Lie-Algebra über K), wenn für alle a, b, $c \in A$ gilt:

(i) $a \circ a = 0_A$,

(ii) $(a \circ b) \circ c + (b \circ c) \circ a + (c \circ a) \circ b = 0_A$.

Das Gesetz (i) wird das Anti-Kommutativgesetz genannt. Es impliziert

(i') $a \circ b = -b \circ a$,

denn nach (i) gilt: $0_A = (a+b)\circ(a+b) = a\circ a + a\circ b + b\circ a + b\circ b = a\circ b + b\circ a$. Falls $2_K a = 0_A$ nur für $a = 0_A$ gilt – insbesondere also, falls K ein Körper von einer Charakteristik $\neq 2$ ist – folgt umgekehrt auch (i) aus (i'), indem man letzteres mit $b = a$ anwendet. Die Gleichung (ii) heißt die Jacobi-Identität. Mit ${}^K\mathfrak{L}$ bezeichnen wir die Klasse aller Lie-(Links-)Algebren über K, und im Falle $K = \mathbb{Z}$ schreiben wir einfach $\mathfrak{L}$.[9] Nach 4.0.2 und 4.0.3 entsteht aus jeder *assoziativen* Algebra A vermöge der Kommutatorbildung als neuer Multiplikation kanonisch eine Lie-Algebra, die die zu A assoziierte Lie-Algebra genannt wird. Trivialerweise gilt:

[9] Aus historischen Gründen nennt man eine Lie-Algebra über $\mathbb{Z}$ in der Literatur häufig einen „Lie-Ring“. Da andererseits üblicherweise und bei uns stets das multiplikative Assoziativgesetz zu den Ring-Axiomen gehört, schließen wir uns dieser Tradition nicht an.

4.1.1 *${}^K\mathfrak{L}$ ist gegen Durchschnittsbildungen und Algebren-Homomorphismen abgeschlossen.* □

Wir zeigen als nächstes, daß es zu jeder Menge X eine von X frei erzeugte Lie-Algebra über K gibt. Dazu ziehen wir 2.16 heran. Das Vorgehen ist analog zu dem in 3.1.

4.2 Proposition *Sei $K \in \mathfrak{R}_1$ kommutativ, X eine Menge. Sei J das kleinste Ideal der freien K-Algebra $KX^{(+)}$, das die Elemente*

$$(*) \qquad (aa), \quad ((ab)c) + ((bc)a) + ((ca)b)$$

mit $a, b, c \in KX^{(+)}$ enthält. Sei $\widehat{\ }$ der kanonische Epimorphismus von $KX^{(+)}$ auf $KX^{(+)}/J$. Dann gilt:

(1) $\widehat{\ }|_X$ *ist injektiv*

(2) $KX^{(+)}/J$ *ist eine von $\widehat{X}$ frei erzeugte K-Lie-Algebra.*

Insbesondere gibt es eine von X frei erzeugte K-Lie-Algebra.

Beweis. (1) Sind $x, y \in X$ mit $\widehat{x} = \widehat{y}$, so gilt $x - y \in J \cap KX^{(1)} = \{0_{KX^{(+)}}\}$; letztere Gleichheit aufgrund der Definition von J. Es folgt: $x = y$.

(2) Nach Definition von J repräsentieren alle in $(*)$ aufgeführten Elemente die Nullrestklasse in $KX^{(+)}/J$. D.h., in der K-Algebra $KX^{(+)}/J$ gelten das Anti-Kommutativgesetz und die Jacobi-Identität, es ist also eine K-Lie-Algebra. Ist φ eine Abbildung von $\widehat{X}$ in eine K-Lie-Algebra L, so läßt sich die Zuordnung $\psi : X \to L$, $x \mapsto \widehat{x}\varphi$, zu einem K-Algebren-Homomorphismus $\overline{\psi}$ von $KX^{(+)}$ in L fortsetzen, da die K-Algebra $KX^{(+)}$ von X ${}^K\mathfrak{A}$-frei erzeugt wird. Da L eine K-Lie-Algebra ist, liegen alle Elemente aus $(*)$ in Kern $\overline{\psi}$. Da Kern $\overline{\psi}$ ein Ideal ist, folgt: $J \subseteq$ Kern $\overline{\psi}$. Also induziert $\overline{\psi}$ einen K-Algebren-Homomorphismus (und damit notwendigerweise einen ${}^K\mathfrak{L}$-Homomorphismus) $\overline{\varphi}$ von $KX^{(+)}/J$ in L, und es gilt: $\widehat{x}\overline{\varphi} = x\overline{\psi} = x\psi = \widehat{x}\varphi$. Wegen $\langle X \rangle_{\kappa_{\mathfrak{A}}} = KX^{(+)}$ ist die Fortsetzung $\overline{\varphi}$ von φ eindeutig bestimmt. Die abschließende Behauptung folgt nun mit Hilfe des Erweiterungsprinzips. □

Über die reine Existenz hinaus liefert 4.2 allerdings keine Einblicke in die Struktur der von X frei erzeugten K-Lie-Algebra. Solche werden vielmehr durch eine berühmte ganz andersartige Beschreibung möglich, die von Witt gefunden wurde und der wir uns in der Folge zuwenden werden. Es bedarf zunächst einiger wichtiger Vorbereitungen. Wir beginnen mit einer Erinnerung an Überlegungen im Anschluß an 2.16.1, in denen wir die freien Algebren

vermöge ihrer Teilmoduln der homogenen Elemente direkt zerlegt hatten. In der folgenden Definition geht es um Verfeinerungen der direkten Zerlegungen

$$KX^{(*)} = \bigoplus_{n\in\mathbb{N}_0} KX^{(n)}, \quad KX^* = \bigoplus_{n\in\mathbb{N}_0} KX^n.$$

4.3 Definition Sei X eine Menge. Wir schreiben $\mathbb{N}_{0,\mathrm{fin}}^X$ für die Menge der Abbildungen ν von X in $\mathbb{N}_0$ mit $x\nu = 0$ für fast alle $x \in X$ und nennen jede solche einen Multigrad bezüglich X. Es sei $X^{(\nu)}$ (bzw. X^ν) die Menge der Elemente von $X^{(*)}$ (bzw. X^*), in denen jedes $x \in X$ genau $x\nu$-mal vorkommt.

Beispiel: Sei $X = \{x, y, z\}$, $\nu : x \mapsto 3$, $y \mapsto 0$, $z \mapsto 1$. Dann gilt: $(((xx)z)x)$, $(x((zx)x))$, $((xz)(xx)), \ldots \in X^{(\nu)}$, (bzw. $xxzx$, $xzxx, \ldots \in X^\nu$).

Für jede Teilmenge T von $X^{(*)}$ (bzw. X^*) setzen wir $T_{(\nu)} := T \cap X^{(\nu)}$ (bzw. $T_\nu := T \cap X^\nu$). Der K-Teilmodul $KX^{(\nu)}$ von $KX^{(*)}$ (bzw. KX^ν von KX^*) heißt der K-Modul der homogenen Elemente vom Multigrad ν.

4.3.1 *Für alle $n \in \mathbb{N}_0$ gilt:*

$$KX^{(n)} = \dot{\bigoplus_{\substack{\nu:X\to\mathbb{N}_0 \\ \sum_{x\in X} x\nu = n}}} KX^{(\nu)}, \quad KX^n = \dot{\bigoplus_{\substack{\nu:X\to\mathbb{N}_0 \\ \sum_{x\in X} x\nu = n}}} KX^\nu.$$

□

Es folgt:

4.3.2 *Ist $\overline{\ }$ ein ${}^K\mathfrak{M}$-Homomorphismus von $KX^{(*)}$ (bzw. von KX^*) in einen K-Modul, so ist Bild $\overline{\ } = \langle \overline{X^{(\nu)}} \mid \nu$ Multigrad bezüglich $X\rangle_K$ (bzw. $= \langle \overline{X^\nu} \mid \nu$ Multigrad bezüglich $X\rangle_K$).* □

Sind ν, ν' Multigrade bezüglich X, so bedeute $\nu \le \nu'$, daß für alle $x \in X$ gilt: $x\nu \le x\nu'$. Ist $v \in X^{(\nu)}$ (bzw. $v \in X^\nu$), so nennen wir ν den Multigrad von v und setzen $v\mu := \nu$. Es gilt:

4.3.3 *μ ist ein Homomorphismus von $X^{(*)}$ (bzw. X^*) in $(\mathbb{N}_{0,\mathrm{fin}}^X, +)$; es gilt $(vw)\mu = v\mu + w\mu$ für alle v, w.* □

Ist X endlich, etwa $X = \{x_1, \ldots, x_n\}$, so wird ein Multigrad bezüglich X gewöhnlich als n-Tupel über $\mathbb{N}_0$ geschrieben, also z.B. (im Falle $n = 3$): $(x_1x_2x_2x_3x_1x_1)\mu = (3, 2, 1)$, $(((x_1x_1)x_3)(x_3x_3))\mu = (2, 0, 3)$.

4.4 Definition Sei X eine Menge, $H \subseteq X^{(+)}$ und $\preceq$ eine vollständige Ordnung auf H. Das Paar $(H, \preceq)$ heißt ein Hall-Gerüst über X, wenn die folgenden Bedingungen erfüllt sind:

(i) $X \subseteq H$.

(ii) Sind $r, s \in X^{(+)}$, so gilt:

$$(rs) \in H \quad \Leftrightarrow \quad \begin{cases} \text{(a) } r, s \in H \text{ und } r \prec s. \\ \text{(b) Ist } r \notin X \text{ und sind } p, q \in H \text{ mit } r = (pq), \\ \qquad \text{so ist } q \succeq s. \end{cases}$$

(iii) Sind $r, s \in H$ mit $(rs) \in H$, so folgt: $(rs) \prec s$.

Zwar machen diese Bedingungen auf den ersten Blick einen unanschaulich-technischen Eindruck, jedoch werden wir sogleich erkennen, daß sie in sehr natürlicher Weise erfüllbar und wir ihnen früher bereits sogar begegnet sind. Vor 1.20 haben wir das „Innenleben“ der Lyndon-Worte im freien Monoid über X studiert und exemplarisch eine vollständige „Innenzerlegung“ eines Lyndon-Worts angegeben. Durch die dort eigentlich nur zur Verdeutlichung der Zerlegungs-Iteration vorgenommene Beklammerung können wir das Ergebnis jedoch als Element des freien Magmas über X lesen, wobei die Beklammerung die Rolle der Verknüpfung übernimmt (siehe vor 2.6.3). Auf diese Weise gewinnen wir vermöge Iteration der Standardzerlegung (siehe 1.17) eine natürliche Abbildung $(\,.\,)$ von $\mathfrak{L}^X$ in $X^{(+)}$:

Sei $\leq$ eine vollständige Ordnung auf X. Induktiv nach der Wortlänge setzen wir $(x) := x$ für alle $x \in X$, und $(w) := ((u)(v))$, wenn $w \in \mathfrak{L}^X \setminus X$ und (u, v) die Standard-Zerlegung von w ist. Bezeichnet ϑ den kanonischen Epimorphismus von $X^{(*)}$ auf X^*, der durch „Weglassen der Klammern“ gegeben ist, so gilt $(w)\vartheta = w$ für alle $w \in \mathfrak{L}^X$ und damit die triviale Bemerkung

4.4.1 *$(\,.\,)$ ist eine injektive Abbildung von $\mathfrak{L}^X$ in $X^{(+)}$.* □

Aus der Definition von $(\,.\,)$ und der Tatsache, daß jedes Element von $X^{(+)} \setminus X$ sich auf genau eine Weise als Produkt zweier Elemente von $X^{(+)}$ schreiben läßt, erhalten wir für beliebige $u, v \in \mathfrak{L}^X$:

4.4.2 *Genau dann gilt $((u)(v)) \in (\mathfrak{L}^X)$, wenn $uv \in \mathfrak{L}^X$ gilt und (u, v) die Standard-Zerlegung von uv ist.* □

Auf $(\mathfrak{L}^X)$ definieren wir durch

$$\forall v, w \in \mathfrak{L}^X \quad (v) \preceq (w) \;:\Leftrightarrow\; v \underset{lex}{\leq} w$$

eine vollständige Ordnung und zeigen:

4.5 Proposition *$((\mathfrak{L}^X), \preceq)$ ist ein Hall-Gerüst über X.*

Beweis. Nach 4.4.1 und der Definition von $\preceq$ ist $(\,.\,)$ ein Isomorphismus der geordneten Menge $(\mathfrak{L}^X, \underset{lex}{\leq})$ auf die geordnete Menge $((\mathfrak{L}^X), \preceq)$. Daher können wir die zu zeigenden Eigenschaften unter Verwendung von 4.4.2 in äquivalente Aussagen über $\mathfrak{L}^X$ verwandeln; einzusehen sind dann:

(1) $X \subseteq \mathfrak{L}^X$.

(2) Sind u, $v \in X^+$, so ist uv genau dann ein Lyndon-Wort mit Standard-Zerlegung (u, v), wenn die Bedingungen

   (a) $u, v \in \mathfrak{L}^X$ und $u \underset{lex}{<} v$,

   (b) Ist $u \notin X$ und t der längste zu $\mathfrak{L}^X$ gehörige echte Rechtsfaktor von u, so gilt: $t \underset{lex}{\geq} v$

   erfüllt sind.

(3) Sind u, $v \in \mathfrak{L}^X$ mit $uv \in \mathfrak{L}^X$, so folgt: $uv \underset{lex}{<} v$.

Es ist (1) trivial, (2) wurde in 1.19 bewiesen, und (3) ist ein Spezialfall von 1.15. □

Wir haben damit gezeigt, daß der kanonische Epimorphismus ϑ ein gewisses Hall-Gerüst über X bijektiv auf das Repräsentantensystem $\mathfrak{L}^X$ der Konjugiertenklassen primitiver Worte über X abbildet. Diese Aussage ordnet sich dem folgenden schönen Resultat unter, das wir hier ohne Beweis angeben:

Satz (Viennot 1978) *Sei X eine Menge und $(H, \preceq)$ ein Hall-Gerüst über X. Dann ist $\vartheta|_H$ eine Bijektion von H auf ein Repräsentantensystem der Konjugiertenklassen primitiver Worte über X.*

Hall-Gerüste sind für uns vor allem wegen der folgenden Aussage wichtig:

4.6 Proposition *Sei $K \in \mathfrak{R}_1$ kommutativ, $L \in {}^K\mathfrak{L}$, X eine Menge und $\overline{}$ ein ${}^K\mathfrak{A}$-Epimorphismus von $KX^{(+)}$ auf L. Sei $(H, \preceq)$ ein Hall-Gerüst über X. Dann ist $\overline{H}$ ein ${}^K\mathfrak{M}$-Erzeugendensystem von L; genauer: Für alle Multigrade ν bezüglich X gilt: $\langle \overline{X^{(\nu)}} \rangle_K = \langle \overline{H_{(\nu)}} \rangle_K$.*

Beweis (M. Schocker). Es wird gezeigt:

$(*)$ Für alle $r, s \in H$ mit $r \prec s$ gilt: $\overline{r} \circ \overline{s} \in \langle \overline{t} \mid t \in H_{((rs)\mu)}, t \prec s \rangle_K$.

Angenommen, dies sei falsch. Dann gibt es einen minimalen Multigrad ν, zu dem es ein Paar $(r, s) \in H \times H$ mit $r \prec s$ und $(rs)\mu = \nu$ gibt, so daß gilt: $\overline{r} \circ \overline{s} \notin \langle \overline{t} \mid t \in H_{(\nu)}, t \prec s \rangle_K$. Da $X^{(\nu)}$ endlich ist, können wir (r, s) unter allen

solchen Paaren so wählen, daß s dabei bezüglich der Ordnung $\preceq$ so klein wie möglich ist.

Wäre $(rs) \in H$, so $(rs) \in H_{(\nu)}$ und $(rs) \prec s$ nach 4.4(iii), also $\overline{r} \circ \overline{s} = \overline{(rs)} \in \langle \overline{t} \,|\, t \in H_{(\nu)}, t \prec s \rangle_K$, Widerspruch. Also gilt $(rs) \notin H$. Da aber die Bedingung (a) aus 4.4(ii) erfüllt ist, muß (b) verletzt sein, d.h. es muß p, $q \in H$ geben mit $r = (pq)$, $q \prec s$. Es gilt $p \prec q$ nach 4.4(ii), und nach 4.3.3 ferner $(ps)\mu$, $(qs)\mu < \nu$. Nach Wahl von ν impliziert letzteres:

$$\overline{p} \circ \overline{s} \in \langle \overline{t} \,|\, t \in H_{((ps)\mu)},\, t \prec s \rangle_K, \quad \overline{q} \circ \overline{s} \in \langle \overline{t} \,|\, t \in H_{((qs)\mu)},\, t \prec s \rangle_K.$$

Sei $t^* \in H_{((qs)\mu)}$, $t^* \prec s$. Es gilt: $(pt^*)\mu = \nu = (t^*p)\mu$, nach Wahl von s im Falle $p \prec t^*$ also $\overline{p} \circ \overline{t^*} \in \langle \overline{t} \,|\, t \in H_{(\nu)}, t \prec t^* \rangle_K \subseteq \langle \overline{t} \,|\, t \in H_{(\nu)}, t \prec s \rangle_K$, und im Falle $t^* \prec p$ ähnlich $-\overline{p} \circ \overline{t^*} = \overline{t^*} \circ \overline{p} \in \langle \overline{t} \,|\, t \in H_{(\nu)}, t \prec p \rangle_K \subseteq \langle \overline{t} \,|\, t \in H_{(\nu)}, t \prec s \rangle_K$. Es folgt:

$$\overline{p} \circ (\overline{q} \circ \overline{s}) \in \langle \overline{t} \,|\, t \in H_{(\nu)},\, t \prec s \rangle_K,$$

sowie unter Vertauschung der Rollen von p und q ebenso

$$\overline{q} \circ (\overline{p} \circ \overline{s}) \in \langle \overline{t} \,|\, t \in H_{(\nu)},\, t \prec s \rangle_K.$$

Unter Verwendung von 4.1(i),(ii) führt unsere Annahme daher zu dem Widerspruch

$$\overline{r} \circ \overline{s} = (\overline{p} \circ \overline{q}) \circ \overline{s} = \overline{p} \circ (\overline{q} \circ \overline{s}) - \overline{q} \circ (\overline{p} \circ \overline{s}) \in \langle \overline{t} \,|\, t \in H_{(\nu)},\, t \prec s \rangle_K,$$

womit $(*)$ bewiesen ist.

Für jedes $w \in X^{(+)}$ ist zu zeigen: $\overline{w} \in \langle \overline{H_{(w\mu)}} \rangle_K$, was wir durch Induktion nach dem Grad von w einsehen: Ist dieser 1, so $w \in X$ und die Behauptung trivial. Sonst gibt es r, $s \in X^{(+)}$ mit $w = (rs)$. Induktiv können wir annehmen: $\overline{r} \in \langle \overline{H_{(r\mu)}} \rangle_K$, $\overline{s} \in \langle \overline{H_{(s\mu)}} \rangle_K$. Also gibt es $r^{(1)}, \ldots, r^{(m)} \in H_{(r\mu)}$, $s^{(1)}, \ldots, s^{(n)} \in H_{(s\mu)}$, $c_1, \ldots, c_m$, $d_1, \ldots, d_n \in K$ mit $\overline{r} = \sum_{i \in \underline{m}} c_i \overline{r^{(i)}}$, $\overline{s} = \sum_{j \in \underline{n}} d_j \overline{s^{(j)}}$. Es gilt:

$$\overline{w} = \sum_{i,j} c_i d_j (\overline{r^{(i)}} \circ \overline{s^{(j)}}),$$

so daß es genügt zu zeigen: $\overline{r^{(i)}} \circ \overline{s^{(j)}} \in \langle \overline{H} \rangle_K$ für alle $i \in \underline{m}$, $j \in \underline{n}$. Im Falle $r^{(i)} = s^{(j)}$ ist dies trivial. Gilt aber $r^{(i)} \prec s^{(j)}$ oder $s^{(j)} \prec r^{(i)}$, so folgt es wegen $\overline{r^{(i)}} \circ \overline{s^{(j)}} = -\overline{s^{(j)}} \circ \overline{r^{(i)}}$ aus $(*)$. $\square$

Wir werden 4.6 auf eine Teilalgebra L der zu einer assoziativen K-Algebra A assoziierten Lie-Algebra (s. 4.1) anwenden, eine Situation, für die wir die Schreibweise „$L \leq_{K_{\mathfrak{L}}} A$“ verwenden. Ist S ein K-Teilmodul von A, $n \in \mathbb{N}$, so bezeichne $S^{\dot{n}}$ in 4.7 den K-Teilmodul von A, der von der Menge aller

n-stelligen Produkte von Elementen aus S im Sinne der *assoziativen* Multiplikation in A erzeugt wird, zur Unterscheidung von der hier nicht gemeinten entsprechenden Bildung bezüglich der Lie-Multiplikation bzw. auch zur Menge der n-Tupel über S.

4.7 Proposition *Sei $K \in \mathfrak{R}_1$, kommutativ, $A \in {}^K\mathfrak{R}$ und $L \leq_{K\mathfrak{L}} A$.*

(1) *Für alle $n \in \mathbb{N}_{>1}$, $t_1, \ldots, t_n \in L$, $\sigma \in S_n$ gilt:*

$$t_{1\sigma} \cdots t_{n\sigma} \in L^{\dot{n-1}} + t_1 \cdots t_n.$$

(2) *Ist T ein ${}^K\mathfrak{M}$-Erzeugendensystem von L und $\preceq$ eine vollständige Ordnung auf T, so gilt für alle $n \in \mathbb{N}$:*

$$L + L^{\dot{2}} + \cdots + L^{\dot{n}} = \langle t_1 \cdots t_k \,|\, k \in \underline{n}, t_i \in T,\, t_1 \succeq \cdots \succeq t_k \rangle_K.$$

(3) *Sind T, $\preceq$ wie in (2) und gilt $\langle L \rangle_{K\mathfrak{R}} = A$, so folgt:*

$$A = \langle t_1 \cdots t_k \,|\, k \in \mathbb{N}, t_i \in T,\, t_1 \succeq \cdots \succeq t_k \rangle_K.$$

Beweis. (1) Seien $n \in \mathbb{N}_{>1}$, $t_1, \ldots, t_n \in L$ und $i \in \underline{n-1}$. Es gilt

$$t_1 \cdots t_{i-1} t_{i+1} t_i t_{i+2} \cdots t_n = t_1 \cdots t_{i-1} [t_{i+1}, t_i] t_{i+2} \cdots t_n + t_1 \cdots t_n,$$

also die Behauptung in dem Fall, daß σ eine Transposition vom Typ $(i, i+1)$ mit $i \in \underline{n-1}$ ist. Die Menge $\mathcal{T}$ dieser Transpositionen ist ein $\mathfrak{G}$-Erzeugendensystem von S_n. Iteriertes Anwenden der eben gemachten Bemerkung, d.h. eine triviale Induktion nach $l_{\mathcal{T}}(\sigma)$, liefert daher die Behauptung für beliebiges $\sigma \in S_n$.

(2) folgt nun durch Induktion nach n: Wegen $\langle T \rangle_K = L$ ist die Behauptung für $n = 1$ trivial. Sei nun $n > 1$, $L + L^{\dot{2}} \cdots + L^{\dot{n-1}} = \langle t_1 \cdots t_k \,|\, k \in \underline{n-1}, t_i \in T, t_1 \succeq \cdots \succeq t_k \rangle_K$ vorausgesetzt und $z \in L^{\dot{n}}$. Dann ist z eine K-Linearkombination von n-stelligen Produkten über der Menge T. Ist $(t_1, \ldots, t_n) \in T^n$, so liegen nach (1) alle Produkte, die sich aus den Faktoren $t_1, \ldots, t_n$ bilden lassen, in $L^{\dot{n-1}} + t_1 \cdots t_n$. Daher dürfen wir hierin annehmen: $t_1 \succeq \cdots \succeq t_n$. Es folgt:

$$L^{\dot{n}} \subseteq L^{\dot{n-1}} + \langle t_1 \cdots t_n | t_i \in T, t_1 \succeq \cdots \succeq t_n \rangle_K,$$

also

$$L + L^{\dot{2}} + \cdots + L^{\dot{n}} \subseteq \langle t_1 \cdots t_k | k \in \underline{n}, t_i \in T, t_1 \succeq \cdots \succeq t_k \rangle_K.$$

Die umgekehrte Inklusion ist trivial.

(3) Die Voraussetzung $\langle L\rangle_{K\mathfrak{A}} = A$ und (2) ergeben:

$$A = \sum_{n\in\mathbb{N}} L^{\dot n} = \langle t_1 \cdots t_k \,|\, k \in \mathbb{N}, t_i \in T,\, t_1 \succeq \cdots \succeq t_k\rangle_K.$$

□

Ist X eine Teilmenge einer beliebigen Lie-Algebra M über einem kommutativen unitären Ring K, so läßt sich die Identität id_X auf genau eine Weise zu einem ${}^K\mathfrak{A}$-Homomorphismus von $KX^{(+)}$ in M fortsetzen. Dessen Bild ist die von X erzeugte (Lie-)Teilalgebra von M. Wir wählen für M nun speziell die zur freien assoziativen Algebra KX^+ assoziierte Lie-Algebra (siehe 4.0.2):

4.8 Definition Sei $K \in \mathfrak{R}_1$, kommutativ, X eine vollständig geordnete Menge. Die von X erzeugte Lie-Teilalgebra von $((KX^+,+),[.,.])$ wird die K-Algebra der Lie-Elemente über X genannt; wir bezeichnen sie mit $L_{X,K}$ $(=\langle X\rangle_{K\mathfrak{L}})$.[10]

4.8.1 *Für alle $n \in \mathbb{N}_0$ gilt:*

$$L_{X,K} = \dot{\bigoplus_{n\in\mathbb{N}}}(L_{X,K} \cap KX^n), \quad L_{X,K} \cap KX^n = \dot{\bigoplus_{\substack{\nu: X\to\mathbb{N}_0 \\ \sum_{x\in X} x\nu = n}}} (L_{X,K} \cap KX^\nu),$$

denn die Summe aller Räume $L_{X,K} \cap KX^\nu$ ist gegenüber $[.,.]$ abgeschlossen und daher eine ${}^K\mathfrak{L}$-Teilalgebra von $L_{X,K}$. Sie enthält X und stimmt daher mit $L_{X,K}$ überein. Das Behauptete folgt nun aus 4.3.1 bzw. der zweiten vor 4.3 angegebenen Gleichung. □

Wir schreiben $\overline{\ }$ für die Fortsetzung von id_X zu einem ${}^K\mathfrak{A}$-Epimorphismus von $KX^{(+)}$ auf $L_{X,K}$ und definieren eine Abbildung $[.]$ von $\mathfrak{L}^X$ in $L_{X,K}$, indem wir für alle $v \in \mathfrak{L}^X$ setzen: $[v] := \overline{(v)}$. (Zur Definition von $(.)$ siehe 4.4.)

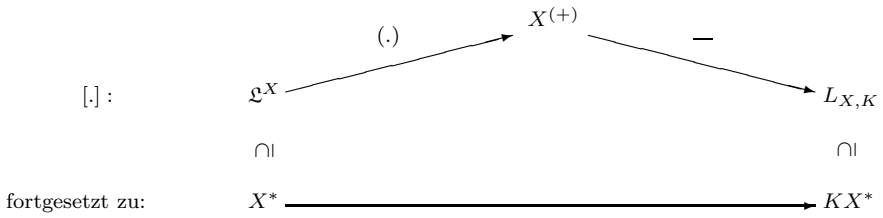
Diese Abbildung setzen wir zu einer Abbildung von X^* in KX^* fort, indem wir für alle $w \in X^*$ festlegen:

$$[w] := [z^{(1)}]\cdots[z^{(k)}], \text{ falls } (z^{(1)},\ldots,z^{(k)}) \text{ die Lyndon-Zerlegung von } w \text{ ist.}$$

[10] Bezeichnet ω den ${}^K\mathfrak{M}$-Homomorphismus von KX^+ in $L_{X,K}$ mit $x\omega = x$, $(ux)\omega = [u\omega, x]$ für alle $x \in X$, $u \in X^+$, so ist $\{w | w \in KX^+, \forall u \in X^+ \quad (uw)\omega = [u\omega, w]\}$ eine Lie-Teilalgebra von KX^+ und enthält X, folglich $L_{X,K}$. Für alle $n \in \mathbb{N}$, $v \in KX^n$ folgern wir $v\omega^2 = nv\omega$: Induktiv erhält man $(ux)\omega^2 = \big((u\omega)x - x(u\omega)\big)\omega = [u\omega^2, x] - [x, u\omega] = [(n-1)u\omega, x] + (ux)\omega = n(ux)\omega$ für $n > 1$, $u \in X^{n-1}$, $x \in X$. Es gilt $KX^+\omega = L_{X,K}$, also

$$KX^n\omega = L_{X,K} \cap KX^n \subseteq \{v | v \in KX^n, v\omega = nv\}.$$

Ist $n \cdot 1_K$ in K invertierbar, so gilt Gleichheit, insbesondere also, wenn K ein Körper der Charakteristik 0 ist (= **Satz von Dynkin-Specht-Wever (1947)**).



Das Bild jedes Wortes w unter [.] ist also das (assoziative) Produkt derjenigen Lie-Kommutatoren, die man aus dem „Innenleben“ der Lyndon-Worte (siehe vor 1.20) aus der Lyndon-Zerlegung von w gewinnt, indem man in diesen jede „runde Klammer“ durch eine Kommutatorbildung (=„eckige Klammer“) ersetzt. Wir illustrieren dies vermöge des Beispiels von S. 24: Ist $X = \{a, b, c\}$, $a < b < c$ und

$$w = bbabbacababc,$$

so ist $(b, b, abbac, ababc)$ die Lyndon-Zerlegung von w. Die vollständige „Innenzerlegung“ durch Iteration der Standardzerlegung haben wir für die aufgetretenen Faktoren bereits früher bestimmt (s.S. 22), so daß sich ergibt:

$$[w] = b\,b\,\Big[\big[[a,b],b\big],[a,c]\Big]\,\Big[[a,b],\big[a,[b,c]\big]\Big].$$

Offensichtlich gilt für jeden Multigrad ν bezüglich X:

4.8.2 $[X^\nu] \subseteq KX^\nu$, $[\mathfrak{L}_\nu^X] \subseteq L_{X,K} \cap KX^\nu$. □

4.9 Lemma *Sei $K \in \mathfrak{R}_1$, kommutativ, und sei X eine vollständig geordnete Menge. Sind $k \in \mathbb{N}$, $w^{(1)}, \ldots, w^{(k)} \in \mathfrak{L}^X$ und $w := w^{(1)} \cdots w^{(k)}$, so gilt:*

$$[w^{(1)}] \cdots [w^{(k)}] \in w + \langle y | y \in X^{w\mu},\ y \underset{lex}{>} w\rangle_K.$$

Beweis. Für jedes $w \in X^*$ sei $M_w := \langle y | y \in X^{w\mu},\ y \underset{lex}{>} w\rangle_K$. Wir beweisen die Behauptung durch Induktion nach k und behandeln den Induktionsanfang $k = 1$ durch Induktion nach der Wortlänge:

Sei $w \in \mathfrak{L}^X$. Im Falle $w \in X$ ist die Behauptung trivial. Sei nun $w \in \mathfrak{L}^X \setminus X$ und (u, v) die Standardzerlegung von w. Induktiv sei vorausgesetzt, daß es Elemente $u' \in M_u$, $v' \in M_v$ gebe mit $[u] = u + u'$, $[v] = v + v'$. Es folgt:

$$\begin{aligned}[w] = [[u],[v]] &= (u+u')(v+v') - (v+v')(u+u') \\ &= uv + uv' + u'[v] - vu - vu' - v'[u] \\ &\in w + M_w,\end{aligned}$$

denn die Elemente uv', $u'[v]$, vu, vu', $v'[u]$ sind sämtlich homogen vom Multigrad $w\mu$ und K-Linearkombinationen von Worten, die lexikographisch größer als w sind; vu ist selbst ein solches wegen $w = uv \in \mathfrak{L}^X$, und für die übrigen gilt das Behauptete aufgrund der Eigenschaften von u' und v'. Im Fall $k = 1$ ist die Behauptung damit bewiesen.

Seien nun $k > 1$ und $w^{(1)}, \ldots, w^{(k)} \in \mathfrak{L}^X$. Wir setzen $w^* := w^{(1)} \cdots w^{(k-1)}$, $w := w^* w^{(k)}$ und nehmen induktiv an, daß es Elemente $t \in M_{w^*}$, $t' \in M_{w^{(k)}}$ gebe mit $[w^{(1)}] \cdots [w^{(k-1)}] = w^* + t$, $[w^{(k)}] = w^{(k)} + t'$. Es folgt:

$$\begin{aligned}[w^{(1)}] \cdots [w^{(k)}] &= (w^* + t)(w^{(k)} + t') \\ &= w + w^* t' + t\,[w^{(k)}] \\ &\in w + M_w,\end{aligned}$$

ebenso wie oben. □

4.10 Hauptsatz *Sei $K \in \mathfrak{R}_1$, kommutativ, und sei X eine vollständig geordnete Menge.*

(1) *Die Abbildung $[.]$ ist ein ${}^K\mathfrak{M}$-Automorphismus von KX^*.*

(2) **(Chen, Fox, Lyndon 1958)** *$[\mathfrak{L}^X]$ ist eine ${}^K\mathfrak{M}$-Basis von $L_{X,K}$.*

(3) **(Basis-Satz für Lyndon'sche Lie-Elemente)** *$[X^*]$ ist eine ${}^K\mathfrak{M}$-Basis von KX^*.*

(4) **(Witt 1937)** *$L_{X,K}$ ist eine von X frei erzeugte Lie-Algebra über K.*

(5) **(Witt'sche Dimensionsformeln, 1937)** *Ist X endlich, $r := |X|$ und $n \in \mathbb{N}$, so gilt:*

$$rk_K(L_{X,K} \cap KX^n) = \frac{1}{n} \sum_{d|n} \mu(d) r^{\frac{n}{d}},$$

$$rk_K(L_{X,K} \cap KX^\nu) = \frac{1}{n} \sum_{d|k_1,\ldots,k_r} \mu(d) \frac{\frac{n}{d}!}{\frac{k_1}{d}! \cdots \frac{k_r}{d}!},$$

wenn ν der (als r-Tupel geschriebene) Multigrad $(k_1, \ldots, k_r)$ und $n = k_1 + \cdots + k_r$ ist.

(6) **(Satz von der universellen Eigenschaft)** *Ist $A \in {}^K\mathfrak{R}_1$, so läßt sich jeder ${}^K\mathfrak{L}$-Homomorphismus von $L_{X,K}$ in $((A, +), [.,.])$ zu genau einem ${}^K\mathfrak{R}_1$-Homomorphismus von KX^* in A fortsetzen.*

Beweis. (1), (2), (3) Sind $w_1, \ldots, w_n \in X^*$, paarweise verschieden, und ist darunter w_n das lexikographisch kleinste Element, so folgt für alle $c_1, \ldots, c_n \in K$ nach 4.9:

$$c_1[w_1] + \cdots + c_n[w_n] \in c_n w_n + \langle y | y \in X^*, y \underset{lex}{>} w_n \rangle_K,$$

also $c_1[w_1] + \cdots + c_n[w_n] \neq 0_{KX^*}$, falls $c_n \neq 0_K$. Es folgt, daß X^* von $[.]$ bijektiv auf eine K-linear unabhängige Teilmenge von KX^* abgebildet wird.

Aus 4.5 und 4.6 folgt $\langle [\mathfrak{L}^X] \rangle_{{}^K\mathfrak{M}} = L_{X,K}$, woraus zunächst (2) folgt. Weiter gilt wegen $X \subseteq L_{X,K}$ trivialerweise $\langle [\mathfrak{L}^X] \rangle_{{}^K\mathfrak{A}} = KX^*$, so daß die Voraussetzungen von 4.7(3) erfüllt sind; diese Aussage wenden wir nun an mit $A := KX^*$, $L := L_{X,K}$, $T := [\mathfrak{L}^X]$ und der durch

$$[u] \preceq [v] \ :\Leftrightarrow \ u \underset{lex}{\leq} v \qquad (u, v \in \mathfrak{L}^X)$$

auf T gegebenen Ordnung. Sie besagt dann, daß $[X^*]$ ein ${}^K\mathfrak{M}$-Erzeugendensystem von KX^* ist, so daß auch (3) bewiesen ist. Es ist also $[.]$ ein ${}^K\mathfrak{M}$-Endomorphismus von KX^*, der eine ${}^K\mathfrak{M}$-Basis (nämlich X^*) bijektiv auf eine ${}^K\mathfrak{M}$-Basis (nämlich $[X^*]$) abbildet; es folgt (1).

(4) Seien J und $\widehat{\ }$ wie in 4.2. Wir wollen zeigen, daß $L_{X,K}$ zu $KX^{(+)}/\underline{J}$ ${}^K\mathfrak{L}$-isomorph ist und betrachten dazu den kanonischen ${}^K\mathfrak{A}$-Epimorphismus $\overline{\ }$ von $KX^{(+)}$ auf $L_{X,K}$ (siehe 4.8). Da $L_{X,K}$ eine Lie-Algebra ist, enthält sein Kern alle in 4.2$(*)$ aufgeführten Elemente von $KX^{(+)}$ und induziert daher einen ${}^K\mathfrak{L}$-Epimorphismus ψ von $KX^{(+)}/J$ auf $L_{X,K}$. Wir müssen nur zeigen, daß ψ injektiv ist; dann folgt die Behauptung unter Verwendung beider Teile von 4.2, denn $\widehat{x}\psi = [x] = x$ für alle $x \in X$. Sei $s \in KX^{(+)}$ mit $\widehat{s} \in \operatorname{Kern}\psi$. Da $(\mathfrak{L}^X)$ nach 4.5 ein Hall-Gerüst über X ist, folgt aus 4.6: $\langle \widehat{(\mathfrak{L}^X)} \rangle_{{}^K\mathfrak{M}} = KX^{(+)}/J$. Also existieren paarweise verschiedene $w_1, \ldots, w_n \in \mathfrak{L}^X$ und $c_1, \ldots, c_n \in K$ mit $\widehat{s} = c_1\widehat{(w_1)} + \cdots + c_n\widehat{(w_n)}$. Es folgt:

$$0_{KX^*} = \widehat{s}\psi = c_1[w_1] + \cdots + c_n[w_n],$$

denn die Hintereinanderausführung von $\widehat{\ }$ und ψ ist die Abbildung $\overline{\ }$. Nach (1) und (2) impliziert dies: $c_1, \ldots, c_n = 0_K$.

(5) Die Formeln folgen aus (1) und (2) in Verbindung mit 1.12(3),(4) denn nach 4.8.1 ist $L_{X,K}$ die direkte Summe der Räume $L_{X,K} \cap KX^n$, $n \in \mathbb{N}$, und gleichfalls die der Räume $L_{X,K} \cap KX^\nu$, $\nu \in \mathbb{N}_0^X$.

(6) Sei $A \in {}^K\mathfrak{R}_1$ und φ ein ${}^K\mathfrak{L}$-Homomorphismus von $L_{X,K}$ in $((A, +), [., .])$. Da KX^* von X ${}^K\mathfrak{R}_1$-frei erzeugt wird, besitzt die Abbildung $\varphi|_X$ genau eine

Fortsetzung ψ zu einem ${}^K\mathfrak{R}_1$-Homomorphismus von KX^* in A. Es ist $\psi|_{L_{X,K}}$ ein ${}^K\mathfrak{L}$-Homomorphismus von $L_{X,K}$ in A mit $\psi|_X = \varphi|_X$. Da $L_{X,K}$ von X nach (4) ${}^K\mathfrak{L}$-frei erzeugt wird, gibt es genau eine Fortsetzung von $\varphi|_X$ zu einem ${}^K\mathfrak{L}$-Homomorphismus von $L_{X,K}$ in A. Also gilt: $\varphi = \psi|_{L_{X,K}}$, d.h. ψ ist eine Fortsetzung von φ zu einem ${}^K\mathfrak{R}_1$-Homomorphismus von KX^* in A. Wegen $X \subseteq L_{X,K}$ ist diese Fortsetzung eindeutig bestimmt. □

Die sogenannte „Chen-Fox-Lyndon-Basis“ $[\mathfrak{L}^X]$ der (nach 4.10(4)) freien Lie-Algebra $L_{X,K}$ ist nur ein Beispiel für eine umfangreiche Familie von Basen: Es läßt sich zeigen, daß nicht nur $[\mathfrak{L}^X]$, sondern allgemein das Bild $\overline{H}$ eines beliebigen Hall-Gerüsts H über X unter dem id_X fortsetzenden Algebren-Epimorphismus von $KX^{(+)}$ auf $L_{X,K}$ eine K-Basis von $L_{X,K}$ ist ([Bo2], II §2.11).

Die in 4.10(6) zum Ausdruck gebrachte „universelle Eigenschaft“ ist keineswegs an die Freiheit von $L_{X,K}$ als Lie-Algebra gebunden: Sei L eine beliebige Lie-Algebra über K. Ein Paar (U, ι), bestehend aus einer assoziativen unitären K-Algebra U und einem ${}^K\mathfrak{L}$-Homomorphismus ι von L in U heißt eine universelle Einhüllende von L, wenn es zu jedem ${}^K\mathfrak{L}$-Homomorphismus φ von L in die zu einer assoziativen unitären K-Algebra A assoziierte Lie-Algebra genau einen ${}^K\mathfrak{R}_1$-Homomorphismus ψ von U in A gibt mit $\varphi = \iota\psi$. Sind (U, ι), (U', ι') universelle Einhüllende zu einer Lie-Algebra L, so gibt es einen ${}^K\mathfrak{R}_1$-Isomorphismus α von U auf U' mit $\iota\alpha = \iota'$. Die Aussage 4.10(6) bedeutet, daß $(KX^*, \mathrm{id}_{L_{X,K}})$ eine universelle Einhüllende der freien Lie-Algebra $L_{X,K}$ ist. Hieraus erhalten wir unschwer die Einsicht, daß es zu *jeder* Lie-Algebra L über K eine universelle Einhüllende gibt: Sei X ein ${}^K\mathfrak{L}$-Erzeugendensystem von L, λ der ${}^K\mathfrak{L}$-Epimomorphismus von $L_{X,K}$ auf L mit $x \mapsto x$ für alle $x \in X$, $J := \operatorname{Kern}\lambda$. Sei $\tilde{J}$ das von J erzeugte Ideal der assoziativen K-Algebra KX^*, ι die Hintereinanderausführung von λ^{-1} und dem kanonischen Homomorphismus von $L_{X,K}/J$ in $KX^*/\tilde{J}$. Ist nun φ ein ${}^K\mathfrak{L}$-Homomorphismus von L in die zu einer assoziativen unitären K-Algebra A assoziierte Lie-Algebra, so ist $\varphi_{|X}$ eindeutig zu einem ${}^K\mathfrak{R}$-Homomorphismus von KX^* in A fortsetzbar. Dessen Kern enthält J, damit aber auch $\tilde{J}$. Es folgt, daß $(KX^*/\tilde{J}, \iota)$ eine universelle Einhüllende von L ist. Im allgemeinen muß zwar ι dabei nicht injektiv sein; dies ist aber jedenfalls dann beweisbar, wenn der Trägerraum der Lie-Algebra eine K-Basis besitzt, insbesondere also im Falle eines Körpers K. Die Einbettung der freien Lie-Algebra über X in die freie unitäre assoziative Algebra über X nach dem Satz von Witt (4.10(4)) illustriert als wichtiger Sonderfall den nachstehenden allgemeinen

Satz (Poincaré, Birkhoff, Witt 1937) *Sei K ein kommutativer unitärer Ring, $L \in {}^K\mathfrak{L}$, (U, ι) eine universelle Einhüllende von L. Es gebe eine K-Basis B von L. Dann gilt:*

(1) *ι ist injektiv.*

(2) **(Allgemeiner Basis-Satz für Lie-Elemente)** *Ist $\preceq$ eine vollständige Ordnung auf B und o. B. d. A. (nach (1)) $L \leq_{K_{\mathfrak{L}}} U$, so ist*

$$\{b_1 \cdots b_k | k \in \mathbb{N}_0, b_1, \ldots, b_k \in B, b_1 \succeq \cdots \succeq b_k\}$$

eine K-Basis von U.

Die Theorie der freien Lie-Algebren berührt nicht nur in überraschender Weise die der freien Monoide und freien assoziativen Algebren, sondern steht vor allem auch in inhaltsreicher Beziehung zu der der freien Gruppen. Diesem gehaltvollen Zusammenhang wenden wir uns in der Folge zu. Wir erinnern zunächst daran, daß eine Verknüpfung auf einer Menge M stets eine Verknüpfung auf der Potenzmenge $\mathfrak{P}(M)$ von M induziert, und zwar vermöge der natürlichen Setzung $ST := \{st | s \in S,\ t \in T\}$ für alle $S, T \subseteq M$.

4.11 Definition Sei $(M, \cdot)$ ein Magma und $s \in \mathbb{N}_0$. Eine Abbildung

$$\mathbb{N}_{\geq s} \to \mathfrak{P}(M), \quad n \mapsto A_n$$

nennen wir subhomomorph[11], wenn gilt:

$$\forall m, n \in \mathbb{N}_{\geq s} \quad A_m \cdot A_n \subseteq A_{m+n}.$$

Unter einer natürlichen Graduierung einer Algebra $(A, +, \cdot)$ verstehen wir eine bezüglich $\cdot$ subhomomorphe Abbildung von $\mathbb{N}_{\geq s}$ auf eine Menge additiver Untergruppen von A, so daß A deren direkte Summe ist und es höchstens im Fall $A_n = \{0_A\}$ ein $m \in \mathbb{N}_{\geq s}$ mit $m \neq n$, $A_m = A_n$ gibt. Bei gegebener Graduierung $(A_n)_{n \in \mathbb{N}_{\geq s}}$ von A heißen die Elemente von $A_n \setminus \{0_A\}$ homogen vom Grad n.

Sei zum Beispiel X eine Menge, K ein kommutativer unitärer Ring und $A := KX^{(*)}$, $A_n := KX^{(n)}$ für alle $n \in \mathbb{N}_0$. Es gilt

$$KX^{(*)} = \dot{\bigoplus_{n \in \mathbb{N}_0}} KX^{(n)}, \qquad (KX^{(m)})(KX^{(n)}) \subseteq KX^{(m+n)} \text{ für alle } m, n \in \mathbb{N}_0,$$

[11] Einen begrifflichen Rahmen in natürlicher Allgemeinheit hierfür erhält man, indem man zwei *beliebige* Magmen $(\mathcal{M}, \cdot)$, $(I, \circ)$ sowie eine teilweise Ordnung $\preceq$ auf $\mathcal{M}$ betrachtet und eine Abbildung $i \mapsto A_i$ von I in $\mathcal{M}$ subhomomorph nennt, wenn für alle $i, j \in I$ gilt: $A_i \cdot A_j \preceq A_{i \circ j}$. Für uns wird aber nur der obige Spezialfall ($\mathcal{M} = \mathfrak{P}(M)$ für ein Magma M, $\preceq = \subseteq$, $I = \mathbb{N}_{\geq s}$, $\circ = +$) eine Rolle spielen. In den Anwendungen in diesem Kapitel wird sogar stets $s \in \{0, 1\}$ gelten. Wir verwenden für unsere subhomomorphen Abbildungen die übliche „Folgen-Schreibweise“ $(A_n)_{n \in \mathbb{N}_{\geq s}}$.

also ist $(A_n)_{n\in\mathbb{N}_0}$ eine natürliche Graduierung von $KX^{(*)}$. Ebenso erhält man eine Graduierung von KX^* durch die Folge $(KX^n)_{n\in\mathbb{N}_0}$. Als weiteres Beispiel sei der Polynomring $K[X]$ genannt, für den die Folge der Teilräume $A_n := \langle x_1^{k_1}\cdots x_r^{k_r} | x_i \in X,\ k_1+\cdots+k_r = n\rangle_K$ $(n \in \mathbb{N}_0)$ eine Graduierung ist. Dieses Beispiel ist als Ursprung der Namensgebung anzusehen. Die in diesem Kapitel wichtigsten graduierten Algebren beruhen jedoch auf einem anderen Konzept:

Sei G eine (hier multiplikativ, wie üblich ohne Verknüpfungssymbol geschriebene) Gruppe und $\bullet$ eine weitere, zunächst ganz beliebige Verknüpfung auf der Trägermenge G. Unter einer natürlichen Filtrierung des Doppelmagmas G verstehen wir eine (bezüglich $\bullet$) subhomomorphe[12] Abbildung von $\mathbb{N}_{\geq s}$ auf eine Menge von Untergruppen der Gruppe G mit

$$G = G_s \geq G_{s+1} \geq G_{s+2} \geq \cdots \quad .$$

Für alle $k \in \mathbb{N}_{\geq s}$ gilt $1_G \in G_k$, also $G_n \bullet 1_G,\ 1_G \bullet G_n \subseteq G_{n+k}$. Es folgt:

$$\forall n \in \mathbb{N}_{\geq s} \quad G_n \bullet 1_G,\ 1_G \bullet G_n \subseteq \bigcap_{m\in\mathbb{N}_{\geq s}} G_m.$$

Wir nennen die Filtrierung $(G_n)_{n\in\mathbb{N}_{\geq s}}$ hausdorffsch, wenn $\bigcap_{n\in\mathbb{N}_{\geq s}} G_n = \{1_G\}$ gilt. Sei nun $(G_n)_{n\in\mathbb{N}_{\geq s}}$ eine Filtrierung von G, die für alle $m, n \in \mathbb{N}_{\geq s}$ die folgenden Bedingungen erfüllt:

(i) $G_{n+1} \trianglelefteq G_n$ und G_n/G_{n+1} ist abelsch,

(ii) $\forall g, g' \in G_m\ \forall h, h' \in G_n \quad \begin{array}{l} G_{m+1}gg' \bullet G_{n+1}h \subseteq G_{m+n+1}(g\bullet h)(g'\bullet h) \\ G_{m+1}g \bullet G_{n+1}hh' \subseteq G_{m+n+1}(g\bullet h)(g\bullet h') \end{array}$.

[12] Im Falle der trivialen Verknüpfung, $g \bullet h := 1_G$ für alle $g, h \in G$, ist *jede* Abbildung von $\mathbb{N}_{\geq s}$ auf eine Menge von Untergruppen von G subhomomorph, so daß dann die Filtrierungen von G durch die absteigenden Untergruppenketten mit Anfangsglied G gegeben sind. Dieser Trivialfall unserer Definition ist stets gemeint, wenn (in der Literatur) von einer zweiten Verknüpfung auf G gar nicht die Rede ist. Eine Filtrierung von G induziert eine Topologie auf der Menge G, erzeugt durch die (Rechts-)Restklassen von G_{n+1} in G_n $(n \in \mathbb{N}_{\geq s})$ als Topologie-Basis. Häufig werden Filtrierungen, in anderer Richtung als hier, zum Zweck topologischer Betrachtungen eingeführt und dann statt $\mathbb{N}_{\geq s}$ andere bzw. allgemeiner definierte „Indexmengen“ betrachtet. In [Bo2] II.4 etwa wird $\mathbb{R}$ (statt unserem $\mathbb{N}_{\geq s}$) gewählt und dann von einer „reellen Filtrierung“ gesprochen. Der Filtrierungsbegriff läßt sich ohne weiteres auch so fassen, daß die betrachtete Untergruppenkette kein erstes Glied besitzt, bei dem sie beginnt. Alle auf möglichst große *Allgemeinheit der Indexmenge* ausgerichteten Fassungen sind für unseren kombinatorisch-algebraischen Kontext ohne Belang, während die Abhängigkeit von der *Wahl einer Verknüpfung* $\bullet$ *auf* G hier fundamental ist. – Schließlich sei erwähnt, daß es dual zu dem Obigen in der Literatur auch den Begriff der Filtrierung für *aufsteigende* Untergruppenketten gibt.

Dann ist nach (i) das cartesische Produkt $\mathcal{P}((G_n/G_{n+1})_{n\in\mathbb{N}_{\geq s}})$ aller Faktorgruppen G_n/G_{n+1} bei komponentenweiser Verknüpfung eine abelsche Gruppe. Wir verwenden $\dot{+}$ als Verknüpfungssymbol dieser Gruppe. In additiver Sprechweise ist $\mathcal{P}((G_n/G_{n+1})_{n\in\mathbb{N}_{\geq s}})$ die direkte Summe der abelschen Gruppen G_n/G_{n+1} mit $n \in \mathbb{N}_{\geq s}$, und $\dot{+}$ setzt die auf den direkten Summanden G_n/G_{n+1} gegebenen Gruppenverknüpfungen auf $\mathcal{P}((G_n/G_{n+1})_{n\in\mathbb{N}_{\geq s}})$ kanonisch fort. Speziell gilt für alle $n \in \mathbb{N}_{\geq s}$, $g, h \in G_n$ also $G_{n+1}g\dot{+}G_{n+1}h = G_{n+1}gh$.

Aufgrund von (ii) ergibt die Setzung

$$\forall g \in G_m\ \forall h \in G_n \quad G_{m+1}g \circ G_{n+1}h := G_{m+n+1}(g \bullet h)$$

durch distributive Fortsetzung eine beide Distributivgesetze erfüllende Multiplikation auf der abelschen Trägergruppe $(\mathcal{P}((G_n/G_{n+1})_{n\in\mathbb{N}_{\geq s}}), \dot{+})$, d. h. das Doppelmagma $(\mathcal{P}((G_n/G_{n+1})_{n\in\mathbb{N}_{\geq s}}), \dot{+}, \circ)$ ist eine Algebra. Es gilt:

$$(*) \qquad \sum_{r\in\underline{k}} G_{m_r+1}g_r \circ \sum_{t\in\underline{l}} G_{n_t+1}h_t = \sum_{r\in\underline{k},\, t\in\underline{l}} G_{m_r+n_t+1}(g_r \bullet h_t),$$

wenn $k, l \in \mathbb{N}$, $m_1, \ldots, m_k, n_1, \ldots, n_l \in \mathbb{N}_{\geq s}$, $g_r \in G_{m_r}$, $h_t \in G_{n_t}$ für alle $r \in \underline{k}$, $t \in \underline{l}$. Nach Konstruktion ist die Zuordnung $n \mapsto G_n/G_{n+1}$ eine Graduierung. Es gilt daher

4.11.1 *Für jede Filtrierung $(G_n)_{n\in\mathbb{N}_{\geq s}}$ von G, die die Bedingungen* (i), (ii) *erfüllt, ist $(\mathcal{P}((G_n/G_{n+1})_{i\in\mathbb{N}_{\geq s}}), \dot{+}, \circ)$ eine graduierte Algebra mit den homogenen Komponenten G_n/G_{n+1} ($n \in \mathbb{N}_{\geq s}$).* □

Mit $(*)$ verifiziert man leicht

4.11.2 *Sei $(G_n)_{n\in\mathbb{N}_{\geq s}}$ eine Filtrierung von G, die die Bedingungen* (i), (ii) *erfüllt. Sei B eine Algebra und für alle $n \in \mathbb{N}_{\geq s}$ ein Gruppenhomomorphismus ψ_n von G_n/G_{n+1} in die additive Gruppe von B gegeben, so daß für alle $m, n \in \mathbb{N}_{\geq s}$ gilt:*

$$\forall g \in G_m\ \forall h \in G_n \quad (G_{m+1}g)\psi_m\,(G_{n+1}h)\psi_n = (G_{m+n+1}(g \bullet h))\psi_{m+n}.$$

Sei ψ der Homomorphismus der additiven Gruppe von $\mathcal{P}((G_n/G_{n+1})_{n\in\mathbb{N}_{\geq s}})$ in die additive Gruppe von B mit $\psi_{|G_n/G_{n+1}} = \psi_n$ für alle $n \in \mathbb{N}_{\geq s}$. Dann ist ψ ein Algebren-Homomorphismus. □

Geht man von einer Algebra A aus und wählt für $\bullet$ die Multiplikation von A, so erfüllt *jede* Filtrierung der additiven Gruppe von A die Bedingungen (i), (ii), führt also zu einer graduierten Algebra nach 4.11.1. Hat dabei A

bereits eine Graduierung $(A_n)_{n\in\mathbb{N}_{\geq s}}$ und setzen wir $\tilde{A}_n := \sum_{m\geq n} A_m$ für alle $n \in \mathbb{N}_{\geq s}$, so ist $(\tilde{A}_n)_{n\in\mathbb{N}_{\geq s}}$ eine Filtrierung der additiven Gruppe von A, und es gilt offensichtlich $\mathcal{P}((\tilde{A}_n/\tilde{A}_{n+1})_{n\in\mathbb{N}_{\geq s}}) \cong A$. Eine wesentlich interessantere Einsicht besteht darin, daß im Fall einer *assoziativen* Algebra (also ausgehend von einem Ring) jedes Ideal in natürlicher Weise eine Filtrierung hervorruft:

Sei J ein echtes Ideal eines unitären Ringes R, $J^0 := R$. Dann ist (bezüglich der Ringmultiplikation als Wahl für $\bullet$) $(J^n)_{n\in\mathbb{N}_0}$ eine Filtrierung der additiven Gruppe von R, die (i) und (ii) erfüllt. Sei $\mathcal{P}(J^{n\geq 0}) := \mathcal{P}((J^n/J^{n+1})_{n\in\mathbb{N}_0})$. Für jedes $n \in \mathbb{N}_0$ ist die additive Gruppe von J^n vermöge Linksmultiplikation (ebenso vermöge Rechtsmultiplikation) ein unitaler R-Modul, damit auch jeder Quotient J^n/J^{n+1} und weiter auch die additive Gruppe der Algebra $\mathcal{P}(J^{n\geq 0})$. Da J bei dieser Operation stets den Modul annulliert, ist nach 2.11(1) letzterer ein unitaler R/J-Modul.[13]

Sei nun K ein kommutativer unitärer Ring. Die Abbildung

$$KG \to K, \sum_g c_g g \mapsto \sum_g c_g$$

(wobei sich die Summation über endlich viele $g \in G$ erstreckt) ist ein ${}^K\mathfrak{R}_1$-Epimorphismus. Sei J dessen Kern; man nennt J das Augmentationsideal von KG. Es gilt: $J\dot{\oplus}K1_G = KG$.

4.11.3 *Sei G eine Gruppe, X ein $\mathfrak{G}$-Erzeugendensystem von G und J das Augmentationsideal von KG. Dann gilt:*

$$J = \langle G - 1_G\rangle_{K\mathfrak{M}} = \langle X - 1_G\rangle_{KG\mathfrak{M}}.$$

Denn aus der für beliebige $g_1, \ldots, g_n \in G$, $c_1, \ldots, c_n \in K$ geltenden Identität

$$c_1g_1 + \cdots + c_ng_n = \big(\sum_{i=1}^{n} c_i\big)1_G + c_1(g_1 - 1_G) + \cdots + c_n(g_n - 1_G)$$

folgt zunächst $J \subseteq \langle G-1_G\rangle_{K\mathfrak{M}}$. Da die Inklusion $\langle X-1_G\rangle_{KG\mathfrak{M}} \subseteq J$ trivial ist, bleibt für jedes $g \in G$ zu zeigen: $g-1_G \in \langle X-1_G\rangle_{KG\mathfrak{M}}$, was wir mittels Induktion nach der $\tilde{X}$-Länge von g einsehen: Sei $k = l_{\tilde{X}}(g)$ und $g = x_1^{\varepsilon_1}\cdots x_k^{\varepsilon_k}$ mit $x_i \in X$, $\varepsilon_i \in \{1, -1\}$. Sei $k > 0$, $h := x_2^{\varepsilon_2}\cdots x_k^{\varepsilon_k}$ und induktiv vorausgesetzt: $h - 1_G \in \langle X - 1_G\rangle_{KG\mathfrak{M}}$. Es gilt

$$m := x_1^{\varepsilon_1}(x_1^{-\varepsilon_1} - 1_G) = 1_G - x_1^{\varepsilon_1} \in \langle X - 1_G\rangle_{KG\mathfrak{M}},$$

[13]Da wir hier von der *Links*multiplikation ausgegangen sind, operiert R/J ebenso wie R vermöge eines *Anti*-Homomorphismus in den Endomorphismenring. In der Folge wird es jedoch stets um Fälle gehen, in denen die Faktoralgebra R/J kommutativ ist, so daß Anti-Homomorphismen und Homomorphismen von R/J dieselbe Bedeutung haben.

da $\varepsilon_1 \in \{1,-1\}$; also auch $g - 1_G = x_1^{\varepsilon_1}(h-1_G) - m \in \langle X - 1_G\rangle_{KG}\mathfrak{M}$. $\square$

Wir betrachten nun zu $R := KG$ die graduierte Algebra $\mathcal{P}(J^{n\geq 0})$, die aufgrund ihrer R/J-Modul-Struktur als K-Modul aufgefaßt werden kann vermöge der Setzung

$$c(J^{n+1}+a) := (J + c1_G) \circ (J^{n+1}+a) = J^{n+1} + ca$$

für alle $n \in \mathbb{N}_0$, $c \in K$, $a \in J^n$. Da offensichtlich das Komplement $K1_G$ von J in KG im Zentrum des Ringes KG liegt, ist $\mathcal{P}(J^{n\geq 0})$ bezüglich dieser K-Modul-Operation eine K-Algebra, und zwar assoziativ und unitär. Im Fall einer *freien* Gruppe erhalten wir das folgende grundlegende Resultat:

4.12 Satz *Sei $K \in \mathfrak{R}_1$, kommutativ, F die freie Gruppe über einer Menge X, $R := KF$ und J das Augmentationsideal von R. Dann gilt:*

(1) *J ist ein von $X - 1_F$ frei erzeugter R-Modul.*

(2) *$\mathcal{P}(J^{n\geq 0})$ ist eine von $\{J^2 + (x-1_F) | x \in X\}$ frei erzeugte assoziative unitäre K-Algebra.*

Beweis. (1) Wie 4.11.3 zeigt, müssen wir nur noch einsehen, daß $X - 1_F$ ${}^R\mathfrak{M}$-unabhängig ist. Dieses geschieht mit Hilfe der sog. Magnus-Einbettung von R in einen Ring von (2×2)-Matrizen, die als nächstes beschrieben wird:

Nach dem Entgiftungssatz (S. 6) gibt es eine zu X gleichmächtige und zu KX^* disjunkte Menge $\overline{X}$. Sei $\overline{\ }$ eine Bijektion von X auf $\overline{X}$ und M ein von $\overline{X}$ frei erzeugter R-Links-Modul (siehe 2.9). (Dann läßt sich M überdies auch noch trivialerweise als K-Rechtsmodul auffassen, indem man setzt: $mc := (c1_F)m$ für alle $m \in M$, $c \in K$.) Sei

$$A := \left\{ \begin{pmatrix} a & m \\ 0 & c \end{pmatrix} \,\middle|\, c \in K, a \in R, m \in M \right\}.$$

Unter Verwendung der Multiplikationen in K und R sowie der (doppelseitigen) Modul-Eigenschaft von M definiert die der gewöhnlichen Matrix-Multiplikation „nachempfundene“ Produktbildung

$$\begin{pmatrix} a & m \\ 0 & c \end{pmatrix}\begin{pmatrix} b & n \\ 0 & d \end{pmatrix} = \begin{pmatrix} ab & an+md \\ 0 & cd \end{pmatrix} \text{ für alle } a,b \in R, c,d \in K, m,n \in M$$

eine assoziative Verknüpfung auf A. Mit komponentenweiser Addition und jener Multiplikation ist A eine assoziative K-Algebra mit dem Einselement $\begin{pmatrix} 1_F & 0_M \\ 0 & 1_K \end{pmatrix}$. Sei $\varphi : X \to A,\ x \mapsto \begin{pmatrix} x & \overline{x} \\ 0 & 1_K \end{pmatrix}$. Es gilt:

$$\begin{pmatrix} x & \overline{x} \\ 0 & 1_K \end{pmatrix}\begin{pmatrix} x^{-1} & -x^{-1}\overline{x} \\ 0 & 1_K \end{pmatrix} = \begin{pmatrix} 1_F & 0_M \\ 0 & 1_K \end{pmatrix} = \begin{pmatrix} x^{-1} & -x^{-1}\overline{x} \\ 0 & 1_K \end{pmatrix}\begin{pmatrix} x & \overline{x} \\ 0 & 1_K \end{pmatrix}$$

für alle $x \in X$, so daß φ eine Abbildung von X in die Einheitengruppe von A ist. Da F von X frei erzeugt ist, ist φ zu einem $\mathfrak{G}$-Homomorphismus $\overline{\varphi}$ von F in die Einheitengruppe von A fortsetzbar. Die auf R definierte K-lineare Fortsetzung $\overline{\overline{\varphi}}$ von $\overline{\varphi}$ ist ein ${}^{K}\mathfrak{R}_1$-Monomorphismus, die „Magnus-Einbettung“ von R in A: Für alle $a \in R$ gilt für geeignete $c \in K$, $m \in M$: $a\overline{\overline{\varphi}} = \begin{pmatrix} a & m \\ 0 & c \end{pmatrix}$.

Sei nun T eine endliche Teilmenge von X und für jedes $x \in T$ ein Element $a_x \in R$ gegeben, so daß gilt: $\sum\limits_{x \in T} a_x(x - 1_F) = 0_R$. Es folgt:

$$\begin{aligned}\begin{pmatrix} 0_R & 0_M \\ 0 & 0_K \end{pmatrix} &= \Big(\sum_{x \in T} a_x(x - 1_F)\Big)\overline{\overline{\varphi}} = \sum_{x \in T} a_x\overline{\overline{\varphi}}(x\overline{\overline{\varphi}} - 1_F\overline{\overline{\varphi}}) \\ &= \sum_{x \in T} \begin{pmatrix} a_x & m_x \\ 0 & z_x \end{pmatrix}\begin{pmatrix} x - 1_F & \overline{x} \\ 0 & 0_K \end{pmatrix} = \sum_{x \in T} \begin{pmatrix} a_x(x - 1_F) & a_x\overline{x} \\ 0 & 0_K \end{pmatrix} \\ &= \begin{pmatrix} \sum_{x \in T} a_x(x - 1_F) & \sum_{x \in T} a_x\overline{x} \\ 0 & 0_K \end{pmatrix},\end{aligned}$$

also $\sum_{x \in T} a_x\overline{x} = 0_M$ und damit $a_x = 0_R$ für alle $x \in T$, da $\overline{X}$ ${}^{R}\mathfrak{M}$-unabhängig ist.

(2) Aus (1) folgt mit 2.8(1),(2) für jedes $n \in \mathbb{N}_0$, daß J^n ein freier R-Modul mit ${}^{R}\mathfrak{M}$-Basis $\{(x_1 - 1_F)\cdots(x_n - 1_F) | x_1 \cdots x_n \in X^n\}$ ist. Aus 2.11(4) erhalten wir nun:

$(*)$ Für jedes $n \in \mathbb{N}_0$ ist J^n/J^{n+1} ein freier K-Modul mit ${}^{K}\mathfrak{M}$-Basis $\{J^{n+1} + (x_1 - 1_F)\cdots(x_n - 1_F) | x_1, \ldots, x_n \in X\}$.

Sei

$$\psi : X \to \mathcal{P}(J^{n \geq 0}),\ x \mapsto J^2 + (x - 1_F)$$

und $\overline{\psi}$ die Fortsetzung von ψ zu einem ${}^{K}\mathfrak{R}_1$-Homomorphismus von KX^* in $\mathcal{P}(J^{n \geq 0})$. Wie die Definition der Multiplikation in $\mathcal{P}(J^{n \geq 0})$ zeigt, gilt dann $(x_1 \cdots x_n)\overline{\psi} = J^{n+1} + (x_1 - 1_F)\cdots(x_n - 1_F)$ für alle $x_1, \ldots, x_n \in X$. Wie $(*)$ lehrt, ist dann $\overline{\psi}|_{X^n}$ für jedes $n \in \mathbb{N}_0$ ein ${}^{K}\mathfrak{M}$-Isomorphismus von KX^n auf J^n/J^{n+1}, also $\overline{\psi}$ eine Bijektion von KX^* auf $\mathcal{P}(J^{n \geq 0})$. Insgesamt folgt: $KX^* \underset{{}^{K}\mathfrak{R}_1}{\cong} \mathcal{P}(J^{n \geq 0})$ vermöge $\overline{\psi}$. □

Wir werden noch eine ganz anders geartete Form von Filtrierung betrachten und benötigen dazu eine Reihe von Kommutatorregeln in Gruppen, die im folgenden zusammengestellt werden. Zunächst werden wir uns in einem rein gruppentheoretischen Kontext befinden, so daß eine Verwechslung mit

dem algebrentheoretischen Kommutator nicht zu befürchten ist. Für Gruppenelemente a, b bedeutet also im folgenden $[a, b]$ das Element $a^{-1}b^{-1}ab$. Für beliebige nichtleere Teilmengen S, T einer Gruppe setzen wir

$$[S, T] := \langle [a, b] | a \in S,\, b \in T \rangle_{\mathfrak{G}}.$$

Induktiv setzen wir für alle $n \in \mathbb{N}_{>1}$ und nichtleere Teilmengen $T_1, \ldots, T_n$ einer Gruppe $[T_1, \ldots, T_n] := [[T_1, \ldots, T_{n-1}], T_n]$, für Gruppenelemente $g_1, \ldots, g_n$ entsprechend $[g_1, \ldots, g_n] := [[g_1, \ldots, g_{n-1}], g_n]$.

4.12.1 *Sei G eine Gruppe. Für alle a, b, $c \in G$ gilt:*

(1) $[a, b]^{-1} = [b, a] = [a, b^{-1}]^b = [a^{-1}, b]^{a^b}$,

(2) $[ab, c] = [a, c]^b[b, c], \quad [a, bc] = [a, c][a, b]^c$,

(3) $[a, b]^c = [a^c, b^c]$.

(4) **(Witt-Identität)** $[a, b^{-1}, c]^b\, [b, c^{-1}, a]^c\, [c, a^{-1}, b]^a = 1_G$.

Folgerung *Für alle Untergruppen H_1, H_2 von G gilt: $[H_2, H_1] = [H_1, H_2] \trianglelefteq \langle H_1, H_2 \rangle_{\mathfrak{G}}$. Sind H_1, H_2 Normalteiler von G, so auch $[H_1, H_2]$.* □

Zum Beweis von (4) genügt es, einmal die Umformung

$$[a, b^{-1}, c]^b = ((a^{-1}bab^{-1})^{-1}c^{-1}a^{-1}bab^{-1}c)^b = a^{-1}b^{-1}ac^{-1}a^{-1} \cdot \underline{bab^{-1}cb}$$

vorzunehmen und diese dann zwei weitere Male anzuwenden (nämlich mit zyklischem Rollentausch von a, b und c), mit dem Ergebnis

$$[b, c^{-1}, a]^c = \underline{b^{-1}c^{-1}ba^{-1}b^{-1}} \cdot \underline{\underline{cbc^{-1}ac}},$$
$$[c, a^{-1}, b]^a = \underline{c^{-1}a^{-1}cb^{-1}c^{-1}} \cdot aca^{-1}ba.$$

Die in gleicher Weise markierten Terme sind jeweils zueinander invers, woraus die Witt-Identität folgt. □

Eine unmittelbare Konsequenz ist:

4.12.2 *Enthält ein Normalteiler der Gruppe G sowohl $[a, b^{-1}, c]$ als auch $[b, c^{-1}, a]$, so enthält er auch $[c, a^{-1}, b]$,*

denn $[c, a^{-1}, b] = (([b, c^{-1}, a]^c)^{-1}([a, b^{-1}, c]^b)^{-1})^{a^{-1}}$ nach 4.12.1(4). □

Hieraus in Verbindung mit 4.12.1(2) erhält man

Folgerung („Drei-Untergruppen-Lemma“) Sind H_1, H_2, H_3 Untergruppen von G mit $[H_1, H_2, H_3] = \{1_G\} = [H_2, H_3, H_1]$, so gilt auch $[H_3, H_1, H_2] = \{1_G\}$. □

4.12.3 *Sei G eine Gruppe, $N \trianglelefteq G$, $N \leq H \leq G$. Dann gilt:*

$$H/N \leq Z(G/N) \Leftrightarrow [H, G] \leq N,$$

denn $hN\, gN = gN\, hN \Leftrightarrow hg \in ghN \Leftrightarrow [h, g] \in N$ für beliebige $g, h \in G$. □

4.13 Definition Eine Folge $(G_n)_{n\in\mathbb{N}}$ von Untergruppen einer Gruppe G heißt eine absteigende Zentralkette von G, wenn gilt: $G = G_1$ und

$$\forall n \in \mathbb{N} \quad G_n \geq G_{n+1} \geq [G_n, G].$$

Induktiv folgt dann $G_n \trianglelefteq G$ für alle $n \in \mathbb{N}$: Ist $G_n \trianglelefteq G$ für ein $n \in \mathbb{N}$, so nach der Folgerung aus 4.12.1 auch $[G_n, G] \trianglelefteq G$, und nach 4.12.3 $G_{n+1}/[G_n, G] \leq Z(G/[G_n, G])$, insbesondere $G_{n+1} \trianglelefteq G$. Gilt $G = G_1$ und sogar

$$\forall n, j \in \mathbb{N} \quad G_n \geq G_{n+j} \geq [G_n, G_j],$$

so heißt die absteigende Zentralkette $(G_n)_{n\in\mathbb{N}}$ stark.

4.13.1 Beispiel *Sei $\gamma_1(G) := G$, $\gamma_{n+1}(G) := [\gamma_n(G), G]$ für alle $n \in \mathbb{N}$. Dann ist $(\gamma_n(G))_{n\in\mathbb{N}}$ offensichtlich eine absteigende Zentralkette von G. Wir zeigen, daß $(\gamma_n(G))_{n\in\mathbb{N}}$ stark ist:*

Für alle $j \in \mathbb{N}$ behaupten wir: $\forall n \in \mathbb{N} \quad [\gamma_n(G), \gamma_j(G)] \subseteq \gamma_{n+j}(G)$. Dies ist für $j = 1$ trivial. Sei nun $j > 1$ und für $j-1$ statt j die Behauptung als wahr angenommen. Dann gilt für alle $n \in \mathbb{N}$:

$$[G, \gamma_n(G), \gamma_{j-1}(G)] = [\gamma_{n+1}(G), \gamma_{j-1}(G)] \subseteq \gamma_{(n+1)+(j-1)}(G) = \gamma_{n+j}(G),$$
$$[\gamma_n(G), \gamma_{j-1}(G), G] \subseteq [\gamma_{n+j-1}(G), G] = \gamma_{n+j}(G),$$

nach der Folgerung aus 4.12.2 also auch

$$[\gamma_n(G), \gamma_j(G)] = [\gamma_j(G), \gamma_n(G)] = [\gamma_{j-1}(G), G, \gamma_n(G)] \subseteq \gamma_{n+j}(G).$$

Die Folge $(\gamma_n(G))_{n\in\mathbb{N}}$ heißt die absteigende Zentralreihe von G.

4.13.2 *Ist $(G_n)_{n\in\mathbb{N}}$ eine absteigende Zentralkette von G, so gilt $\gamma_n(G) \subseteq G_n$ für alle $n \in \mathbb{N}$.*

Denn es gilt $\gamma_1(G) = G = G_1$, und für beliebiges $n \in \mathbb{N}$ mit $\gamma_n(G) \subseteq G_n$ gilt auch $\gamma_{n+1}(G) = [\gamma_n(G), G] \leq [G_n, G] \leq G_{n+1}$. □

Die Gruppe G heißt nilpotent, wenn es ein $m \in \mathbb{N}$ gibt mit $\gamma_m(G) = \{1_G\}$. Nach 4.13.2 ist dies äquivalent dazu, daß es eine absteigende Zentralkette

$(G_n)_{n\in\mathbb{N}}$ von G gibt, bei der $G_k = \{1_G\}$ für ein $k \in \mathbb{N}$ gilt. Für jede Gruppe G gilt nach 4.13.2

$$\bigcap_{n\in\mathbb{N}} \gamma_n(G) = \bigcap_{\substack{N\trianglelefteq G \\ G/N \text{ nilpotent}}} N \ .$$

Dieser Durchschnitt heißt das Nilpotenz-Residuum von G. Ist dieses gleich $\{1_G\}$, so heißt G residuell nilpotent.

Analog heißt eine assoziative Algebra A nilpotent, wenn es ein $m \in \mathbb{N}$ gibt mit $A^m = \{0_A\}$. Sie heißt residuell nilpotent, wenn gilt: $\bigcap_{n\in\mathbb{N}} A^n = \{0_A\}$. Wir bemerken:

4.13.3 *Untergruppen (residuell) nilpotenter Gruppen bzw. Teilalgebren (residuell) nilpotenter Algebren sind (residuell) nilpotent.* □

4.13.4 *Jede starke absteigende Zentralkette der Gruppe G ist eine Filtrierung von G bezüglich der Verknüpfung $[.,.]$, die die Bedingungen* (i), (ii) *in* 4.11 *erfüllt.*[14]

Beweis. Ist $(G_n)_{n\in\mathbb{N}}$ eine starke absteigende Zentralkette von G und sind $m, n \in \mathbb{N}$, $g, g' \in G_m$, $h, h' \in G_n$, so gilt $[G_m, G_n] \leq G_{m+n}$ und nach 4.12.1(2)

$$[G_{m+1}gg', G_{n+1}h] \subseteq G_{m+n+1}[gg', h] = G_{m+n+1}[g, h]^{g'}[g', h] = G_{m+n+1}[g, h][g', h],$$

vermöge 4.12.3; ähnlich $[G_{m+1}g, G_{n+1}hh'] \subseteq G_{m+n+1}[g, h][g, h']$. □

Gemäß 4.11.1 erhalten wir somit zu jeder starken absteigenden Zentralkette $(G_n)_{n\in\mathbb{N}}$ von G eine graduierte Algebra $(\mathcal{P}((G_n)_{n\in\mathbb{N}}), \dot{+}, \circ)$ mit

$$G_{m+1}g \circ G_{n+1}h = G_{m+n+1}[g, h] \quad \text{für alle } m, n \in \mathbb{N}, g \in G_m, h \in G_n.$$

4.14 Proposition *Sei G eine Gruppe und $(G_n)_{n\in\mathbb{N}}$ eine starke absteigende Zentralkette von G. Dann gilt:*

(1) *$(\mathcal{P}((G_n)_{n\in\mathbb{N}}), \dot{+}, \circ) \in \mathcal{L}$, und es gibt einen $\mathcal{L}$-Homomorphismus*

$$\acute{\alpha} : \mathcal{P}((\gamma_n)_{n\in\mathbb{N}}) \to \mathcal{P}((G_n)_{n\in\mathbb{N}}) \text{ mit } \gamma_{n+1}(G)g \mapsto G_{n+1}g \text{ für alle } g \in \gamma_n(G).$$

(2) *Ist X ein $\mathfrak{G}$-Erzeugendensystem von G, so ist $\{G'x | x \in X\}$ ein $\mathcal{L}$-Erzeugendensystem von $\mathcal{P}((\gamma_n)_{n\in\mathbb{N}})$.*

[14] Offenbar ist es nur ein Wechsel der Ausdrucksweise, ob man sagt, daß die Filtrierung $(\gamma_n(G))_{n\in\mathbb{N}}$ hausdorffsch oder daß G residuell nilpotent sei.

Beweis. (1) Seien $r, n_1, \ldots, n_r \in \mathbb{N}$ und $g_j \in G_{n_j}$ für alle $j \in \underline{r}$. Es gilt $G_{n_j+1}g_j \circ G_{n_j+1}g_j = G_{2n_j+1}[g_j, g_j] = G_{2n_j+1}$, also

$$\begin{aligned}\sum_{j\in\underline{r}} G_{n_j+1}g_j \circ \sum_{j\in\underline{r}} G_{n_j+1}g_i &= \sum_{i<j}(G_{n_i+1}g_i \circ G_{n_j+1}g_j \dot{+} G_{n_j+1}g_j \circ G_{n_i+1}g_i)\\ &= \sum_{i<j} G_{n_i+n_j+1}[g_i, g_j][g_j, g_i] = 0_{\mathcal{P}((G_n)_{n\in\mathbb{N}})},\end{aligned}$$

nach 4.12.1(1). Also gilt in $\mathcal{P}((G_n)_{n\in\mathbb{N}})$ das Antikommutativgesetz. Zum Nachweis der Jacobi-Identität seien $m_1, \ldots, m_q, t_1, \ldots, t_s \in \mathbb{N}$, $f_i \in G_{m_i}$ für $i \in \underline{q}$, $h_k \in G_{t_k}$ für $k \in \underline{s}$. Es gilt

$$\sum_{i\in\underline{q}} G_{m_i+1}f_i \circ \sum_{j\in\underline{r}} G_{n_j+1}g_j \circ \sum_{k\in\underline{s}} G_{t_k+1}h_k = \sum_{i,j,k} G_{m_i+n_j+t_k+1}[f_i, g_j, h_k]$$

und $[f_i, g_j, h_k]^{-1} \equiv [[f_i, g_j]^{-1}, h_k] \equiv [f_i, g_j^{-1}, h_k] \mod G_{m_i+n_j+t_k+1}$, wie man aus 4.12.1(1) und 4.12.3 erschließt. Vermöge der Witt-Identität 4.12.1(4) folgt

$$[f_i, g_j, h_k][g_j, h_k, f_i][h_k, f_i, g_j] \in G_{m_i+n_j+t_k+1} \text{ für alle } i \in \underline{q}, j \in \underline{r}, k \in \underline{s}$$

und damit

$$\begin{aligned}\sum_{i\in\underline{q}} G_{m_i+1}f_i \circ \sum_{j\in\underline{r}} G_{n_j+1}g_j \circ \sum_{k\in\underline{s}} G_{t_k+1}h_k\\ \dot{+}\sum_{j\in\underline{r}} G_{n_j+1}g_j \circ \sum_{k\in\underline{s}} G_{t_k+1}h_k \circ \sum_{i\in\underline{q}} G_{m_i+1}f_i\\ \dot{+}\sum_{k\in\underline{s}} G_{t_k+1}h_k \circ \sum_{i\in\underline{q}} G_{m_i+1}f_i \circ \sum_{j\in\underline{r}} G_{n_j+1}g_j = 0_{\mathcal{P}((G_n)_{n\in\mathbb{N}})}.\end{aligned}$$

Für jedes $i \in \mathbb{N}$ ist die (nach 4.13.2 wohldefinierte) Abbildung

$$\alpha_n : \gamma_n(G)/\gamma_{n+1}(G) \to G_n/G_{n+1}, \quad \gamma_{n+1}(G)g \mapsto G_{n+1}g \quad (\text{für } g \in \gamma_n(G))$$

ein Gruppenhomomorphismus. Bezeichnen wir die Algebrenmultiplikation in $\mathcal{P}((G_n)_{n\in\mathbb{N}})$ (zur Unterscheidung von der in $\mathcal{P}((\gamma_n)_{n\in\mathbb{N}})$) mit ó, so gilt für alle $g \in \gamma_m(G)$, $h \in \gamma_n(G)$

$$\begin{aligned}(\gamma_{m+1}(G)g)\alpha_m \acute{o} (\gamma_{n+1}(G)h)\alpha_n &= [G_{m+1}g, G_{n+1}h]\\ &= G_{m+n+1}[g, h] = (\gamma_{m+n+1}(G)[g, h])\alpha_{m+n},\end{aligned}$$

also ist der Homomorphismus $\acute{\alpha}$ der additiven Gruppe von $\mathcal{P}((\gamma_n)_{n\in\mathbb{N}})$ in die additive Gruppe von $\mathcal{P}((G_n)_{n\in\mathbb{N}})$ mit $\acute{\alpha}_{|\gamma_n(G)/\gamma_{n+1}(G)} = \alpha_n$ für alle $n \in \mathbb{N}$ ein

Algebren-Homomorphismus (4.11.2).

(2) Für alle $n \in \mathbb{N}$ sei $X(n) := \{[x_1, \ldots, x_n] | x_i \in X\}$. Wir zeigen induktiv für alle $n \in \mathbb{N}$:

$$\gamma_n(G)/\gamma_{n+1}(G) = \langle X(n)\rangle_{\mathfrak{G}}\gamma_{n+1}(G)/\gamma_{n+1}(G).$$

Für $n = 1$ ist dies trivial. Sei $n > 1$, und es gelte die Behauptung für $n-1$ statt n. Wegen $\gamma_n(G) = [\gamma_{n-1}(G), G] = \langle [a, g] | a \in \gamma_{n-1}(G),\ g \in G\rangle_{\mathfrak{G}}$ genügt es, für beliebiges $a \in \gamma_{n-1}(G)$, $g \in G$ zu zeigen: $[a, g] \in \langle X(n)\rangle_{\mathfrak{G}}\gamma_{n+1}(G)$. Seien $y_1, \ldots, y_r \in X(n-1)$, $\varepsilon_i \in \{1, -1\}$, $b \in \gamma_n(G)$ mit $a = y_1^{\varepsilon_1} \cdots y_r^{\varepsilon_r} b$, und seien $x_1, \ldots, x_s \in X$, $\delta_j \in \{1, -1\}$ mit $g = x_1^{\delta_1} \cdots x_s^{\delta_s}$. Dann gilt modulo $\gamma_{n+1}(G)$ unter Benutzung von 4.12.1(1),(2)

$$[a, g] \equiv [y_1^{\varepsilon_1} \cdots y_r^{\varepsilon_r}, x_1^{\delta_1} \cdots x_s^{\delta_s}] \equiv \prod_{i,j} [y_i^{\varepsilon_i}, x_j^{\delta_j}] \equiv \prod_{i,j} [y_i, x_j]^{\varepsilon_i \delta_j} \in \langle X(n)\rangle_{\mathfrak{G}}.$$

Die Behauptung folgt. □

4.15 Proposition *Sei R ein unitärer Ring, $J \lhd R$ und G eine Untergruppe des (multiplikativen) Monoids $1_R + J$, $J^n(G) := (1_R + J^n) \cap G$ für alle $n \in \mathbb{N}$. Dann ist $(J^n(G))_{n\in\mathbb{N}}$ eine starke absteigende Zentralkette von G, und es gibt einen $\mathfrak{L}$-Monomorphismus*

$$\rho : \mathcal{P}((J^n(G))_{n\in\mathbb{N}}) \to \mathcal{P}(J^{n\geq 0}) \text{ mit } J^{n+1}(G)g \mapsto J^{n+1} + g - 1_G$$

für alle $n \in \mathbb{N}$ und $g \in J^n(G)$.

Beweis. Seien $m, n \in \mathbb{N}$, $g \in J^m(G)$, $h \in J^n(G)$, $u := g - 1_R$, $v := h - 1_R$. Dann gilt $u \in J^m$, $v \in J^n$ und $gh - hg = uv - vu \in J^{m+n}$, nach 4.0.1 also $g^{-1}h^{-1}gh \in J^{m+n}(G)$. Es folgt

$$J^m(G) = (1_R + J^m) \cap G \geq (1_R + J^{m+n}) \cap G = J^{m+n}(G) \geq [J^m(G), J^n(G)].$$

Für alle g, $g' \in G$ gilt die Äquivalenzenkette

$$\begin{aligned} &J^{n+1}(G)g = J^{n+1}(G)g' \Leftrightarrow \exists v \in J^{n+1} \quad g' = (1_G + v)g \\ &\quad \Leftrightarrow \exists v \in J^{n+1} \quad g' - 1_G = g - 1_G + vg \Leftrightarrow J^{n+1} + g - 1_G = J^{n+1} + g' - 1_G. \end{aligned}$$

Insbesondere erhalten wir für jedes $n \in \mathbb{N}$ durch

$$\rho_n : J^n(G)/J^{n+1}(G) \to J^n/J^{n+1},\ J^{n+1}(G)g \mapsto J^{n+1} + g - 1_G \ \ (g \in J^n(G)).$$

eine injektive Abbildung. Für alle g, $g' \in J^n(G)$ gilt $gg' - g - g' + 1_G = (g - 1_G)(g' - 1_G) \in J^{2n}$, also $gg' - 1_G \equiv (g - 1_G) + (g' - 1_G) \mod J^{2n}$, somit

$$\begin{aligned} &(J^{n+1}(G)g\, J^{n+1}(G)g')\rho_n = (J^{n+1}(G)gg')\rho_n = J^{n+1} + gg' - 1_G \\ &\quad = (J^{n+1} + g - 1_G) + (J^{n+1} + g' - 1_G) = (J^{n+1}(G)g)\rho_n + (J^{n+1}(G)g')\rho_n, \end{aligned}$$

d. h. die Abbildungen ρ_n sind Monomorphismen abelscher Gruppen. Sei nun ρ der $\mathfrak{M}$-Homomorphismus von $\mathcal{P}((J^n(G))_{n\in\mathbb{N}})$ in $\mathcal{P}(J^{n\geq 0})$ mit $\rho_{|J^n(G)/J^{n+1}(G)} = \rho_n$ für alle $n \in \mathbb{N}$. Da $\mathcal{P}(J^{n\geq 0})$ direkte Summe der Gruppen J^n/J^{n+1} ist, ist ρ injektiv. Wir zeigen, daß ρ ein $\mathfrak{L}$-Homomorphismus ist. Zur Unterscheidung bedeute $[.,.]_{\mathfrak{L}}$ den algebrentheoretischen, $[.,.]_{\mathfrak{G}}$ den gruppentheoretischen Kommutator. Sind $g \in J^m(G)$, $h \in J^n(G)$, so folgt:

$$[g,h]_{\mathfrak{L}} = (g-1_G)(h-1_G) - (h-1_G)(g-1_G) \in J^{m+n},$$

also

$$\begin{aligned}[g,h]_{\mathfrak{G}} - 1_G &= g^{-1}h^{-1}[g,h]_{\mathfrak{L}} \\ &= (g^{-1}h^{-1} - 1_G)[g,h]_{\mathfrak{L}} + [g,h]_{\mathfrak{L}} \in J^{m+n+1} + [g,h]_{\mathfrak{L}},\end{aligned}$$

und damit

$$\begin{aligned}[(J^{m+1}(G)g)\rho, (J^{n+1}(G)h)\rho]_{\mathfrak{L}} &= [J^{m+1} + (g-1_G), J^{n+1} + (h-1_G)]_{\mathfrak{L}} \\ &= J^{m+n+1} + [g,h]_{\mathfrak{L}} = J^{m+n+1} + ([g,h]_{\mathfrak{G}} - 1_G) = (J^{m+n+1}[g,h]_{\mathfrak{G}})\rho.\end{aligned}$$

Mit 4.11.2 folgt die Behauptung. □

Ist speziell K ein kommutativer unitärer Ring und J_K das Augmentationsideal von $(R =)KG$, so heißt $J_K^n(G)$ die n-te Dimensionsuntergruppe von G zu K. Ohne Nennung eines kommutativen unitären Ringes K ist stets $\mathbb{Z}$ gemeint, und man setzt üblicherweise $D_n(G) := J_{\mathbb{Z}}^n(G)$ für alle $n \in \mathbb{N}$.

Sei X eine beliebige Menge und J das Ideal von $\mathbb{Z}\langle\langle X\rangle\rangle$, das aus den Potenzreihen mit absolutem Glied 0 besteht (siehe 2.17.1). Offensichtlich ist J residuell nilpotent und nach 4.15 daher auch die Gruppe $1+J$ (siehe 2.17.3). Aus 2.18 folgt, daß die freie Gruppe über X zu der Untergruppe $F := \langle 1+X\rangle_{\mathfrak{G}}$ von $1+J$ isomorph ist.[15] Daher gilt mit 4.13.3:

4.15.1 (Magnus 1935) *Jede freie Gruppe ist residuell nilpotent.* □

Sei nun X eine Menge, $L(X)$ die freie ($\mathbb{Z}$-)Lie-Algebra über X und G eine von X $\mathfrak{G}$-erzeugte Gruppe. Dann gibt es nach 4.14(1) einen (nach 4.14(2) surjektiven) $\mathfrak{L}$-Homomorphismus

$$\chi : L(X) \to \mathcal{P}((\gamma_n(G))_{n\in\mathbb{N}}) \text{ mit } x\chi = G'x \text{ für alle } x \in X.$$

[15] Setzen wir $\mathcal{D}_n(F) := (1+J^n) \cap F$ für alle $n \in \mathbb{N}$, so ist $(\mathcal{D}_n(F))_{n\in\mathbb{N}}$ nach 4.15 eine starke absteigende Zentralkette der freien Gruppe F. Es läßt sich jedoch zeigen, daß für alle $n \in \mathbb{N}$ gilt: $\mathcal{D}_n(F) = D_n(F)$.

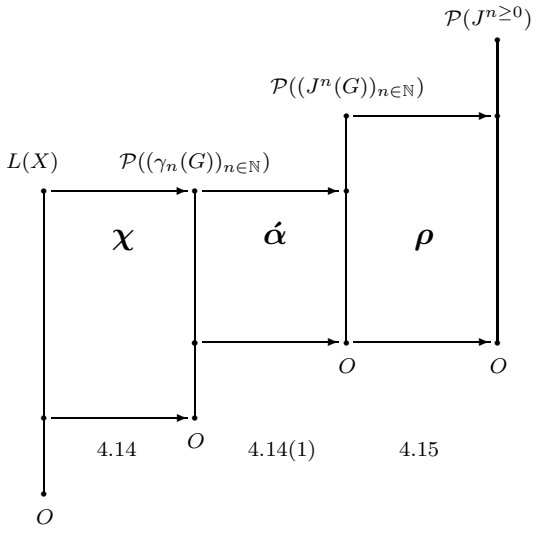
Ist R ein unitärer Ring, $J \lhd R$, G Untergruppe von $1_R + J$ und setzen wir $J^n(G) := (1_R + J^n) \cap G$ für alle $n \in \mathbb{N}$, so erhalten wir aus 4.14(1) einen $\mathfrak{L}$-Homomorphismus

$$\acute{\alpha} : \mathcal{P}((\gamma_n(G))_{n\in\mathbb{N}}) \to \mathcal{P}((J^n(G))_{n\in\mathbb{N}})$$
$$\text{mit } \gamma_{n+1}(G)g \mapsto J^{n+1}(G)g \text{ für alle } g \in \gamma_n(G)$$

und aus 4.15 einen $\mathfrak{L}$-Monomorphismus

$$\rho : \mathcal{P}((J^n(G))_{n\in\mathbb{N}}) \to \mathcal{P}((J^n)_{n\in\mathbb{N}_0})$$
$$\text{mit } J^{n+1}(G)g \mapsto J^{n+1} + g - 1_G \text{ für alle } g \in J^n(G).$$

Für alle $x \in X$ erhalten wir: $x\chi\acute{\alpha}\rho = J^2 + x - 1_G$.



Wir betrachten nun den Spezialfall des Augmentationsideals J im ganzzahligen Gruppenring $\mathbb{Z}F$ der freien Gruppe F über X. Es ist $J^n(F) = D_n(F)$ die n-te Dimensionsuntergruppe von F. Nach 4.12(2) gibt es einen $\mathfrak{R}$-Isomorphismus

$$\iota : \mathcal{P}((J^n)_{n\in\mathbb{N}_0}) \to \mathbb{Z}X^* \text{ mit } J^2 + x - 1_F \mapsto x \text{ für alle } x \in X.$$

Nach 4.10(4) ist $L_{X,\mathbb{Z}}$ eine von X frei erzeugte ($\mathbb{Z}$)-Lie-Algebra. Also gibt es einen $\mathfrak{L}$-Isomorphismus

$$\varsigma : L_{X,\mathbb{Z}} \to L(X) \text{ mit } x \mapsto x \text{ für alle } x \in X.$$

Also ist die Hintereinanderausführung $\chi\acute{\alpha}\rho\iota\varsigma$ ein $\mathfrak{L}$-Endomorphismus von $L(X)$ mit $x \mapsto x$ für alle $x \in X$. Es folgt: $\chi\acute{\alpha}\rho\iota\varsigma = \mathrm{id}_{L(X)}$. Dies hat unmittelbar eine wichtige Konsequenz:

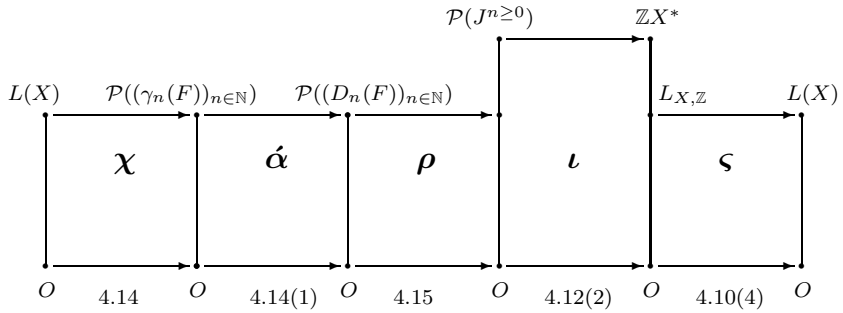
4.15.2 *Im Fall der freien Gruppe über X (und der Wahl $K = \mathbb{Z}$) sind die Homomorphismen χ und $\acute{\alpha}$ injektiv.*

4.16 Hauptsatz *Sei X eine Menge und F die freie Gruppe über X. Dann gilt:*

(1) *$\mathcal{P}((\gamma_n(F))_{n\in\mathbb{N}})$ ist eine von $\{F'x|x \in X\}$ frei erzeugte $\mathbb{Z}$-Lie-Algebra. Ihre homogenen Komponenten sind die Moduln $\gamma_n(F)/\gamma_{n+1}(F)$ $(n \in \mathbb{N})$. Insbesondere ist $\gamma_n(F)/\gamma_{n+1}(F)$ für jedes $n \in \mathbb{N}$ eine freie abelsche Gruppe. Ist X endlich, so ist ihr Rang $\frac{1}{n}\sum_{d|n}\mu(d)|X|^{\frac{n}{d}}$.*

(2) *Sei $\overline{}$ der Homomorphismus von $X^{(*)}$ in die Struktur $(F, [.,.])$ mit $\overline{\overline{x}} = x$ für alle $x \in X$. Bezüglich einer beliebigen Ordnung auf X sei $[w]_{\mathfrak{G}} := \overline{\overline{(w)}}$ für alle $w \in \mathfrak{L}^X$ und $(.)$ wie in 4.4. Dann ist $[\mathfrak{L}^X \cap X^n]_{\mathfrak{G}}$ ein Repräsentantensystem einer $\mathfrak{M}$-Basis von $\gamma_n(F)/\gamma_{n+1}(F)$, für alle $n \in \mathbb{N}$.*

(3) **Dimensionsuntergruppensatz (Magnus, Grün 1937)** *Für alle $n \in \mathbb{N}$ gilt: $\gamma_n(F) = D_n(F)$.*



Beweis. (1) Nach 4.15.2 ist χ injektiv, also ein $\mathfrak{L}$-Isomorphismus von $L(X)$ auf $\mathcal{P}((\gamma_n(F))_{n\in\mathbb{N}})$. Er bildet X auf $\{F'x|x \in X\}$ und die homogene Komponente zum Grad n von $L(X)$ auf $\gamma_n(F)/\gamma_{n+1}(F)$ ab. Die Rangaussage folgt aus 4.10(5).

(2) folgt direkt aus (1) und 4.10(2).

(3) Für $n = 1$ ist die Behauptung trivial. Gilt $\gamma_n(F) = D_n(F)$ für ein $n \in \mathbb{N}$ und ist $g \in D_{n+1}(F)$, so jedenfalls $g \in \gamma_n(F)$ und $(\gamma_{n+1}(F)g)\acute{\alpha} = D_{n+1}(F)g = 0_{\mathcal{P}((D_n(F))_{n\in\mathbb{N}})}$, also $g \in \gamma_{n+1}(F)$ nach 4.15.2. □

Wir schließen mit einigen Kommentaren:

Sei $Y = \{x, y\}$ mit $x \neq y$ und $L(Y)$ die freie $\mathbb{Z}$-Lie-Algebra über Y. Sei $L := \langle 2x, y, x \circ y\rangle_{\mathfrak{L}}$. Dann gilt: $x \circ y \notin L \circ L$ und $2(x \circ y) \in L \circ L$; also enthält die additive Gruppe $L/L \circ L$ ein Element der Ordnung 2. Gäbe es

eine $\mathfrak{L}$-Basis X von L, so gälte $L \underset{\mathfrak{L}}{\cong} \mathcal{P}((\gamma_n(F))_{n\in\mathbb{N}})$ nach 4.16(1) (wobei F die freie Gruppe über X ist), und es wäre $L/L\circ L \underset{\mathfrak{M}}{\cong} F/F'$ eine freie abelsche Gruppe, ein Widerspruch, da $L/L\circ L$ nicht torsionsfrei ist. Also besitzt L keine $\mathfrak{L}$-Basis. Das Beispiel zeigt, daß Teilalgebren einer freien Lie-Algebra (im Falle des Skalarbereichs $\mathbb{Z}$) nicht notwendig frei sein müssen. Umso bemerkenswerter ist, daß das Analogon von 3.11 für Lie-Algebren tatsächlich gilt, wenn der Skalarbereich ein Körper ist:

Satz (Shirshov 1953) *Sei K ein Körper. Dann ist jede K-Teilalgebra einer freien K-Lie-Algebra frei.*

Nach 4.10(4) darf man dabei von einer Teilalgebra L der Algebra $L_{X,K}$ der Lie-Elemente von KX^* ausgehen, deren Freiheit dann nachzuweisen ist. Für alle $n \in \mathbb{N}_0$ sei V_n der K-Teilraum von L, der aus 0_L und den Elementen vom Grad $\leq n$ von L besteht, weiter $W_n := V_n \cap \langle V_{n-1}\rangle_{\mathfrak{L}}$ für alle $n \in \mathbb{N}$. Ist dann $Y_n \subseteq L$ ein Repräsentantensystem einer K-Basis von V_n/W_n, so ist $\bigcup_{n\in\mathbb{N}} Y_n$ eine $\mathfrak{L}$-Basis von L, wie sich mit Hilfe des Satzes von Poincaré, Birkhoff, Witt (siehe S. 89) einsehen läßt ([Reu] 2.2).

Es ist sehr reizvoll, Analoga zu aus der Theorie der freien Gruppen bekannten Aussagen in der Theorie der freien Lie-Algebren auf ihre Gültigkeit zu prüfen, insbesondere wenn sich dabei nichttriviale *kontrastierende* Einsichten ergeben. Das ist zum Beispiel der Fall bezüglich der Aussage 3.14.1: Ist K ein Körper, so hat nämlich *jedes* nichttriviale Ideal einer freien K-Lie-Algebra eine unendliche $\mathfrak{L}$-Basis. (Einen Beweis dazu findet man z.B. in [Bah], 2.4.)

Nach 4.16(2) ist $B := [\mathfrak{L}^X]$ eine aus (höheren) Kommutatoren bestehende, durch Übertragung der Ordnung von $\mathfrak{L}^X$ vollständig geordnete Teilmenge von F mit folgender Eigenschaft:

Zu jedem $g \in F$ gibt es genau eine Abbildung ζ von B in $\mathbb{Z}$, so daß für jedes $n \in \mathbb{N}$ gilt:

$$g \in \gamma_{n+1}(F) b_1^{\zeta(b_1)} \cdots b_{l(n,g)}^{\zeta(b_{l(n,g)})}, \quad b_1 > b_2 > \cdots,$$

wobei $\{b_1, \ldots, b_{l(n,g)}\} \subseteq B \setminus \gamma_{n+1}(F)$, $\zeta(b_j) \neq 0$. Die Existenz einer solchen Menge B ist die Aussage des sog. *Hall'schen Basissatzes* (P. Hall 1933). Allgemeiner läßt sich für B das Bild $\overline{\overline{H}}$ eines *beliebigen* Hallgerüstes H unter $\overline{\overline{\ }}$ wählen. Da wir nach 4.6 bereits wissen, daß $\overline{\overline{H}}$ ein $\mathfrak{M}$-Erzeugendensystem von $\mathcal{P}((\gamma_n(F))_{n\in\mathbb{N}})$ sein muß, fehlt uns dazu nur noch die Einsicht, daß $\overline{\overline{H}}$ eine $\mathfrak{M}$-unabhängige Teilmenge von $\mathcal{P}((\gamma_n(F))_{n\in\mathbb{N}})$ ist. Gäbe es ein Hallgerüst H, bei dem $\overline{\overline{H}}$ $\mathfrak{M}$-abhängig wäre, so gäbe es offenbar auch ein solches für eine *endliche* Menge X, und für ein geeignetes $n \in \mathbb{N}$ wäre bereits $\overline{\overline{H \cap X^{(n)}}}$ $\mathfrak{M}$-abhängig. Vermöge 2.12(3) brauchen wir daher nur zu zeigen,

daß $|H \cap X^{(n)}| = |\mathfrak{L}^X \cap X^n|$ gilt. In der Tat läßt sich eine Bijektion von $H \cap X^{(n)}$ auf die Menge der Konjugiertenklassen primitiver Worte der Länge n über X (mit etwas Aufwand) direkt angeben. Unter Verwendung des hier nicht bewiesenen Satzes von Viennot (S. 81) ist man jedoch ohne dieses sofort am Ziel.

Nach 4.13.2 gilt $\gamma_n(G) \subseteq D_n(G)$ für *jede* Gruppe G. Jahrzehnte bestand die Vermutung, daß hier stets – wie im Falle einer *freien* Gruppe (4.16(3)) – Gleichheit gelte. Im Jahr 1972 wurde jedoch von E. Rips ein aufsehenerregendes Beispiel einer Gruppe G (von 2-Potenz-Ordnung) konstruiert mit $\gamma_4(G) = \{1_G\}$, $|D_4(G)| = 2$. Damit erhob sich die Frage nach einer Strukturbeschreibung von $D_n(G)/\gamma_n(G)$. Ein Hauptresultat in dieser Richtung ist der folgende

Satz (Sjögren 1979) *Für alle $n \in \mathbb{N}$ sei $s_n := \prod_{k \in \underline{n-2}} (kgV(\underline{k}))^{\binom{n-2}{k}}$. Ist G eine Gruppe und $g \in D_n(G)$, so gilt: $g^{s_n} \in \gamma_n(G)$. Insbesondere ist $D_n(G)/\gamma_n(G)$ eine Gruppe von endlichem Exponenten.*

Wegen $s_1 = s_2 = s_3 = 1$ gilt also $\gamma_n(G) = D_n(G)$ für $n \leq 3$ bei jeder Gruppe G, d.h. der Fall $n = 4$ ist der kleinste, bei dem eine Verschiedenheit von $\gamma_n(G)$ und $D_n(G)$ auftreten kann – was nach Rips' Beispiel auch tatsächlich vorkommt. Es gilt: $s_5 = 48$. K.-I. Tahara hat jedoch im Jahr 1981 bewiesen, daß für $g \in D_5(G)$ bereits g^6 in $\gamma_5(G)$ liegt; dies zeigt, daß die im Satz von Sjögren angegebene Zahl s_n keineswegs bestmöglich ist. Der Satz von Sjögren liefert u.a. die folgenden unmittelbaren Korollare:

Für jede Gruppe G gilt $\gamma_3(G) = D_3(G)$ (Higman/Rees), *$g^2 \in \gamma_4(G)$ für alle $g \in D_4(G)$* (Losey).

Ist G eine Gruppe, für die $\gamma_n(G)/\gamma_{n+1}(G)$ für jedes $n \in \mathbb{N}$ torsionsfrei ist, so folgt $\gamma_n(G) = D_n(G)$ für jedes $n \in \mathbb{N}$ (Hall/Jennings).

Für endliche Gruppen G sind in den zurückliegenden Jahren erhebliche Fortschritte im Verständnis der Struktur von $D_n(G)/\gamma_n(G)$ gemacht worden: Es genügt offensichtlich, G als nilpotent anzunehmen. Eine weitere Reduktion führt auf den Fall einer Gruppe G von p-Potenz-Ordnung für eine Primzahl p. Nach N. Gupta [Gup2] gilt für $p \neq 2$ dann stets $\gamma_n(G) = D_n(G)$, also die Dimensions-Untergruppen-Vermutung. Für $p = 2$ aber ist der Exponent von $D_n(G)/\gamma_n(G)$ ein Teiler der kleinsten Potenz von 2, die größer oder gleich n ist. Dennoch ist man von einer allgemeinen befriedigenden Beschreibung dieser Faktorgruppe auch heute noch weit entfernt. Für eine umfassende Übersicht über neuere Entwicklungen dieses Forschungsgebiets konsultiere man [MP].

Einige Bezeichnungen

Sei K ein kommutativer unitärer Ring, R ein unitärer Ring. (Das Weglassen von K bzw. R in einer der folgenden Schreibweisen kommt der Wahl von $\mathbb{Z}$ für K bzw. R gleich.) Es bedeutet

${}^{K}\mathfrak{A}_1$	die Klasse der unitären K-(Links-)Algebren,
${}^{K}\mathfrak{A}$	die Klasse der K-(Links-)Algebren,
$\mathfrak{G}$	die Klasse der Gruppen,
${}^{K}\mathfrak{L}$	die Klasse der K-(Links-)Lie-Algebren,
$\mathfrak{L}$	die Klasse der Lie-Algebren über $\mathbb{Z}$ (der „Lie'schen Ringe"),
${}^{R}\mathfrak{M}$	die Klasse der R-(Links-)Moduln,
$\mathfrak{M}$	die Klasse der abelschen Gruppen,
${}^{K}\mathfrak{R}_1$	die Klasse der unitären assoziativen K-(Links-)Algebren,
${}^{K}\mathfrak{R}$	die Klasse der assoziativen K-(Links-)Algebren,
$\mathfrak{R}_1$	die Klasse der unitären Ringe,
$\mathfrak{R}$	die Klasse der Ringe,
$\mathfrak{S}_1$	die Klasse der Monoide (der unitären Semigruppen),
$\mathfrak{S}$	die Klasse der Semigruppen.

Sei X eine beliebige Menge. Es bedeutet

$X^{(+)}$	von X frei erzeugtes Magma,
$X^{(*)}$	von X frei erzeugtes unitäres Magma,
X^{+}	von X frei erzeugte Semigruppe,
X^{*}	von X frei erzeugtes freies Monoid,
$K[X]$	(kommutativer) Polynomring über K in der Variablenmenge X,
$K[[X]]$	(kommutative) Potenzreihenalgebra über K in der Variablenmenge X,
$K\langle X\rangle$	Polynomring über K in der nichtkommutierenden Variablenmenge X,
$K\langle\langle X\rangle\rangle$	Potenzreihenalgebra über K in der nichtkommutierenden Variablenmenge X.

Literaturverzeichnis

[Bah] Yu. A. Bahturin, *Identical Relations in Lie Algebras*. Utrecht 1987

[Bau] G. Baumslag, *Topics in Combinatorial Group Theory*. Basel 1993

[Bo1] N. Bourbaki, *Elements of Mathematics, Algebra I, Ch. 1-3*. Berlin–Heidelberg–New York 1989

[Bo2] N. Bourbaki, *Elements of Mathematics, Lie Groups and Lie Algebras, Ch. 1-3*. Berlin–Heidelberg–New York 1989

[CM] B. Chandler, W. Magnus, *The History of Combinatorial Group Theory: A Case Study in the History of Ideas*. New York 1982

[Gup1] N. Gupta, *Free Group Rings*. Providence–Rhode Island 1987

[Gup2] N. Gupta, The dimension subgroup conjecture holds for odd order groups. *J. Group Theory* 5:481-491, 2002

[Hal] M. Hall, *The Theory of Groups*. New York 1959

[Jac] N. Jacobson, *Lie Algebras*. New York 1962

[Kur] A. G. Kurosch, *Gruppentheorie I*. Berlin 1970

[Lot] M. Lothaire, *Combinatorics on Words*. Reading, Mass. 1983

[LS] R. C. Lyndon, P. E. Schupp, *Combinatorial Group Theory*. Berlin–Heidelberg 1977

[MKS] W. Magnus, A. Karrass, D. Solitar, *Combinatorial Group Theory*. New York 1966

[MP] R. Mikhailov, I. B. S. Passi, *Lower Central and Dimension Series of Groups*. Berlin-Heidelberg 2009

[Reu] C. Reutenauer, *Free Lie Algebras*. Oxford 1993

Index