Algebra: Gruppen, Ringe, Körper

Skript zur Vorlesung im Wintersemester 2020/21

Prof. Meinolf Geck, Lehrstuhl für Algebra, Universität Stuttgart
https://pnp.mathematik.uni-stuttgart.de/iaz/iaz2/geckmf

Dieses Skript zur Vorlesung Algebra im Wintersemester 2020/21 (V4Ü2, 13 Wochen) ist eine etwas schlankere, leicht modifizierte und umorganisierte Version des Textes:

M. Geck, *Algebra: Gruppen, Ringe, Körper – Mit einer Einführung in die Darstellungstheorie endlicher Gruppen*. Edition Delkhofen, 2014. (Erhältlich z.B. bei Wittwer/Thalia.)

In diesem Skript ist insbesondere das erste Kapitel eine ausführlichere Einleitung, in der an Definitionen und Beispiele erinnert wird, die vermutlich schon zumindest ansatzweise in der Vorlesung Lineare Algebra I (oder Analysis I) eine Rolle spielten. Außerdem werden hier bereits einige grundlegende Begriffe und Konstruktionen eingeführt, bevor sie in den späteren drei Kapiteln zu Gruppen, Ringen und Körpern vertieft werden.

Die "Galois-Theorie" wird in obigem Text behandelt (und auch dort bereits in einer gegenüber vielen Lehrbüchern vereinfachten Form), kommt aber hier nicht mehr vor. Alternativ ist dieser Teil des Stoffes auch zum Beispiel gut geeignet für 4-5 Vorträge in einem Hauptseminar in einem der folgenden Semester. Der obige Text enthält ansonsten einige Ergänzungskapitel, die weitere Algebra-Themen behandeln und auch zum Selbststudium geeignet sind (z.B. eine Einführung in die Charaktertheorie endlicher Gruppen). Umgekehrt findet sich bis auf wenige Ausnahmen alles, was in diesem Skript vorkommt, auch in obigem Text (manchmal mit alternativen Beweisen). Eine solche Ausnahme ist die Beschreibung des RSA-Verfahrens in §5; eine weitere der in §19 vorgestellte Beweis des Fundamentalsatzes der Algebra, der auf dem Hauptsatz über symmetrische Polynome (siehe §14) beruht.

Es gibt zwei Anhänge, die nicht in der Vorlesung behandelt wurden (aber beide sind möglich als Stoff für eine 14. Vorlesungswoche): Anhang A enthält einen vollständigen Beweis, dass es keine allgemeinen Lösungsformeln für Polynome vom Grad $\geqslant 5$ geben kann ("Satz von Abel–Ruffini"); Anhang B diskutiert die Existenz eines algebraischen Abschlusses eines Körpers.

Im Durchschnitt werden pro Woche etwa 6 Seiten dieses Skriptes behandelt (am Anfang, bei den Wiederholungen aus der Linearen Algebra etwas mehr). **Kommentare sehr willkommen!** (Insbesondere Druckfehler, sonstige Unklarheiten, Verbesserungsvorschläge etc.)

Stuttgart, Februar 2021

LITERATUR

[Ar] M. ARTIN, Algebra. Aus dem Englischen übersetzt von Annette A'Campo. Birkhäuser Verlag, 1993.

[Bo] N. BOURBAKI, Éléments de Mathématiques. Algèbre. Chap. 1 à 3, Masson, Paris, 1970; Chap. 4 à 7, Masson, Paris, 1981.

[Eb] H. D. EBBINGHAUS, H. HERMES, F. HIRZEBRUCH, M. KOECHER, K. MAINZER, J. NEUKIRCH, A. PRESTEL UND R. REMMERT, Zahlen. Grundwissen Mathematik, vol. 1, Springer-Verlag, Berlin, 1983.

[Fi] G. FISCHER, Lehrbuch der Algebra, Springer Spektrum, 2008.

[FS] G. FISCHER UND R. SACHER, Einführung in die Algebra (Teubner Studienbücher Mathematik). Vieweg + Teubner Verlag; 3. Auflage 1983.

[Fr] J. B. FRALEIGH, A first course in abstract algebra. Pearson, 7th edition, 2002.

[GAP] THE GAP GROUP, GAP - Groups, Algorithms, and Programming, Version 4.11.0, 2020. Frei verfügbares Computer-Algebra-System, siehe **http://www.gap-system.org**.

[G14] M. GECK, Algebra: Gruppen, Ringe, Körper – Mit einer Einführung in die Darstellungstheorie endlicher Gruppen. Edition Delkhofen, 2014.

[Ha] P. R. HALMOS, Naive Mengenlehre, Vandenhoeck & Ruprecht, 5. Auflage, 1994.

[Ja] N. JACOBSON, Basic Algebra I, II. 2nd Edition. H. Freeman and Company, New York, 1985 und 1989.

[KM] C. KARPFINGER UND K. MEYBERG, Algebra: Gruppen - Ringe - Körper. Spektrum Akademischer Verlag, 2008.

[KS] H. KURZWEIL UND B. STELLMACHER, The theory of finite groups. Springer–Verlag, 2004.

[Lo] F. LORENZ, Einführung in die Algebra (2 Bände). Spektrum Akademischer Verlag, 1996 und 1997.

[Pe] N. PERRIN, Cours d'algèbre. Ellipses, Paris, 1996.

[Ro] M. I. ROSEN, Niels Hendrik Abel and equations of the fifth degree, Amer. Math. Monthly **102** (1995), 495–505.

[So] R. SOLOMON, A brief history of the classification of the finite simple groups. Bull. Amer. Math. Soc. **38** (2001), 315–352; siehe auch **http://en.wikipedia.org/wiki/List_of_finite_simple_groups**.

[St] J. STILLWELL, Elements of algebra. Geometry, numbers, equations. Undergraduate Texts in Mathematics, Springer–Verlag, New York, 1994.

Kapitel I: Grundlagen

In der Linearen Algebra haben Sie vermutlich bereits die Definitionen der grundlegenden algebraischen Strukturen gesehen: Gruppen, Ringe, Körper und natürlich Vektorräume. In diesem Einleitungskapitel beginnen wir mit einigen Erinnerungen, Beispielen und grundlegenden Konstruktionen. Alle Argumente werden hier etwas ausführlicher als später ausgeführt, weil sie tatsächlich wesentlich zum Verständnis des weiteren Stoffes sind.

1. *Algebraische Strukturen*

Eine nicht-leere Menge G zusammen mit einer Verknüpfung $\star\colon G \times G \to G$ heißt ***Gruppe***, wenn die Verknüpfung assoziativ ist (also $a \star (b \star c) = (a \star b) \star c$ für alle $a, b, c \in G$), es ein neutrales Element gibt (also ein $e \in G$ mit $a \star e = e \star a = a$ für alle $a \in G$) und jedes Element von G ein Inverses besitzt (es also zu jedem $a \in G$ ein $a' \in G$ gibt mit $a \star a' = a' \star a = e$). Wie bei Vektorräumen zeigt man dann leicht, dass das neutrale Element e eindeutig bestimmt ist, und dass das Inverse eines Elements ebenfalls eindeutig bestimmt ist. Beachte: Für das Inverse eines Produktes gilt die Regel

$$(a \star b)' = b' \star a' \qquad \text{für alle } a, b \in G.$$

(Denn: Setze $c := b' \star a'$. Dann gilt $b \star c = b \star (b' \star a') = (b \star b') \star a' = e \star a' = a'$ und damit $(a \star b) \star c = a \star (b \star c) = a \star a' = e$; analog sieht man $c \star (a \star b) = e$.)

Beispiele von Gruppen:

• Sei $n \geqslant 1$. Die ***symmetrische Gruppe*** S_n ist die Menge aller bijektiven Abbildungen $\sigma\colon \{1, \dots, n\} \to \{1, \dots, n\}$ (auch Permutationen genannt), zusammen mit der Hintereinanderausführung "$\circ$" als Verknüpfung. Das neutrale Element ist die identische Abbildung $\mathrm{id} \in S_n$ (mit $\mathrm{id}(i) = i$ für $1 \leqslant i \leqslant n$) und das Inverse von $\sigma \in S_n$ die Umkehrabbildung σ^{-1}. Elemente von S_n schreiben wir in der Form

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix} \in S_n.$$

Sei etwa $n = 3$ und $\sigma = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{smallmatrix}\right) \in S_3$, $\tau = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{smallmatrix}\right) \in S_3$; dann ist $\sigma \circ \tau = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{smallmatrix}\right)$ (denn $(\sigma \circ \tau)(1) = \sigma(\tau(1)) = \sigma(1) = 2$, $(\sigma \circ \tau)(2) = \sigma(\tau(2)) = \sigma(3) = 3$, $(\sigma \circ \tau)(3) = \sigma(\tau(3)) = \sigma(2) = 1$).

• Etwas allgemeiner: Sei $X \neq \varnothing$ eine Menge und S_X die Menge aller bijektiven Abbildungen $\sigma\colon X \to X$. Dann ist S_X eine Gruppe mit der Hintereinanderausführung "$\circ$" von Abbildungen als Verknüpfung. Das neutrale Element ist die identische Abbildung id_X; das Inverse von $\sigma \in S_X$ ist wiederum die Umkehrabbildung von σ. Diese Gruppe heißt ***symmetrische Gruppe*** auf X. Dann gilt also $S_n = S_X$ mit $X = \{1, \dots, n\}$.

• Sei K ein Körper und $n \geqslant 1$. Mit $M_n(K)$ bezeichnen wir den Vektorraum aller $n \times n$-Matrizen mit Einträgen in K. Aus der Linearen Algebra ist bekannt: $A \in M_n(K)$ ist invertierbar genau dann, wenn $\det(A) \neq 0$ gilt; außerdem: Für $A, B \in M_n(K)$ gilt $\det(A \cdot B) = \det(A)\det(B)$. Damit ist

$$\mathrm{GL}_n(K) := \{A \in M_n(K) \mid \det(A) \neq 0\}$$

eine Gruppe mit der üblichen Matrixmultiplikation als Verknüpfung; das neutrale Element ist die Einheitsmatrix I_n. Diese Gruppe heißt die ***allgemeine lineare Gruppe***.

Definition 1.1. Eine Gruppe G heißt ***abelsch*** (zu Ehren des Mathematikers H. N. Abel), wenn die Multiplikation kommutativ ist, also $a \star b = b \star a$ für alle $a, b \in G$. In diesem Fall wird die Verknüpfung auch oft als Addition $a + b$ geschrieben, das neutrale Element mit 0 bezeichnet und das Inverse von $a \in G$ mit $-a$.

Beispiele: Die Gruppe der ganzen Zahlen $\mathbb{Z}$, mit der üblichen Addition, ist abelsch. (Bezüglich der Multiplikation ist $\mathbb{Z}$ keine Gruppe, weil nicht jedes Element ein Inverses besitzt.) Die symmetrischen Gruppen S_1 und S_2 sind ebenfalls abelsch. Für $n = 3$ gilt:

$$\sigma := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad \tau := \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \quad \Rightarrow \quad \sigma \circ \tau = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \neq \tau \circ \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

Also ist S_3 nicht abelsch; genauso sieht man, dass S_n für $n \geqslant 3$ niemals abelsch ist. Für $n = 1$ ist $\mathrm{GL}_1(K) = \{(a) \mid 0 \neq a \in K\}$ abelsch; für $n \geqslant 2$ ist $\mathrm{GL}_n(K)$ nicht abelsch. (Finden Sie selbst zwei invertierbare Matrizen, die nicht vertauschbar sind.)

Abelsche Gruppen sind bereits aus der Linearen Algebra vertraut: Ein Vektorraum V über einem Körper K ist *zunächst* eine abelsche Gruppe $(V, +)$, auf der zusätzlich eine "äußere" Verknüpfung $K \times V \to V$ (Multiplikation von Vektoren mit Skalaren aus einem Körper K) definiert ist, so dass die bekannten Bedingungen gelten.

Ein ***Ring*** R ist eine abelsche Gruppe $(R, +)$, auf der zusätzlich eine Multiplikation $\bullet\colon R \times R \to R$ definiert ist, die assoziativ ist und so dass die Distributivregeln gelten: $a \bullet (b + c) = a \bullet b + a \bullet c$ und $(a + b) \bullet c = a \bullet c + b \bullet c$ für alle $a, b, c \in R$. Aus diesen Regeln folgt zum Beispiel sofort:

$$a \bullet 0 = 0 \bullet a = 0 \qquad \text{für alle } a \in R.$$

Wir bezeichnen R als kommutativen Ring, wenn die Multiplikation kommutativ ist, also $a \bullet b = b \bullet a$ für alle $a, b \in R$.

Standard-Beispiele: $R = \mathbb{Z}$ ist ein kommutativer Ring, mit der üblichen Addition und Multiplikation von ganzen Zahlen. Außerdem ist $R = M_n(K)$ ein Ring mit der üblichen Addition und Multiplikation von Matrizen. Dieser Ring ist kommutativ genau dann, wenn $n = 1$.

Definition 1.2. Sei R ein Ring, der ein neutrales Element 1_R bezüglich der Multiplikation besitzt (kurz: R ist ein Ring mit 1). Ein Element $a \in R$ heißt ***Einheit***, wenn a bezüglich der Multiplikation ein Inverses besitzt, es also ein $b \in R$ gibt mit $a \bullet b = b \bullet a = 1_R$. Dann heißt

$$R^\times := \{a \in R \mid a \text{ Einheit}\}$$

die ***Einheitengruppe*** von R. Zum Beispiel ist $1_R \in R^\times$. Man sieht sofort, dass $R^\times$ tatsächlich eine Gruppe bezüglich der Multiplikation in R ist; das neutrale Element ist 1_R.

Beachte: Der Extremfall $1_R = 0$ ist in der Definition nicht ausgeschlossen. Aber aus $1_R = 0$ folgt $a = a \bullet 1_R = a \bullet 0 = 0$ für alle $a \in R$; also $R = R^\times = \{0\}$.

Ein ***Körper*** K ist damit ein kommutativer Ring mit $1 \neq 0$, so dass $K^\times = K \setminus \{0\}$ gilt. Dann heißt $K^\times$ auch die multiplikative Gruppe von K. Standard-Beispiele sind $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$; für jede Primzahl p gibt es auch einen Körper $\mathbb{F}_p$ mit p Elementen (den Sie vielleicht schon in der Linearen Algebra gesehen haben; ansonsten siehe §4).

Beispiele zu Einheitengruppen: Für $R = \mathbb{Z}$ ist $\mathbb{Z}^\times = \{1, -1\}$. Für $R = M_n(K)$ (mit K Körper) ist $M_n(K)^\times = \{A \in M_n(K) \mid A \text{ invertierbar}\} = GL_n(K)$.

Eine Methode, um neue algebraische Strukturen zu erhalten, besteht darin, Unterstrukturen zu betrachten (analog zu Unterräumen von Vektorräumen).

Definition 1.3. (a) Sei $(G, \star)$ eine Gruppe und $U \subseteq G$ eine Teilmenge. Dann heißt U eine ***Untergruppe*** von G (in Zeichen: $U \leqslant G$), wenn gilt:

$$1_G \in U, \qquad a \star b \in U \qquad \text{und} \quad a' \in U \qquad \text{für alle } a, b \in U.$$

(Hier ist a' das Inverse von a.) Dann ist U zusammen mit der Einschränkung der Verknüpfung auf G auch selbst eine Gruppe. Zum Beispiel sind $\{1_G\}$ und G immer Untergruppen.

(b) Sei $(R, +, \bullet)$ ein Ring und $S \subseteq R$ eine Teilmenge. Dann heißt S ein ***Teilring*** von R, wenn S eine Untergruppe von R bezüglich der Addition $+$ ist und außerdem $a \bullet b \in S$ für alle $a, b \in S$ gilt. Wiederum ist S selbst zusammen mit den Einschränkungen der Verknüpfung auf R auch selbst ein Ring. Zum Beispiel sind $\{0\}$ und R immer Teilringe.

Bemerkung 1.4. Sei $\{U_i\}_{i \in I}$ eine Familie von Untergruppen von G (mit beliebiger Indexmenge I). Dann ist auch $\bigcap_{i \in I} U_i$ eine Untergruppe (wie Sie sofort nachprüfen), aber $\bigcup_{i \in I} U_i$ ist im Allgemeinen keine Untergruppe (siehe Übungen). Analoge Aussagen gelten für Durchschnitte von Teilringen.

Beispiel 1.5. Sei $G = GL_n(K)$ die allgemeine lineare Gruppe. Sei $B_n(K) \subseteq G$ die Teilmenge, die aus allen oberen Dreiecksmatrizen mit Einträgen ungleich Null auf der Diagonalen besteht. Dann ist $B_n(K)$ eine Untergruppe. Denn erstens gilt $I_n \in B_n(K)$. Sind außerdem

$A, B \in B_n(K)$, so folgt leicht aus der Definition des Matrixprodukts, dass auch $A \cdot B$ eine obere Dreiecksmatrix ist; sind außerdem $a_1, \ldots, a_n$ die Diagonaleinträge von A und $b_1, \ldots, b_n$ die Diagonaleinträge von B, so sind $a_1 b_1, \ldots, a_n b_n$ die Diagonaleinträge von $A \cdot B$. Also sind auch diese alle ungleich Null und damit $A \cdot B \in B_n(K)$. Schließlich gilt $\det(A) = a_1 \cdots a_n \neq 0$, also ist A invertierbar. Dann muss man sich noch überzeugen, dass A^{-1} ebenfalls eine obere Dreiecksmatrix ist, mit Diagonaleinträgen $a_1^{-1}, \ldots, a_n^{-1}$. Also ist auch $A^{-1} \in B_n(K)$.

Beispiel 1.6. Die ***Gauß'schen Zahlen*** $\mathbb{Z}[i] := \{n + mi \mid n, m \in \mathbb{Z}\} \subseteq \mathbb{C}$ bilden einen Teilring des Körpers $\mathbb{C}$ (wobei wie üblich $i = \sqrt{-1} \in \mathbb{C}$). Denn sind $n + mi \in \mathbb{Z}[i]$ und $n' + m'i \in \mathbb{Z}[i]$, so gilt

$$(n + mi) + (n' + m'i) = (n + n') + (m + m')i \in \mathbb{Z}[i],$$
$$(n + mi) \cdot (n' + m'i) = (nn' - mm') + (nm' + n'm)i \in \mathbb{Z}[i].$$

Daran sieht man sofort, dass die Teilring-Bedingungen gelten. Da $\mathbb{C}$ ein Körper ist, ist $\mathbb{Z}[i]$ ein kommutativer Ring. Die Zahl 1 ist das neutrale Element bezüglich der Multiplikation.

Behauptung: $\mathbb{Z}[i]^\times = \{1, -1, i, -i\}$.

Dazu: Es ist klar, dass ± 1 und $\pm i$ Einheiten sind. Sei umgekehrt $0 \neq n + mi \in \mathbb{Z}[i]^\times$. Dann gibt es also $n', m' \in \mathbb{Z}$ mit $1 = (n + mi)(n' + m'i)$. Anwenden von komplexer Konjugation ergibt $1 = (n - mi)(n' - m'i)$, also ist auch $n - mi \in \mathbb{Z}[i]^\times$ und damit $n^2 + m^2 = (n+mi)(n-mi) \in \mathbb{Z}[i]^\times$. Also gibt es $n'', m'' \in \mathbb{Z}$ mit $1 = (n^2+m^2)(n''+m''i) = (n^2+m^2)n''+(n^2+m^2)m''i$; dann muss aber $1 = (n^2+m^2)n''$ gelten, also $n^2+m^2 = \pm 1$. Dies ist nur möglich für $n = 0$, $m = \pm 1$ oder für $n = \pm 1$, $m = 0$. — Wie man an diesem Beispiel erahnt, kann es für beliebige Ringe durchaus schwierig sein, die Einheiten zu bestimmen.

Bemerkung 1.7. Sei R ein Ring und $S \subseteq R$ ein Teilring. Besitzt R ein Eins-Element, so muss S nicht unbedingt ein Eins-Element besitzen. Und selbst wenn dies der Fall ist, so können R und S verschiedene Eins-Elemente besitzen. Beispiele:

• Sei $S := 2\mathbb{Z} = \{\text{alle geraden ganzen Zahlen}\} \subseteq R = \mathbb{Z}$. Dann ist S ein Teilring; dieser besitzt aber kein Eins-Element.

• Sei $S := \{\left(\begin{smallmatrix} a & 0 \\ 0 & 0 \end{smallmatrix}\right) \mid a \in \mathbb{Q}\} \subseteq R = M_2(\mathbb{Q})$. Dann ist S ein Teilring mit Eins-Element $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right) \neq I_2$.

Schließlich: Ist $0 \neq s \in S$ eine Einheit in R, so gibt es also ein Inverses $s^{-1} \in R$, aber es muss nicht unbedingt $s^{-1} \in S$ gelten. Beispiel: $S = \mathbb{Z} \subseteq R = \mathbb{Q}$, $2^{-1} = \frac{1}{2} \in \mathbb{Q} \setminus \mathbb{Z}$.

Ein weiteres allgemeines Konstruktionsprinzip sind direkte Produkte.

Definition 1.8. Seien $(G_1, \star_1)$ und $(G_2, \star_2)$ Gruppen. Wir betrachten das kartesische Produkt $G := G_1 \times G_2 = \{(a_1, a_2) \mid a_1 \in G_1, a_2 \in G_2\}$ und definieren eine Verknüpfung auf G

durch $(a_1, a_2) \star (b_1, b_2) := (a_1 \star_1 b_1, a_2 \star_2 b_2)$ für alle $a_1, b_1 \in G_1$ und $a_2, b_2 \in G_2$. Dann sieht man leicht, dass $(G, \star)$ auch eine Gruppe ist, die als ***direktes Produkt*** bezeichnet wird. Das neutrale Element ist $e := (e_1, e_2) \in G$ wobei e_1 das neutrale Element in G_1 und $e_2 \in G_2$ das neutrale Element in G_2 ist. Das Inverse von $(a_1, a_2) \in G$ ist gegeben durch $(a_1, a_2)' = (a_1', a_2')$, wobei jeweils $a_i' \in G_i$ das Inverse zu a_i in G_i ist. Sind G_1 und G_2 abelsch, so ist auch G abelsch.

Analog definiert man auch das direkte Produkt von zwei Ringen $(R_1, +_1, \bullet_1)$ und $(R_2, +_2, \bullet_2)$. Man bildet das kartesische Produkt $R := R_1 \times R_2$ und definiert Verknüpfungen

$$(a_1, a_2) + (b_1, b_2) := (a_1 +_1 b_1, a_2 +_2 b_2) \quad \text{und} \quad (a_1, a_2) \bullet (b_1, b_2) := (a_1 \bullet_1 b_1, a_2 \bullet_2 b_2)$$

für alle $a_1, b_1 \in R_1$ und $a_2, b_2 \in R_2$. Dann sieht man leicht, dass $(R, +, \bullet)$ wieder ein Ring ist. Sind R_1 und R_2 kommutativ, so ist auch R kommutativ. Besitzen R_1 und R_2 Eins-Elemente 1_{R_1} und 1_{R_2}, so ist $1_R = (1_{R_1}, 1_{R_2})$ ein Eins-Element von R. In diesem Fall folgt dann auch sofort, dass $R^\times = R_1^\times \times R_2^\times$ gilt.

Schließlich erwähnen wir, dass direkte Produkte nicht nur mit zwei Faktoren, sondern auch mit endlich vielen Faktoren gebildet werden können.

2. *Faktorstrukturen*

Sei M eine nicht-leere Menge und $\sim$ eine Äquivalenzrelation auf M (also eine reflexive, symmetrische, transitive Relation). Für $m \in M$ bezeichne $[m] := \{m' \in M \mid m' \sim m\}$ die Äquivalenzklasse von m. Sei $M/\sim$ die Menge der Äquivalenzklassen. Ist M nicht nur eine Menge, sondern eine algebraische Struktur, und ist $\sim$ mit dieser Struktur verträglich, so besteht eine gute Chance, dass auch $M/\sim$ wieder eine (neue) algebraische Struktur ist. Dies ist ein sehr vielseitiges und schlagkräftiges Konstruktionsprinzip.

Als erstes und bekanntes Beispiel behandeln wir die rationalen Zahlen, also Brüche $\frac{n}{m} \in \mathbb{Q}$ mit $n, m \in \mathbb{Z}$ wobei $m \neq 0$. Ein Bruch wie $\frac{2}{3} \in \mathbb{Q}$ kann auf verschiedene Weise dargestellt werden, etwa $\frac{2}{3} = \frac{4}{6} = \frac{-10}{-15}$. Der formal korrekte Hintergrund dieser verschiedenen Darstellungsweisen ist letztlich eine Äquivalenzrelation.

Beispiel 2.1. (Konstruktion von $\mathbb{Q}$ aus $\mathbb{Z}$). Sei $M := \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$, also die Menge aller Paare (a, b) mit $a, b \in \mathbb{Z}$ und $b \neq 0$. Wir definieren eine Relation $\sim$ auf M wie folgt: Für $(a, b) \in M$ und $(c, d) \in M$ sei $(a, b) \sim (c, d)$, wenn $ad = bc$. Diese Relation ist reflexiv, denn es gilt $(a, b) \sim (a, b)$ wegen $ab = ba$. Sie ist symmetrisch, denn aus $(a, b) \sim (c, d)$ folgt $ad = bc$ und damit auch $cb = da$, also $(c, d) \sim (a, b)$. Schließlich ist $\sim$ auch transitiv, denn seien $(a, b) \sim (c, d)$ und $(c, d) \sim (e, f)$; dann gilt $ad = bc$, $cf = de$ und damit $daf =$

$(ad)f = (bc)f = b(cf) = b(de) = dbe$. Nun ist $d \neq 0$, also kann man d auf beiden Seiten kürzen. Also gilt auch $af = be$ und $(a, b) \sim (e, f)$. Damit ist $\sim$ eine Äquivalenzrelation. Die Äquivalenzklasse von $(a, b) \in M$ bezeichnen wir mit a/b; die Menge aller Äquivalenzklassen definieren wir als $\mathbb{Q} := \{a/b \mid a, b \in \mathbb{Z}, b \neq 0\}$.

Wir wollen nun Verknüpfungen auf $\mathbb{Q}$ definieren. Für $(a, b) \in M$ und $(c, d) \in M$ setzen wir:

$$a/b + c/d := (ad + bc)/bd \qquad \text{und} \qquad a/b \cdot c/d := (ac)/(bd).$$

(Beachte: Wegen $b \neq 0$ und $d \neq 0$ ist auch $bd \neq 0$.) Damit diese Operationen überhaupt sinnvoll sind, muss zuerst gezeigt werden, dass sie ***wohl-definiert*** sind, d.h., nicht von der Wahl der Repräsentanten (a, b) und (c, d) abhängen. Konkret heißt das in diesem Fall: Seien $(a', b') \in M$ und $(c', d') \in M$ mit $(a, b) \sim (a', b')$ und $(c, d) \sim (c', d')$; dann muss man zeigen, dass auch

$$(ad + bc, bd) \sim (a'd' + b'c', b'd') \qquad \text{und} \qquad (ac, bd) \sim (a'c', b'd')$$

gilt, also bei der Berechnung der Addition oder Multiplikation jeweils das gleiche Ergebnis herauskommt. Diese einfache Rechnung sei als Übung überlassen. Genauso sieht man dann leicht, dass obige Addition und Multiplikation kommutativ sind; außerdem rechnet man nach, dass die Operationen assoziativ sind und die Distributivregeln gelten. Das neutrale Element bezüglich der Addition ist dann die Äquivalenzklasse $0/1 \in \mathbb{Q}$. Denn für $(c, d) \in M$ gilt $0/1 + c/d = (0 \cdot d + 1 \cdot c)/(1 \cdot d) = c/d$. Für $(a, b) \in M$ gilt $(a, b) \sim (0, 1) \Leftrightarrow a = a \cdot 1 = b \cdot 0 = 0$, also ist $0/1 = \{0/b \mid b \in \mathbb{Z}, b \neq 0\}$. Genauso sieht man, dass $1/1 = \{a/a \mid 0 \neq a \in \mathbb{Z}\}$ das neutrale Element bezüglich der Multiplikation ist. Also ist $\mathbb{Q}$ ein kommutativer Ring mit 1. Weiterhin erhält man: Ist $a/b \in \mathbb{Q}$ und $a/b \neq 0/1$, so folgt $a \neq 0$ und $b \neq 0$; also ist auch $b/a \in \mathbb{Q}$ und $a/b \cdot b/a = 1/1$. Folglich ist $\mathbb{Q}$ ein Körper, denn alle Elemente ungleich $0/1$ in $\mathbb{Q}$ sind Einheiten. Schließlich können wir $\mathbb{Z}$ als eine Teilmenge von $\mathbb{Q}$ auffassen, indem wir $n \in \mathbb{Z}$ mit dem Bruch $n/1 \in \mathbb{Q}$ identifizieren. Die Abbildung $n \mapsto n/1$ ist injektiv, denn $n/1 = m/1$ (für $n, m \in \mathbb{Z}$) impliziert $n = m$ nach Definition von $\sim$. Mit dieser Identifikation entsprechen dann auch die übliche Addition $n + m$ und Multiplikation nm in $\mathbb{Z}$ den Operationen $n/1 + m/1$ und $n/1 \cdot m/1$. — Man sieht hier, dass im Detail viele Regeln verifiziert werden müssen, aber dies meistens Routine-Aufgaben sind nachdem einmal gezeigt ist, dass die Operationen "wohl-definiert" sind.

Beispiel 2.2. In einer Vorlesung (oder einem Lehrbuch) zur Analysis werden Sie vielleicht die Konstruktion von $\mathbb{R}$ aus $\mathbb{Q}$ mit Hilfe von Cauchy-Folgen gesehen haben. Auch dies ist ein Beispiel für obiges allgemeines Konstruktionsprinzip: Man erhält $\mathbb{R}$ als Menge der Äquivalenzklassen von Cauchy-Folgen $(a_n)_{n\in\mathbb{N}}$ (mit $a_n \in \mathbb{Q}$ für alle $n \in \mathbb{N}$), wobei $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ in Relation stehen, wenn $(a_n - b_n)_{n\in\mathbb{N}}$ eine Nullfolge ist. Genauso wie in obigem

Beispiel muss man dann zahlreiche Regeln im Detail nachweisen, um zu sehen, dass die Menge der Äquivalenzklassen ein Körper ist (also $\mathbb{R}$), in dem jede Cauchy-Folge konvergiert. Schließlich fassen wir $\mathbb{Q}$ als Teilmenge von $\mathbb{R}$ auf, indem wir $x \in \mathbb{Q}$ mit der Äquivalenzklasse der Folge $(x, x, x, \ldots)$ identifizieren.

Für das folgende Beispiel wiederholen wir einige Grundtatsachen zu ganzen Zahlen. Zunächst sei an die Bezeichnungen $\mathbb{N} = \{1, 2, 3, \ldots\}$ und $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ erinnert. (In manchen Büchern ist $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$.) Seien $m, n \in \mathbb{Z}$. Wie üblich schreiben wir $m \mid n$, wenn m ein Teiler von n ist, es also ein $k \in \mathbb{Z}$ mit $n = km$ gibt. Wir setzen außerdem die ***Division mit Rest*** als bekannt voraus. Sind also $m \in \mathbb{N}$ und $n \in \mathbb{Z}$ gegeben, so gibt es $q, r \in \mathbb{Z}$ mit $n = qm + r$ und $0 \leqslant r < m$; hierbei sind q, r eindeutig bestimmt. Zur Erinnerung: Ist $n \geqslant 0$, so finde das größte $q \geqslant 0$ mit $qm \leqslant n$; dann ist $0 \leqslant r := n - qm < m$. Ist $n < 0$, so schreibe $-n = q'm + r'$ mit $0 \leqslant r' < m$; dann ist $n = (-q')m - r'$. Ist $r' = 0$, so sind wir fertig mit $r = 0$. Ist $r' > 0$, so gilt $n = (-q' - 1)m + m - r'$ mit $0 \leqslant r := m - r' < m$. Zum Beispiel für $m = 5$ und $n = 17$: Es gilt $3 \cdot 5 \leqslant 17$ aber $4 \cdot 5 > 17$; dies ergibt $17 = 3 \cdot 5 + 2$, also $q = 3$, $r = 2$; damit folgt $-17 = (-3) \cdot 5 - 2 = (-3 - 1) \cdot 5 + 3$, also $q = -4$, $r = 3$.

Beispiel 2.3. (Die Ringe $\mathbb{Z}/m\mathbb{Z}$). Sei $m \in \mathbb{N}$ fest. Wir definieren eine Relation $\sim$ auf $\mathbb{Z}$ durch $a \sim b$, wenn $m \mid b - a$ (für $a, b \in \mathbb{Z}$). Man sieht leicht, dass dies eine Äquivalenzrelation ist. Die Äquivalenzklasse von $a \in \mathbb{Z}$ bezeichnen wir mit $\overline{a}$ und die Menge aller Äquivalenzklassen mit $\mathbb{Z}/m\mathbb{Z}$. Weitere Schreibweise: $a \equiv b \bmod m$ falls $m \mid b - a$.

Durch Division mit Rest erhalten wir $a = qm + r$ mit $q, r \in \mathbb{Z}$ und $0 \leqslant r < m$. Also ist $a \equiv r \bmod m$ und damit $\overline{a} = \overline{r}$. Es folgt $\mathbb{Z}/m\mathbb{Z} = \{\overline{0}, \overline{1}, \ldots, \overline{m-1}\}$ und $|\mathbb{Z}/m\mathbb{Z}| = m$. Wir wollen nun eine Addition und eine Multiplikation auf $\mathbb{Z}/m\mathbb{Z}$ definieren. Für $a, b \in \mathbb{Z}$ setze

$$\overline{a} + \overline{b} := \overline{a + b} \qquad \text{und} \qquad \overline{a} \cdot \overline{b} := \overline{ab}.$$

Wie üblich müssen wir zuerst zeigen, dass dies ***wohl-definiert*** ist. Seien also $a, b, a', b' \in \mathbb{Z}$ mit $\overline{a} = \overline{a'}$ und $\overline{b} = \overline{b'}$, d.h., $m \mid a' - a$ und $m \mid b' - b$. Dann ist auch $(a' + b') - (a + b) = (a' - a) + (b' - b)$ ein Vielfaches von m; und ebenso $a'b' - ab = a'b' - ab' + ab' - ab = (a' - a)b' + a(b' - b)$. Also gilt $\overline{a+b} = \overline{a'+b'}$ und $\overline{ab} = \overline{a'b'}$, wie gewünscht. Aufgrund der obigen Definition ist klar, dass $\overline{0}$ neutrales Element bezüglich "+" und $\overline{1}$ neutrales Element bezüglich "$\cdot$" ist. Jedes $\overline{a} \in \mathbb{Z}/m\mathbb{Z}$ hat ein Inverses bezüglich "+", nämlich $\overline{-a}$. Die weiteren Ring-Axiome folgen sofort aus den entsprechenden Regeln in $\mathbb{Z}$, zum Beispiel:

$$\overline{a} \cdot (\overline{b} + \overline{c}) = \overline{a} \cdot \overline{(b + c)} = \overline{a(b + c)} = \overline{ab + ac} = \overline{ab} + \overline{ac} = \overline{a} \cdot \overline{b} + \overline{a} \cdot \overline{c},$$

wobei beim 3. Gleichheitszeichen die Regel $a(b + c) = ab + ac$ für $a, b, c \in \mathbb{Z}$ verwendet wurde. Der Beweis der anderen Regeln verläuft analog und sei als Übung überlassen. Damit ist $\mathbb{Z}/m\mathbb{Z}$ ein kommutativer Ring mit 1.

Für $m = 1$ ist $\mathbb{Z}/1\mathbb{Z} = \{\overline{0}\}$. Für $m = 2$ ist $\mathbb{Z}/2\mathbb{Z} = \{\overline{0}, \overline{1}\}$ mit $\overline{1} + \overline{1} = \overline{0}$. Für $m = 3, 4$ sind die Verknüpfungstabellen wie folgt gegeben:

$m = 3$:

$+$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{0}$	$\overline{1}$

$\cdot$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{1}$

(also $\overline{2}^{-1} = \overline{2}$)

$m = 4$:

$+$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{0}$	$\overline{1}$	$\overline{2}$

$\cdot$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{0}$	$\overline{2}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

(kein Körper)

In $\mathbb{Z}/5\mathbb{Z} = \{\overline{0}, \overline{1}, \overline{2}, \overline{3}, \overline{4}\}$ gilt: $\overline{1}^{-1} = \overline{1}$, $\overline{2}^{-1} = \overline{3}$, $\overline{3}^{-1} = \overline{2}$, $\overline{4}^{-1} = \overline{4}$.

In §4 werden wir sehen, dass $\mathbb{Z}/m\mathbb{Z}$ genau dann ein Körper ist, wenn m eine Primzahl ist.

Zum Abschluss dieses Abschnittes wollen wir die obige Konstruktion von $\mathbb{Z}/m\mathbb{Z}$ in einen etwas allgemeineren Kontext versetzen.

Definition 2.4. Sei $(R, +, \cdot)$ ein Ring. Eine Teilmenge $I \subseteq R$ heißt ein ***Ideal*** in R (in Zeichen $I \trianglelefteq R$), wenn I eine Untergruppe von $(R, +)$ ist und für alle $a \in I$, $b \in R$ auch $a \cdot b \in I$ und $b \cdot a \in I$ gilt. Wir definieren eine Relation $\sim$ auf R wie folgt. Seien $a, b \in R$; dann ist $a \sim b$ falls $b - a \in I$. Dies ist eine Äquivalenzrelation. Dazu: Die Relation ist reflexiv, denn $a - a = 0 \in I$; sie ist symmetrisch, weil mit $b - a \in I$ auch $a - b = -(b - a) \in I$ gilt; sie ist transitiv, weil mit $b - a \in I$ und $c - b \in I$ auch $c - a = (c - b) + (b - a) \in I$ gilt. (Jedesmal benutzen wir, dass I eine Untergruppe bezüglich der Addition ist.) Die Äquivalenzklasse von $a \in R$ bezeichnen wir mit $\overline{a}$ und die Menge aller Äquivalenzklassen mit R/I. Es gilt:

$$\overline{a} = \{b \in R \mid b - a \in I\} = \{b \in R \mid b - a = c \text{ für ein } c \in I\} = \{a + c \mid c \in I\};$$

wir schreiben dies auch kurz als $a + I$. Wir wollen nun eine Addition und eine Multiplikation auf R/I definieren. Für $a, b \in R$ setze

$$\overline{a} + \overline{b} := \overline{a + b} \qquad \text{und} \qquad \overline{a} \cdot \overline{b} := \overline{a \cdot b}.$$

Wie zuvor müssen wir zuerst zeigen, dass dies ***wohl-definiert*** ist. Seien also $a, b, a', b' \in R$ mit $\overline{a} = \overline{a'}$ und $\overline{b} = \overline{b'}$, d.h., $a' - a \in I$ und $b' - b \in I$. Dann ist auch $(a' + b') - (a + b) = (a' - a) + (b' - b) \in I$, weil I eine Untergruppe von R bezüglich der Addition ist; und ebenso $a' \cdot b' - a \cdot b = a' \cdot b' - a \cdot b' + a \cdot b' - a \cdot b = (a' - a) \cdot b' + a \cdot (b' - b) \in I$ aufgrund der weiteren Bedingung an I. Also gilt $\overline{a+b} = \overline{a'+b'}$ und $\overline{a \cdot b} = \overline{a' \cdot b'}$, d.h., die Verknüpfungen sind

wohl-definiert. Sobald dies gezeigt ist, folgt wiederum sofort aus den Ring-Axiomen für R, dass die Axiome auch für R/I gelten. Der Ring R/I heißt ***Faktorring*** von R nach I. Ist R kommutativ, so auch R/I; besitzt R ein Eins-Element 1_R, so ist $\bar{1}_R$ ein Eins-Element in R/I.

Beispiel 2.5. Sei $(R, +, \cdot)$ ein kommutativer Ring. Für $a \in R$ setzen wir

$$(a) := \{a \cdot b \mid b \in R\} \subseteq R.$$

Dann sieht man sofort, dass (a) ein Ideal ist. Ideale dieser Form heißen ***Hauptideale***. Wir können also den Faktorring $R/(a)$ für jedes $a \in R$ bilden.
Ist $R = \mathbb{Z}$ und $m \in \mathbb{N}$, so ist $(m) = m\mathbb{Z}$ die Menge aller $n \in \mathbb{Z}$ mit $m \mid n$. Damit gilt $b - a \in (m) \Leftrightarrow m \mid b - a \Leftrightarrow a \equiv b \bmod m$ für alle $a, b \in \mathbb{Z}$. Also ist $\mathbb{Z}/m\mathbb{Z}$ (wie in Beispiel 2.3) das Gleiche wie $\mathbb{Z}/(m)$.

Ab hier Woche 2

3. *Der Satz von Lagrange*

In Analogie zu obiger Definition eines Faktorringes werden wir in §7 auch Faktorgruppen nach bestimmten Untergruppen bilden. Hier zeigen wir nun zunächst den Satz von Lagrange, der grundlegend für viele Aussagen über endliche Gruppen ist.

Sei G eine Gruppe. Meistens schreiben wir nun die Verknüpfung in G als $a \cdot b$ oder einfach als ab. Üblicherweise wird dann auch das neutrale Element mit 1_G (oder einfach 1) bezeichnet, sowie das inverse Element mit a^{-1}. Die Mächtigkeit $|G|$ wird auch als ***Ordnung*** von G bezeichnet. Wir werden sowohl endliche als auch unendliche Gruppen betrachten. Für beliebige Teilmengen $A, B \subseteq G$ schreiben wir

$$A \cdot B := \{a \cdot b \mid a \in A, b \in B\} \qquad \text{und} \qquad A^{-1} := \{a^{-1} \mid a \in A\}.$$

Ist $g \in G$, so sei $g \cdot A := \{g \cdot a \mid a \in A\}$ und $A \cdot g := \{a \cdot g \mid a \in A\}$. Damit gilt für eine Teilmenge $U \subseteq G$: $\quad U$ ist eine Untergruppe $\iff 1_G \in U, \quad U \cdot U \subseteq U, \quad U^{-1} \subseteq U$.

Satz 3.1. *Sei $U \leqslant G$ eine Untergruppe. Wir definieren wie folgt eine Relation $\sim_U$ auf G. Für $a, b \in G$ schreibe $a \sim_U b$, falls $a^{-1} \cdot b \in U$. Dann gilt:*

(a) *Die Relation $\sim_U$ ist eine Äquivalenzrelation. Sei G/U die Menge der Äquivalenzklassen. Diese sind von der Form $a \cdot U$ mit $a \in G$ und heißen **Linksnebenklassen**. Ist also T ein Vertretersystem der Äquivalenzklassen, so gilt*

$$G = \bigcup_{t \in T} t \cdot U \qquad \textit{(disjunkte Vereinigung).}$$

(b) *Für jedes feste $a \in G$ ist $U \to a \cdot U$, $u \mapsto a \cdot u$, eine Bijektion. Also gilt $|a \cdot U| = |U|$.*

Beweis. (a) Die Relation $\sim_U$ ist reflexiv: Für $a \in G$ ist $1_G = a^{-1} \cdot a \in U$ also $a \sim_U a$.
Die Relation $\sim_U$ ist symmetrisch: Für $a, b \in G$ mit $a \sim_U b$ ist $a^{-1} \cdot b \in U$ also auch $b^{-1} \cdot a = (a^{-1} \cdot b)^{-1} \in U$ und damit $b \sim_U a$.
Die Relation $\sim_U$ ist transitiv: Seien $a, b, c \in G$ mit $a \sim_U b$ und $b \sim_U c$. Dann ist $a^{-1} \cdot b \in U$ und $b^{-1} \cdot c \in U$, also auch $a^{-1} \cdot c = (a^{-1} \cdot b) \cdot (b^{-1} \cdot c) \in U$ und damit $a \sim_U c$.
Damit ist $\sim_U$ eine Äquivalenzrelation. Sei $a \in G$ fest. Dann gilt $a \sim_U b \Leftrightarrow a^{-1} \cdot b \in U \Leftrightarrow a^{-1} \cdot b = u$ für ein $u \in U \Leftrightarrow b = a \cdot u$ für ein $u \in U \Leftrightarrow b \in a \cdot U$. Also sind alle Äquivalenzklassen von der Form $a \cdot U$ mit einem $a \in G$.

(b) Die Abbildung ist offenbar surjektiv; sie ist injektiv da $a \in G$ ein Inverses besitzt. Aus $a \cdot u = a \cdot u'$ folgt also $u = 1_G \cdot u = (a^{-1} \cdot a) \cdot u = a^{-1} \cdot (a \cdot u) = a^{-1} \cdot (a \cdot u') = \ldots = u'$. □

Bemerkung 3.2. Wir können auch eine Relation $\sim'_U$ wie folgt definieren: $a \sim'_U b \overset{\text{def}}{\Longleftrightarrow} a \cdot b^{-1} \in U$ für alle $a, b \in G$. (Beachte: Ist G abelsch, so sind $\sim_U$ und $\sim'_U$ gleich; aber im Allgemeinen ist dies nicht der Fall.) Dann gelten die analogen Aussagen wie oben, also:

(a) $\sim'_U$ ist eine Äquivalenzrelation. Sei $U\backslash G$ die Menge der Äquivalenzklassen. Diese sind von der Form $U \cdot a$ mit $a \in G$ und heißen ***Rechtsnebenklassen***. Ist also T' ein Vertretersystem der Äquivalenzklassen, so gilt $G = \bigcup_{t \in T'} U \cdot t$, wobei die Vereinigung disjunkt ist.

(b) Für jedes feste $a \in G$ ist $U \to U \cdot a$, $u \mapsto u \cdot a$, eine Bijektion. Also gilt $|U| = |U \cdot a|$.

Die Beziehung zwischen $\sim_U$ und $\sim'_U$ ist wie folgt gegeben:

(c) Ist T ein Vertretersystem der Äquivalenzklassen von $\sim_U$, so ist $T' := T^{-1}$ ein Vertretersystem der Äquivalenzklassen von $\sim'_U$. Es gilt also $|G/U| = |U\backslash G|$.

Denn: Sei $a \in G$. Dann gibt es ein $t \in T$ mit $a^{-1} \in t \cdot U$, also $a^{-1} = t \cdot u$ mit $u \in U$, und damit $a = (t \cdot u)^{-1} = u^{-1} \cdot t^{-1} \in U \cdot t^{-1}$. Also folgt bereits $G = \bigcup_{t \in T} U \cdot t^{-1}$. Diese Vereinigung ist disjunkt, denn seien $t_1, t_2 \in T$ mit $U \cdot t_1^{-1} \cap U \cdot t_2^{-1} \neq \varnothing$. Dann gibt es $u, v \in U$ mit $u \cdot t_1^{-1} = v \cdot t_2^{-1}$, also $t_1^{-1} \cdot t_2 = u^{-1} \cdot v \in U$, d.h., $t_1 \sim_U t_2$ und damit $t_1 = t_2$.

Satz 3.3 (Lagrange)**.** *Sei* $|G| < \infty$*. Ist* $U \leqslant G$ *eine Untergruppe, so gilt* $|G| = |U| \cdot |G/U|$*; insbesondere ist* $|U|$ *ein Teiler von* $|G|$*. Dann heißt* $[G : U] := |G/U|$ *der* ***Index*** *von* U *in* G*. (Beachte: Wegen Bemerkung 3.2(c) gilt auch* $[G : U] = |U\backslash G|$*.)*

Beweis. Sei $T = \{t_1, \ldots, t_r\}$ ein Vertretersystem wie in Satz 3.1. Dann gilt $|G/U| = r$ und $|G| = \sum_{i=1}^{r} |U \cdot t_i|$. Wegen $|U| = |U \cdot t_i|$ für alle i folgt also $|G| = r|U|$. □

Sei $g \in G$. Für jedes $m \in \mathbb{Z}$ können wir die Potenz $g^m \in G$ bilden (mit den Konventionen: $g^3 = g \cdot g \cdot g$, $g^0 = 1_G$, $g^{-5} = (g^{-1})^5$, usw.). Es gilt dann $g^{n+m} = g^n \cdot g^m$ und $g^{nm} = (g^n)^m$ für alle $m, n \in \mathbb{Z}$, und man sieht sofort, dass $\langle g \rangle := \{g^m \mid m \in \mathbb{Z}\}$ eine Untergruppe von G ist; diese heißt die von g erzeugte ***zyklische Untergruppe*** von G. Gilt $G = \langle g \rangle$ für ein

$g \in G$, so bezeichnen wir G als ***zyklische Gruppe*** und sagen, dass G von g erzeugt wird. Dies sind gewissermaßen die Gruppen mit der einfachst möglichen Struktur; offensichtlich sind zyklische Gruppen abelsch.
Standardbeispiele von zyklischen Gruppen: $(\mathbb{Z}, +)$ und $(\mathbb{Z}/m\mathbb{Z}, +)$ (mit $m \in \mathbb{N}$); diese werden von 1 bzw. $\bar{1}$ erzeugt (beachte, dass die Verknüpfungen hier additiv geschrieben werden).

Definition 3.4. Sei $g \in G$. Gibt es ein $m \in \mathbb{N}$ mit $g^m = 1_G$, so definieren wir

$$o(g) := \min\{m \in \mathbb{N} \mid g^m = 1_G\};$$

gibt es kein solches m, so setzen wir $o(g) = \infty$. Dann heißt $o(g)$ die ***Ordnung von*** g. Diese kann sowohl endlich oder unendlich sein.

Ist zum Beispiel $G = \mathrm{GL}_2(\mathbb{Q})$ und $g = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in G$, so gilt $g^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$ für alle $n \in \mathbb{Z}$, also ist $o(g) = \infty$. Ist dagegen $g = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$, so erhalten wir $o(g) = 6$, denn

$$g^2 = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}, \quad g^3 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad g^4 = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}, \quad g^5 = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}, \quad g^6 = I_2.$$

Lemma 3.5. *Sei $g \in G$ fest.*
(a) *Sei $m \in \mathbb{N}$ mit $g^m = 1_G$; dann gilt $o(g) \mid m$.*
(b) *Ist $o(g) < \infty$, so gilt $o(g) = |\langle g \rangle|$ und $\langle g \rangle = \{1_G, g, g^2, \ldots, g^{o(g)-1}\}$.*
(c) *Ist $o(g) = \infty$, so gilt auch $|\langle g \rangle| = \infty$ und $g^i \neq g^j$ für alle $i, j \in \mathbb{Z}$, $i \neq j$.*

Beweis. (a) Sei $d := o(g) < \infty$. Division mit Rest ergibt $m = qd + r$ mit $q, r \in \mathbb{Z}$ und $0 \leqslant r < d$. Damit ist $g^m = g^{qd+r} = (g^d)^q \cdot g^r = 1_G^q \cdot g^r = g^r$. Da $g^m = 1_G$ gilt, ist also auch $g^r = 1_G$ und damit $r = 0$ wegen der Minimalität von d. Also gilt $o(g) = d \mid m$.

(b) Sei wieder $d := o(g)$. Sei $i \in \mathbb{Z}$ beliebig. Division mit Rest ergibt $i = dq + r$ mit $q, r \in \mathbb{Z}$ und $0 \leqslant r < d$. Wie oben folgt $g^i = g^{dq+r} = g^r$. Damit ist $\langle g \rangle = \{g^i \mid i \in \mathbb{Z}\} = \{1, g, g^2, \ldots, g^{d-1}\}$. Für $0 \leqslant i < j \leqslant d - 1$ ist $g^i \neq g^j$, denn sonst wäre $g^{j-i} = 1_G$ mit $1 \leqslant j - i < d$, Widerspruch zur Minimalität von d. Also ist $|\langle g \rangle| = d = o(g)$.

(c) Ist $|\langle g \rangle| < \infty$, so kann die Abbildung $\mathbb{Z} \to \langle g \rangle$, $m \mapsto g^m$, nicht injektiv sein. Daher gibt es $i > j$ in $\mathbb{Z}$ mit $g^i = g^j$, also $g^{i-j} = 1_G$ und damit $o(g) < \infty$. □

Folgerung 3.6. *Sei G eine endliche Gruppe, also $|G| < \infty$. Für alle $g \in G$ ist dann $o(g) < \infty$ ein Teiler von $|G|$ und es gilt $g^{|G|} = 1$.*

Beweis. Wäre $o(g) = \infty$, so auch $|\langle g \rangle| = \infty$ nach Lemma 3.5(c), und damit $|G| = \infty$, Widerspruch. Nach Lemma 3.5(b) hat die Untergruppe $\langle g \rangle \leqslant G$ genau $o(g)$ Elemente, also ist $o(g)$ ein Teiler von $|G|$ nach dem Satz von Lagrange. Schreiben wir $|G| = o(g)m$ mit $m \in \mathbb{N}$, so folgt also $g^{|G|} = (g^{o(g)})^m = 1_G^m = 1_G$. □

Bemerkung 3.7. Sei $g \in G$ mit $o(g) < \infty$. Behauptung: Für $d \in \mathbb{N}$ mit $d \mid o(g)$ gilt $o(g^d) = o(g)/d$. Dazu: Sei $m := o(g^d)$ und $n := o(g)/d$. Dann ist $(g^d)^n = g^{dn} = g^{o(g)} = 1_G$, also $m = o(g^d) \mid n$ nach Lemma 3.5. Umgekehrt ist $1_G = (g^d)^m = g^{dm}$, also $o(g) \mid dm$ (wiederum nach Lemma 3.5) und damit $n = o(g)/d \mid m$. Also folgt $n = m$.

Allein mit dem Satz von Lagrange und den obigen Folgerungen können wir bereits eine Reihe von Beispielen untersuchen.

Beispiel 3.8. (a) Sei $|G| = p$ eine Primzahl. Dann ist G zyklisch, und G besitzt überhaupt nur die Untergruppen $\{1_G\}$ und G. Denn ist $U \leqslant G$, so ist $|U|$ ein Teiler von $p = |G|$, also $|U| = 1$ oder $|U| = p$, also $U = \{1_G\}$ oder $U = G$. Ist $g \neq 1_G$, so folgt damit $\{1_G\} \neq \langle g \rangle = G$.

(b) Sei $|G| = 4$. Dann ist $o(g) \in \{1, 2, 4\}$ für alle $g \in G$. Gibt es ein $g \in G$ mit $o(g) = 4$, so ist $G = \langle g \rangle$ zyklisch. Sonst gilt $g^2 = 1_G$ für alle $g \in G$. Damit ist $G = \{1_G, x, y, z\}$ mit folgender Multiplikationstabelle:

	1_G	x	y	z
1_G	1_G	x	y	z
x	x	1_G	z	y
y	y	z	1_G	x
z	z	y	x	1_G

Dazu: Betrachte $xy \in G$. Dies muss also gleich 1_G, x, y oder z sein. Ist $xy = 1_G$, so $x = y^{-1} = y$ (wegen $y^2 = 1_G$), Widerspruch; ist $xy = x$, so $y = 1_G$, Widerspruch; ist $xy = y$, so $x = 1_G$, Widerspruch. Also muss $xy = z$ gelten. Analog findet man alle anderen Produkte. Insbesondere sieht man, dass G abelsch ist. Ein solches G heißt ***Klein'sche Vierergruppe***.

Beispiel 3.9. Sei $G = S_3$. Dann besteht G genau aus den 6 Permutationen:

$$\mathrm{id} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix},\ \sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},\ \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix},\ \sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},\ \pi = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix},\ \pi' = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

Wir berechnen $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \mathrm{id}$, $\pi^2 = \pi'$ und $\pi^3 = \pi'^3 = \mathrm{id}$. Also gilt $o(\sigma_1) = o(\sigma_2) = o(\sigma_3) = 2$ und $o(\pi) = o(\pi') = 3$. Damit erhalten wir folgende zyklische Untergruppen:

$$U_i = \langle \sigma_i \rangle = \{\mathrm{id}, \sigma_i\} \quad \text{für } i = 1, 2, 3, \qquad V = \langle \pi \rangle = \langle \pi' \rangle = \{\mathrm{id}, \pi, \pi'\}.$$

Wir behaupten, dass dies alle echten Untergruppen von G sind. Sei also $U' \leqslant G$ eine beliebige Untergruppe mit $\{\mathrm{id}\} \neq U' \neq G$. Nach Lagrange ist $|U'|$ ein Teiler von 6, also $|U'| = 2$ oder 3. Nach Beispiel 3.8(a) ist U' zyklisch und muss daher eine der obigen Untergruppen sein.

Beispiel 3.10. Wir betrachten die folgenden Elemente in $G = \mathrm{GL}_2(\mathbb{C})$:

$$E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad J = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad K = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},$$

(wobei wie üblich $i = \sqrt{-1}$). Man rechnet sofort nach, dass $I^2 = J^2 = K^2 = I \cdot J \cdot K = -E$ gilt; daraus erhält man $I^{-1} = -I$, $J^{-1} = -J$, $K^{-1} = -K$ sowie $I \cdot J = -K^{-1} = K$, $J \cdot K = -I^{-1} = I$, $I \cdot K = -J$. Damit sieht man auch leicht, dass $Q_8 := \{\pm E, \pm I, \pm J, \pm K\}$ eine Untergruppe von $GL_2(\mathbb{C})$ ist. Diese Gruppe wird als ***Quaternionengruppe*** bezeichnet; sie hat genau 8 Elemente und ist nicht abelsch (zum Beispiel ist $I \cdot J = K$ und $J \cdot I = -K$). Die obigen Rechnungen zeigen, dass $o(-E) = 2$ und $o(\pm I) = o(\pm J) = o(\pm K) = 4$ gilt. Wir erhalten also die folgenden zyklischen Untergruppen:

$$U_1 = \langle I \rangle = \{\pm E, \pm I\}, \quad U_2 = \langle J \rangle = \{\pm E, \pm J\}, \quad U_3 = \langle K \rangle = \{\pm E, \pm K\}, \quad Z = \langle -E \rangle = \{\pm E\}.$$

Wir behaupten wiederum, dass dies alle echten Untergruppen von G sind. Sei also $U' \leqslant Q_8$ eine beliebige Untergruppe mit $\{\mathrm{id}\} \neq U' \neq Q_8$. Nach Lagrange ist $|U'|$ ein Teiler von 8, also $|U'| = 2$ oder 4. Ist $|U'| = 2$, so ist U' zyklisch, also muss $U' = Z$ gelten, denn es gibt nur ein Element der Ordnung 2 (nämlich $-E$). Sei nun $|U'| = 4$. Ist U' zyklisch, so muss $U' = U_i$ für $i \in \{1, 2, 3\}$ gelten, denn jedes Element der Ordnung 4 liegt in einer dieser Untergruppen. Bleibt also noch die Möglichkeit, dass U' nicht zyklisch ist. Nach Beispiel 3.8(b) enthält dann U' aber 3 Elemente der Ordnung 2, Widerspruch.

Beispiel 3.11. Sei G eine zyklische Gruppe, also $G = \langle g \rangle$ mit einem $g \in G$. Dann gilt:
(a) Ist $U \leqslant G$ beliebige Untergruppe, so ist U zyklisch; es gibt ein $m \in \mathbb{N}_0$ mit $U = \langle g^m \rangle$.
(b) Ist $n := |G| < \infty$ und $d \in \mathbb{N}$ mit $d \mid n$, so gibt es genau eine Untergruppe $U_d \leqslant G$ mit $|U_d| = d$; diese ist gegeben durch $U_d = \langle g^{n/d} \rangle$.

Zu (a): Ist $U = \{1_G\}$, so gilt die Behauptung mit $m = 0$. Sei nun $U \neq \{1_G\}$; es gibt also ein $i \in \mathbb{Z}$, $i \neq 0$, mit $g^i \in U$. Ist $i < 0$, so gilt auch $g^{-i} = (g^{-1})^i \in U$. In jedem Fall existiert also ein $m \in \mathbb{N}$ mit $g^m \in U$; sei m minimal mit dieser Eigenschaft. Wegen $g^m \in U$ ist auch $\langle g^m \rangle \subseteq U$. Umgekehrt sei $u \in U$ beliebig. Schreibe $u = g^j$ mit $j \in \mathbb{Z}$. Division mit Rest ergibt $j = qm + r$ mit $q, r \in \mathbb{Z}$ und $0 \leqslant r < m$. Dann ist $g^r = g^{j-qm} = g^j \cdot (g^m)^{-q} \in U$. Ist $r > 0$, so erhalten wir einen Widerspruch zur Minimalität von m. Also muss $r = 0$ gelten und damit $u = g^j = (g^m)^q \in \langle g^m \rangle$. Also gilt auch $U \subseteq \langle g^m \rangle$, und damit Gleichheit.

Zu (b): Nach Lemma 3.5 ist $n = o(g)$. Sei nun $d \in \mathbb{N}$ mit $d \mid n$. Nach Bemerkung 3.7 ist $o(g^d) = n/d$, also ist $U_d := \langle g^{n/d} \rangle \leqslant G$ eine Untergruppe mit $|U_d| = d$ (wiederum nach Lemma 3.5). Sei nun $U \leqslant G$ eine beliebige Untergruppe mit $|U| = d$. Nach (a) ist $U = \langle g^m \rangle$ mit einem $m \in \mathbb{N}_0$. Nach Folgerung 3.6 gilt $1_G = (g^m)^d = g^{md}$, also folgt $n \mid md$ mit Lemma 3.5, d.h., $md = an$ mit $a \in \mathbb{N}_0$ und damit $m = (n/d)a$. Also ist $g^m = (g^{n/d})^a \in U_d$ und damit $U = \langle g^m \rangle \subseteq U_d$. Wegen $|U| = |U_d| = d$ folgt schließlich $U = U_d$.

4. *Die Eulersche Phi-Funktion*

In diesem Abschnitt erinnern wir zunächst an einige Begriffe und Definitionen aus der elementaren Zahlentheorie. Seien $d, n \in \mathbb{Z}$ mit $d \neq 0$ und $n \neq 0$. Gilt $d \mid n$, so folgt natürlich auch $(-d) \mid n$. Um alle Teiler d von n zu bestimmen, brauchen wir also nur den Fall $d > 0$ zu betrachten. Sei nun $d > 0$. Aus $d \mid n$ folgt offenbar auch $d \leqslant |n|$; also hat n nur endlich viele Teiler. Sind $n, m \in \mathbb{Z}$ mit $n \neq 0$ oder $m \neq 0$ gegeben, so definieren wir

$$\mathrm{ggT}(n, m) := \max\{a \in \mathbb{N} \mid \text{es gilt } a \mid n \text{ und } a \mid m\}$$

als den ***größten gemeinsamen Teiler*** von n und m. Gilt $\mathrm{ggT}(n, m) = 1$, so bezeichnen wir m und n als ***teilerfremd***.

Lemma 4.1 (Lemma von Bézout)**.** *Gegeben seien $n, m \in \mathbb{Z}$ mit $n \neq 0$ oder $m \neq 0$. Dann gibt es $a, b \in \mathbb{Z}$ mit $\mathrm{ggT}(n, m) = an + bm$. Ist auch $d' \in \mathbb{Z}$ ein gemeinsamer Teiler von n und m, so folgt $d' \mid \mathrm{ggT}(n, m)$.*

Beweis. Ist $n = 0$ oder $m = 0$, so ist die Aussage sehr einfach zu sehen. (Ist z.B. $n = 0$ und $m < 0$, so ist $-m = \mathrm{ggT}(n, m) = 0 \cdot n + (-1) \cdot m$.) Sei also jetzt $n \neq 0$ und $m \neq 0$. Wir beschreiben einen Algorithmus, genannt (erweiterter) ***Euklidischer Algorithmus***, zur Bestimmung von $\mathrm{ggT}(n, m)$ und $a, b \in \mathbb{Z}$ mit $\mathrm{ggT}(n, m) = an + bm$. Dazu berechnen wir rekursiv eine endliche Folge von Tripeln (r_k, a_k, b_k) für $k = 0, 1, 2, 3, \ldots$, wie folgt. Ist $n > 0$ und $m > 0$, so initialisieren wir

$$r_0 := n,\ a_0 := 1,\ b_0 := 0 \qquad \text{und} \qquad r_1 := m,\ a_1 := 0,\ b_1 := 1.$$

(Ist $n < 0$, so setze $r_0 := -n, a_0 := -1, b_0 := 0$; ist $m < 0$, so setze $r_1 := -m$, $a_1 := 0$, $b_1 := -1$.) In jedem Fall gilt dann $r_0 = a_0 n + b_0 m \geqslant 1$ und $r_1 = a_1 n + b_1 m \geqslant 1$. Sei nun $k \geqslant 1$ und r_i, a_i, b_i bereits konstruiert für $0 \leqslant i \leqslant k$, wobei jeweils $r_i = a_i n + b_i m \geqslant 1$ gelte. Division mit Rest liefert

$$r_{k-1} = q_k r_k + r_{k+1} \qquad \text{mit} \qquad q_k, r_{k+1} \in \mathbb{Z} \quad \text{und} \quad 0 \leqslant r_{k+1} < r_k;$$

dies definiert r_{k+1}; dann setze $a_{k+1} := a_{k-1} - q_k a_k$ und $b_{k+1} := b_{k-1} - q_k b_k$. Damit gilt wieder

$$r_{k+1} = r_{k-1} - q_k r_k = (a_{k-1} n + b_{k-1} m) - q_k (a_k n + b_k m) = a_{k+1} n + b_{k+1} m.$$

Dieses Verfahren wird solange fortgesetzt, bis $r_{k+1} = 0$ gilt. (Wegen $r_1 > r_2 > \ldots \geqslant 0$ muss es ein solches k geben.) Dann ist $r_k > 0$ und $r_{k-1} = q_k r_k$. Im vorherigen Schritt ist $r_{k-2} = q_{k-1} r_{k-1} + r_k$; wegen $r_k \mid r_{k-1}$ folgt also auch $r_k \mid r_{k-2}$. Dies setzt sich entsprechend in alle vorherigen Schritte fort, also gilt $r_k \mid r_i$ für $0 \leqslant i \leqslant k-1$. Insbesondere ist $r_k \mid n = \pm r_0$ und $r_k \mid m = \pm r_1$, also $r_k \leqslant \mathrm{ggT}(n, m)$. Wegen $r_k = a_k n + b_k m$ folgt aber auch $d \mid r_k$ für

jeden gemeinsamen Teiler d von n und m. Also ist $\mathrm{ggT}(n,m) = r_k = a_k n + b_k m$. (Für mehr Details siehe `https://en.wikipedia.org/wiki/Extended_Euclidean_algorithm`). □

Sei zum Beispiel $n = 1071$ und $m = 462$. Dann initialisieren wir $r_0 = 1071$, $a_0 = 1$, $b_0 = 0$ und $r_1 = 462$, $a_1 = 0$, $b_1 = 1$. Das obige Verfahren liefert nun nacheinander:

$$\begin{aligned}
r_0 &= 1071 = 2 \cdot 462 + 147 = q_1 r_1 + r_2, && \text{also } q_1 = 2,\ r_2 = 147 \text{ und } a_2 = 1,\ b_2 = -2,\\
r_1 &= \ \ 462 = 3 \cdot 147 + 21 \ \ = q_2 r_2 + r_3, && \text{also } q_2 = 3,\ r_3 = 21 \ \ \text{und } a_3 = -3,\ b_3 = 7,\\
r_2 &= \ \ 147 = 7 \cdot 21 + 0 \ \ \ \ = q_3 r_3 + r_4, && \text{also } q_3 = 7,\ r_4 = 0.
\end{aligned}$$

Damit bricht das Verfahren bei $k = 3$ mit $r_3 = 21$, $a_3 = -3$, $b_3 = 7$ ab und wir erhalten $21 = \mathrm{ggT}(1071, 462) = (-3) \cdot 1071 + 7 \cdot 462$. Versuchen Sie, dieses Verfahren möglichst effizient zu programmieren (in Python oder einer beliebigen anderen Programmiersprache).

Bemerkung 4.2. Seien $m_1, m_2, n \in \mathbb{Z}$, $m_1 \neq 0$ und $m_2 \neq 0$. Dann gilt:

(a) $\quad \mathrm{ggT}(m_i, n) = 1 \quad \text{für } i = 1,2 \quad \Rightarrow \quad \mathrm{ggT}(m_1 m_2, n) = 1,$

(b) $\quad \mathrm{ggT}(m_1, m_2) = 1 \quad \text{und} \quad m_i \mid n \text{ für } i = 1,2 \quad \Rightarrow \quad m_1 m_2 \mid n.$

Diese Aussagen folgen sofort aus der eindeutigen Zerlegung von ganzen Zahlen in Primfaktoren, aber man kann sie auch sehr leicht direkt mit Bézouts Lemma zeigen.
Zu (a): Es gibt $a_i, b_i \in \mathbb{Z}$ mit $a_i m_i + b_i n = 1$ für $i = 1,2$. Dann erhalten wir auch eine Gleichung $1 = (a_1 m_1 + b_1 n)(a_2 m_2 + b_2 n) = a_1 a_2 m_1 m_2 + (a_1 m_1 b_2 + b_1 a_2 m + b_1 b_2 n)n$. Ist $d \in \mathbb{N}$ ein Teiler von $m_1 m_2$ und n, so teilt d die rechte Seite der Gleichung und damit $d = 1$.
Zu (b): Es gibt $a, b \in \mathbb{Z}$ mit $1 = a m_1 + b m_2$; außerdem ist $n = c_i m_i$ mit $c_i \in \mathbb{Z}$ für $i = 1,2$. Dann erhalten wir $n = a m_1 n + b m_2 n = a m_1 c_2 m_2 + b m_2 c_1 m_1$, also $m_1 m_2 \mid n$.

Satz 4.3. *Sei $m \in \mathbb{N}$ und $a \in \mathbb{Z}$. Dann ist $\bar{a} \in (\mathbb{Z}/m\mathbb{Z})^\times$ genau dann, wenn $\mathrm{ggT}(a,m) = 1$ gilt. Insbesondere ist $\mathbb{Z}/m\mathbb{Z}$ genau dann ein Körper, wenn m eine Primzahl ist.*
Ist p eine Primzahl, so schreiben wir auch $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$.

Beweis. Für $a \in \mathbb{Z}$ gelten die folgenden Äquivalenzen:

$$\begin{aligned}
\bar{a} \in (\mathbb{Z}/m\mathbb{Z})^\times \quad &\Leftrightarrow \quad \bar{1} = \bar{a} \cdot \bar{b} = \overline{ab} \text{ für ein } b \in \mathbb{Z} \quad \Leftrightarrow \quad m \mid ab - 1 \text{ für ein } b \in \mathbb{Z}\\
&\Leftrightarrow \quad 1 = ab + cm \text{ für ein } b \in \mathbb{Z} \text{ und ein } c \in \mathbb{Z}.
\end{aligned}$$

Nach Bézouts Lemma ist aber die letzte Bedingung dazu äquivalent, dass $\mathrm{ggT}(a,m) = 1$ ist. Sei nun m eine Primzahl. Ist $\bar{a} \neq \bar{0}$, so gilt $m \nmid a$ und damit $\mathrm{ggT}(a,m) = 1$, also $\bar{a} \in (\mathbb{Z}/m\mathbb{Z})^\times$. D.h., jedes Element ungleich $\bar{0}$ in $\mathbb{Z}/m\mathbb{Z}$ ist eine Einheit. Also ist $\mathbb{Z}/m\mathbb{Z}$ ist ein Körper. Umgekehrt: Ist m keine Primzahl, d.h., $m = m_1 m_2$ mit $0 < m_1, m_2 < m$, so

gilt $\bar{m}_1 \cdot \bar{m}_2 = \bar{m} = \bar{0}$, aber $\bar{m}_1 \neq \bar{0}$ und $\bar{m}_2 \neq \bar{0}$. Wäre $\mathbb{Z}/m\mathbb{Z}$ ein Körper, so existiert $\bar{m}_1^{-1} \in \mathbb{Z}/m\mathbb{Z}$. Aber dann folgt $\bar{0} = \bar{m}_1^{-1} \cdot \bar{0} = \bar{m}_1^{-1} \cdot (\bar{m}_1 \cdot \bar{m}_2) = \bar{m}_2$, Widerspruch. □

Definition 4.4. Für $m \in \mathbb{N}$ setze $\phi(m) := |(\mathbb{Z}/m\mathbb{Z})^\times|$. Wegen $\mathbb{Z}/m\mathbb{Z} = \{\bar{0}, \bar{1}, \ldots, \overline{m-1}\}$ ist $\phi(m)$ also nach Satz 4.3 die Anzahl der $i \in \{0, 1, \ldots, m-1\}$ mit $\mathrm{ggT}(i, m) = 1$. Diese Funktion heißt ***Eulersche Phi-Funktion***. Hier sind einige Werte:

m	1	2	3	4	5	6	7	8	9	10	11	12	...
$\phi(m)$	1	1	2	2	4	2	6	4	6	4	10	4	...

Es gilt stets $\phi(m) > 0$ (denn $\bar{1} \in (\mathbb{Z}/m\mathbb{Z})^\times$, und dies gilt auch für $m = 1$). Ist p eine Primzahl, so gilt offenbar $\phi(p) = p - 1$.

Folgerung 4.5 (Satz von Euler)**.** *Sei* $m \in \mathbb{N}$. *Dann gilt* $a^{\phi(m)} \equiv 1 \bmod m$ *für alle* $a \in \mathbb{Z}$ *mit* $\mathrm{ggT}(a, m) = 1$.

Beweis. Sei $a \in \mathbb{Z}$ und $\mathrm{ggT}(a, m) = 1$. Nach Satz 4.3 ist also $\bar{a} \in (\mathbb{Z}/m\mathbb{Z})^\times$. Wegen $|(\mathbb{Z}/m\mathbb{Z})^\times| = \phi(m)$ erhalten wir $\bar{a}^{\phi(m)} = \bar{1}$ mit Folgerung 3.6, also die Behauptung. □

Folgerung 4.6 (Kleiner Satz von Fermat)**.** *Ist* p *eine Primzahl, so gilt* $a^p \equiv a \bmod p$ *für alle* $a \in \mathbb{Z}$.

Beweis. Sei zuerst $p \nmid a$, also $\mathrm{ggT}(a, p) = 1$. Wegen $\phi(p) = p - 1$ folgt $a^{p-1} \equiv 1 \bmod p$ aus dem Satz von Euler, also auch $a^p \equiv a \bmod p$. Ist $p \mid a$, so gilt auch $p \mid a^p$ und damit ebenfalls $a^p \equiv 0 \equiv a \bmod p$. □

Beispiel 4.7. Eine ***Mersenne-Zahl*** ist eine Zahl der Form $2^n - 1$ mit $n \in \mathbb{N}$. Diese Zahlen sind besonders geeignet, um große Primzahlen zu finden. Ist nämlich p eine Primzahl, so gilt $q > p$ für alle Primzahlen q mit $q \mid 2^p - 1$. (Insbesondere liefert dies einen weiteren Beweis, dass es keine größte Primzahl gibt.)
Beweis: Ist $q \mid 2^p - 1$, so gilt also $\bar{2}^p = \bar{1}$ in $\mathbb{Z}/q\mathbb{Z}$. Damit ist $\bar{2}$ eine Einheit in $\mathbb{Z}/q\mathbb{Z}$. Wegen $\bar{2}^p = \bar{1}$ folgt $o(\bar{2}) \mid p$; siehe Lemma 3.5. Wegen $\bar{2} \neq \bar{1}$ und weil p eine Primzahl ist, gilt $o(\bar{2}) = p$. Nach Satz 4.3 ist $\mathbb{Z}/q\mathbb{Z}$ ein Körper, also $|(\mathbb{Z}/q\mathbb{Z})^\times| = q - 1$. Mit Folgerung 3.6 folgt dann $p = o(\bar{2}) \mid q - 1$ und damit $p < q$.

Wir zeigen nun noch eine weitere Charakterisierung von $\phi(n)$, wobei wir die Aussagen zu zyklischen Gruppen aus dem letzten Abschnitt verwenden. Im Folgenden werden wir mehrfach Lemma 3.5 verwenden, ohne es jeweils stets explizit zu zitieren.

Lemma 4.8. *Sei* G *eine Gruppe und* $g \in G$ *mit* $d := o(g) < \infty$. *Sei* $i \in \mathbb{Z}$. *Dann gilt*

$$\langle g \rangle = \langle g^i \rangle \quad \Leftrightarrow \quad o(g) = o(g^i) \quad \Leftrightarrow \quad \mathrm{ggT}(i, d) = 1.$$

Beweis. Wegen $g^i \in \langle g \rangle$ ist $\langle g^i \rangle \subseteq \langle g \rangle$. Also folgt $\langle g \rangle = \langle g^i \rangle \Leftrightarrow |\langle g \rangle| = |\langle g^i \rangle| \Leftrightarrow o(g) = o(g^i)$. Es bleibt also noch die Äquivalenz mit $\mathrm{ggT}(i, d) = 1$ zu zeigen. Sei zuerst $\langle g \rangle = \langle g^i \rangle$. Dann ist $g = (g^i)^m = g^{im}$ für ein $m \in \mathbb{Z}$. Daraus folgt $g^{im-1} = 1_G$, also $d = o(g) \mid im - 1$ und damit $im - 1 = ad$ mit $a \in \mathbb{Z}$. Dann ist $1 = im - ad$, also $\mathrm{ggT}(i, d) = 1$. Sei nun $\mathrm{ggT}(i, d) = 1$. Nach Bézout gibt es $a, b \in \mathbb{Z}$ mit $1 = ai + bd$. Dann folgt $g = g^1 = g^{ai+bd} = (g^i)^a \cdot (g^d)^b = (g^i)^a \cdot 1_G = (g^i)^a \in \langle g^i \rangle$, also auch $\langle g \rangle \subseteq \langle g^i \rangle$. Die umgekehrte Inklusion ist klar, also gilt Gleichheit. □

Folgerung 4.9. *Sei* G *eine zyklische Gruppe mit* $n := |G| < \infty$*. Dann ist* $\phi(n)$ *gleich der Anzahl der* $x \in G$ *mit* $G = \langle x \rangle$*.*

Beweis. Sei $g \in G$ fest mit $G = \langle g \rangle$. Dann ist $G = \{g^i \mid 0 \leqslant i \leqslant n-1\}$ und $n = o(g)$. Sei $x \in G$, also $x = g^i$ mit $0 \leqslant i \leqslant n-1$. Nach Lemma 4.8 gilt $G = \langle x \rangle = \langle g^i \rangle \Leftrightarrow \mathrm{ggT}(i, n) = 1$. Also ist die Anzahl der $x \in G$ mit $G = \langle x \rangle$ genau gleich der Anzahl der $i \in \{0, 1, \ldots, n-1\}$ mit $\mathrm{ggT}(i, n) = 1$, also gleich $\phi(n)$. □

Ab hier Woche 3

Schließlich können wir die folgende Umkehrung von Beispiel 3.11 zeigen, die auch noch viel später in Kapitel III (siehe dort Satz 12.8) eine Anwendung finden wird.

Satz 4.10. *Sei* G *eine Gruppe mit* $n := |G| < \infty$*. Gibt es zu jedem* $d \in \mathbb{N}$ *mit* $d \mid n$ *höchstens eine Untergruppe* $U \leqslant G$ *mit* $|U| = d$*, so ist* G *zyklisch.*

Beweis. Sei D die Menge aller $d \in \mathbb{N}$ mit $d \mid n$. Für $d \in D$ setze $X_d := \{g \in G \mid o(g) = d\}$ (eine Teilmenge von G). Da $o(g) \mid n$ für alle $g \in G$, folgt $G = \bigcup_{d \in D} X_d$; außerdem ist natürlich $X_d \cap X_{d'} = \varnothing$ für $d \neq d'$ in D. Damit folgt

$$n = |G| = \sum_{d \in D} |X_d|.$$

Sei nun $d \in D$ fest. Nehmen wir an, es gilt $X_d \neq \varnothing$; dann gibt es also ein $g \in X_d$, d.h., $U := \langle g \rangle$ ist eine zyklische Untergruppe der Ordnung $d = o(g)$.
Behauptung: $X_d = \{x \in U \mid U = \langle x \rangle\}$, und damit $|X_d| = \phi(d)$ (nach Folgerung 4.9).
Dazu: Für $x \in X_d$ ist $U' := \langle x \rangle$ ebenfalls eine Untergruppe der Ordnung d; nach Voraussetzung gilt also $U' = U$, d.h., $x \in U$ und $U = \langle x \rangle$. Also gilt "$\subseteq$". Ist umgekehrt $x \in U$ mit $U = \langle x \rangle$, so ist $o(x) = |\langle x \rangle| = |U| = d$ also $x \in X_d$. Also gilt auch "$\supseteq$".
Aus der obigen Summenformel und der obigen Behauptung erhalten wir daher

$$(*) \qquad n = \sum_{d \in D'} \phi(d) \qquad \text{wobei} \qquad D' := \{d \in D \mid X_d \neq \varnothing\}.$$

Wir wenden nun die ganze obige Diskussion auf eine spezielle Gruppe an, nämlich $(\mathbb{Z}/n\mathbb{Z}, +)$. Diese wird von $\bar{1}$ erzeugt, ist also zyklisch. Außerdem gibt es nach Beispiel 3.11 zu jedem $d \in D$ genau eine Untergruppe der Ordnung d, und diese ist sogar selbst zyklisch. Also ist $X_d \neq \varnothing$ für alle $d \in D$, und damit $D' = D$ für die Gruppe $(\mathbb{Z}/n\mathbb{Z}, +)$, d.h., aus $(*)$ erhalten wir die Formel $n = \sum_{d \in D} \phi(d)$.

Kehren wir nun zu unserer gegebenen Gruppe G zurück. Dann gilt also $n = \sum_{d \in D} \phi(d)$ und $n = \sum_{d \in D'} \phi(d)$. Wäre $D' \subsetneq D$, so erhielten wir $\sum_{d \in D \setminus D'} \phi(d) = 0$, Widerspruch weil $\phi(d) > 0$ für alle $d \in \mathbb{N}$ gilt. Also ist $D' = D$ und damit $X_d \neq \varnothing$ für alle $d \in D$. Inbesondere ist $X_n \neq \varnothing$, es gibt also ein $g \in G$ mit $o(g) = n$ und damit $G = \langle g \rangle$. □

Bemerkung 4.11. Sei $n \in \mathbb{N}$. Wir halten fest, dass wir im obigen Beweis sozusagen als Nebenprodukt die Summenformel $n = \sum_{d \in \mathbb{N} : d \mid n} \phi(d)$ für die Phi-Funktion gezeigt haben.

5. *Eine Anwendung: Das RSA-Verfahren*

Als eine Anwendung von Bézouts Lemma und dem Satz von Euler beschreiben wir nun das ***RSA-Verschlüsselungsverfahren***, das 1977 von R. Rivest, A. Shamir und L. Adleman entwickelt wurde; siehe auch `https://en.wikipedia.org/wiki/RSA_(cryptosystem)`.

Alice erwartet von Bob (übliche Namen in Texten zur Verschlüsselung) eine Nachricht, aber sie möchte vermeiden, dass andere auch diese Nachricht verstehen können (selbst wenn sie die Nachricht sehen können). Eine Nachricht ist hier einfach eine Zahl $m \in \mathbb{N}$, die man nach einem bestimmten Verfahren aus Buchstaben oder sonstigen Daten bildet.

Konkretes Beispiel: Alice erwartet von Bob einen Tipp, ob der VfB Stuttgart am nächsten Bundesliga-Spieltag gewinnt, verliert oder unentschieden spielt. Um dies zu vereinfachen, soll Bob einfach nur ein G, V oder U senden, also als Zahl 7, 22 oder 21 (Nummer im Alphabet).

Idee der Verschlüsselung. Alice wählt zwei Primzahlen $p \neq q$ und bildet das Produkt $N := pq$. Hierbei muss $m < N$ gelten, für alle m die man als Nachricht senden möchte. (Und generell wird die Verschlüsselung umso sicherer, je größer p, q und damit N sind; siehe die Diskussion weiter unten.) Dann berechnet Alice den Wert von Eulers Funktion $\phi(N)$, mit folgender einfacher Formel[1]:

Lemma 5.1. *Sei $N = pq$ mit Primzahlen $p \neq q$ wie oben. Dann gilt $\phi(N) = (p-1)(q-1)$.*

Beweis. Die einzigen $d \in \mathbb{N}$ mit $d \mid N$ sind $1, p, q$ und N. Mit der Summenformel in Bemerkung 4.11 erhalten wir also $pq = N = \phi(1) + \phi(p) + \phi(q) + \phi(N)$. Wegen $\phi(1) = 1$, $\phi(p) = p - 1$ und $\phi(q) = q - 1$ folgt $\phi(N) = pq - p - q + 1 = (p-1)(q-1)$. □

[1] Auf dem dritten Übungsblatt wird eine allgemeine Formel für n beliebig gezeigt.

Schließlich wählt Alice eine weitere Zahl $e \in \mathbb{N}$ mit

$$1 < e < \phi(N) \qquad \text{und} \qquad \mathrm{ggT}(e, \phi(N)) = 1,$$

also zum Beispiel eine weitere Primzahl, die $\phi(N)$ nicht teilt. Nun veröffentlicht Alice das Paar (e, N); dieses heißt daher auch "***public key***". (Aber Alice hält die Primzahlen p, q geheim.) Alle, die Alice eine Nachricht verschlüsselt schicken wollen, können nun wie folgt vorgehen: Ist $m \in \mathbb{N}$ die Nachricht ($1 \leqslant m < N$), so berechne m^e und dividiere dies mit Rest durch N; dies liefert ein eindeutiges $c \in \mathbb{Z}$ mit

$$m^e \equiv c \bmod N \qquad \text{und} \qquad 0 \leqslant c < N.$$

Dann verschicke c als verschlüsselte Nachricht. Wie kann Alice aus c, e, N die Zahl m (also die tatsächlich interessierende Nachricht) zurückberechnen? Nun, es gilt $\bar{m}^e = \bar{c}$ in $\mathbb{Z}/N\mathbb{Z}$; also müsste sie (oder auch jemand sonst) die e-te Wurzel von $\bar{c}$ in $\mathbb{Z}/N\mathbb{Z}$ berechnen — aber dafür ist kein effizientes Verfahren bekannt! Sie könnte zum Beispiel einfach alle $0 \leqslant a < N$ darauf testen, ob $a^e \equiv c \bmod N$ gilt, aber für große N ist dies nicht praktikabel (zu viele Multiplikationen und Divisionen mit Rest).

Idee der Entschlüsselung. Hier kommt nun die Wahl von e ins Spiel. Alice hatte diese Zahl so gewählt, dass $\mathrm{ggT}(e, \phi(N)) = 1$ gilt, d.h., $\bar{e}$ ist eine Einheit in $\mathbb{Z}/\phi(N)\mathbb{Z}$. Also gibt es ein eindeutiges $d \in \mathbb{Z}$ mit $1 \leqslant d < \phi(N)$ und $\bar{d} \cdot \bar{e} = \bar{1}$ in $\mathbb{Z}/\phi(N)\mathbb{Z}$. Dieses d wird als "***private key***" bezeichnet, und wird von Alice geheim gehalten. Sie kann d leicht wie folgt berechnen. Nach Bézouts Lemma gibt es $a, b \in \mathbb{Z}$ mit $1 = \mathrm{ggT}(e, \phi(N)) = ae + b\phi(N)$; dann folgt $\bar{a} \cdot \bar{e} = \bar{1}$ in $\mathbb{Z}/\phi(N)\mathbb{Z}$. Division mit Rest liefert ein $d \in \mathbb{Z}$ mit $a \equiv d \bmod \phi(N)$ und $0 \leqslant d < \phi(N)$; dann ist $\bar{a} = \bar{d}$ und $d \neq 0$, also ist d der gewünschte "private key". Die folgende Aussage zeigt nun, dass wir damit eine effiziente Methode erhalten, um die e-te Wurzel einer Zahl modulo N zu bestimmen.

Lemma 5.2. *Mit obigen Bezeichnungen gilt* $(m^e)^d \equiv m \bmod N$.

Beweis. Wir zeigen zuerst $(m^e)^d \equiv m \bmod p$, d.h., $p \mid (m^e)^d - m$. Ist $p \mid m$, so auch $p \mid (m^e)^d$ und damit $(m^e)^d \equiv 0 \equiv m \bmod p$, wie gewünscht. Sei nun $p \nmid m$, also $\mathrm{ggT}(p, m) = 1$. Nun ist $\phi(p) = p - 1$. Aus dem Satz von Euler folgt also $m^{p-1} \equiv 1 \bmod p$. Die Zahl d war so definiert, dass $\bar{e} \cdot \bar{d} = \bar{1}$ in $\mathbb{Z}/\phi(N)\mathbb{Z}$ gilt, also $\phi(N) \mid ed - 1$. Nach Lemma 5.1 ist $\phi(N) = (p-1)(q-1)$. Also ist auch $p - 1 \mid ed - 1$ und wir schreiben $ed - 1 = a(p-1)$ mit $a \in \mathbb{N}$. Damit erhalten wir

$$(m^e)^d \equiv m^{ed} \equiv m^{ed-1+1} \equiv m^{a(p-1)+1} \equiv (m^{p-1})^a m \equiv 1^a m \equiv m \bmod p,$$

wie gewünscht. Auf völlig analoge Weise zeigt man dann auch $(m^e)^d \equiv m \bmod q$, d.h., $q \mid (m^e)^d - m$. Da p, q verschiedene Primzahlen sind, gilt $\mathrm{ggT}(p, q) = 1$ und mit Bemerkung 4.2(b) folgt schließlich auch $pq \mid (m^e)^d - m$, also die Behauptung. □

Mit dem "private key" d kann Alice nun leicht die empfangene Nachricht c entschlüsseln: Sie bildet c^d und dividiert dies mit Rest durch N; wegen $c \equiv m^e$ ist der Rest genau die Nachricht m; siehe Lemma 5.2.

Wo liegt die Sicherheit dieses Verfahrens? Alice hält den private key d geheim, sowie die Primzahlen p und q, mit denen N gebildet wurde; sie kann d leicht mit dem Euklidischen Algorithmus und Bézouts Lemma ausrechnen, weil sie $\phi(N) = (p-1)(q-1)$ kennt. Aber für jemanden, der nur N und e kennt, ist es praktisch extrem schwierig (jedenfalls bei großen N) den Wert $\phi(N)$ (und damit dann auch den private key d) zu berechnen. Entweder testet man direkt alle $i \in \{1, \dots, N-1\}$ darauf, ob $\mathrm{ggT}(i, N) = 1$ gilt, oder man versucht die Faktorisierung $N = pq$ zu finden. — In Anwendungen werden tatsächlich Primzahlen p und q mit Hunderten von Ziffern verwendet, und selbst die schnellsten Computer der Welt würden zu lange brauchen, um $\phi(N)$ direkt zu berechnen oder die Faktorisierung $N = pq$ zu finden. (Es gibt allerdings keinen formalen Beweis, dass dieses Problem nicht doch eines Tages eine effiziente Lösung findet.) Für weitere Details siehe auch
C. Karpfinger und H. Kiechle, Kryptologie, Algebraische Methoden und Algorithmen, Vieweg+Teubner Verlag, 2010.

Zurück zum konkreten Beispiel. Alice wählt $p = 101$ und $q = 103$. Damit erhält man $N = pq = 10403$ und $\phi(N) = (p-1)(q-1) = 10200$. Außerdem wählt sie $e = 1001$ und berechnet $d = 2201$. (Dann gilt in der Tat $ed \equiv 1 \bmod 10200$.) Also:

$$\text{"Public key"}: \quad (e, N) = (1001, 10403), \qquad \text{"private key"}: \quad d = 2201.$$

Alice erhält von Bob die Nachricht: $c = 2532$. Welchen Tipp (also welche ursprüngliche Nachricht m) hat Bob ihr geschickt?

[Es gilt $c^d = 2532^{2201} \equiv 7 \bmod 10403$, also natürlich $m = 7$, "Gewinn". Zum Berechnen von c^d (vor allem für große d) benutzt man am besten folgende Rekursion: $c^d \equiv \begin{cases} (c^2)^{d/2} \bmod N & \text{falls } d \text{ gerade,} \\ c(c^2)^{(d-1)/2} \bmod N & \text{falls } d \text{ ungerade.} \end{cases}$]

Am Beispiel des RSA-Verfahrens zeigt sich, dass Betrachtungen zu Primzahlen, also Jahrhunderte alten Themen der reinen Mathematik, schließlich doch zu konkreten Anwendungen führen können.[2]

[2] Und die Geschichte hinsichtlich Verschlüsselungsverfahren ist hiermit noch keineswegs beendet, Stichwort "Elliptic Curve Cryptography"; siehe zum Beispiel die Diskussion in §6.3.4 im Buch von Karpfinger–Meyberg, oder Kapitel 13 im oben genannten Buch von Karpfinger–Kiechle.

Kapitel II: Gruppen

Wir haben bereits einige Aussagen zu Gruppen gezeigt, aber meistens zu abelschen oder zyklischen Gruppen. Hier geht es nun um den Fall allgemeiner (nicht unbedingt abelscher) Gruppen. Wie zuvor schreiben wir die Verknüpfung in einer Gruppe G als $a \cdot b$ oder einfach als ab; das neutrale Element wird mit 1_G bezeichnet, das zu $a \in G$ inverse Element mit a^{-1}.

6. *Erzeugendensysteme*

Sei G eine Gruppe und $S \subseteq G$ eine Teilmenge. Wir definieren das ***Erzeugnis von*** S als

$$\langle S \rangle := \bigcap_{U \leqslant G \text{ mit } S \subseteq U} U.$$

Der Schnitt wird über eine nicht-leere Menge gebildet, denn $U = G$ kommt immer in dieser Menge vor. Wir haben bereits in §1 bemerkt, dass beliebige Durchschnitte von Untergruppen wieder Untergruppen sind. Also ist $\langle S \rangle$ eine Untergruppe von G. Für $S = \varnothing$ ist $\langle \varnothing \rangle = \{1_G\}$. Ist $S = \{s_1, \ldots, s_n\}$ eine endliche Menge, so schreiben wir auch einfach $\langle S \rangle = \langle s_1, \ldots, s_n \rangle$. Die obige Konstruktion liefert einerseits ein praktisches Verfahren, um Untergruppen einer gegebenen Gruppe G zu definieren. Andererseits kann man sich für G selbst die Frage stellen, eine möglichst einfache oder kleine Teilmenge $S \subseteq G$ mit $G = \langle S \rangle$ zu finden.

Lemma 6.1. *Sei* $\varnothing \neq S \subseteq G$. *Dann ist* $\langle S \rangle = \{1_G\} \cup \{s_1 \cdots s_r \mid r \geqslant 1, s_i \in S \text{ oder } s_i^{-1} \in S\}$. *Für* $S = \{g\}$ *ist* $\langle \{g\} \rangle = \langle g \rangle = \{g^m \mid m \in \mathbb{Z}\}$ *(wie bereits in §3 definiert).*

Beweis. Sei H die rechte Seite obiger Gleichung. Es gilt $1_G \in H$ und man sieht sofort, dass Produkte und Inverse von Elementen in H wieder in H sind. Also ist $H \leqslant G$. Nun ist $S \subseteq H$, also kommt H im Durchschnitt in der Definition von $\langle S \rangle$ vor. Damit gilt $\langle S \rangle \subseteq H$. Umgekehrt sei $U \leqslant G$ beliebig mit $S \subseteq U$. Dann ist auch $H \subseteq U$, also ist H auch im Durchschnitt in der Definition von $\langle S \rangle$ enthalten. Also gilt $H = \langle S \rangle$. Die Aussage über $\langle \{g\} \rangle$ ist dann klar. □

Beispiel 6.2. (a) Sei $G = S_3$. Dann gilt $G = \langle \sigma_1, \sigma_2 \rangle$, wobei $\sigma_1 = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{smallmatrix}\right)$ und $\sigma_2 = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{smallmatrix}\right)$. Beweis: Sei $U := \langle \sigma_1, \sigma_2 \rangle \leqslant G$. Wie in Beispiel 3.9 gilt $o(\sigma_1) = o(\sigma_2) = 2$. Wir berechnen $\sigma_1 \circ \sigma_2 = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{smallmatrix}\right) = \pi$ und $o(\pi) = 3$. Wegen $\sigma_1 \in U$ und $\pi \in U$ ist $|U|$ durch 2 und durch 3 teilbar, also durch 6. Also muss $U = G$ gelten.

(b) Sei $G = Q_8$ die Quaternionengruppe in Beispiel 3.10. Dann gilt $G = \langle I, J \rangle$. Denn sei $U := \langle I, J \rangle$. Es gilt nun $-E = I^2 \in U$, $K = I \cdot J \in U$, also schließlich $\pm I, \pm J, \pm K \in U$.

(c) Sei $S = \{g_1, \ldots, g_n\} \subseteq G$ und es gelte $g_i g_j = g_j g_i$ für alle i, j. Dann kann man Faktoren in Produkten der g_i beliebig vertauschen. Also folgt sofort dass $\langle S \rangle$ abelsch ist und es gilt $\langle S \rangle = \{g_1^{m_1} \cdots g_n^{m_n} \mid m_i \in \mathbb{Z} \text{ für } 1 \leqslant i \leqslant n\}$.

Beispiel 6.3. Die Gruppe G heißt eine ***Diedergruppe***, wenn G von 2 Elementen der Ordnung 2 erzeugt wird. Es ist also $G = \langle s, t\rangle$ mit $s \neq t$, $s \neq 1$, $t \neq 1$ und $s^2 = t^2 = 1$. In den Übungen werden Sie zeigen: Ist $3 \leqslant m := o(st) < \infty$, so gilt $|G| = 2m$ und G ist nicht abelsch. Zum Beispiel ist $G = S_3$ eine Diedergruppe mit $m = 3$, denn die beiden Erzeuger σ_1, σ_2 im obigen Beispiel haben Ordnung 2. Die Quaternionengruppe ist keine Diedergruppe, denn es gibt in Q_8 überhaupt nur ein Element der Ordnung 2. In den Übungen haben Sie gesehen, dass die folgende Menge von Matrizen eine Diedergruppe der Ordnung 8 ist:

$$D_8 := \left\{ \begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}, \begin{pmatrix} 0 & \pm 1 \\ \pm 1 & 0 \end{pmatrix} \right\} \subseteq \mathrm{GL}_2(\mathbb{Q}).$$

Beispiel 6.4. Sei $n \in \mathbb{N}$ und $G = S_n$ die symmetrische Gruppe. Wir führen einige nützliche Bezeichnungen ein. Seien $r \geqslant 1$ und $i_1, \ldots, i_r$ paarweise verschiedene Ziffern in $\{1, \ldots, n\}$. Dann definiere eine Permutation $\sigma \in S_n$ durch $\sigma(j) := j$ für $j \notin \{i_1, \ldots, i_r\}$, und

$$\sigma(i_1) := i_2, \qquad \sigma(i_2) := i_3, \qquad \ldots, \qquad \sigma(i_{r-1}) := i_r, \qquad \sigma(i_r) := i_1.$$

Eine solche Permutation heißt r-***Zykel*** (oder einfach Zykel); wir schreiben dann einfach $\sigma = (i_1\ i_2\ \ldots\ i_r) \in S_n$. Beachte: Die Reihenfolge der Ziffern $i_1, \ldots, i_r$ ist wichtig, aber es ist egal, wo man den Zykel beginnt; es gilt zum Beispiel auch $\sigma = (i_2\ \ldots\ i_r\ i_1)$, und so fort. Für $r = 1$ ist $\sigma = \mathrm{id}$. Für $r = 2$ ist $\sigma = (i_1\ i_2)$ die Permutation, die i_1 und i_2 vertauscht und alle anderen Ziffern festlässt; ein solcher 2-Zykel heißt auch ***Transposition***. Es gilt nun:

(a) Ist $r \geqslant 1$ und $\sigma \in S_n$ ein r-Zykel, so ist $o(\sigma) = r$.

Denn: Es gilt $\sigma^2(i_1) = \sigma(\sigma(i_1)) = \sigma(i_2) = i_3$ und dann analog $\sigma^d(i_1) = i_{d+1} \neq i_1$ für $1 \leqslant d < r$; also ist $o(\sigma) \geqslant r$. Wegen $\sigma^{r-1}(i_1) = i_r$ ist $\sigma^r(i_1) = i_1$. Analog findet man $\sigma^r(i_j) = i_j$ für $1 \leqslant j \leqslant r$, also $\sigma^r = \mathrm{id}$. Damit ist (a) gezeigt.

Sei nun auch $\tau = (j_1\ j_2\ \ldots\ j_s) \in S_n$ ein s-Zykel, wobei $s \geqslant 1$ und $j_1, \ldots, j_s$ paarweise verschieden sind. Dann heißen σ und τ disjunkte Zykeln, wenn $\{i_1, \ldots, i_r\} \cap \{j_1, \ldots, j_s\} = \varnothing$ gilt. Man sieht leicht, dass disjunkte Zykeln vertauschbar sind. Außerdem gilt:

(b) Jede Permutation $\pi \in S_n$ lässt sich als Produkt von disjunkten Zykeln schreiben[3].

Anstatt einen formalen Beweis zu geben, illustrieren wir dies mit einem Beispiel. Sei

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 8 & 4 & 5 & 1 & 2 & 7 & 6 \end{pmatrix} \in S_8 \qquad \rightsquigarrow \qquad \pi = (1\ 3\ 4\ 5) \circ (2\ 8\ 6) \circ (7).$$

Dazu beginnt man mit der Ziffer 1 und wendet wiederholt π darauf an, bis man wieder 1 erhält, also $1 \mapsto 3 \mapsto 4 \mapsto 5 \mapsto 1$. (Beachte: Wegen $o(\pi) < \infty$ muss es ein $d \geqslant 1$ mit $\pi^d(1) = 1$ geben.) Dies definiert den ersten Zykel (1 3 4 5). Dann verfährt man genauso mit

[3] Für einen formalen Beweis der Zykel-Zerlegung siehe zum Beispiel §9.1 im Buch von Karpfinger–Meyberg. Auf diese Weise werden Permutationen auch in einem Computer-Algebra-System, z.B. GAP, dargestellt.

der kleinsten Ziffer, die nicht in diesem Zykel vorkommt, in diesem Fall also 2. Man erhält $2 \mapsto 8 \mapsto 6 \mapsto 2$; dies definiert den zweiten Zykel (2 8 6). Die kleinste Ziffer, die noch nicht in diesen beiden Zykeln vorkommt, ist 7. Nun erhält man $7 \mapsto 7$, also einen 1-Zykel, d.h., die Identität, die man dann auch in der Produktdarstellung weglassen kann. Weiterhin gilt:

(c) Ist $r \geqslant 1$ und $\sigma \in S_n$ ein r-Zykel, so ist σ ein Produkt von $r-1$ Transpositionen.

Denn es ist $\sigma = (i_1\ i_2) \circ (i_2\ i_3) \circ \cdots \circ (i_{r-1}\ i_r)$, wie man sofort verifiziert (indem man beide Seiten auf eine beliebige Ziffer $j \in \{1, \ldots, n\}$ anwendet). Zum Beispiel ist (1 3 4 5) = (1 3) ∘ (3 4) ∘ (4 5) und (2 8 6) = (2 8) ∘ (8 6), wobei (8 6) = (6 8). Für obiges Element $\pi \in S_8$ erhalten wir also $\pi = (1\ 3) \circ (3\ 4) \circ (4\ 5) \circ (2\ 8) \circ (6\ 8)$. Oder allgemein:

Satz 6.5. *Es gilt* $S_n = \langle (i\ j) \mid 1 \leqslant i < j \leqslant n \rangle$*, d.h.,* S_n *wird von Transpositionen erzeugt.*

Beweis. Sei $\pi \in S_n$ beliebig. Wie oben beschrieben ist $\pi = \pi_1 \circ \cdots \circ \pi_k$ mit disjunkten Zykeln $\pi_1, \ldots, \pi_k$. Weiterhin ist jedes π_j ein Produkt von Transpositionen, also insgesamt π ein Produkt von Transpositionen und damit $\pi \in \langle (i\ j) \mid 1 \leqslant i < j \leqslant n \rangle$. □

Ab hier Woche 4

Schließlich zeigen wir, dass abelsche Gruppen spezielle Erzeugendensysteme besitzen. (Im nächsten Abschnitt werden wir dies noch etwas umformulieren; siehe Beispiel 7.14.)

Satz 6.6 (Hauptsatz über endlich erzeugte abelsche Gruppen)**.** *Sei* G *eine abelsche Gruppe, die von endlich vielen Elementen erzeugt wird. (Also zum Beispiel eine beliebige endliche abelsche Gruppe.) Dann gibt es endlich viele zyklische Untergruppen* $U_1, \ldots, U_r \leqslant G$*, so dass sich jedes* $g \in G$ *auf eindeutige Weise schreiben lässt als* $g = u_1 \cdots u_r$ *mit* $u_i \in U_i$*.*

Wählen wir also Erzeuger $g_i \in U_i$ und setzen $d_i := o(g_i)$ für $1 \leqslant i \leqslant r$, so gilt $G = \langle g_1, \ldots, g_r \rangle$ und wir können jedes $g \in G$ eindeutig schreiben als $g = g_1^{m_1} \cdots g_r^{m_r}$, wobei entweder $m_i \in \mathbb{Z}$ beliebig (falls $d_i = \infty$) oder $0 \leqslant m_i < d_i$ (falls $d_i < \infty$).

Beweis. Nach Voraussetzung gibt es ein $n \in \mathbb{N}$ und $a_1, \ldots, a_n \in G$ mit $G = \langle a_1, \ldots, a_n \rangle$. Da G abelsch ist, ist jedes $g \in G$ darstellbar als $g = a_1^{m_1} \cdots a_n^{m_n}$ mit $m_i \in \mathbb{Z}$ für $1 \leqslant i \leqslant n$; siehe Beispiel 6.2(c). Wir führen nun den Beweis mit Induktion nach n. Ist $n = 1$, so ist $G = \langle a_1 \rangle$ zyklisch, die Behauptung gilt also mit $r = 1$ und $U_1 = G$. Sei nun $n > 1$. Sei $\mathcal{M}$ die Menge aller Tupel $S = (g_1, \ldots, g_n)$ mit $g_1, \ldots, g_n \in G$ und $G = \langle g_1, \ldots, g_n \rangle$. Sei $S = (g_1, \ldots, g_n) \in \mathcal{M}$. Gibt es ein $i \in \mathbb{N}$ mit $g_1^i \in \langle g_2, \ldots, g_n \rangle$, so definiere

$$\mu(S) := \min\{i \in \mathbb{N} \mid g_1^i \in \langle g_2, \ldots, g_n \rangle\};$$

gibt es kein solches i, so setze $\mu(S) := \infty$. Wir zeigen nun, dass es ein $S = (g_1, \ldots, g_n) \in \mathcal{M}$ gibt mit $\langle g_1 \rangle \cap \langle g_2, \ldots, g_n \rangle = \{1_G\}$. Dazu unterscheiden wir mit Hilfe von $\mu(S)$ zwei Fälle.

1. Fall: Es gibt ein $S = (g_1 \ldots, g_n) \in \mathcal{M}$ mit $\mu(S) = \infty$. Dann ist also $g_1^m \notin H := \langle g_2, \ldots, g_n \rangle$ für alle $m \in \mathbb{N}$. Behauptung: Es gilt $\langle g_1 \rangle \cap H = \{1_G\}$. Dazu: Sei $g \in \langle g_1 \rangle \cap H$; dann ist $g = g_1^m$ mit $m \in \mathbb{Z}$. Wegen $g_1^m \notin H$ muss also $m \leqslant 0$ gelten. Aber es ist auch $g^{-1} = g_1^{-m} \in \langle g_1 \rangle \cap H$. Also muss ebenfalls $-m \leqslant 0$ gelten und damit $m = 0$, also $g = 1_G$.

2. Fall: Es gilt $\mu(S) < \infty$ für alle $S \in \mathcal{M}$. Dann sei $m := \min\{\mu(S) \mid S \in \mathcal{M}\} \in \mathbb{N}$ und wir wählen ein $S = (g_1, \ldots, g_n) \in \mathcal{M}$ mit $m = \mu(S)$. Wegen $g_1^m \in \langle g_2, \ldots, g_n \rangle$ gibt es $m_2, \ldots, m_n \in \mathbb{Z}$ mit $g_1^m = g_2^{m_2} \cdots g_n^{m_n}$. Division mit Rest ergibt $-m_i = q_i m + r_i$ mit $q_i, r_i \in \mathbb{Z}$ und $0 \leqslant r_i < m$ für alle $i \geqslant 2$. Hier ist nun der entscheidende Trick: Setze

$$h := g_1 g_2^{q_2} \cdots g_n^{q_n} \in G \qquad \text{und} \qquad S' := (h, g_2, \ldots, g_n).$$

Dann ist $g_1 = h g_2^{-q_2} \cdots g_n^{-q_n} \in \langle S' \rangle$; da auch $g_2, \ldots, g_n \in S'$ gilt, folgt also $\langle S' \rangle = G$, d.h., $S' \in \mathcal{M}$. Außerdem ist $h^m = g_1^m g_2^{q_2 m} \cdots g_n^{q_n m}$ und damit

$$h^m g_2^{r_2} \cdots g_n^{r_n} = g_1^m g_2^{q_2 m + r_2} \cdots g_n^{q_n m + r_n} = g_1^m g_2^{-m_2} \cdots g_n^{-m_n} = 1_G.$$

Annahme, es gibt ein $i \in \{2, \ldots, n\}$ mit $r_i > 0$. Verschieben wir dann g_i von seiner i-ten Position in S' an den Anfang, so erhalten wir $S'' := (g_i, h, g_2, \ldots, g_{i-1}, g_{i+1}, \ldots, g_n)$. Da S'' lediglich durch Umordnung der Elemente von S' entsteht, ist natürlich auch $S'' \in \mathcal{M}$. Aus obiger Gleichung $h^m g_2^{r_2} \cdots g_n^{r_n} = 1_G$ erhält man aber

$$g_i^{r_i} = h^{-m} g_2^{-r_2} \cdots g_{i-1}^{-r_{i-1}} g_{i+1}^{-r_{i+1}} \cdots g_n^{-r_n} \in \langle h, g_2, \ldots, g_{i-1}, g_{i+1}, \ldots, g_n \rangle,$$

also ist $\mu(S'') \leqslant r_i < m$, Widerspruch zur Minimalität von m. Also gilt $r_i = 0$ für $2 \leqslant i \leqslant n$, und damit $h^m = 1_G$. Kehren wir zurück zum Tupel $S' = (h, g_2, \ldots, g_n)$. Wegen $h^m = 1_G$ ist $\langle h \rangle = \{1_G, h, h^2, \ldots, h^{m-1}\}$. Gäbe es ein $l \in \{1, \ldots, m-1\}$ mit $h^l \in \langle g_2, \ldots, g_n \rangle$, so wäre $\mu(S') \leqslant l < m$, Widerspruch zur Minimalität von m. Also gilt $\langle h \rangle \cap \langle g_2, \ldots, g_n \rangle = \{1_G\}$, d.h., S' ist das gewünschte Tupel.

In beiden Fällen haben wir also ein $S = (g_1, \ldots, g_n) \in \mathcal{M}$ gefunden mit $\langle g_1 \rangle \cap \langle g_2, \ldots, g_n \rangle = \{1_G\}$. Setze damit $U_1 := \langle g_1 \rangle$ und $H := \langle g_2, \ldots, g_n \rangle$. Nach Induktion gibt es zyklische Untergruppen $U_2, \ldots, U_r \leqslant H$, so dass sich jedes $h \in H$ eindeutig schreiben lässt als $h = u_2 \cdots u_r$ mit $u_i \in U_i$ für $2 \leqslant i \leqslant r$. Sei nun $g \in G$. Wir schreiben g als Produkt von Potenzen der g_i; die Terme mit $i \geqslant 2$ bilden ein Element in H; also erhalten wir $g = u_1 h$ mit $u_1 \in U_1$ und $h \in H$. Schreibe $h = u_2 \cdots u_n$ mit $u_i \in U_i$ für $2 \leqslant i \leqslant r$. Dann haben wir also eine Darstellung $g = u_1 u_2 \cdots u_r$ mit $u_i \in U_i$ für $1 \leqslant i \leqslant r$. Sei auch $g = v_1 v_2 \cdots v_n$ mit $v_i \in U_i$ mit $1 \leqslant i \leqslant r$. Dann ist $h' := v_2 \cdots v_r \in H$ und $g = u_1 h = v_1 h'$. Hieraus folgt $v_1^{-1} u_1 = h' h^{-1} \in U_1 \cap H = \{1_G\}$, also $v_1^{-1} u_1 = 1_G = h' h^{-1}$ und damit $u_1 = v_1$, $h = h'$. Nach Induktion folgt außerdem $u_i = v_i$ für $2 \leqslant i \leqslant r$. □

7. *Normalteiler und Homomorphismen*

Sei $U \leqslant G$ Untergruppe. Dann sieht man leicht, dass folgende Bedingungen äquivalent sind.

(1) $gug^{-1} \in U$ für alle $u \in U$ und $g \in G$.

(2) $gU \subseteq Ug$ für alle $g \in G$.

(3) $gU = Ug$ für alle $g \in G$.

[Denn: Es gelte (1). Für $u \in U$ und $g \in G$ ist dann $gu = (gug^{-1})g \in Ug$, also folgt (2). Es gelte (2). Seien $u \in U$ und $g \in G$. Dann ist $(ug)^{-1} = g^{-1}u^{-1} \in g^{-1}U \subseteq Ug^{-1}$ nach (2), also $(ug)^{-1} = vg^{-1}$ mit einem $v \in U$. Damit folgt $ug = (vg^{-1})^{-1} = gv^{-1} \in gU$, also (3). Schließlich gelte (3). Für $u \in U$ und $g \in G$ ist dann $gu = vg$ mit einem $v \in U$, also folgt $gug^{-1} = (vg)g^{-1} = v \in U$ und damit (1).]

Definition 7.1. Eine Untergruppe $U \leqslant G$ heißt ***Normalteiler*** (in Zeichen: $U \trianglelefteq G$), wenn die obigen drei äquivalenten Bedingungen gelten.
Zum Beispiel sind die Untergruppen $\{1_G\}$ und G stets Normalteiler. Ist G abelsch, so ist jede Untergruppe automatisch Normalteiler.

Beispiel 7.2. Sei $U \leqslant G$ mit $[G : U] = 2$. Dann ist $U \trianglelefteq G$.
Denn: Sei $g \in G$ beliebig. Ist $g \in U$, so gilt $gU = U = Ug$. Sei nun $g \notin U$. Wegen $[G : U] = 2$ ist dann $G = U \dot\cup gU$. Mit Satz 3.3 gilt auch $|U\backslash G| = 2$ also $G = U \dot\cup Ug$. Damit folgt $gU = Ug$. Also gilt obige Bedingung (3) für alle $g \in G$.

Beispiel 7.3. Sei $G = S_3$. Mit den Bezeichnungen in Beispiel 3.9 sind die echten Untergruppen von S_3 gegeben durch $U_1 = \langle\sigma_1\rangle, U_2 = \langle\sigma_2\rangle, U_3 = \langle\sigma_3\rangle$ und $V = \langle\pi\rangle$, wobei $\sigma_1, \sigma_2, \sigma_3$ Transpositionen sind und π ein Element der Ordnung 3. Wir rechnen sofort nach: $\sigma_2\sigma_1\sigma_2^{-1} = \sigma_3 \in U_3$ und $\sigma_1\sigma_3\sigma_1 = \sigma_2 \in U_2$, also sind U_1, U_2, U_3 keine Normalteiler. Nun ist $[G : V] = 2$, also ist $V \trianglelefteq G$ nach Beispiel 7.2.

Beispiel 7.4. Mit $Z(G) := \{g \in G \mid xg = gx \text{ für alle } x \in G\}$ bezeichnen wir das ***Zentrum*** von G. Man sieht leicht, dass $Z(G)$ eine Untergruppe ist. Beachte auch: $G = Z(G) \Leftrightarrow G$ abelsch. Behauptung: Es gilt $Z(G) \trianglelefteq G$. Denn sei $z \in Z(G)$ und $g \in G$. Zu zeigen ist $gzg^{-1} \in Z(G)$. Dazu sei $x \in G$ beliebig. Wegen $z \in Z(G)$ ist $zx = xz$ und auch $zg^{-1}xg = g^{-1}xgz$ gilt. Daraus folgt $gzg^{-1}x = xgzg^{-1}$ also $gzg^{-1} \in Z(G)$.

Beispiel 7.5. Sei $G = Q_8$. Nach Beispiel 3.10 sind die echten Untergruppen gegeben durch $\langle I\rangle$, $\langle J\rangle$, $\langle K\rangle$ und $\langle -E\rangle$. Die ersten drei Untergruppen haben Index 2, sind also Normalteiler nach Beispiel 7.2. Man stellt leicht fest, dass $\pm E$ mit allen Elementen von Q_8 vertauschbar ist, also ist auch $\langle -E\rangle$ ein Normalteiler. (Es gilt sogar $Z(Q_8) = \langle -E\rangle$.) Wir haben hier also ein Beispiel, in dem alle Untergruppen Normalteiler sind, aber G selbst nicht abelsch ist.

Analog zur Konstruktion des Faktorrings eines Ringes nach einem Ideal (siehe Definition 2.4) erhalten wir nun die Faktorgruppe einer Gruppe nach einem Normalteiler.

Satz 7.6. *Sei* G *eine Gruppe und* $U \trianglelefteq G$ *ein Normalteiler. Dann ist die Menge der Nebenklassen* $G/U = \{gU \mid g \in G\}$ *eine Gruppe mit Multiplikation* $aU \cdot bU = (ab)U$ *für alle* $a, b \in G$*. Diese Gruppe heißt die* ***Faktorgruppe*** *von* G *nach* U*.*

Beweis. Wir wollen eine Verknüpfung $G/U \times G/U \to G/U$, $(aU, bU) \mapsto (ab)U$, definieren, müssen also zuerst zeigen, dass dies wohl-definiert ist. Seien also $a, a', b, b' \in G$ so, dass $aU = a'U$ und $bU = b'U$ gilt. Dann ist zu zeigen, dass $(ab)U = (a'b')U$ gilt. Aufgrund der obigen Äquivalenzen ist $b'U = Ub'$. Damit erhalten wir

$$(ab)U \subseteq (aU)(bU) \subseteq (a'U)(b'U) \subseteq a'(Ub')U \subseteq a'(b'U)U \subseteq (a'b')U,$$

also gilt auch Gleicheit (weil Nebenklassen entweder gleich oder disjunkt sind). Also ist die Verknüpfung wohl-definiert. Die Assoziativität in G impliziert dann auch die Assoziativität für die Verknüpfung auf G/U. Die Nebenklasse U ist das Einselement dieser Multiplikation und es gilt $(gU)^{-1} = g^{-1}U$ für alle $g \in G$. □

Überzeugen Sie sich selbst, dass die "Wohl-Definiertheit" im obigen Beweis nicht funktioniert, wenn U kein Normalteiler ist, zum Beispiel bereits bei $U = \langle (1\ 2) \rangle \leqslant G = S_3$.

Beispiel 7.7. Sei $G = Q_8$. Dann ist $Z = \langle -E \rangle \trianglelefteq Q_8$ und $[Q_8 : Z] = 4$. Nebenklassenvertreter sind gegeben durch $\{E, I, J, K\}$, also ist $Q_8/Z = \{\bar{E}, \bar{I}, \bar{J}, \bar{K}\}$ wobei $\bar{g} = gZ$ für alle $g \in G$. Wie rechnen wir in Q_8/Z? Es gilt $\bar{I} \cdot \bar{J} = \overline{IJ} = \bar{K}$ und $\bar{J} \cdot \bar{I} = \overline{JI} = \overline{-K} = \bar{K}$, weil $-E \in Z$. Genauso: $\bar{J} \cdot \bar{K} = \bar{I} = \bar{K} \cdot \bar{J}$ und $\bar{K} \cdot \bar{I} = \bar{J} = \bar{I} \cdot \bar{K}$. Also ist Q_8/Z abelsch. Ausserdem $\bar{I}^2 = \bar{J}^2 = \bar{K}^2 = \overline{-E} = \bar{E}$. Damit ist also Q_8/Z eine abelsche Gruppe der Ordnung 4, die nicht zyklisch ist. Also ist Q_8/Z eine ***Klein'sche Vierergruppe***, wie in Beispiel 3.8(b).

Definition 7.8. Seien $(G, \cdot)$ und $(H, \star)$ Gruppen. Eine Abbildung $\varphi\colon G \to H$ heißt ein ***Gruppen-Homomorphismus***, wenn $\varphi(g_1 \cdot g_2) = \varphi(g_1) \star \varphi(g_2)$ für alle $g_1, g_2 \in G$ gilt. Ist φ bijektiv, so heißt φ ein ***Isomorphismus*** und wir sagen, dass G und H ***isomorph*** sind (in Zeichen: $G \cong H$). In diesem Fall kann man in G genauso rechnen wie in H; die beiden Gruppen sind also in diesem Sinne "gleich".

Bemerkung: a) Ist $\varphi\colon G \to H$ ein Gruppen-Homomorphismus, so gilt $\varphi(1_G) = 1_H$ und $\varphi(g^{-1}) = \varphi(g)^{-1}$ für alle $g \in G$.

[Denn: $1_H = \varphi(1_G)^{-1} \star \varphi(1_G) = \varphi(1_G)^{-1} \star \varphi(1_G \cdot 1_G) = \varphi(1_G)^{-1} \star (\varphi(1_G) \star \varphi(1_G)) = 1_H \star \varphi(1_G) = \varphi(1_G)$. Ausserdem: $1_H = \varphi(1_G) = \varphi(g \cdot g^{-1}) = \varphi(g) \star \varphi(g^{-1})$, also $\varphi(g)^{-1} = \varphi(g^{-1})$.]

(b) Ist $\varphi\colon G \to H$ ein bijektiver Gruppen-Homomorphismus, so ist die Umkehrabbildung $\varphi^{-1}\colon H \to G$ ebenfalls ein Gruppen-Homomorphismus.

[Denn: Seien $h_1, h_2 \in H$ und $g_1, g_2 \in G$ mit $\varphi(g_i) = h_i$ für $i = 1, 2$. Dann ist $\varphi\big(\varphi^{-1}(h_1 \star h_2)\big) = h_1 \star h_2 = \varphi(g_1) \star \varphi(g_2) = \varphi(g_1 \cdot g_2)$ also $\varphi^{-1}(h_1 \star h_2) = g_1 \cdot g_2 = \varphi^{-1}(h_1) \cdot \varphi^{-1}(h_2)$.]

Lemma 7.9. *Sei* $\varphi\colon G \to H$ *ein Gruppen-Homomorphismus, wie oben. Dann gilt:*

(a) *Ist* $V \leqslant H$ *Untergruppe, so ist auch* $\varphi^{-1}(V) \leqslant G$ *Untergruppe. Ist* $V \trianglelefteq H$*, so gilt auch* $\varphi^{-1}(V) \trianglelefteq G$*. Insbesondere ist* $\mathrm{Kern}(\varphi) := \{g \in G \mid \varphi(g) = 1_H\} = \varphi^{-1}(\{1_H\}) \trianglelefteq G$*. Ausserdem gilt:* φ *injektiv* $\Leftrightarrow \mathrm{Kern}(\varphi) = \{1_G\}$*.*

(b) *Ist* $U \leqslant G$ *Untergruppe, so ist auch* $\varphi(U) \leqslant H$ *Untergruppe. Insbesondere ist* $\mathrm{Bild}(\varphi) = \varphi(G) \leqslant H$ *eine Untergruppe. Ausserdem: Ist* $U \trianglelefteq G$ *und* φ *surjektiv, so gilt* $\varphi(U) \trianglelefteq H$*.*

Beweis. (a) Sei $V \leqslant H$. Wegen $\varphi(1_G) = 1_H \in V$ ist $1_G \in \varphi^{-1}(V)$. Sind $g_1, g_2 \in \varphi^{-1}(V)$, so gilt $\varphi(g_1) \in V$ und $\varphi(g_2) \in V$, also auch $\varphi(g_1 \cdot g_2) = \varphi(g_1) \star \varphi(g_2) \in V$ und damit $g_1 \cdot g_2 \in \varphi^{-1}(V)$. Analog wird gezeigt: $g_1^{-1} \in \varphi^{-1}(V)$. Also ist $\varphi^{-1}(V) \leqslant G$. Sei nun zusätzlich $V \trianglelefteq H$. Sei $g \in \varphi^{-1}(V)$ und $x \in G$. Dann ist $\varphi(x \cdot g \cdot x^{-1}) = \varphi(x) \star \varphi(g) \star \varphi(x)^{-1} \in V$ wegen $V \trianglelefteq H$. Also ist auch $\varphi^{-1}(V) \trianglelefteq G$.
Es bleibt noch zu zeigen: φ injektiv $\Leftrightarrow \mathrm{Kern}(\varphi) = \{1_G\}$. Die Richtung "$\Rightarrow$" ist wegen $\varphi(1_G) = 1_H$ klar. Sei nun umgekehrt $\mathrm{Kern}(\varphi) = \{1_G\}$. Gilt $\varphi(g_1) = \varphi(g_2)$, so auch $\varphi(g_1 g_2^{-1}) = \varphi(g_1) \star \varphi(g_2)^{-1} = 1_H$, also $g_1 g_2^{-1} = 1_G$, d.h., $g_1 = g_2$.

(b) Die erste Aussage wird analog wie in (a) bewiesen. Sei $U \trianglelefteq G$ und φ surjektiv. Sei $x \in H$ und $h \in \varphi(U)$, also $h = \varphi(g)$ mit $g \in U$. Wegen φ surjektiv ist $x = \varphi(y)$ mit $y \in G$. Wegen $U \trianglelefteq G$ folgt $y \cdot g \cdot y^{-1} \in U$, also $x \star h \star x^{-1} = \varphi(y) \star \varphi(g) \star \varphi(y)^{-1} = \varphi(y \cdot g \cdot y^{-1}) \in \varphi(U)$. □

Beispiel 7.10. Sei $U \trianglelefteq G$ Normalteiler und G/U die zugehörige Faktorgruppe. Dann ist

$$\pi_U\colon G \to G/U, \qquad g \mapsto gU,$$

ein Homomorphismus, denn $\pi_U(gg') = gg'U = (gU)(g'U) = \pi_U(g)\pi_U(g')$ für alle $g, g' \in G$. Es gilt $\mathrm{Kern}(\pi_U) = \{g \in G \mid gU = 1_G U\} = \{g \in G \mid g \in U\} = U$ und π_U ist offensichtlich surjektiv. Wir bezeichnen π_U als den durch U definierten ***kanonischen Homomorphismus***. Dies zeigt auch, dass jeder Normalteiler Kern eines Homomorphismus ist.

Beispiel 7.11. (a) $\exp\colon (\mathbb{R}, +) \to (\mathbb{R}^\times, \cdot)$ ist ein Homomorphismus, denn $\exp(a + b) = \exp(a) \cdot \exp(b)$ für alle $a, b \in \mathbb{R}$.

(b) Sei $n \geqslant 1$ und $G = S_n$. Für $\pi \in S_n$ bezeichnet $N(\pi) := \{1 \leqslant i < j \leqslant n \mid \pi(i) > \pi(j)\}$ die Menge der Fehlstände von π. Dann heißt $\varepsilon(\pi) := (-1)^{|N(\pi)|}$ das ***Signum*** von π. (Dieses erscheint in der Leibniz-Formel für die Determinante einer Matrix.) In der Linearen Algebra werden Sie sicherlich gesehen haben, dass $\varepsilon(\pi \circ \pi') = \varepsilon(\pi)\varepsilon(\pi')$ für alle $\pi, \pi' \in S_n$ gilt, also ist $\varepsilon\colon S_n \to \{\pm 1\} = \mathbb{Z}^\times$ ein Gruppen-Homomorphismus. Dann heißt $A_n := \mathrm{Kern}(\varepsilon) \trianglelefteq S_n$ die ***alternierende Gruppe*** vom Grad n. Ist $n \geqslant 2$ und $\tau \in S_n$ eine Transposition, so gilt $\varepsilon(\tau) = -1$, also ist $\tau \notin A_n$. Ist σ ein r-Zykel mit $r \geqslant 2$, so lässt sich σ als Produkt von $r - 1$

Transpositionen schreiben (siehe Beispiel 6.4(c)) und es folgt $\varepsilon(\sigma) = (-1)^{r-1}$. Also liegen zum Beispiel alle 3-Zykel in A_n.

(c) Sei K ein Körper und $G = GL_n(K)$. Mit den üblichen Eigenschaften der Determinante folgt, dass $\det\colon G \to K^\times$ ein Homomorphismus ist mit

$$SL_n(K) := \{A \in M_n(K) \mid \det(A) = 1\} = \operatorname{Kern}(\det) \trianglelefteq G;$$

diese Gruppe heißt ***spezielle lineare Gruppe***. Der Homomorphismus ist surjektiv, denn für $0 \neq a \in K$ ist $\det(A) = a$, wobei A die Diagonalmatrix mit $a, 1, \dots, 1$ auf der Diagonalen ist.

(d) Sei K Körper und V ein K-Vektorraum mit $n = \dim V < \infty$. Dann ist

$$GL(V) := \{\varphi\colon V \to V \mid \varphi \text{ linear und bijektiv}\},$$

eine Gruppe (mit der Hintereinanderausführung von Abbildungen als Verknüpfung). Sei $B = \{v_1, \dots, v_n\}$ eine Basis von V. Für $\varphi \in GL(V)$ sei $M_B(\varphi) \in GL_n(K)$ die Matrix von φ bezüglich B. Dann ist die Abbildung $\beta\colon GL(V) \to GL_n(K)$, $\varphi \mapsto M_B(\varphi)$, ein Isomorphismus. Es gilt also $GL(V) \cong GL_n(K)$, aber dieser Isomorphismus hängt von der Wahl von B ab.

(e) Seien X, Y nicht-leere Mengen mit $|X| = |Y|$, d.h., es gibt eine Bijektion $f\colon X \to Y$. Dann erhalten wir einen Isomorphismus $\varphi\colon S_X \to S_Y$, $\sigma \mapsto f \circ \sigma \circ f^{-1}$.
Insbesondere: Ist $|X| = n < \infty$, so gibt es eine Bijektion $f\colon X \to \{1, \dots, n\}$, also gilt $S_X \cong S_n$.

Satz 7.12 (Homomorphie-Satz)**.** *Seien G, H Gruppen und $\varphi\colon G \to H$ ein Homomorphismus. Sei $N \trianglelefteq G$ ein Normalteiler mit $N \subseteq \operatorname{Kern}(\varphi)$. Dann gibt es genau einen Homomorphismus $\overline{\varphi}\colon G/N \to H$ mit $\varphi = \overline{\varphi} \circ \pi_U$, wobei $\pi_U\colon G \to G/N$ der kanonische Homomorphismus ist. Es gilt $\operatorname{Bild}(\overline{\varphi}) = \operatorname{Bild}(\varphi)$. Ist $N = \operatorname{Kern}(\varphi)$, so ist $\overline{\varphi}$ injektiv, und $G/\operatorname{Kern}(\varphi) \cong \operatorname{Bild}(\varphi)$.*

Beweis. Wir wollen $\overline{\varphi}\colon G/N \to H$ wie folgt definieren: Für $gN \in G/N$ mit $g \in G$ setze $\overline{\varphi}(gN) := \varphi(g) \in H$. Wie üblich müssen wir zuerst zeigen, dass dies wohl-definiert ist. Seien also $g_1, g_2 \in G$ mit $g_1N = g_2N$. Dann ist $g_2^{-1}g_1 \in N \subseteq \operatorname{Kern}(\varphi)$, also $\varphi(g_2)^{-1}\varphi(g_1) = \varphi(g_2^{-1}g_1) = 1_H$ und damit $\varphi(g_1) = \varphi(g_2)$. Also gilt in der Tat $\overline{\varphi}(g_1N) = \overline{\varphi}(g_2N)$. Sobald dies gezeigt ist, folgt sofort, dass $\overline{\varphi}$ ein Homomorphismus ist. Nach Definition gilt $(\overline{\varphi} \circ \pi_U)(g) = \overline{\varphi}(\pi_U(g)) = \overline{\varphi}(gN) = \varphi(g)$ für alle $g \in G$, also $\overline{\varphi} \circ \pi_U = \varphi$. Diese Gleichung zeigt auch, dass $\operatorname{Bild}(\overline{\varphi}) = \operatorname{Bild}(\varphi)$ gilt. Sei auch $\psi\colon G/N \to H$ ein Homomorphismus mit $\varphi = \psi \circ \pi_U$. Dann ist aber $\psi(gN) = (\psi \circ \pi_U)(g) = \varphi(g) = \overline{\varphi}(gN)$ für alle $g \in G$, also $\psi = \overline{\varphi}$.
Sei schließlich $N = \operatorname{Kern}(\varphi)$. Sei $gN \in \operatorname{Kern}(\overline{\varphi})$, also $\varphi(g) = \overline{\varphi}(gN) = 1_H$. Es folgt $g \in \operatorname{Kern}(\varphi) = N$ und damit $gN = N = 1_{G/H}$. □

Beispiel 7.13. (a) Sei K ein Körper. Dann ist $\det\colon GL_n(K) \to K^\times$ ein surjektiver Homomorphismus mit $\operatorname{Kern}(\det) = SL_n(K)$. Also gilt $GL_n(K)/SL_n(K) \cong K^\times$.

(b) Sei $n \geqslant 2$. Dann ist die Signum-Abbildung $\varepsilon\colon S_n \to \{\pm 1\}$ surjektiv mit $\mathrm{Kern}(\varepsilon) = A_n$. Also gilt $S_n/A_n \cong \{\pm 1\}$ und damit $|A_n| = \frac{1}{2}n!$.

(c) Sei G eine zyklische Gruppe, also $G = \langle g\rangle$ für ein $g \in G$. Dann ist $\varphi\colon \mathbb{Z} \to G$, $m \mapsto g^m$, ein surjektiver Homomorphismus. Da $(\mathbb{Z}, +)$ zyklisch ist (es gilt $\mathbb{Z} = \langle 1\rangle$), ist auch jede Untergruppe von $(\mathbb{Z}, +)$ zyklisch; siehe Beispiel 3.11(a). Damit folgt $\mathrm{Kern}(\varphi) = \langle d\rangle = d\mathbb{Z}$ mit einem $d \in \mathbb{N}_0$. Also gilt $G \cong \mathbb{Z}/d\mathbb{Z}$ nach Satz 7.12. Ist $|G| = \infty$, so ist $d = 0$ und damit $G \cong \mathbb{Z}$. Ist $|G| = d < \infty$, so gilt $G \cong \mathbb{Z}/d\mathbb{Z}$ und $d = o(g)$. Damit haben wir gezeigt:

Jede zyklische Gruppe ist entweder isomorph zu $\mathbb{Z}$ oder zu $\mathbb{Z}/d\mathbb{Z}$ für ein $d \in \mathbb{N}$.

Ab hier Woche 5

Beispiel 7.14. Sei G eine abelsche Gruppe, die von endlich vielen Elementen erzeugt wird. Nach Satz 6.6 gibt es zyklische Untergruppen $U_1, \ldots, U_r \leqslant G$ so dass sich jedes $g \in G$ eindeutig schreiben lässt als $g = u_1 \cdots u_r$ mit $u_i \in U_i$. Betrachte nun die Abbildung

$$\varphi\colon U_1 \times \ldots \times U_r \to G, \qquad (u_1, \ldots, u_r) \mapsto u_1 \cdots u_r,$$

(wobei das direkte Produkt auf der linken Seite nach Definition 1.8 gebildet wird). Weil G abelsch ist, folgt sofort, dass φ ein Homomorphismus ist. Die eindeutige Darstellbarkeit von jedem $g \in G$ als $g = u_1 \cdots u_r$ wie oben impliziert, dass φ bijektiv ist; also ist φ ein Isomorphismus. Wählen wir nun die Numerierung so, dass $d_i := |U_i| < \infty$ für $1 \leqslant i \leqslant s$ und $|U_i| = \infty$ für $i > s$ gilt (wobei $0 \leqslant s \leqslant r$). Nach Beispiel 7.13(c) ist dann $U_i \cong \mathbb{Z}/d_i\mathbb{Z}$ für $1 \leqslant i \leqslant s$ und $U_i \cong \mathbb{Z}$ für $i > s$. Also folgt insgesamt: $G \cong (\mathbb{Z}/d_1\mathbb{Z}) \times \ldots \times (\mathbb{Z}/d_s\mathbb{Z}) \times \mathbb{Z}^{r-s}$. Insbesondere: $G \cong U \times \mathbb{Z}^d$, wobei $d \geqslant 0$ und U eine endliche abelsche Gruppe ist. (Die Struktur von U wird weiter geklärt in Satz 10.7.)

8. *Operation von Gruppen auf Mengen*

Sei G eine Gruppe und X eine nicht-leere Menge. Wir sagen, dass G ***auf*** X ***operiert***, oder dass X eine G-***Menge*** ist, wenn es eine Abbildung $\mu\colon G \times X \to X$, $(g, x) \mapsto g.x$, mit folgenden Eigenschaften gibt:

(a) Es gilt $1_G.x = x$ für alle $x \in X$.

(b) Es gilt $(gh).x = g.(h.x)$ für alle $g, h \in G$ und alle $x \in X$.

Genauso wie "Normalteiler" ist dies ein fundamental wichtiger Begriff. Die meisten Anwendungen von Gruppen haben mit Operationen auf geeigneten Mengen zu tun, vor allem in der Geometrie; viele Strukturaussagen über G folgen durch die Betrachtung von Operationen.

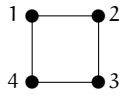
Beispiel 8.1. (a) Die Gruppe $G = S_n$ operiert auf $X = \{1, \ldots, n\}$ durch $\sigma.i = \sigma(i)$ für alle $\sigma \in S_n$ und $i \in \{1, \ldots, n\}$. Daraus ergeben sich auch weitere Operationen. Sei zum Beispiel

$\mathcal{P}(X)$ die Potenzmenge von X. Für $S \in \mathcal{P}(X)$ ist dann auch $\sigma.S := \{\sigma.i \mid i \in S\} \in \mathcal{P}(X)$ mit $|\sigma.X| = |X|$ und man prüft sofort nach, dass dies eine Operation von S_n auf $\mathcal{P}(X)$ definiert.

(b) Sei K ein Körper. Die Gruppe $G = GL_n(K)$ operiert auf $X = K^n$ (Spaltenvektoren) durch $A.v = A \cdot v$ für alle $A \in G$ und $v \in K^n$ (wobei $A \cdot v =$ Multiplikation von A mit v). Auch hier erhält man weitere Operationen, zum Beispiel auf den Teilräumen von X.

(c) Ist $U \leqslant G$ eine Untergruppe, so operiert G auf $X = G/U$ durch $g.(aU) = gaU$ für alle $a, g \in G$. [Dies ist wohl-definiert, denn ist $aU = bU$, so gilt $a^{-1}b = u$ mit einem $u \in U$, also auch $(ga)^{-1}gb = a^{-1}b \in U$ und damit $gaU = gbU$.]

Bemerkung 8.2. Gruppen-Operationen hängen eng zusammen mit dem Konzept der "Symmetrie" von gegebenen Objekten (etwa geometrischen Figuren in der Ebene oder im $\mathbb{R}^3$). Beispiel: Sei $\mathcal{G}$ ein Graph mit Ecken nummeriert durch $\{1, \ldots, n\}$ und Kantenmenge V, d.h., V ist eine Menge von 2-elementigen Teilmengen von $\{1 \ldots, n\}$. (Ist $v = \{i, j\} \in V$, so sind i und j die beiden Endpunkte einer Kante in $\mathcal{G}$.) Dann erhalten wir die ***Symmetrie-Gruppe*** $S_{\mathcal{G}} := \{\pi \in S_n \mid \pi.v \in V \text{ für alle } v \in V\} \leqslant S_n$ des Graphen $\mathcal{G}$ (wobei $\pi.v$ für $v = \{i, j\} \in V$ wie in Beispiel 8.1(a) definiert ist; man prüft sofort nach, dass $S_{\mathcal{G}}$ eine Untergruppe von S_n ist). Sei etwa $\mathcal{G}$ gegeben durch das Quadrat mit Eckpunkten $\{1, 2, 3, 4\}$ und Kantenmenge $V = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{1, 4\}\}$. Dann erhalten wir als Symmetrie-Gruppe



$$S_{\mathcal{G}} = \{\mathrm{id}, (2\ 4), (1\ 4\ 3\ 2), (1\ 4)\circ(2\ 3), (1\ 2)\circ(3\ 4), (1\ 2\ 3\ 4), (1\ 3), (1\ 3)\circ(2\ 4)\} \leqslant S_4$$

(eine Diedergruppe der Ordnung 8). — Weitere Beispiele in den Übungen.

Satz 8.3. *Sei X eine G-Menge und $x \in X$ fest. Dann ist $G_x = \mathrm{Stab}_G(x) := \{g \in G \mid g.x = x\}$ eine Untergruppe von G, genannt der **Stabilisator** von x.*

Beweis. Wegen $1_G.x = x$ ist $1_G \in G_x$. Seien $g, h \in G_x$, also $g.x = x$ und $h.x = x$. Dann ist $(gh).x = g.(h.x) = g.x = x$, also auch $gh \in G_x$. Außerdem gilt $x = 1_G.x = (g^{-1}g).x = g^{-1}.(g.x) = g^{-1}.x$ und damit auch $g^{-1} \in G_x$. □

Satz 8.4. *Sei X eine G-Menge und $x \in X$ fest. Dann heißt $\mathcal{O}_x := \{g.x \mid g \in G\} \subseteq X$ die **Bahn** von x. Die folgende Abbildung ist wohl-definiert und bijektiv:*

$$\mu_x \colon \mathcal{O}_x \to G/G_x, \qquad g.x \mapsto gG_x.$$

Beweis. Ist $g.x = h.x$ mit $g, h \in G$, so gilt $(g^{-1}h).x = g^{-1}.(g.x) = 1_G.x = x$ also $g^{-1}h \in G_x$ und damit $gG_x = hG_x$. Dies zeigt, dass μ_x wohl-definiert ist. Die Abbildung ist offensichtlich surjektiv. Sei schließlich $gG_x = hG_x$, Dann folgt $g^{-1}h \in G_x$ also $g.x = g.((g^{-1}h).x) = (gg^{-1}h).x = h.x$; damit ist die Abbildung auch injektiv. □

Satz 8.5 (Bahnensatz)**.** *Sei X eine G-Menge. Dann sind je zwei Bahnen entweder gleich oder disjunkt. Die Menge X ist also eine disjunkte Vereinigung von Bahnen, d.h., ist X_0 ein*

Vertretersystem der Bahnen, so gilt $X = \bigcup_{x \in X_0} \mathcal{O}_x$*, wobei die Vereinigung disjunkt ist. Ist* $|G| < \infty$*, so gilt* $|G| = |\mathcal{O}_x||G_x|$ *für alle* $x \in X$*.*

Beweis. Seien $x, y \in X$. Dann schreiben wir $x \sim y$, wenn es ein $g \in G$ mit $y = g.x$ gibt. Dies definiert eine Relation auf X. Wegen $1_G.x = x$ ist $\sim$ reflexiv. Ist $x \sim y$, also $y = g.x$ für ein $g \in G$, so gilt $g^{-1}.y = g^{-1}.(g.x) = (g^{-1}g).x = 1_G.x = x$, also $y \sim x$; damit ist $\sim$ symmetrisch. Seien schließlich $x \sim y$ und $y \sim z$, also $y = g.x$ und $z = h.y$ mit $g, h \in G$. Dann ist $z = h.y = h.(g.x) = (hg).x$, also $x \sim z$. Damit gilt: $\sim$ ist eine Äquivalenzrelation.

Für ein festes $x \in X$ ist $y \in X$ in der gleichen Äquivalenzklasse wie x genau dann, wenn $y = g.x$ für ein $g \in G$ gilt, also genau dann, wenn y in der Bahn von x liegt. Damit ist gezeigt, dass X eine disjunkte Vereinigung von Bahnen ist, also je zwei Bahnen entweder gleich oder disjunkt sind. Die Aussage $|G| = |\mathcal{O}_x||G_x|$ folgt aus der Bijektion in Satz 8.4. □

Beispiel 8.6. (a) Sei $G = S_n$ und $X = \{1, \ldots, n\}$ wie in Beispiel 8.1(a). Dann gibt es genau eine Bahn, denn zu jedem $i \in \{1, \ldots, n\}$ gibt es ein $\sigma_i \in S_n$ mit $\sigma_i(1) = i$ (zum Beispiel $\sigma_i = (1, i)$). Eine solche Operation (d.h., eine Operation mit nur einer Bahn) heißt eine ***transitive Operation***. Sei nun $x := n$. Dann ist X die Bahn von n und $\mathrm{Stab}_{S_n}(n)$ besteht aus allen Permutationen $\sigma \in S_n$ mit $\sigma(n) = n$. Diese Untergruppe besteht aber genau aus allen Permutationen von $\{1, \ldots, n-1\}$, also den Elementen in S_{n-1}. Die Formel $|G| = |\mathcal{O}_x||G_x|$ besagt hier also $|S_n| = n \cdot |S_{n-1}|$. Mit Induktion nach n folgt damit sofort die Formel $|S_n| = n!$.

(b) Sei $G = \mathrm{GL}_n(K)$ und $X = K^n$ wie in Beispiel 8.1(b). Dann ist sicher $\mathcal{O}_0 = \{0\}$ eine Bahn. (Diese Operation ist also nicht transitiv.) Andererseits gibt es zu jedem $0 \neq v \in K^n$ stets eine invertierbare Matrix $A \in \mathrm{GL}_n(K)$ mit v als erste Spalte. Dann ist $Ae_1 = v$, wobei $e_1 \in K^n$ der Standard-Basisvektor mit 1 in der 1. Komponente und 0 sonst ist. Also ist $\mathcal{O}_{e_1} = K^n \setminus \{0\}$ eine Bahn, und $K^n = \mathcal{O}_0 \dot\cup \mathcal{O}_{e_1}$ ist die Zerlegung von K^n in Bahnen. Es gilt

$$G_{e_1} = \{A \in \mathrm{GL}_n(K) \mid A \cdot e_1 = e_1\} = \left\{ \left(\begin{array}{c|ccc} 1 & a_2 & \ldots & a_n \\ \hline 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{array} \right) \;\middle|\; a_i \in K, A' \in \mathrm{GL}_{n-1}(K) \right\}.$$

Ist also $|K| = q < \infty$, so gilt $|\mathcal{O}_{e_1}| = q^n - 1$ und $|G_{e_1}| = q^{n-1}|\mathrm{GL}_{n-1}(K)|$. Mit dem Bahnensatz folgt also $|\mathrm{GL}_n(K)| = |\mathcal{O}_{e_1}||G_{e_1}| = (q^n - 1)q^{n-1}|\mathrm{GL}_{n-1}(K)|$. Durch Induktion nach n erhalten wir damit sofort einen Beweis für die folgende Formel:

$$|\mathrm{GL}_n(K)| = q^{n(n-1)/2}(q^n - 1)(q^{n-1} - 1) \cdots (q - 1).$$

Beispiel 8.7. Sei K Körper und $G = \mathrm{GL}_n(K)$. Dann operiert G auf $X = M_n(K)$ (der Menge aller $n \times n$-Matrizen mit Einträgen in K) durch $g.A = gAg^{-1}$ wobei $g \in G$ und $A \in M_n(K)$. Zwei Matrizen $A, A' \in M_n(K)$ liegen also genau dann in der gleichen Bahn, wenn sie ähnlich

sind. Das Problem, ein Vertretersystem der Bahnen anzugeben, ist also äquivalent dazu, Normalformen für Matrizen zu finden (was in der Linearen Algebra behandelt wird).

Beispiel 8.8. Die Gruppe G operiert auf $X = G$ durch

$$G \times X \to X, \qquad (g, x) \mapsto gxg^{-1}, \qquad \text{“}\boldsymbol{Konjugation}\text{”}.$$

(Überzeugen Sie sich selbst davon, dass dies eine Operation ist.) Die Bahn von $x \in G$ ist hier gegeben durch $C_x = \{gxg^{-1} \mid g \in G\}$ und heißt ***Konjugiertenklasse*** von x. Der Stabilisator von $x \in G$ ist hier gegeben durch

$$C_G(x) := \{g \in G \mid gxg^{-1} = x\} = \{g \in G \mid gx = xg\}$$

und heißt ***Zentralisator*** von x. Ist $G \neq \{1_G\}$, so ist diese Operation sicher nicht transitiv, denn $C_1 = \{1_G\}$ ist immer eine Konjugiertenklasse (mit $C_G(1_G) = G$). Ist also $|G| < \infty$, so ist G nach dem Bahnensatz eine disjunkte Vereinigung von Konjugiertenklassen:

$$G = C_1 \cup \ldots \cup C_r \qquad \text{mit} \qquad |C_i| = [G : C_G(g_i)] \quad \text{wobei } g_i \in C_i.$$

Beachte: Für $x \in G$ gilt $x \in Z(G) \Leftrightarrow G = C_G(x) \Leftrightarrow \{x\}$ ist eine Konjugiertenklasse.

Sei zum Beispiel $G = S_3$. Dann gibt es 3 Konjugiertenklassen, nämlich $C_1 = \{\mathrm{id}\}$ und

$$C_2 = \{(1\ 2), (2\ 3), (1\ 3)\} \qquad \text{mit} \qquad C_{S_3}\big((1\ 2)\big) = \{\mathrm{id}, (1\ 2)\};$$
$$C_3 = \{(1\ 2\ 3), (1\ 3\ 2)\} \qquad \text{mit} \qquad C_{S_3}\big((1\ 2\ 3)\big) = \{\mathrm{id}, (1\ 2\ 3), (1\ 3\ 2)\}.$$

Allgemein besteht ein enger Zusammenhang zwischen Konjugiertenklassen und Normalteilern. Ist nämlich $N \trianglelefteq G$ und $x \in N$, so gilt $gxg^{-1} \in N$ für alle $g \in G$. Damit folgt $C \subseteq N$, wobei C die Konjugiertenklasse von x bezeichnet. Da dies für alle $x \in N$ gilt, ist also N eine Vereinigung von Konjugiertenklassen von G. Hier ist eine erste Illustration dieses Prinzips.

Beispiel 8.9. Sei p eine Primzahl und G eine endliche Gruppe mit $|G| = p^n$, $n \geqslant 1$. Behauptung: $Z(G) \neq \{1_G\}$ (wobei $Z(G)$ das Zentrum von G ist, siehe Beispiel 7.4).

Dazu betrachten wir wie oben die Konjugiertenklassen: $G = C_1 \cup \ldots \cup C_r$; sei $g_i \in C_i$ und $g_1 = 1_G$, also $C_1 = \{1_G\}$. Es gilt $|C_i| = [G : C_G(g_i)]$, nach Lagrange also $|C_i| = p^{n_i}$ mit $0 \leqslant n_i \leqslant n$. Dies ergibt $p^n = |G| = \sum_{i=1}^r |C_i| = \sum_{i=1}^r p^{n_i}$. Nun ist $n_1 = 0$, also erhalten wir

$$p^n = 1 + p^{n_2} + \ldots + p^{n_r}.$$

Wäre $n_i \geqslant 1$ für alle $i \geqslant 2$, so erhielten wir $p \mid 1 = p^n - (p^{n_2} + \ldots + p^{n_r})$, Widerspruch. Also gibt es ein $i \geqslant 2$ mit $n_i = 0$. Dies bedeutet aber $C_i = \{g_i\}$ und damit $1_G \neq g_i \in Z(G)$.

Folgerung 8.10. *Sei $|G| = p^2$ mit einer Primzahl p. Dann ist G abelsch.*

Beweis. Nach Beispiel 8.9 ist $Z(G) \neq \{1_G\}$. Sei $1_G \neq z \in Z(G)$. Dann ist $o(z) = p$ oder p^2. Ist $o(z) = p^2$, so ist $G = \langle z \rangle$ und wir sind fertig. Sei nun $o(z) = p$, also $\langle z \rangle \subsetneqq G$. Sei $y \in G \setminus \langle z \rangle$.

Dann ist $\{1_G\} \subsetneqq \langle z\rangle \subsetneqq \langle y, z\rangle$, also muss $G = \langle y, z\rangle$ gelten (wegen Lagrange und $|G| = p^2$). Aber aus $yz = zy$ folgt dann wiederum, dass $G = \langle y, z\rangle$ abelsch ist; siehe Beispiel 6.2(c). □

Beispiel 8.11. Sei $X \neq \varnothing$ eine G-Menge, d.h., G operiert auf X. Sei $g \in G$ fest. Dann ist

$$\pi_g \colon X \to X, \qquad x \mapsto g.x,$$

eine Bijektion. (Denn es gilt $\pi_g \circ \pi_{g^{-1}} = \pi_{g^{-1}} \circ \pi_g = \mathrm{id}_X$, also hat π_g eine Umkehrabbildung, ist damit bijektiv.) Also erhalten wir eine Abbildung $\pi \colon G \to S_X$, $g \mapsto \pi_g$. Dies ist ein Gruppen-Homomorphismus, denn

$$\pi(gg')(x) = (gg').x = g.(g'.x) = \pi_g(\pi_{g'}(x)) = (\pi_g \circ \pi_{g'})(x) = (\pi(g) \circ \pi(g'))(x)$$

für alle $g, g' \in G$ und $x \in X$, also gilt $\pi(gg') = \pi(g) \circ \pi(g')$. Damit gilt:

Zu jeder G-Menge X gehört ein Gruppen-Homomorphismus $\pi \colon G \to S_X$.

Hier ist $\mathrm{Kern}(\pi) = \{g \in G \mid \pi_g = \mathrm{id}_X\} = \{g \in G \mid g.x = x \text{ für alle } x \in X\} = \bigcap_{x\in X} \mathrm{Stab}_G(x)$.

Betrachten wir folgenden Spezialfall: Sei $U \leqslant G$. Dann ist $X = G/U$ eine G-Menge, mit Operation $G \times G/U \to G/U, (g, xU) \mapsto gxU$. (Siehe Beispiel 8.1(c).) Also erhalten wir einen Homomorphismus $\pi \colon G \to S_X$. Für $x \in G$ ist

$$\begin{aligned}\mathrm{Stab}_G(xU) &= \{g \in G \mid gxU = xU\} = \{g \in G \mid x^{-1}gx \in U\}\\ &= \{g \in G \mid g \in xUx^{-1}\} = xUx^{-1}.\end{aligned}$$

Insbesondere zeigt dies, dass xUx^{-1} eine Untergruppe von G ist für jedes $x \in G$. Außerdem folgt $\mathrm{Kern}(\pi) = \bigcap_{x\in G} xUx^{-1} \subseteq U$.

Satz 8.12 (Cayley)**.** *Sei G eine endliche Gruppe. Dann gibt es ein $n \geqslant 1$ und einen injektiven Homomorphismus $\tilde\pi \colon G \to S_n$. Also ist G isomorph zu einer Untergruppe von S_n.*

Beweis. Sei $n = |G|$. Wir wenden die Konstruktion in Beispiel 8.11 an mit $U = \{1_G\}$, also $|X| = |G/U| = n$. Wir erhalten einen Homomorphismus $\pi \colon G \to S_X$ mit $\mathrm{Kern}(\pi) \subseteq U = \{1_G\}$, d.h., π ist injektiv. Wegen $|X| = n$ ist $S_X \cong S_n$, also erhalten wir auch einen injektiven Homomorphismus $\tilde\pi \colon G \to S_n$. Dann ist G isomorph zu $\tilde G = \tilde\pi(G) \leqslant S_n$. □

Definition 8.13. Sei K Körper und $n \geqslant 1$. Für $\sigma \in S_n$ definieren wir eine Matrix

$$A^\sigma = (a_{ij}^\sigma)_{1\leqslant i,j\leqslant n} \in M_n(K) \qquad \text{wobei} \qquad a_{ij}^\sigma := \begin{cases} 1 & \text{falls } i = \sigma(j),\\ 0 & \text{sonst.}\end{cases}$$

Die Matrix A^σ heißt die zu σ gehörige ***Permutationsmatrix***. Einige Beispiele für $n = 3$:

$$A^{(1,2)} = \begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix}, \qquad A^{(1,3,2)} = \begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}, \qquad A^{(1,3)} = \begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}.$$

(Wir erhalten A^σ, indem wir die Zeilen der Einheitsmatrix gemäß σ permutieren.)

Lemma 8.14. *Mit den Bezeichnungen in obiger Definition ist* $A^\sigma \in \mathrm{GL}_n(K)$ *für alle* $\sigma \in S_n$. *Die Abbildung* $\rho\colon S_n \to \mathrm{GL}_n(K)$, $\sigma \mapsto A^\sigma$, *ist ein injektiver Homomorphismus.*

Beweis. Ist $\sigma = \mathrm{id}$, so ist $A^\sigma = I_n$ die Einheitsmatrix. Seien nun $\sigma, \tau \in S_n$. Der (i,j)-Eintrag von $A^\sigma \cdot A^\tau$ berechnet sich dann als $(A^\sigma \cdot A^\tau)_{ij} = \sum_{k=1}^n a^\sigma_{ik} a^\tau_{kj}$. Nun ist $a^\tau_{kj} = 1$ falls $k = \tau(j)$, und 0 sonst. Also reduziert sich die Summe zu

$$(A^\sigma \cdot A^\tau)_{ij} = a^\sigma_{i\tau(j)} = \begin{cases} 1 & \text{falls } i = \sigma(\tau(j)), \\ 0 & \text{sonst.} \end{cases}$$

Dies ist gleich $(A^{\sigma\circ\tau})_{ij}$, also gilt $A^\sigma \cdot A^\tau = A^{\sigma\circ\tau}$. Damit folgt auch $A^\sigma \cdot A^{\sigma^{-1}} = A^{\mathrm{id}} = I_n$, also ist $A^\sigma \in \mathrm{GL}_n(K)$. Die Gleichung $A^\sigma \cdot A^\tau = A^{\sigma\circ\tau}$ zeigt dann, dass ρ ein Homomorphismus ist. Ist $\sigma \neq \mathrm{id}$, so ist $\sigma(i) \neq i$ für ein i, also $A^\sigma \neq I_n$. Damit ist ρ injektiv. □

Folgerung 8.15 (Matrixversion von Cayley's Satz)**.** *Sei* G *eine endliche Gruppe und* K *ein Körper. Dann gibt es ein* $n \geqslant 1$, *so dass* G *isomorph zu einer Untergruppe von* $\mathrm{GL}_n(K)$ *ist.*

Beweis. Nach Cayley's Satz 8.12 gibt es ein $n \geqslant 1$ und einen injektiven Homomorphismus $\tilde\pi\colon G \to S_n$. Mit ρ wie in Lemma 8.14 erhalten wir einen injektiven Homomorphismus $\rho \circ \tilde\pi\colon G \to \mathrm{GL}_n(K)$. Also ist G isomorph zu einer Untergruppe von $\mathrm{GL}_n(K)$. □

9. *Einfache Gruppen und auflösbare Gruppen*

Besitzt eine Gruppe G einen Normalteiler $N \trianglelefteq G$ mit $\{1_G\} \neq N \neq G$, so kann man sich G aufgebaut denken aus N und der Faktorgruppe G/N, was oft hilfreich bei der weiteren Untersuchung von G ist. Dies führt einerseits zu folgender Definition.

Definition 9.1. Eine Gruppe $G \neq \{1_G\}$ heißt eine ***einfache Gruppe***, wenn $\{1_G\}$ und G die einzigen Normalteiler sind. — Ist zum Beispiel $|G| = p$ mit einer Primzahl p, so gibt es nur die Untergruppen $\{1_G\}$ und G (siehe Beispiel 3.8), also ist G einfach.

Andererseits führt es zu der Frage, wie man Normalteiler finden kann. Wir kennen bisher im Wesentlichen zwei allgemeine Methoden: Es gilt stets $Z(G) \trianglelefteq G$; außerdem: Ist X eine G-Menge, so gibt es einen Homomorphismus $\pi\colon G \to S_X$, mit $\mathrm{Kern}(\pi) \trianglelefteq G$. Wir stellen nun noch eine dritte Methode vor (die auch später, in Kapitel IV, von Bedeutung sein wird).

Definition 9.2. Für $g, h \in G$ heißt $[g,h] := g^{-1}h^{-1}gh \in G$ der ***Kommutator*** von g, h. Die ***Kommutator-Untergruppe***[4] von G ist definiert als

$$G' := \langle [g,h] \mid g, h \in G \rangle \leqslant G.$$

[4] Im Allgemeinen ist $\{[g,h] \mid g,h \in G\}$ keine Untergruppe, siehe I. M. ISAACS, Commutators and the commutator subgroup, Amer. Math. Monthly 84 (1977), 720–722; das kleinste solche Beispiel hat Ordnung 96.

Beachte: Es gilt $[g,h]=1 \Leftrightarrow gh=hg$, also ist $G'=\{1_G\}$ genau dann, wenn G abelsch ist. Wir können also die Größe von G' als ein Maß dafür ansehen, wie "nicht-abelsch" G ist.

Satz 9.3. (a) *Es gilt* $G' \trianglelefteq G$ *und* G/G' *ist abelsch.*
(b) *Ist* $H \trianglelefteq G$ *beliebig mit* G/H *abelsch, so ist* $G' \subseteq H$.
(c) *Ist* $\varphi\colon G \to H$ *surjektiver Homomorphismus, so gilt* $\varphi(G')=H'$.

Beweis. Zuerst (c). Zunächst gilt $\varphi(G') \subseteq H'$, denn für alle $g_1, g_2 \in G$ ist

$$\varphi([g_1,g_2]) = \varphi(g_1^{-1}g_2^{-1}g_1g_2) = \varphi(g_1)^{-1}\varphi(g_2)^{-1}\varphi(g_1)\varphi(g_2) = [\varphi(g_1),\varphi(g_2)] \in H'.$$

Nun ist φ surjektiv. Sind also $h_1, h_2 \in H$, so gibt es $g_1, g_2 \in G$ mit $\varphi(g_i)=h_i$ für $i=1,2$ und dann zeigt obige Rechnung, dass $[h_1,h_2] \in \varphi(G')$ gilt. Also ist auch $H' \subseteq \varphi(G')$.

(a) Sei $g \in G$. Man sieht leicht, dass die Abbildung $\gamma_g\colon G \to G$, $x \mapsto gxg^{-1}$, ein Gruppen-Isomorphismus ist (mit $\gamma_g^{-1}=\gamma_{g^{-1}}$; siehe auch Ü5). Mit (c) folgt $gG'g^{-1}=\gamma_g(G')=G'$, also ist $G' \trianglelefteq G$. Betrachte nun den kanonischen Homomorphismus $\pi\colon G \to G/G'$. Nach (c) gilt $(G/G')'=\pi(G')=\{1_{G/G'}\}$, also ist nach obiger Bemerkung G/G' abelsch.

(b) Sei $\pi\colon G \to G/H$ der kanonische Homomorphismus. Mit (c) folgt $\pi(G')=(G/H)'=\{1_{G/H}\}$, weil G/H abelsch ist. Also ist $G' \subseteq \mathrm{Kern}(\pi)=H$. □

Ab hier Woche 6

Definition 9.4. Wir definieren Untergruppen $G^{(i)} \leqslant G$ wie folgt für alle $i=0,1,2,3,\ldots$. Seien $G^{(0)} := G$ und $G^{(1)} := G'$. Dann setze $G^{(i)} := (G^{(i-1)})'$ für $i=2,3,\ldots$. Die Gruppe G heißt ***auflösbar***, wenn $G^{(r)}=\{1_G\}$ gilt für ein $r \geqslant 1$.
Zum Beispiel ist jede abelsche Gruppe G auflösbar, weil hier bereits $G^{(1)}=G'=\{1_G\}$ gilt.

Lemma 9.5. *Sei* G *auflösbar. Dann ist jede Untergruppe von* G *auflösbar. Ist also* $\{1_G\} \neq U \leqslant G$*, so folgt* $U' \subsetneq U$.

Beweis. Sei $U \leqslant G$. Aufgrund der Definition der Kommutator-Untergruppe ist $U' \subseteq G'$ und dann auch $U^{(i)} \subseteq G^{(i)}$ für alle $i \geqslant 1$. Nach Voraussetzung gibt es ein $r \geqslant 1$ mit $G^{(r)}=\{1_G\}$; also ist auch $U^{(r)}=\{1_G\}$ und damit U auflösbar. Sei nun $U \neq \{1_G\}$. Wäre $U'=U$, so auch $U^{(i)}=U \neq \{1_G\}$ für alle $i \geqslant 1$, Widerspruch. □

Lemma 9.6. *Sei* $N \trianglelefteq G$ *ein Normalteiler. Sind* N *und* G/N *auflösbar, so ist auch* G *auflösbar.*

Beweis. Sei $H := G/N$ und $\pi\colon G \to H$ der kanonische Homomorphismus. Nach Satz 9.3 ist $\pi(G')=H'$. Durch Einschränkung erhalten wir auch einen surjektiven Homomorphismus $\pi|_{G'}\colon G' \to H'$, und es folgt wiederum $\pi((G')')=\pi|_{G'}((G')')=(H')'$. Durch weiteres

Wiederholen dieses Argumentes folgt also $\pi(G^{(i)}) = H^{(i)}$ für alle $i \geqslant 1$. Nach Voraussetzung gibt es ein $r \geqslant 1$ mit $H^{(r)} = \{1_H\}$. Dann ist $G^{(r)} \subseteq \mathrm{Kern}(\pi) = N$. Ebenfalls nach Voraussetzung gibt es ein $s \geqslant 1$ mit $N^{(s)} = \{1_G\}$. Wie im obigen Beweis folgt dann $(G^{(r)})^{(s)} \subseteq N^{(s)} = \{1_G\}$. Schließlich beachte: Es gilt $G^{(r+1)} = (G^{(r)})' = (G^{(r)})^{(1)}$, sodann $G^{(r+2)} = (G^{(r+1)})' = ((G^{(r)})^{(1)})' = (G^{(r)})^{(2)}$ und so fort, also $G^{(r+s)} = (G^{(r)})^{(s)} = \{1_G\}$, d.h., G ist auflösbar. □

Beispiel 9.7. Sei $|G| = p^n$ mit $n \geqslant 0$ und einer Primzahl p. Dann ist G auflösbar.

Beweis mit Induktion nach n. Ist $n = 0$, so ist $G = \{1_G\}$ auflösbar. Sei nun $n > 0$. Dann betrachte das Zentrum $N := Z(G) \trianglelefteq G$. Nach Beispiel 8.9 ist $N \neq \{1_G\}$. Nach Lagrange also $|G/N| = p^m$ mit $0 \leqslant m < n$. Nach Induktion ist G/N auflösbar. Da N abelsch ist, folgt mit Lemma 9.6 dann auch, dass G selbst auflösbar ist.

Lemma 9.8. *Die Gruppe G ist auflösbar $\iff$ Es gibt eine Folge von Untergruppen $\{1_G\} = U_0 \leqslant U_1 \leqslant U_2 \leqslant \ldots \leqslant U_r = G$ mit $U_{i-1} \trianglelefteq U_i$ und U_i/U_{i-1} abelsch für $1 \leqslant i \leqslant r$.*

Beweis. Sei zuerst G auflösbar und $G^{(r)} = \{1_G\}$ mit $r \geqslant 1$. Dann ist $G \supseteq G' = G^{(1)} \supseteq G^{(2)} \supseteq \ldots \supseteq G^{(r)} = \{1_G\}$ eine Folge von Untergruppen wie gewünscht, denn $G^{(i)}/G^{(i+1)}$ ist abelsch nach Satz 9.3(a). Sei umgekehrt eine Folge von Untergruppen $\{1\} = U_0 \leqslant U_1 \leqslant U_2 \leqslant \ldots \leqslant U_r = G$ wie oben gegeben. Wir zeigen mit Induktion nach r, dass G auflösbar ist. Ist $r = 1$, so ist G abelsch, also auflösbar. Sei nun $r \geqslant 2$ und $N := U_{r-1} \trianglelefteq U_r = G$. Dann ist G/N abelsch, also auflösbar. Wegen $\{1_G\} = U_0 \leqslant U_1 \leqslant \ldots \leqslant U_{r-1} = N$ ist N nach Induktion auflösbar. Mit Lemma 9.6 ist also auch G selbst auflösbar. □

Beispiel 9.9. (a) S_3, S_4, A_3, A_4 sind auflösbar. Denn in jedem Fall gibt es eine Folge von Untergruppen wie in Lemma 9.8. Nach Beispiel 7.3 hat S_3 den Normalteiler $\{\mathrm{id}\} \trianglelefteq V = A_3 \trianglelefteq S_3$, wobei $|V| = 3$ und $|S_3/V| = 2$. Nach Ü4A5 hat S_4 die Untergruppen $\{\mathrm{id}\} \leqslant V_4 \leqslant A_4 \leqslant S_4$, wobei $|V_4| = 4$, $V_4 \trianglelefteq A_4$ und $A_4 \trianglelefteq S_4$, wobei die Faktorgruppen jeweils abelsch sind.

(b) Sei $G = \langle s, t\rangle$ eine endliche Diedergruppe, mit $s \neq t$ der Ordnung 2. Nach Beispiel 6.3 ist $|G| = 2m$ wobei st Ordnung m hat. Nach Beispiel 7.2 ist $N = \langle st\rangle \trianglelefteq G$. Wegen $|G/N| = 2$ ist $\{1_G\} \subseteq N \subseteq G$ eine Folge von Untergruppen wie in Lemma 9.8, also G auflösbar.

Lemma 9.10. *Sei $n \geqslant 3$. Dann wird die alternierende Gruppe A_n von 3-Zykeln erzeugt (und alle 3-Zykeln gehören zu A_n).*

Beweis. In Beispiel 7.11(b) haben wir bereits gesehen, dass alle 3-Zykel in A_n liegen. Sei nun $\mathrm{id} \neq \pi \in A_n$. Nach Satz 6.5 gibt es Transpositionen $\tau_1, \ldots, \tau_r \in S_n$ mit $\pi = \tau_1 \circ \ldots \circ \tau_r$. Es gilt $\varepsilon(\tau_i) = -1$ für alle i. Weil $\varepsilon(\pi) = 1$ ist, muss r gerade sein, also $r = 2s$ mit $s \geqslant 1$. Also gilt auch $\pi = \sigma_1 \circ \ldots \circ \sigma_s$, wobei $\sigma_i = \tau_{2i-1} \circ \tau_{2i}$ für $1 \leqslant i \leqslant s$. Es genügt also zu zeigen, dass sich ein Produkt von zwei Transpositionen stets als Produkt von 3-Zykeln schreiben

lässt. Dazu: Seien $i, j, k, l \in \{1, \ldots, n\}$ gegeben mit $i \neq j$ und $k \neq l$; wir betrachten das Produkt $\sigma := (i\ j) \circ (k\ l)$. Ist $\{i, j\} = \{k, l\}$, so gilt $(i\ j) = (k\ l)$ und damit $\sigma = \mathrm{id}$. Sei nun $\{i, j\} \cap \{k, l\} = \varnothing$. Wegen $(j\ k) \circ (j\ k) = \mathrm{id}$ ist dann $\sigma = (i\ j) \circ (k\ l) = (i\ j) \circ (j\ k) \circ (j\ k) \circ (k\ l)$. Nun rechnet man leicht nach, dass $(i\ j) \circ (j\ k) = (i\ j\ k)$ und $(j\ k) \circ (k\ l) = (j\ k\ l)$ gilt, also ist $\sigma = (i\ j\ k) \circ (j\ k\ l)$. Sei schließlich $|\{i, j\} \cap \{k, l\}| = 1$; wir wählen die Bezeichnungen so, dass $j = k$ ist. Dann folgt $\sigma = (i\ j) \circ (k\ l) = (i\ j) \circ (j\ l) = (i\ j\ l)$. □

Satz 9.11. *Sei $n \geqslant 5$. Dann gilt $S_n' = A_n = A_n'$; damit sind S_n und A_n nicht auflösbar.*

Beweis. Weil S_n/A_n abelsch ist, folgt $S_n' \subseteq A_n$; und natürlich ist $A_n' \subseteq S_n'$. Wir zeigen nun, dass alle 3-Zykel in A_n' liegen. Nach Lemma 9.10 folgt dann $A_n \subseteq A_n'$ und wir sind fertig. Sei also $\sigma = (i\ j\ k) \in A_n$ ein beliebiger 3-Zykel. Wegen $n \geqslant 5$ gibt es noch Ziffern $l, m \in \{1, \ldots, n\} \setminus \{i, j, k\}$ und $l \neq m$. Wir setzen $\pi := (j\ k) \circ (m\ l) \in A_n$. Weil die Ziffernmengen jeweils disjunkt sind, ist die Transposition $(m\ l)$ mit σ und mit $(j\ k)$ vertauschar. Insbesondere ist $\pi^{-1} = \pi$; außerdem $\sigma^{-1} = (i\ k\ j)$. Damit erhalten wir

$$[\sigma, \pi] = \sigma^{-1} \circ \pi^{-1} \circ \sigma \circ \pi = (i\ k\ j) \circ (j\ k) \circ (m\ l) \circ (i\ j\ k) \circ (j\ k) \circ (m\ l) = \ldots = (i\ j\ k),$$

wobei die letzte Gleichheit mit einer einfachen Rechnung folgt. Also ist $\sigma = [\sigma, \pi] \in A_n'$. □

Das wohl stärkste bekannte Auflösbarkeits-Kriterium ist der folgende Satz:

Feit und Thompson (1963): *Ist $|G| < \infty$ ungerade, so ist G auflösbar.*

Siehe `http://en.wikipedia.org/wiki/Feit-Thompson_theorem` für weitere Informationen dazu. Der Original-Beweis ist 255 Seiten lang!

Beachte auch, dass wir bisher noch keine einzige nicht-abelsche einfache Gruppe kennen! (Im nächsten Abschnitt werden wir wenigstens zeigen, dass A_5 einfach ist.) Solche Gruppen sind im Allgemeinen sehr schwierig zu finden. Umso bemerkenswerter ist es, dass die komplette Liste aller endlichen einfachen Gruppen seit etwa 1981 bekannt ist. Dies ist generell eines der bedeutendsten Ergebnisse der Mathematik des 20. Jahrhunderts, und außerordentlich schwierig zu beweisen; siehe Solomon's Überblicksartikel [So].

Viele der endlichen einfachen Gruppen werden nach ihrem Entdecker benannt; so wurde etwa die einfache Gruppe $M_{11} := \langle\ (1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 10\ 11),\ (3\ 7\ 11\ 8) \circ (4\ 10\ 5\ 6)\ \rangle \leqslant S_{11}$ von É. L. Mathieu 1873 entdeckt und heißt die ***Mathieu-Gruppe*** M_{11}. (Mit GAP können Sie leicht die Ordnung von M_{11} bestimmen.) Siehe auch **`https://de.wikipedia.org/wiki/Sporadische_Gruppe`**. — Solche endlichen einfachen Gruppen zu finden ist wie das sprichwörtliche Finden einer Nadel in einem Heuhaufen!

10. *Sylow-Untergruppen*

In diesem Abschnitt sei G stets eine endliche Gruppe. Wir erhoffen uns, dass sich die Struktur von G irgendwie in der Primfaktorzerlegung von $|G|$ widerspiegelt. Dazu werden wir verschiedene Operationen von G betrachten. Wir erinnern vor Allem an die Operation von G auf sich durch Konjugation; siehe Beispiel 8.8. Diese Operation lässt sich wie folgt erweitern. Sei $U \leqslant G$ und $g \in G$. Dann ist $gUg^{-1} = \{gug^{-1} \mid u \in U\}$ wiederum eine Untergruppe (siehe Beispiel 8.11) und es gilt $|U| = |gUg^{-1}|$. (Die Abbildung $u \mapsto gug^{-1}$ ist ein Isomorphismus von U auf gUg^{-1}.) Eine solche Untergruppe heißt eine zu U ***konjugierte Untergruppe***. Die Gruppe G operiert dann auf $X := \{U \subseteq G \mid U \text{ Untergruppe}\}$ durch

$$G \times X \to X, \qquad (g, U) \mapsto gUg^{-1}.$$

(Überzeugen Sie sich selbst davon, dass dies eine Operation ist.) Hier heißt

$$N_G(U) := \mathrm{Stab}_G(U) = \{g \in G \mid gUg^{-1} = U\} \leqslant G$$

der ***Normalisator*** von U in G.
Beachte: Es gilt stets $U \subseteq N_G(U)$; außerdem ist $U \trianglelefteq G$ genau dann, wenn $G = N_G(U)$.

Definition 10.1. Sei p eine Primzahl und $|G| = p^a m$ mit $a \geqslant 0$ und $p \nmid m$. Eine Untergruppe $P \leqslant G$ heißt p-***Sylow-Untergruppe*** wenn $|P| = p^a$ gilt. Wir bezeichnen mit $\mathrm{Syl}_p(G)$ die Menge aller p-Sylow-Untergruppen von G.
Beachte auch: Ist $P \in \mathrm{Syl}_p(G)$ und $g \in G$, so ist auch $gPg^{-1} \in \mathrm{Syl}_p(G)$.

Beispiel 10.2. (a) Sei $G = S_4$. Dann ist $|G| = 2^3 \cdot 3$. Es gilt

$$P := \langle (1\,2), (1\,3)\circ(2\,4)\rangle \in \mathrm{Syl}_2(S_4) \qquad \text{und} \qquad Q := \langle (1\,2\,3)\rangle \in \mathrm{Syl}_3(S_4).$$

[Denn: $(1\,2)\circ((1\,3)\circ(2\,4)) = (1\,3\,2\,4)$ hat Ordnung 4, also ist P eine Diedergruppe der Ordnung 8.]

(b) Sei $K = \mathbb{Z}/p\mathbb{Z}$ Körper mit p Elementen und $G = \mathrm{GL}_n(K)$. Nach Beispiel 8.6(b) ist

$$|G| = p^{n(n-1)/2}(p^n - 1)(p^{n-1} - 1)\cdots(p - 1),$$

also $|G| = p^{n(n-1)/2}m$ mit $p \nmid m$. Wie in Beispiel 1.5 sieht man, dass die Menge $P \subseteq G$ aller oberen Dreiecksmatrizen mit 1 auf der Diagonalen eine Untergruppe ist. Da man alle $n(n-1)/2$ Einträge der Matrizen in P oberhalb der Diagonalen beliebig mit den p Elementen aus K besetzen kann, gilt $|P| = p^{n(n-1)/2}$, also $P \in \mathrm{Syl}_p(G)$.

(c) Es muss nicht zu jedem Teiler d von $|G|$ eine Untergruppe $U \leqslant G$ mit $|U| = d$ geben. Zum Beispiel gilt $A_n = A_n'$ für $n \geqslant 5$; es kann hier also keine Untergruppe vom Index 2 geben (denn sonst wäre diese ein Normalteiler mit abelscher Faktorgruppe).

Lemma 10.3. *Sei* $H \leqslant G$. *Dann gilt:* $\mathrm{Syl}_p(G) \neq \varnothing \;\Rightarrow\; \mathrm{Syl}_p(H) \neq \varnothing$.
Genauer: Ist $P \in \mathrm{Syl}_p(G)$, *so gibt es ein* $g \in G$ *mit* $gPg^{-1} \cap H \in \mathrm{Syl}_p(H)$.

Beweis. Die Gruppe H operiert auf $X = G/P$ durch Linksmultiplikation; siehe Beispiel 8.1(c). Sei $G/P = \mathcal{O}_1 \cup \ldots \cup \mathcal{O}_r$ die Zerlegung in H-Bahnen, wobei die Vereinigung disjunkt ist. Dann ist $|G/P| = |\mathcal{O}_1| + \ldots + |\mathcal{O}_r|$. Wegen $p \nmid |G/P|$ muss es ein i geben mit $p \nmid |\mathcal{O}_i|$. Sei $g_iP \in \mathcal{O}_i$ mit $g_i \in G$. Dann ist

$$\begin{aligned} Q &:= \mathrm{Stab}_H(g_iP) = \{h \in H \mid hg_iP = g_iP\} = \{h \in H \mid g_i^{-1}hg_i \in P\} \\ &= \{h \in H \mid h \in g_iPg_i^{-1}\} = g_iPg_i^{-1} \cap H \leqslant g_iPg_i^{-1}. \end{aligned}$$

Nach Lagrange ist wegen $|g_iPg_i^{-1}| = |P|$ also auch die Ordnung von Q eine p-Potenz. Andererseits ist nach dem Bahnensatz ("Länge der Bahn gleich Index des Stabilisators") $[H : Q] = |\mathcal{O}_i|$ und dies ist nicht durch p teilbar. D.h., $|Q|$ ist eine p-Potenz, und dies ist genau der p-Anteil in $|H|$; damit gilt $Q = g_iPg_i^{-1} \cap H \in \mathrm{Syl}_p(H)$. □

Satz 10.4 (Sylow)**.** *Sei* $|G| = p^a m$ *mit* $a \geqslant 0$ *und* $p \nmid m$ *(wobei* p *Primzahl).*

(a) *Es gilt* $\mathrm{Syl}_p(G) \neq \varnothing$.
(b) *Sei* $U \leqslant G$ *mit* $|U| = p^b$ *und* $0 \leqslant b \leqslant a$. *Dann gibt es ein* $P \in \mathrm{Syl}_p(G)$ *mit* $U \leqslant P$.
(c) *Sind* $P, Q \in \mathrm{Syl}_p(G)$, *so gibt es ein* $g \in G$ *mit* $P = gQg^{-1}$.
(d) *Sei* $n_p(G) := |\mathrm{Syl}_p(G)|$. *Dann gilt* $n_p(G) \equiv 1 \bmod p$ *und* $n_p(G) \mid m$.

Beweis. (a) Sei $K = \mathbb{Z}/p\mathbb{Z}$. Nach Folgerung 8.15 gibt es ein $n \geqslant 1$ und einen injektiven Gruppen-Homomorphismus $\varphi\colon G \to \mathrm{GL}_n(K)$. Sei $\tilde{G} = \varphi(G) \leqslant \mathrm{GL}_n(K)$; dann ist $G \cong \tilde{G}$. Nach Beispiel 10.2(b) ist $\mathrm{Syl}_p(\mathrm{GL}_n(K)) \neq \varnothing$. Also gibt es nach Lemma 10.3 ein $\tilde{P} \in \mathrm{Syl}_p(\tilde{G})$. Sei $P := \varphi^{-1}(\tilde{P}) \leqslant G$. Weil φ injektiv ist, gilt $|P| = |\tilde{P}| = p^a$, also folgt $P \in \mathrm{Syl}_p(G)$.

(b) Sei $Q \in \mathrm{Syl}_p(G)$. Lemma 10.3 (angewandt mit $H = U$) zeigt, dass es ein $g \in G$ gibt mit $gQg^{-1} \cap U \in \mathrm{Syl}_p(U)$. Wegen $|U| = p^b$ ist aber $\mathrm{Syl}_p(U) = \{U\}$, also gilt $gQg^{-1} \cap U = U$ und damit $U \subseteq P := gQg^{-1}$. Wegen $|Q| = |gQg^{-1}|$ folgt $P \in \mathrm{Syl}_p(G)$.

(c) Da $Q \in \mathrm{Syl}_p(G)$, können wir wieder Lemma 10.3 anwenden (diesmal mit $H = P$) und erhalten ein $g \in G$ mit $gQg^{-1} \cap P \in \mathrm{Syl}_p(P) = \{P\}$, also $P \subseteq gQg^{-1}$. Wegen $|P| = |Q| = |gQg^{-1}|$ folgt $P = gQg^{-1}$.

(d) G operiert auf $X = \mathrm{Syl}_p(G)$ durch Konjugation, denn für $Q \in \mathrm{Syl}_p(G)$ und $g \in G$ ist auch $gQg^{-1} \in \mathrm{Syl}_p(G)$. Nach (b) ist diese Operation transitiv, d.h., es gibt genau eine Bahn. Sei $P \in \mathrm{Syl}_p(G)$ fest. Dann ist $N_G(P) = \{g \in G \mid P = gPg^{-1}\}$ der Stabilisator von P unter dieser Operation. Mit dem Bahnensatz folgt also $|\mathrm{Syl}_p(G)| = [G : N_G(P)]$. Wegen $P \subseteq N_G(P)$ folgt außerdem $[G : N_G(P)] \mid [G : P] = m$. Damit gilt $n_p(G) \mid m$.

Um auch $n_p(G) \equiv 1 \bmod p$ zu zeigen, schränken wir die obige Operation auf P ein, d.h., wir lassen P auf $\mathrm{Syl}_p(G)$ durch Konjugation operieren. Diese Operation wird im Allgemeinen nicht mehr transitiv sein, also betrachten wir die Zerlegung in P-Bahnen:

$$\mathrm{Syl}_p(G) = \mathcal{O}_1 \cup \mathcal{O}_2 \cup \ldots \cup \mathcal{O}_r \qquad \text{(disjunkte Vereinigung)}.$$

Sei $P_i \in \mathcal{O}_i$ wobei wir die Bezeichnungen so wählen, dass $P_1 = P$ gilt. Nach dem Bahnensatz ist $|\mathcal{O}_i| = [P : N_P(P_i)] = p^{k_i}$ mit $k_i \geqslant 0$. Für $i = 1$ ist $gPg^{-1} = P$ für alle $g \in P$, also $\mathcal{O}_1 = \{P\}$ und $k_1 = 0$. Sei nun $i \geqslant 2$. Wäre auch hier $k_i = 0$, d.h., $\mathcal{O}_i = \{P_i\}$, so folgt $gP_ig^{-1} = P_i$ für alle $g \in P$, also $P \subseteq N_G(P_i)$. Damit sind P und P_i in $\mathrm{Syl}_p(N_G(P_i))$ enthalten, nach (b) gibt es also ein $h \in N_G(P_i)$ mit $P = hP_ih^{-1} = P_i$, Widerspruch zu $i \geqslant 2$. Also ist $k_i \geqslant 1$ und damit $p \mid |\mathcal{O}_i|$ für alle $i \geqslant 2$. Insgesamt folgt $n_p(G) = |\mathrm{Syl}_p(G)| = |\mathcal{O}_1| + \sum_{i=2}^{r} |\mathcal{O}_i| \equiv 1 \bmod p$. □

Folgerung 10.5 (Cauchy). *Ist p (Primzahl) ein Teiler von $|G|$, so gibt es ein $g \in G$ der Ordnung p.*

Beweis. Sei wie oben $|G| = p^a m$ mit $p \nmid m$; nach Voraussetzung ist $a > 0$. Nach Satz 10.4(a) gibt es ein $P \in \mathrm{Syl}_p(G)$. Sei $x \in P$ mit $x \neq 1_G$. Dann ist $o(x) = p^i$ mit $1 \leqslant i \leqslant a$; siehe Folgerung 3.6. Nach Bemerkung 3.7 gilt $o(g) = p$ für $g := x^{p^{i-1}} \in G$. □

Folgerung 10.6. *Sei p Primzahl und $P \in \mathrm{Syl}_p(G)$. Dann gilt: $P \trianglelefteq G \Leftrightarrow \mathrm{Syl}_p(G) = \{P\}$.*

Beweis. Nach Satz 10.4(b) sind alle Gruppen in $\mathrm{Syl}_p(G)$ konjugiert zu P. Ist also $P \trianglelefteq G$, dann folgt $\mathrm{Syl}_p(G) = \{P\}$. Sei umgekehrt $\mathrm{Syl}_p(G) = \{P\}$. Für $g \in G$ ist auch $gPg^{-1} \in \mathrm{Syl}_p(G)$, also folgt $gPg^{-1} = P$ und damit $P \trianglelefteq G$. □

Der Satz von Sylow ist von grundlegender Bedeutung für die Theorie der endlichen Gruppen. Nun zu einigen Anwendungen. Nach Beispiel 7.14 ist eine endlich erzeugte abelsche Gruppe isomorph zu einem direkten Produkt $U \times \mathbb{Z}^d$ wobei $d \geqslant 0$ und U eine endliche abelsche Gruppe ist. Für den endlichen Anteil U erhalten wir:

Satz 10.7 (Hauptsatz über endliche abelsche Gruppen). *Sei $G \neq \{1_G\}$ eine endliche abelsche Gruppe. Dann ist G isomorph zu einem direkten Produkt von zyklischen Gruppen der Form $\mathbb{Z}/p^n\mathbb{Z}$, wobei p eine Primzahl ist und $n \geqslant 1$.*[5]

Beweis. Sei $|G| = p_1^{n_1} \cdots p_r^{n_r}$ mit paarweise verschiedenen Primzahlen p_i und $n_i \geqslant 1$ für $1 \leqslant i \leqslant r$. Sei $P_i \in \mathrm{Syl}_{p_i}(G)$ für $1 \leqslant i \leqslant r$. Weil G abelsch ist, sieht man genauso wie in Beispiel 7.14, dass die Abbildung $\varphi\colon P_1 \times \ldots \times P_r \to G$, $(g_1, \ldots, g_r) \mapsto g_1 \cdots g_r$, ein Gruppen-Homomorphismus ist. Sei $U := \mathrm{Bild}(\varphi) \leqslant G$. Offenbar ist $P_i \subseteq U$ für $1 \leqslant i \leqslant r$, also ist

[5] Es gilt dann sogar, dass diese Darstellung von G als direktes Produkt eindeutig ist bis auf die Reihenfolge der Faktoren, aber das werden wir hier nicht zeigen; siehe dazu §10.2 im Buch von Karpfinger–Meyberg [KM].

$p_i^{n_i} = |P_i|$ ein Teiler von $|U|$. Wegen $p_i \neq p_j$ für $i \neq j$ folgt also auch $|G| = p_1^{n_1} \cdots p_r^{n_r} \mid |U|$ und damit $U = G$, d.h., φ is surjektiv. Wegen $|P_1 \times \ldots \times P_r| = |P_1| \cdots |P_r| = p_1^{n_1} \cdots p_r^{n_r} = |G|$ ist dann φ auch automatisch injektiv, also bijektiv. Es gilt damit $G \cong P_1 \times \ldots \times P_r$.
Jedes P_i ist natürlich selbst wieder eine endliche abelsche Gruppe. Nach Beispiel 7.14 ist P_i isomorph zu einem direkten Produkt von zyklischen Gruppen der Form $\mathbb{Z}/d\mathbb{Z}$ wobei $d \in \mathbb{N}$. Dabei muss nach Lagrange d ein Teiler von $|P_i|$ sein, also ist d eine Potenz von p_i. □

Beispiel 10.8. Sei $|G| = 2p$ mit einer Primzahl $p \geqslant 3$. Behauptung: Entweder ist G zyklisch oder eine Diedergruppe. Dazu: Nach dem Satz von Cauchy gibt es ein $x \in G$ mit $o(x) = p$ und ein $y \in G$ mit $o(y) = 2$. Wir unterscheiden zwei Fälle.

1. Fall: Es gilt $g := xy = yx$. Nun ist $o(g) \mid 2p$. Wegen $xy = yx$ folgt $g^2 = x^2y^2 = x^2 \neq 1_G$ und $g^p = x^py^p = y^p = y \neq 1_G$. Also muss $o(g) = 2p$ gelten, d.h., $G = \langle g \rangle \cong \mathbb{Z}/2p\mathbb{Z}$.

2. Fall: Es gilt $xy \neq yx$. Setze $z := xyx^{-1} \neq y$ und $U := \langle y, z \rangle$. Es gilt $z^2 = xyx^{-1}xyx^{-1} = xy^2x^{-1} = xx^{-1} = 1_G$ und $z \neq 1_G$, also $o(z) = 2$. Wegen $y \in U$ und $z \neq y$ ist $|U|$ gerade und $|U| > 2$. Nach Lagrange bleibt nur $|U| = 2p$ übrig, also ist $G = U = \langle y, z \rangle$ eine Diedergruppe.

Schließlich: Im Spezialfall $p = 3$ folgt damit $G \cong \mathbb{Z}/6\mathbb{Z}$ oder $G \cong S_3$. Denn: Im 2. Fall setzen wir $H := \langle y \rangle$. Dann operiert G auf $X = G/H$ und es gibt einen Homomorphismus $\pi\colon G \to S_X$ mit $\mathrm{Kern}(\pi) \subseteq H$; siehe Beispiel 8.11. Da H wegen $xyx^{-1} \notin H$ kein Normalteiler ist, folgt $\mathrm{Kern}(\pi) \subsetneqq H$, also $\mathrm{Kern}(\pi) = \{1_G\}$. Also ist G isomorph zu einer Untergruppe von S_X. Wegen $|X| = 3$ ist $S_X \cong S_3$ (siehe Beispiel 7.11(e)) und $|S_X| = |G|$, also $G \cong S_3$.

Ab hier Woche 7

Es folgen eine Reihe von Beispielen, in denen der Sylow-Satz benutzt wird, um die Auflösbarkeit von Gruppen zu zeigen.

Beispiel 10.9. (a) Sei $|G| = p^nq$ mit Primzahlen $p \neq q$ und $n \geqslant 1$. Ist $q + 1 > p^n$, so ist G auflösbar. Denn: Es ist $n_q(G) \mid p^n$ und $n_q(G) \equiv 1 \bmod q$, also bleibt wegen $q + 1 > p^n$ nur $n_q(G) = 1$ übrig. Sei $Q \in \mathrm{Syl}_q(G)$. Dann ist $Q \trianglelefteq G$ nach Folgerung 10.6. Wegen $|Q| = q$ und $|G/Q| = p^n$ sind Q und G/Q nach Beispiel 9.7 auflösbar, also auch G selbst (Lemma 9.6).

(b) Sei $|G| = 3p^n$ mit einer Primzahl $p \neq 3$ und $n \geqslant 1$. Dann ist G auflösbar. Dazu: Sei $P \in \mathrm{Syl}_p(G)$. Dann operiert G auf $X = G/P$ und wir erhalten einen Homomorphismus $\pi\colon G \to S_X \cong S_3$ mit $N := \mathrm{Kern}(\pi) \subseteq P$; siehe Beispiel 8.11. Dann ist $|N| = p^m$ mit $m \leqslant n$, also N auflösbar (Beispiel 9.7). Nach dem Homomorphiesatz ist $G/N \cong \mathrm{Bild}(\pi) \leqslant S_X$. Wegen $|X| = 3$ ist $S_X \cong S_3$ auflösbar, also auch $G/N \cong \mathrm{Bild}(\pi)$ auflösbar (Lemma 9.5). Also ist wiederum G selbst auflösbar (Lemma 9.6).

Beispiel 10.10. Ist $|G| = 30 = 2 \cdot 3 \cdot 5$ (der kleinste Fall mit mindestens drei Primzahlen in $|G|$), so ist G auflösbar. Dazu: Es ist $n_5(G) \mid 6$ und $n_5(G) \equiv 1 \bmod 5$, also bleibt nur

$n_5(G) \in \{1, 6\}$ übrig. Ist $n_5(G) = 1$, so sei $P \in \mathrm{Syl}_5(G)$. Dann ist $P \trianglelefteq G$. Wegen $|P| = 5$ ist $P \cong \mathbb{Z}/5\mathbb{Z}$ auflösbar; wegen $|G/P| = 6$ ist G/P auflösbar nach Beispiel 10.9(a). Also ist G selbst auflösbar. Sei nun $n_5(G) = 6$ und $\mathrm{Syl}_5(G) = \{P_1, \ldots, P_6\}$. Wegen $P_i \cong \mathbb{Z}/5\mathbb{Z}$ gilt $P_i \cap P_j = \{1_G\}$ für $i \neq j$, also gibt es $6 \cdot (5-1) = 24$ Elemente der Ordnung 5 in G. Weiterhin ist $n_3(G) \mid 10$ und $n_3(G) \equiv 1 \bmod 3$. Wäre $n_3(G) > 1$, so $n_3(G) \geqslant 4$, also gäbe es mit einem analogen Argument mindestens $4 \cdot (3-1) = 8$ Elemente der Ordnung 3, also hätte G bereits $24 + 8 > 30$ Elemente, Widerspruch. Also ist $n_3(G) = 1$. Sei $P \in \mathrm{Syl}_3(G)$; dann ist $P \trianglelefteq G$. Wegen $|G/P| = 10 = 2 \cdot 5$ ist G/P auflösbar nach Beispiel 10.9(a); also ist auch G auflösbar.

Beispiel 10.11. Ist $|G| = 2020$, so ist G auflösbar.
Denn: $|G| = 2^2 \cdot 5 \cdot 101$. Es folgt $n_{101}(G) \mid 20$ und $n_{101}(G) \equiv 1 \bmod 101$, also bleibt nur $n_{101}(G) = 1$ übrig. Sei $P \in \mathrm{Syl}_{101}(G)$. Dann ist $P \trianglelefteq G$ mit Folgerung 10.6. Nun ist $|G/P| = 2^2 \cdot 5$, also G/P auflösbar nach Beispiel 10.9(a). Also ist G selbst auflösbar.
(In manch anderen Jahren, in denen diese Vorlesung stattfindet, ist dieses Beispiel nicht so einfach ...)

Folgerung 10.12. (a) *Alle Gruppen der Ordnung* $\leqslant 30$ *sind auflösbar.*
(b) *Die alternierende Gruppe* A_5 *ist einfach.*

Beweis. (a) Sei G eine Gruppe mit $|G| \leqslant 30$. Ist $|G|$ die Potenz einer Primzahl, so ist G auflösbar nach Beispiel 9.7. Es bleibt noch $|G| \in \{6, 10, 12, 14, 15, 18, 20, 21, 22, 24, 26, 28, 30\}$ zu betrachten. Mit den Beispielen 10.9 und 10.10 erledigt man alle diese Fälle bis auf $|G| = 18$. Ist $|G| = 18 = 2 \cdot 3^2$, so sei $P \in \mathrm{Syl}_3(G)$; dann ist $[G : P] = 2$, also $P \trianglelefteq G$. Nun sind P und G/P auflösbar, also auch G.

(b) Sei $N \trianglelefteq A_5$. Annahme: $\{1_G\} \neq N \neq A_5$. Wegen Lagrange gilt dann $|N| \leqslant 30$ und $|A_5/N| \leqslant 30$, also sind N und A_5/N auflösbar nach (a). Nach Lemma 9.6 ist dann aber auch A_5 selbst auflösbar, Widerspruch zu Satz 9.11. □

Bemerkung 10.13. Mit ähnlichen Methoden lässt sich zeigen:
(a) Ist $|G| < 60$, so ist G auflösbar.
(b) Ist $|G| = 60$ und G einfach, so folgt $G \cong A_5$.
Damit ist A_5 (bis auf Isomorphie) die kleinste nicht-abelsche einfache Gruppe. Siehe Ü7 und das Buch Kurzweil–Stellmacher [KS] für weitere Einzelheiten und Anwendungen.

Versuchen Sie selbst ähnliche Aufgaben wie in Beispiel 10.11 oder 10.9, wobei Sie als $|G|$ Ihr Geburtsjahr oder sonstige Zahlen nehmen, die Ihnen gerade einfallen. Sie werden feststellen, dass es bei einigen Zahlen gar nicht so leicht ist weiterzukommen, zum Beispiel $|G| = 168$, 360, 504, 660, 1092, 2448, ... Tippen Sie diese Ziffern in die ***On-Line Encyclopedia of Integer Sequences*** ein (siehe `http://oeis.org`), um den Grund dafür zu finden.

Kapitel III: Ringe

Der ***Hauptsatz der elementaren Arithmetik*** besagt, dass sich jede natürliche Zahl $\neq 1$ schreiben lässt als ein Produkt von Primzahlen, wobei die Faktoren bis auf die Reihenfolge eindeutig bestimmt sind. Wir wollen nun als Erstes untersuchen, unter welchen allgemeinen Bedingungen eine solche Aussage richtig bleibt.

11. *Faktorielle Ringe*

Es sei stets $(R, +, \cdot)$ ein kommutativer Ring mit $1 \neq 0$ (zum Beispiel $R = \mathbb{Z}$). Sind $a, b \in R$, so schreiben wir $a \mid b$, wenn es ein $c \in R$ gibt $b = a \cdot c$. Ein Element $a \in R$ heißt ***Nullteiler***, wenn es ein $0 \neq c \in R$ gibt mit $a \cdot c = 0$. Gibt es keine Nullteiler außer 0 in R, so heißt R ein ***Integritätsring***. In einem solchen Ring gilt also für alle $a, b \in R$:

$$a \cdot b = 0 \qquad \Rightarrow \qquad a = 0 \quad \text{oder} \quad b = 0.$$

Standardbeispiel: $R = \mathbb{Z}$; dagegen ist $R = \mathbb{Z}/4\mathbb{Z}$ kein Integritätsring, wegen $\bar{2} \cdot \bar{2} = \bar{0}$. Weitere Bemerkungen: Ist $0 \neq a \in R^\times$, so gilt $a \cdot b \neq 0$ für alle $0 \neq b \in R$. [Denn sonst $a \cdot b = 0 \Rightarrow b = a^{-1} \cdot (a \cdot b) = a^{-1} \cdot 0 = 0$.] Insbesondere: Jeder Körper ist ein Integritätsring; oder allgemeiner: Jeder Teilring (mit $1 \neq 0$) eines Körpers ist ein Integritätsring.

Bemerkung 11.1. Sei R ein Integritätsring.
(1) Es gilt folgende Kürzungsregel: Sind $a, b, c \in R$ mit $a \cdot c = b \cdot c$ und $c \neq 0$, so gilt $a = b$.
[Denn $a \cdot c = b \cdot c \Rightarrow (a - b) \cdot c = 0$, also $a - b = 0$ weil $c \neq 0$.]
(2) Seien $a, b \in R \setminus \{0\}$. Gilt $a \mid b$ und $b \mid a$, so folgt $b = u \cdot a$ mit einer Einheit $u \in R^\times$.
[Denn schreibe $b = u \cdot a$ und $a = v \cdot b$ mit $u, v \in R$. Dann ist $b = u \cdot a = u \cdot v \cdot b$. Wegen $b \neq 0$ können wir b auf beiden Seiten kürzen, also folgt $1 = u \cdot v$.]
(3) Seien $a, b \in R \setminus \{0\}$. Dann gilt $(a) = (b) \Leftrightarrow b = u \cdot a$ mit einem $u \in R^\times$.
[Dazu: Ist $b = u \cdot a$ mit $u \in R^\times$, so ist $b \in (a)$, also auch $(b) \subseteq (a)$. Ebenso $a = u^{-1} \cdot b \in (b)$, also $(a) \subseteq (b)$ und damit $(a) = (b)$. Umgekehrt sei $(a) = (b)$. Dann ist $a \in (b)$, also $b \mid a$; genauso $b \in (a)$, also $a \mid b$. Nach (2) gilt $b = u \cdot a$ mit einem $u \in R^\times$.]

Definition 11.2. Sei R ein Integritätsring. Sei $0 \neq p \in R$ ein Element, das keine Einheit ist.
(a) p heißt ***irreduzibel***, wenn gilt: Ist $p = a \cdot b$ mit $a, b \in R$, so ist a oder b eine Einheit.
(b) p heißt ***Primelement***, wenn gilt: Sind $a, b \in R$ mit $p \mid a \cdot b$, so ist $p \mid a$ oder $p \mid b$.

Beachte: Ist $u \in R^\times$ und $p \in R$ irreduzibel (oder ein Primelement), so ist auch $u \cdot p$ irreduzibel (bzw. ein Primelement).

Zum Beispiel sind in $\mathbb{Z}$ die irreduziblen Elemente gegeben durch $\pm p$ mit einer Primzahl p. Für diese Elemente gilt auch die Bedingung in (b). Denn sei $p \in \mathbb{Z}$ irreduzibel und $p \mid ab$

mit $a, b \in \mathbb{Z}$. Annahme: $p \nmid a$. Dann ist $\mathrm{ggT}(p, a) = 1$, also gibt es $r, s \in \mathbb{Z}$ mit $1 = rp + sa$ (Bézout). Dann folgt $b = rbp + sab$; nun teilt p die rechte Seite, also auch b.

Beispiel 11.3. Wir betrachten den Teilring $R = \mathbb{Z}[\sqrt{-5}] = \{n + m\sqrt{-5} \mid n, m \in \mathbb{Z}\} \subseteq \mathbb{C}$; siehe Ü2A1. Hier gilt: $2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$.
Behauptung: $R^\times = \{\pm 1\}$; außerdem sind $2, 3, 1 \pm \sqrt{-5}$ irreduzibel, aber keine Primelemente.

Dazu: Für $z \in \mathbb{C}$ sei wie üblich $\bar{z} \in \mathbb{C}$ die konjugiert-komplexe Zahl und $|z| = \sqrt{z\bar{z}}$ der Absolutbetrag. Wir setzen $N(z) := |z|^2 = z\bar{z}$. Für beliebiges $a = n + m\sqrt{-5} \in R$ gilt also:

$$N(a) = (n + m\sqrt{-5})(n - m\sqrt{-5}) = n^2 + 5m^2, \quad \text{und dies ist eine positive ganze Zahl.}$$

Außerdem gilt für beliebige $u, v \in \mathbb{C}$ auch $|uv| = |u||v|$, also ist hier $N(ab) = N(a)N(b)$ für alle $a, b \in R$. Daraus folgt sofort: Ist $a \in R^\times$, so muss $N(a) = 1$ gelten. (Denn ist $1 = ab$ für ein $b \in R$, so $N(a)N(b) = N(ab) = N(1) = 1$, also $N(a) = N(b) = 1$.) Aber für $N(a) = 1$ gibt es nur die Lösung $n = \pm 1$ und $m = 0$, d.h., $a = \pm 1 \in R^\times$. Damit ist $R^\times = \{\pm 1\}$ gezeigt.

Sei nun $2 = ab$ mit $a, b \in R$. Dann ist $N(a)N(b) = N(ab) = N(2) = 4$, also $N(a) \in \{1, 2, 4\}$. Die Formel für $N(a)$ zeigt, dass niemals $N(a) = 2$ sein kann. Ist $N(a) = 1$, so ist $a = \pm 1 \in R^\times$. Ist $N(a) = 4$, so ist $N(b) = 1$, also $b = \pm 1 \in R^\times$. Damit ist 2 irreduzibel. Angenommen, 2 wäre ein Primelement. Dann müsste 2 ein Teiler von $1+\sqrt{-5}$ oder $1-\sqrt{-5}$ sein, also $4 = N(2)$ ein Teiler von $N(1 \pm \sqrt{-5}) = 6$, Widerspruch. Damit ist 2 irreduzibel, aber kein Primelement. Der Beweis für 3 und $1 \pm \sqrt{-5}$ ist analog.

Der Einfachheit halber schreiben wir ab jetzt ab für $a \cdot b$ (wobei $a, b \in R$).

Lemma 11.4. *Sei R ein Integritätsring und $0 \neq p \in R$. Dann gilt:*

$$R/(p) \text{ Integritätring} \quad \Leftrightarrow \quad p \text{ Primelement} \quad \Rightarrow \quad p \text{ irreduzibel.}$$

Beweis. Sei p Primlement. Seien $a, b \in R$ mit $p = ab$. Dann ist $p \mid ab$ also $p \mid a$ oder $p \mid b$. Nehmen wir an, es wäre $p \mid a$, also $a = pc$ mit $c \in R$. Dann ist aber $p = ab = pbc$. Wegen $p \neq 0$ können wir p auf beiden Seiten kürzen, also folgt $1 = bc$, d.h., $b \in R^\times$. Ist $p \mid b$, so folgt analog $a \in R^\times$. Also ist p irreduzibel.
Betrachte nun $R/(p)$. Für $a \in R$ sei $\bar{a} := a+(p) \in R/(p)$. Sei zuerst $R/(p)$ ein Integritätsring; dann ist $\bar{1} \neq \bar{0}$, also $(p) \subsetneq R$ und damit $p \notin R^\times$ (denn sonst $1 = up$ mit einem $u \in R$ und dann $a = aup \in (p)$ für alle $a \in R$, Widerspruch). Seien nun $a, b \in R$ mit $p \mid ab$. Dann folgt $ab \in (p)$ und damit $\bar{a} \cdot \bar{b} = \overline{ab} = \bar{0}$. Da $R/(p)$ ein Integritätsring ist, folgt $\bar{a} = \bar{0}$ oder $\bar{b} = \bar{0}$. Im ersten Fall ist $a \in (p)$, also $p \mid a$; im zweiten Fall ist analog $p \mid b$.
Sei umgekehrt p ein Primelement. Dann ist $p \notin R^\times$ und damit $(p) \subsetneq R$ (denn sonst $1 = up$ mit einem $u \in R$), also $\bar{0} \neq 1$ in $R/(p)$. Seien $a, b \in R$ mit $\bar{a}\bar{b} = \bar{0}$. Dann ist $\overline{ab} = \bar{0}$ also $ab \in (p)$ und damit $p \mid ab$. Da p Primelement, folgt $p \mid a$ oder $p \mid b$, d.h., $a \in (p)$ oder $b \in (p)$, also schließlich $\bar{a} = \bar{0}$ oder $\bar{b} = \bar{0}$. □

Satz 11.5. *Sei* R *ein Integritätsring. Seien* $n, m \geqslant 1$ *und* $p_1, \ldots, p_n, q_1, \ldots, q_m \in R$ *Primelemente mit* $p_1 \cdots p_n = uq_1 \cdots q_m$ *wobei* $u \in R^\times$. *Dann gilt* $n = m$ *und es gibt* $u_1, \ldots, u_n \in R^\times$ *mit* $q_{\pi(i)} = u_i p_i$ *für* $1 \leqslant i \leqslant n$, *wobei* $\pi\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ *eine Bijektion ist (die also lediglich eine Umordnung der Reihenfolge der Faktoren bewirkt).*

Beweis. Induktion nach $n + m \geqslant 2$. Ist $n + m = 2$, so $n = m = 1$ und nach Voraussetzung $p_1 = uq_1$, also Behauptung klar. Sei nun $n + m > 2$. Es gilt $p_1 \mid p_1 \cdots p_n = uq_1 \cdots q_m$, also auch $p_1 \mid q_1 \cdots q_m$ (weil $u \in R^\times$). Behauptung: Es gibt ein $i_1 \in \{1, \ldots, m\}$ mit $p_1 \mid q_{i_1}$.
Dazu: Wegen $p_1 \mid q_1(q_2 \cdots q_m)$ und p_1 Primelement folgt $p_1 \mid q_1$ oder $p_1 \mid q_2 \cdots q_m$. Ist $p_1 \mid q_1$, so sind wir fertig (mit $i_1 = 1$). Andernfalls teilt p_1 also $q_2 \cdots q_m$. Wiederum folgt $p_1 \mid q_2$ (und wir sind fertig mit $i_1 = 2$) oder $p_1 \mid q_3 \cdots q_m$. Wir fahren auf diese Weise fort, bis wir schließlich ein $i_1 \in \{1, \ldots, m\}$ erhalten mit $p_1 \mid q_{i_1}$.

Schreibe nun $q_{i_1} = u_1 p_1$ mit $u_1 \in R$. Nun ist q_{i_1} ein Primelement, also auch irreduzibel nach Lemma 11.4. Weil p_1 keine Einheit ist, muss also $u_1 \in R^\times$ gelten. Damit folgt $p_1 \cdots p_n = uq_1 \cdots q_{i_1-1} u_1 p_1 q_{i_1+1} \cdots q_m = (uu_1) q_1 \cdots q_{i_1-1} p_1 q_{i_1+1} \cdots q_m$. Wegen $p_1 \neq 0$ können wir p_1 auf beiden Seiten kürzen, also erhalten wir eine neue Gleichung

$$p_2 \cdots p_n = u' q_1 \cdots q_{i_1-1} q_{i_1+1} \cdots q_m \qquad \text{mit } u' = uu_1 \in R^\times.$$

Nach Induktion ist $n - 1 = m - 1$, also $n = m$, und es gibt $u_2, \ldots, u_n \in R^\times$ mit $q_{\pi(i)} = u_i p_i$ für $2 \leqslant i \leqslant n$, wobei $\pi\colon \{2, \ldots, n\} \to \{1, \ldots, i_1 - 1, i_1 + 1, \ldots, m\}$ eine Bijektion ist. Setze $\pi(1) = i_1$. Dann ist $\pi\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ die gewünschte Bijektion. □

Definition 11.6. Sei R Integritätsring. Dann heißt R ein ***faktorieller Ring***, wenn folgende beiden Bedingungen gelten:

(1) Sei $0 \neq a \in R$ keine Einheit. Dann lässt sich a schreiben als $a = p_1 p_2 \cdots p_r$ mit $r \geqslant 1$ und $p_i \in R$ irreduzibel für alle i.
(2) Jedes irreduzible Element in R ist ein Primelement.

Nach Satz 11.5 ist dann jede Darstellung wie in (1) bis auf die Reihenfolge der Faktoren und Multiplikation mit Einheiten eindeutig. Faktorielle Ringe sind also genau die Ringe, in denen sinngemäß der ***Hauptsatz der elementaren Arithmetik*** gilt. Wir wollen nun noch Kriterien finden, mit denen man (leicht) zeigen kann, dass ein Integritätsring faktoriell ist.

Definition 11.7. Der Ring R heißt ***Hauptidealring***, wenn jedes Ideal $I \trianglelefteq R$ ein Hauptideal ist, also von der Form $I = (a)$ mit einem $a \in R$ ist (siehe Beispiel 2.5).

Bemerkung 11.8. Sei R ein Hauptidealring. Dann gilt eine analoge Version des ***Lemmas von Bézout***. Seien $a, b \in R$ mit $a \neq 0$ oder $b \neq 0$. Dann sieht man leicht, dass die Teilmenge $I := \{ra + sb \mid r, s \in R\} \subseteq R$ ein Ideal ist. Also gibt es ein $0 \neq d \in R$ mit $I = (d)$; insbesondere

gibt es $r, s \in R$ mit $d = ra + sb$. Wegen $a, b \in I = (d)$ gilt $d \mid a$ und $d \mid b$; also ist d ein gemeinsamer Teiler von a und b. Und für jeden weiteren gemeinsamen Teiler $0 \neq d' \in R$ von a und b gilt dann wegen $d = ra + sb$ auch $d' \mid d$.

Satz 11.9. *Sei R ein Integritätsring, der ein Hauptidealring ist.*
(a) *Ist $p \in R$ irreduzibel, so ist $R/(p)$ ein Körper.*
(b) *Jedes irreduzible Element in R ist auch ein Primelement.*

Beweis. (a) Für $a \in R$ schreiben wir $\bar{a} := a + (p) \in R/(p)$. Wegen $p \notin R^\times$ folgt zunächst $(p) \subsetneqq R$ (denn sonst $1 = up$ mit einem $u \in R$), also $\bar{0} \neq \bar{1}$ in $R/(p)$. Sei nun $a \in R$ mit $\bar{a} \neq \bar{0}$. Wir müssen zeigen, dass $\bar{a} \in R/(p)$ eine Einheit ist.
Nach Bemerkung 11.8 ist $I := \{rp + sa \mid r, s \in R\} \trianglelefteq R$ ein Ideal und es gibt $d, r, s \in R$ mit $I = (d)$ und $d = rp + sa$. Behauptung: $d \in R^\times$. Denn: Wegen $p \in I = (d)$ ist $p = cd$ mit $c \in R$. Wäre $d \notin R^\times$, so müsste $c \in R^\times$ sein, also $d = c'p$ mit $c' = c^{-1} \in R$ und damit $p \mid d$. Nun ist auch $a \in I = (d)$, also $c'p = d \mid a$ und damit $p \mid a$, Widerspruch zu $\bar{a} \neq \bar{0}$.
Also ist in der Tat $d \in R^\times$ und wir erhalten $1 = r'p + s'a$ mit $r' = d^{-1}r \in R$ und $s' = d^{-1}s \in R$. Dann ist $\bar{1} = \bar{r}'\bar{p} + \bar{s}'\bar{a} = \bar{s}'\bar{a}$, also $\bar{a}$ eine Einheit in $R/(p)$.
(b) Sei $p \in R$ irreduzibel. Nach (a) ist $R/(p)$ ein Körper, also insbesondere ein Integritätsring. Mit Lemma 11.4 folgt, dass p Primelement ist. □

Definition 11.10. Sei R ein Integritätsring. Dann heißt R ein ***Euklidischer Ring***, wenn es eine "Norm"-Funktion $\nu\colon R \setminus \{0\} \to \mathbb{N}_0$ mit folgender Eigenschaft gibt: Zu $a, b \in R$ mit $b \neq 0$ gibt es $q, r \in R$ mit $a = bq + r$, wobei entweder $r = 0$ oder $r \neq 0$ und $\nu(r) < \nu(b)$ gilt.

Beispiel 11.11. (a) $R = \mathbb{Z}$ ist euklidisch mit $\nu(n) = |n|$ (Absolutbetrag). Denn seien $a, b \in \mathbb{Z}$ mit $b \neq 0$. Ist $b > 0$, so teile mit Rest $a = qb + r$ wobei $q, r \in \mathbb{Z}$ und $0 \leqslant r < b$. Dann ist $r = 0$ oder $\nu(r) = r < b = \nu(b)$. Ist $b < 0$, so teile $-a$ mit Rest durch $-b$, also $-a = -q'b + r'$ wobei $q', r' \in \mathbb{Z}$ mit $0 \leqslant r' < -b$. Dann gilt $a = qb + r$ mit $q := -q'$, $r := -r'$, wobei $r = 0$ oder $\nu(r) = \nu(-r') = r' < -b = \nu(b)$.

(b) $R = \mathbb{Z}[i] = \{n + mi \mid n, m \in \mathbb{Z}\}$ ist euklidisch mit $\nu(n + mi) = n^2 + m^2$; siehe Ü8A1.

(c) Ist K ein Körper, so ist der Polynomring $K[X]$ euklidisch; mehr dazu in §12.

In einem Euklidischen Ring gibt es dann auch einen (erweiterten) ***Euklidischen Algorithmus***, völlig analog zum Verfahren im Beweis vom Lemma von Bézout.

Satz 11.12. *Sei R ein Integritätsring. Dann gelten die Implikationen:*

$$R \textit{ Euklidischer Ring} \quad \Rightarrow \quad R \textit{ Hauptidealring} \quad \Rightarrow \quad R \textit{ faktoriell.}$$

Beweis. Sei R euklidisch und $I \trianglelefteq R$ ein Ideal. Ist $I = \{0\}$, so ist natürlich $I = (0)$ ein Hauptideal. Sei nun $I \neq \{0\}$. Dann ist $X := \{\nu(a) \mid 0 \neq a \in I\}$ eine nicht-leere Teilmenge

von $\mathbb{N}_0$, also existiert $d := \min X \in \mathbb{N}_0$. Sei $0 \neq a_0 \in I$ mit $\nu(a_0) = d$. Behauptung: $I = (a_0)$. Dazu: Sei $b \in I$ beliebig. Wegen R euklidisch ist $b = qa_0 + r$ mit $q, r \in R$, wobei entweder $r = 0$ oder $r \neq 0$ und $\nu(r) < \nu(a_0)$ gilt. Wäre $r \neq 0$, so $r = b - qa_0 \in I$ und $\nu(r) < \nu(a_0) = d$, Widerspruch zur Minimalität von d. Also ist $r = 0$ und damit $b \in (a_0)$, d.h., $I = (a_0)$. Damit ist gezeigt, dass R ein Hauptidealring ist. (Dieses Argument is analog zu Beispiel 3.11(a)).

Sei nun R ein Hauptidealring. Nach Satz 11.9(b) gilt Bedingung (2) in Definition 11.6. Wir bezeichnen mit $R^\#$ die Menge aller $a \in R$ mit $a \neq 0$ und $a \notin R^\times$. Wir müssen zeigen, dass auch Bedingung (1) in Definition 11.6 gilt. Annahme: Dies ist nicht der Fall. Dann gibt es ein $a \in R^\#$, so dass sich a nicht als Produkt von endlich vielen irreduziblen Elementen schreiben lässt. Insbesondere ist a selbst nicht irreduzibel, also gilt $a = a_1 b_1$ mit $a_1, b_1 \in R^\#$. Ist sowohl a_1 als auch b_1 ein Produkt von endlich vielen irreduziblen Elementen, so gilt dies natürlich auch für $a = a_1 b_1$, Widerspuch. Also gilt die Bedingung (1) nicht für wenigstens eines der beiden Elemente a_1, b_1; wir wählen die Bezeichnungen so, dass dies a_1 ist. Dann ist $a_1 \mid a$, also $(a) \subseteq (a_1)$. Wäre $(a) = (a_1)$, so $a = ua_1$ mit einem $u \in R^\times$ (siehe Bemerkung 11.1(3)), und damit $b_1 = u \in R^\times$, Widerspruch. Also ist $(a) \subsetneq (a_1)$. Wir können nun das ganze Argument mit a_1 anstelle von a wiederholen. Also gibt es ein $a_2 \in R^\#$ mit $(a_1) \subsetneq (a_2)$ und so, dass Bedingung (1) nicht für a_2 gilt. Durch weitere Wiederholungen dieses Argumentes folgt die Existenz einer unendlichen Folge $(a_n)_{n\in\mathbb{N}}$ in $R^\#$ mit $(a) \subsetneq (a_1) \subsetneq (a_2) \subsetneq \ldots$. Nun setze $I := \bigcup_{n\in\mathbb{N}}(a_n)$. Genauso wie in Ü1A6(c) sieht man, dass I ein Ideal ist, also nach Voraussetzung $I = (d)$ für ein $d \in R$. Wegen $d \in I = \bigcup_{n\in\mathbb{N}}(a_n)$ ist $d \in (a_N)$ für ein $N \in \mathbb{N}$, also $(d) \subseteq (a_N)$. Andererseits ist $(a_{N+1}) \subseteq I = (d) \subseteq (a_N) \subsetneq (a_{N+1})$, Widerspruch.[6] □

Ab hier Woche 8

Bemerkung 11.13. (a) Sei $R = \mathbb{Z}$. Nach Beispiel 11.11(a) ist $\mathbb{Z}$ ein Euklidischer Ring. Nach Satz 11.12 ist $\mathbb{Z}$ also sowohl ein Hauptidealring als auch ein faktorieller Ring. Damit haben wir als Spezialfall auch noch einmal den Hauptsatz der elementaren Arithmetik bewiesen.

(b) Die Umkehrungen im obigen Satz gelten nicht. Beispiel: Es ist bekannt[7], dass

$$R = A_{-19} = \{n + m\omega \mid n, m \in \mathbb{Z}\} \subseteq \mathbb{C}, \qquad \text{wobei} \qquad \omega := \tfrac{1}{2}(1 + \sqrt{-19}),$$

ein Hauptidealring ist, aber kein Euklidischer Ring. Siehe Bemerkung 12.3 und Beispiel 13.12 weiter unten für Beispiele von faktoriellen Ringen, die keine Hauptidealringe sind.

[6] Für die Fans der Mengentheorie: Egal wie man dieses Argument organisiert, man braucht für die Implikation "R Hauptidealring ⇒ R faktoriell" das Auswahlaxiom; siehe Corollary 10 in W. HODGES, Läuchli's algebraic closure of $\mathbb{Q}$, Math. Proc. Cambridge Philos. Soc. 79 (1976), 289–297. Man kann aber die Implikation "R Euklidisch ⇒ R faktoriell" direkt zeigen, siehe zum Beispiel Satz 9.7 im Algebra-Skript [G14].

[7] Siehe zum Beispiel Chap. II, §5 im Buch von Perrin [Pe], oder R. A. WILSON, An elementary proof that not all principal ideal domains are Euclidean domains, The Mathematical Gazette 101 (2017), 289–293. Siehe auch A. BEVELACQUA, A family of non-Euclidean PIDs, Amer. Math. Monthly 123 (2016), 936–939.

Bemerkung 11.14. Ist R faktoriell, so können wir für jedes $0 \neq a \in R$ eine ***Länge*** $\ell(a)$ definieren: Ist $a \in R^\times$, so sei $\ell(a) = 0$; ist $a \notin R^\times$, so schreibe $a = p_1 \cdots p_r$ mit irreduziblen Elementen $p_i \in R$ und setze $\ell(a) := r$. Dies ist wohl-definiert nach Satz 11.5. Dieser Satz zeigt auch, dass $\ell(ab) = \ell(a) + \ell(b)$ für alle $a, b \in R$ mit $a \neq 0$, $b \neq 0$ gilt; außerdem gilt für $0 \neq a \in R$: $\quad a \in R^\times \Leftrightarrow \ell(a) = 0$. Dies wird nützlich in Induktionsbeweisen sein.

12. *Polynomringe*

Aus der Linearen Algebra sind Sie sicherlich mit dem Ring der Polynome $K[X]$ mit Koeffizienten in einem Körper K vertraut (sowie der Unterscheidung zwischen Polynomen und Polynomfunktionen). Sei nun R ein kommutativer Ring mit 1. Genauso kann man dann den ***Polynomring*** $R[X]$ in einer Unbestimmten X mit Koeffizienten in R bilden. Es gilt $R \subseteq R[X]$. Jedes $0 \neq f \in R[X]$ lässt sich eindeutig schreiben in der Form

$$f = a_0 + a_1 X + \cdots + a_n X^n \qquad \text{mit } n \geqslant 0,\ a_i \in R \text{ und } a_n \neq 0.$$

In diesem Fall heißt n der ***Grad*** von f und a_n der ***Leitkoeffizient***. Ist $a_n = 1$, so heißt f ***normiert***. (Als Konvention setzen wir auch $\mathrm{Grad}(0) = -\infty$.) Ist $g = b_0 + b_1 X + \ldots + b_m X^m \in R[X]$ ein weiteres Polynom, mit $m \geqslant 0$ und $b_m \neq 0$, so gilt

$$f \cdot g = a_0 b_0 + (a_1 b_0 + a_0 b_1) X + \cdots + a_n b_m X^{n+m};$$

ist also $a_n b_m \neq 0$, so hat fg den Grad $n + m$ und $a_n b_m$ ist der Leitkoeffizient von $f \cdot g$. Dies zeigt: Ist R ein Integritätsring, so auch $R[X]$; außerdem ist dann $(R[X])^\times = R^\times$.
(Denn ist $f \in (R[X])^\times$, so gibt es $g \in R[X]$ mit $1 = fg$, also $0 = \mathrm{Grad}(1) = \mathrm{Grad}(f) + \mathrm{Grad}(g)$ und damit $\mathrm{Grad}(f) = \mathrm{Grad}(g) = 0$, d.h., $f, g \in R$ und damit $f \in R^\times$.)

[Zur Erinnerung: Wie konstruiert man $R[X]$? Sei $\mathcal{F}$ die Menge aller Folgen $(a_n)_{\geqslant 0}$ mit $a_n \in R$, so dass es ein $n_0 \geqslant 0$ gibt (welches von f abhängt) mit $a_n = 0$ für alle $n \geqslant n_0$. Wir definieren eine Addition und Multiplikation wie folgt: $(a_n)_{n\geqslant 0} + (b_n)_{n\geqslant 0} := (a_n + b_n)_{\geqslant 0}$ und $(a_n)_{n\geqslant 0} \cdot (b_n)_{n\geqslant 0} := (c_n)_{n\geqslant 0}$ wobei $c_n = \sum_{i=0}^{n} a_i b_{n-i}$ (Cauchy-Produkt). Dann zeigt man, dass $(\mathcal{F}, +, \cdot)$ ein kommutativer Ring ist, mit Eins-Element gegeben durch die Folge $(1, 0, 0, \ldots) \in \mathcal{F}$. Wir können R als Teilring von $\mathcal{F}$ au assen, indem wir $a \in R$ mit der Folge $(a, 0, 0, \ldots)$ identifizieren. Nun setze $X := (0, 1, 0, 0, \ldots) \in \mathcal{F}$. Die Definition der Multiplikation zeigt dann $X^2 = (0, 0, 1, 0, \ldots)$, $X^3 = (0, 0, 0, 1, 0, \ldots)$ und so fort. Ist also $f = (a_0, a_1, \ldots, a_n, 0, 0, \ldots) \in \mathcal{F}$, so können wir f eindeutig darstellen als $f = a_0 + a_1 X + \cdots + a_n X^n$.]

Bemerkung 12.1. Sei R ein Integritätsring. Sei $0 \neq g \in R[X]$ wie oben, mit $\mathrm{Grad}(g) = m \geqslant 0$ und Leitkoeffizient $0 \neq b_m \in R$. Ist $b_m \in R^\times$, dann ist für ein beliebiges $f \in R[X]$ die ***Division mit Rest*** durch g möglich, d.h., es gibt Polynome $h, r \in R[X]$ mit

$$f = gh + r \quad \text{wobei entweder } r = 0, \text{ oder } r \neq 0 \text{ und } \mathrm{Grad}(r) < \mathrm{Grad}(g) \text{ gilt.}$$

Für $f = 0$ ist $f = 0 = g \cdot 0 + 0$ die gewünschte Darstellung. Für $f \neq 0$ benutzen wir Induktion nach $\mathrm{Grad}(f) = n$. Ist $n < m$, so ist $f = g \cdot 0 + f$ die gewünschte Darstellung. Ist $n \geqslant m$, so setzen wir $f' := f - a_n b_m^{-1} X^{n-m} g$, wobei $0 \neq a_n \in R$ der Leitkoeffizient von f ist. Ist $f' = 0$, so sind wir fertig (mit $r = 0$ und $q = a_n b_m^{-1} X^{n-m}$). Nun ist

$$f' = a_n X^n + a_{n-1} X^{n-1} + \ldots + a_0 - (a_n X^n + a_n b_m^{-1} b_{m-1} X^{n-1} + \ldots + a_n b_m^{-1} b_0 X^{n-m}).$$

Hier fällt der Term $a_n X^n$ weg. Ist also $f' \neq 0$, so folgt $\mathrm{Grad}(f') < n$; also nach Induktion $f' = gh' + r$ mit $r = 0$ oder $r \neq 0$ und $\mathrm{Grad}(r) < m$. Damit erhalten wir $f = f' + a_n b_m^{-1} X^{n-m} g = gh' + r + a_n b_m^{-1} X^{n-m} g = gh + r$ wobei $h = h' + a_n b_m^{-1} X^{n-m}$. □

Hier sind h, r eindeutig bestimmt: Ist auch $f = gh' + r'$, so $g(h - h') = r' - r$. Wäre $r' \neq r$, dann $g \mid r' - r$, was wegen $\mathrm{Grad}(r' - r) < \mathrm{Grad}(g)$ unmöglich ist. Also gilt $r = r'$. Dann ist aber $g(h - h') = 0$. Weil $R[X]$ keine Nullteiler hat, ist also auch $h = h'$.

Beispiel: Sei $R = \mathbb{Z}$, $f = X^4 + 2X^2 - 4$ und $g = X^3 + X^2 - 4X - 4$, also $n = 4$, $a_4 = 1$, $m = 3$, $b_m = 1$. Im ersten Schritt bilden wir $f' := f - a_n b_m^{-1} X^{n-m} g = f - Xg = -X^3 + 6X^2 + 4X - 4$. Danach wird f' mit Rest durch g dividiert; das Ergebnis ist $f' = g \cdot (-1) + 7X^2 - 8$. Einsetzen in $f = f' + Xg$ ergibt $f = gh + r$ mit $h = X - 1$ und $r = 7X^2 - 8$.

Folgerung 12.2. *Sei K ein Körper. Dann ist $K[X]$ ein Euklidischer Ring mit Grad-Funktion $\nu(f) = \mathrm{Grad}(f)$ für $0 \neq f \in K[X]$. Folglich ist $K[X]$ ein Hauptidealring und faktoriell. Ist $\{0\} \neq I \trianglelefteq K[X]$ ein Ideal, so gibt es ein eindeutiges normiertes $0 \neq f \in K[X]$ mit $I = (f)$.*

Beweis. Sei $0 \neq g \in K[X]$ mit $\mathrm{Grad}(g) = m \geqslant 0$ und Leitkoeffizient $0 \neq b_m \in K$. Wegen $K^\times = K \setminus \{0\}$ ist $b_m \in K^\times$, also können wir wie oben mit Rest dividieren. Aus Satz 11.12 folgt dann, dass $K[X]$ ein Hauptidealring und faktoriell ist. Sei nun $\{0\} \neq I \trianglelefteq K[X]$. Da $K[X]$ ein Hauptidealring ist, gibt es ein $0 \neq f \in K[X]$ mit $I = (f)$. Indem wir f mit dem Inversen seines Leitkoeffzienten multiplizieren, können wir annehmen, dass f normiert ist. Ist auch $0 \neq f' \in K[X]$ normiert mit $I = (f')$, so folgt $f' = cf$ mit einem $c \in (K[X])^\times = K^\times$; siehe Bemerkung 11.1(c). Also muss $c = 1$ gelten. □

Bemerkung 12.3. Sei R ein Integritätsring. In Ü8 wird gezeigt: $R[X]$ ist nur dann ein Hauptidealring, wenn R ein Körper ist. Konkretes Beispiel: $\mathbb{Z}[X]$ ist kein Hauptidealring; zum Beispiel ist das Ideal $I = \{2f + Xg \mid f, g \in \mathbb{Z}[X]\}$ kein Hauptideal. Aber wir werden sehen, dass $\mathbb{Z}[X]$ ein faktorieller Ring ist (siehe Satz von Gauß im nächsten Abschnitt).

Definition 12.4. Sind R, S Ringe, so heißt eine Abbildung $\varphi\colon R \to S$ ein ***Ring-Homomorphismus***, wenn φ ein Gruppenhomomorphismus von $(R, +)$ nach $(S, +)$ ist und zusätzlich noch $\varphi(r_1 \cdot r_2) = \varphi(r_1) \cdot \varphi(r_2)$ für alle $r_1, r_2 \in R$ gilt. Ist $R \neq \{0\}$ und sind sowohl R als auch S Ringe mit 1, so verlangen wir zusätzlich, dass $\varphi(1) = 1$ gilt.

Wir sagen, dass R und S isomorph sind (in Zeichen: $R \cong S$), wenn es einen bijektiven Ring-Homomorphismus $\varphi\colon R \to S$ (auch Isomorphismus genannt) gibt. Wie bei Gruppen sehen wir leicht, dass die Umkehrabbildung eines bijektiven Ring-Homomorphismus wieder ein Ring-Homomorphismus ist.

Satz 12.5 (Homomorphiesatz für Ringe)**.** *Seien R, S Ringe und $\varphi\colon R \to S$ ein Ring-Homomorphismus. Dann ist $I := \mathrm{Kern}(\varphi) = \{a \in R \mid \varphi(a) = 0\}$ ein Ideal von R und $\mathrm{Bild}(\varphi)$ ein Unterring von S. Es gilt $I = \{0\}$ genau dann wenn φ injektiv ist. Im Allgemeinen gibt es einen eindeutigen Ring-Homomorphismus $\bar{\varphi}\colon R/I \to S$ mit $\varphi = \bar{\varphi} \circ \pi$, wobei $\pi\colon R \to R/I$, $a \mapsto \bar{a}$, der "kanonische Homomorphismus" ist. Es gilt $R/I \cong \mathrm{Bild}(\varphi)$.*

Beweis. Weil φ ein Homomorphismus von abelschen Gruppen von $(R, +)$ nach $(S, +)$ ist, folgt bereits, dass $I \subseteq R$ ein Normalteiler von $(R, +)$ ist und $\mathrm{Bild}(\varphi)$ eine Untergruppe von $(S, +)$; außerdem ist $R/I \cong \mathrm{Bild}(\varphi)$ (als abelsche Gruppen bezüglich Addition) nach Satz 7.12. Man muss dann noch nachrechnen, dass alles mit der Multiplikation zusammenpasst, also $I = \mathrm{Kern}(\varphi)$ ein Ideal ist und $\bar{\varphi}$ ein Ring-Homomorphismus. (Details selbst als Übung.) □

Satz 12.6 (Universelle Eigenschaft von Polynomringen)**.** *Sei $R[X]$ der Polynomring über einem kommutativem Ring R mit 1. Sei auch S ein kommutativer Ring mit 1 und $y \in S$ fest. Ist $\varphi\colon R \to S$ ein Ring-Homomorphismus, so erhalten wir einen Ring-Homomorphismus $\tilde{\varphi}_y\colon R[X] \to S$ (genannt **Einsetzungs-Homomorphismus**) durch*

$$\tilde{\varphi}_y(f) := \varphi(a_0) + \varphi(a_1)y + \cdots + \varphi(a_n)y^n \quad \text{wobei} \quad f = a_0 + a_1X + \cdots + a_nX^n \in R[X].$$

Wir schreiben dann auch einfach $f(y)$ anstelle von $\tilde{\varphi}_y(f)$, für alle $f \in R[X]$.

Beweis. Man rechnet einfach nach, dass $\tilde{\varphi}_y(f+g) = \tilde{\varphi}_y(f)+\tilde{\varphi}_y(g)$ und $\tilde{\varphi}_y(fg) = \tilde{\varphi}_y(f)\tilde{\varphi}_y(g)$ gilt für alle $f, g \in R[X]$. □

Folgender Spezialfall ist bereits aus der Linearen Algebra vertraut. Sei $S = R$ und $\varphi = \mathrm{id}$ die Identität. Sei $f = a_0 + a_1X + \ldots + a_nX^n \in R[X]$ fest. Dann erhalten wir die zugehörige ***Polynomfunktion*** $\tilde{f}\colon R \to R$ durch $\tilde{f}(y) := \tilde{\varphi}_y(f) = a_0 + a_1y + \ldots + a_ny^n$ für alle $y \in R$.

Lemma 12.7. *Sei K ein Körper und $0 \neq f \in K[X]$ ein Polynom vom Grad $n \geqslant 0$. Dann hat f höchstens n Nullstellen in K. (Eine **Nullstelle** von f ist ein Element $\alpha \in K$ mit $f(\alpha) = 0$.)*

Beweis. Induktion nach n. Ist $n = 0$, so ist f eine Konstante ungleich 0, hat also keine Nullstellen. Sei nun $n > 0$. Gibt es überhaupt keine Nullstellen von f in K, so ist nichts weiter zu zeigen. Nehmen wir jetzt an, es gibt eine Nullstelle $\alpha \in K$ von f. Dann gilt

$$(*) \qquad f = (X - \alpha)g \qquad \text{mit } 0 \neq g \in K[X] \text{ und } \mathrm{Grad}(g) = n - 1.$$

Denn: Division mit Rest ergibt $f = (X-\alpha)g + r$ mit $r = 0$ oder $\mathrm{Grad}(r) < \mathrm{Grad}(X-\alpha) = 1$, also $r \in K$. Weil Einsetzen ein Homomorphismus ist, folgt $0 = f(\alpha) = (\alpha - \alpha)g(\alpha) + r = r$. Nun ist $\mathrm{Grad}(f) = \mathrm{Grad}(g) + \mathrm{Grad}(X - \alpha)$, also $\mathrm{Grad}(g) = n - 1$.

Nach Induktion hat g höchstens $n-1$ Nullstellen. Sei nun $\beta \in K$ eine Nullstelle von f mit $\beta \neq \alpha$. Dann ist $0 = f(\beta) = (\beta - \alpha)g(\beta)$ (weil Einsetzen in $(*)$ ein Homomorphismus ist). Wegen $\beta - \alpha \neq 0$ muss $g(\beta) = 0$ gelten (da K Körper), d.h., β ist eine der höchstens $n-1$ Nullstellen von g. Also hat f insgesamt höchstens $(n-1)+1 = n$ Nullstellen. □

Die Aussage wird falsch im Allgemeinen, wenn der Grundring kein Körper ist. Ist zum Beispiel $R = \mathbb{Z}/8\mathbb{Z}$, so hat $f = X^2 - 1 \in R[X]$ die 4 (!) Nullstellen $\bar{1}$, $\bar{3}$, $\bar{5}$ und $\bar{7}$ in R.

Satz 12.8. *Sei K ein Körper und G eine endliche Untergruppe von $K^\times$. Dann ist G zyklisch. Insbesondere ist also die ganze multiplikative Gruppe $K^\times$ eines endlichen Körpers zyklisch.*

Beweis. Sei $d \in \mathbb{N}$ ein Teiler von $|G|$. Nach Satz 4.10 genügt es zu zeigen, dass G nur höchstens eine Untergruppe $U \leqslant G$ hat mit $|U| = d$. Annahme: Es gibt Untergruppen $U_1, U_2 \leqslant G$ mit $|U_1| = |U_2| = d$ aber $U_1 \neq U_2$. Nach Folgerung 3.6 gilt $g^d = 1_G$ für alle $g \in U_1 \cup U_2$. Alle Elemente in $U_1 \cup U_2$ sind also Nullstellen des Polynoms $f = X^d - 1 \in K[X]$. Dieses Polynom hat nach Lemma 12.7 aber höchstens d Nullstellen in K, Widerspruch weil $|U_1 \cup U_2| > d$. □

Beispiel 12.9. Sei $p \in \mathbb{Z}$ eine Primzahl. Dann ist $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ ein Körper (siehe Satz 4.3) also $\mathbb{F}_p^\times$ eine zyklische Gruppe der Ordnung $p-1$. Einige Beispiele:

$$\mathbb{F}_2^\times = \{\bar{1}\}, \qquad \mathbb{F}_3^\times = \langle\bar{2}\rangle, \qquad \mathbb{F}_5^\times = \langle\bar{2}\rangle, \qquad \mathbb{F}_7^\times = \langle\bar{3}\rangle.$$

Im Allgemeinen ist keine Formel bekannt, die ein $m \in \mathbb{Z}$ angibt mit $p \nmid m$ und $\mathbb{F}_p^\times = \langle\bar{m}\rangle$.

Definition 12.10. Sei K Körper, $n \geqslant 1$ und $f = X^n - 1 \in K[X]$. Dann heißt

$$E_n := \{\alpha \in K \mid f(\alpha) = 0\} = \{\alpha \in K \mid \alpha^n = 1\}$$

die Menge der n-ten ***Einheitswurzeln*** in K; es gilt $|E_n| \leqslant n$ nach Lemma 12.7. Sind $\alpha, \beta \in E_n$ so auch $\alpha \cdot \beta \in E_n$ und $\alpha^{-1} \in E_n$. Also ist E_n eine Untergruppe von $K^\times$. Nach Satz 12.8 ist E_n zyklisch. Damit können wir das n-te ***Kreisteilungspolynom*** über K definieren als

$$\Phi_n := \prod_{\alpha \in E_n^*} (X - \alpha) \quad \in K[X], \qquad \text{wobei} \qquad E_n^* := \{\alpha \in E_n \mid E_n = \langle\alpha\rangle\}.$$

Die Elemente in E_n^* heißen auch ***primitive n-te Einheitswurzeln***.

Beispiel 12.11. (a) Sei $n = p$ und $K = \mathbb{Z}/p\mathbb{Z}$ mit p Primzahl. Nach dem Kleinen Satz von Fermat gilt $\bar{a}^p = \bar{a}$ für alle $\bar{a} \in E_p$, also ist in diesem Fall $E_p = \{\bar{1}\}$ und $\Phi_p = X - 1$.

(b) Sei $n \geqslant 1$ beliebig und $K = \mathbb{C}$. Dann ist $\zeta_n = \cos(2\pi/n) + i\sin(2\pi/n) \in E_n$ und ζ_n hat Ordnung n, wie sofort mit den Additionstheoremen für sin und cos folgt (oder der Formel

$\zeta_n = \exp(2\pi i/n)$). Also gilt $E_n = \langle\zeta_n\rangle$ und $|E_n| = n$. Nach Lemma 4.8 ist $E_n = \langle\zeta_n^i\rangle$ genau dann, wenn $\mathrm{ggT}(i,n) = 1$. Also folgt

$$\mathrm{Grad}(\Phi_n) = \phi(n) \qquad \text{und} \qquad \Phi_n = \prod_{1\leqslant i\leqslant n \text{ mit } \mathrm{ggT}(i,n)=1} (X - \zeta_n^i) \in \mathbb{C}[X].$$

Wir begegnen hier also wieder der Eulerschen Phi-Funktion. Die ersten Beispiele:

$$\begin{array}{lllll} n=1: & E_1=\{1\}, & \zeta_1 = 1 & \text{und} & \Phi_1 = X-1; \\ n=2: & E_2=\{\pm 1\}, & \zeta_2 = -1 & \text{und} & \Phi_2 = X+1; \\ n=3: & E_3=\{1, \frac{1}{2}(-1\pm i\sqrt{3})\}, & \zeta_3 = \frac{1}{2}(-1+i\sqrt{3}) & \text{und} & \Phi_2 = X^2+X+1; \\ n=4: & E_4=\{\pm 1, \pm i\} & \zeta_4 = i & \text{und} & \Phi_4 = X^2+1. \end{array}$$

Satz 12.12. *Sei* $K = \mathbb{C}$ *und* $n \geqslant 1$. *Dann gilt* $\Phi_n \in \mathbb{Z}[X]$ *und* $X^n - 1 = \prod_{d\in\mathbb{N}:\, d|n} \Phi_d$.

Beweis. Ist $d \mid n$, so ist $E_d \subseteq E_n$ und damit $E_d^* \subseteq E_n$. Sei umgekehrt $\alpha \in E_n$ beliebig und d die Ordnung von α. Dann ist $\alpha^d = 1$, also $\alpha \in E_d$ und dann auch $\alpha \in E_d^*$; außerdem gilt $d \mid n$ nach Lagrange. Damit ist gezeigt: $E_n = \bigcup_{d|n} E_d^*$. Diese Vereinigung ist disjunkt (denn sonst gäbe es ein $\alpha \in E_n$ mit $o(g) = d$ und $o(g) = d'$, wobei $d \neq d'$). Also folgt:

$$X^n - 1 = \prod_{\alpha\in E_n} (X-\alpha) = \prod_{d\in\mathbb{N}:\, d|n} \Big(\prod_{\alpha\in E_d^*} (X-\alpha)\Big) = \prod_{d\in\mathbb{N}:\, d|n} \Phi_d.$$

Wir zeigen nun $\Phi_n \in \mathbb{Z}[X]$ mit Induktion nach n. Ist $n = 1$, so ist $\Phi_1 = X - 1 \in \mathbb{Z}[X]$. Sei jetzt $n > 1$. Nach Induktion ist $\Phi_d \in \mathbb{Z}[X]$ für alle $d < n$. Also gilt $X^n - 1 = \Phi_n g$, wobei $g := \prod_{d|n, d<n} \Phi_d \in \mathbb{Z}[X]$. Weil g normiert ist, können wir wie in Bemerkung 12.1 mit Rest dividieren und erhalten $q, r \in \mathbb{Z}[X]$ mit $X^n - 1 = gq + r$, wobei $r = 0$ oder $\mathrm{Grad}(r) < \mathrm{Grad}(g)$. Dann folgt $r = \Phi_n g - qg = (\Phi_n - q)g$. Wäre $\Phi_n \neq q$, so wäre $r \neq 0$ und $\mathrm{Grad}(r) = \mathrm{Grad}((\Phi_n - q)g) \geqslant \mathrm{Grad}(g)$, Widerspruch. Also ist $\Phi_n = q \in \mathbb{Z}[X]$. □

Beispiel 12.13. (a) Sei $p \geqslant 2$ eine Primzahl. Dann zeigt obige Formel $X^p - 1 = \Phi_1\Phi_p$. Wegen $\Phi_1 = X - 1$ und $X^p - 1 = (X-1)(X^{p-1} + X^{p-2} + \ldots + X + 1)$ folgt damit

$$\Phi_p = X^{p-1} + X^{p-2} + \ldots + X + 1.$$

(b) Für $n = 105 = 3\cdot 5\cdot 7$ erhält man $\Phi_{105} = X^{48} + X^{47} + X^{46} - X^{43} - X^{42} - 2X^{41} \pm \ldots + X + 1$. (Dies ist das kleinste Beispiel mit Koeffizienten ungleich ± 1, siehe `http://en.wikipedia.org/wiki/Cyclotomic_polynomial`.)

Einheitswurzeln und Kreisteilungspolynome haben erstaunlich viele Anwendungen. Zum Beispiel erhält man damit auf relativ einfache Weise Spezialfälle des berühmten Primzahlsatzes von Dirichlet (`http://en.wikipedia.org/wiki/Dirichlet's_theorem_on_arithmetic_progressions`). Siehe Satz 11.5 im Algebra-Skript [G14] oder die allgemeinere Diskussion in:

S. Gueron and R. Tessler, Infinitely many primes in artithmetic progressions: The cyclotomic polynomial method, The Mathematical Gazette **86** (2002), 110–114.

Ein weiteres interessantes Thema ist die Darstellung von Wurzeln $\sqrt{\pm n}$ (für $n \in \mathbb{N}$) durch Summen von Einheitswurzeln (sogenannte quadratische ***Gauß'sche Summen***), siehe dazu:
M. Ram Murty and S. Pathak, Evaluation of the quadratic Gauss sums, The Mathematics Student (Indian Math. Soc.) **86** (2017), 139–150.

13. *Irreduzibilität in Polynomringen*

Wie zeigen wir, ob ein Polynom $f \in \mathbb{Q}[X]$ irreduzibel ist, zum Beispiel

$$f = X^7 - 5X^5 + 3X^4 + X^3 - 5X + 3 \qquad \text{oder} \qquad g = X^{24} - 18X^7 + 15 \ ?$$

Wir stellen hier einige grundlegende Verfahren vor, um diese Frage zu entscheiden.

Bemerkung 13.1. Sei R ein Integritätsring und $p \in R$ ein Primelement. Dann ist auch $R/(p)$ ein Integritätsring; siehe Lemma 11.4. Für $a \in R$ sei $\bar{a} := a + (p) \in R/(p)$. Wir betrachten die Polynomringe $R[X]$, $R/(p)[X]$ und erhalten eine surjektive Abbildung

$$\pi_p \colon R[X] \to R/(p)[X], \qquad f = \sum_{i=0}^{n} a_i X^i \mapsto f^* := \sum_{i=0}^{n} \bar{a}_i X^i.$$

Dies ist ein Ring-Homomorphismus. (Beweis durch Nachrechnen; oder weil π_p ein Einsetzungs-Homomorphismus ist.) Für $f = \sum_{i=0}^{n} a_i X^i \in R[X]$ gilt $f \in \mathrm{Kern}(\pi_p) \Leftrightarrow p \mid a_i$ für alle i $\Leftrightarrow f = pg$ mit $g \in R[X]$. Also ist $\mathrm{Kern}(\pi_p)$ das von p erzeugte Hauptideal in $R[X]$.
Beachte: Hier ist sowohl $R[X]$ als auch $R/(p)[X]$ ein Integritätsring.

Definition 13.2. Sei R Integritätsring und $0 \neq f \in R[X]$. Dann heißt f ein ***primitives Polynom***, wenn es keine Nicht-Einheit $0 \neq a \in R$ gibt, so dass alle Koeffizienten von f durch a teilbar sind. Ist zum Beispiel irgendein Koeffizient von f gleich ± 1, so ist f primitiv.

Lemma 13.3 (Reduktions-Kriterium)**.** *Sei R Integritätsring und $0 \neq f \in R[X]$ primitiv mit $\mathrm{Grad}(f) = n \geqslant 1$ und Leitkoeffizient $a_n \in R$. Sei $p \in R$ Primelement mit $p \nmid a_n$. Ist f^* irreduzibel in $R/(p)[X]$, so ist auch f irreduzibel in $R[X]$.*

Beweis. Wegen $p \nmid a_n$ ist $0 \neq f^* \in R/(p)[X]$ und $\mathrm{Grad}(f^*) = n$. Annahme, wir hätten $f = gh$ mit $g, h \in R[X]$. Da R Integritätsring ist, gilt $n = \mathrm{Grad}(g) + \mathrm{Grad}(h)$. Weil a_n gleich dem Produkt der Leitkoeffizienten von g und h ist, sind diese auch nicht durch p teilbar. Also ist $g^* \neq 0$, $h^* \neq 0$ und $\mathrm{Grad}(g^*) = \mathrm{Grad}(g)$, $\mathrm{Grad}(h^*) = \mathrm{Grad}(h)$. Nun ist $f^* = g^* h^*$. Da f^* als irreduzibel vorausgesetzt ist, folgt $\mathrm{Grad}(g) = \mathrm{Grad}(g^*) = 0$ oder $\mathrm{Grad}(h) = \mathrm{Grad}(h^*) = 0$. Da f primitiv ist, muss also $g \in R^\times$ oder $h \in R^\times$ sein. Also ist f irreduzibel. □

Ab hier Woche 9, also Woche 11.1.–15.1.2021

Satz 13.4 (Eisenstein-Kriterium)**.** *Sei R Integritätsring und $0 \neq f = \sum_{i=0}^{n} a_i X^i \in R[X]$ primitiv mit $n \geqslant 1$ und $a_n \neq 0$. Es gebe ein Primelement $p \in R$ mit $p \nmid a_n$, $p \mid a_i$ für alle $0 \leqslant i \leqslant n-1$ aber $p^2 \nmid a_0$. Dann ist f irreduzibel in $R[X]$.*

Beweis. Annahme, wir hätten $f = gh$ mit $g, h \in R[X]$. Wie im obigen Beweis sind $f^*, g^*, h^* \in R/(p)[X]$ ungleich 0 und es gilt $\mathrm{Grad}(f^*) = \mathrm{Grad}(f) = n = m + k$ mit $m = \mathrm{Grad}(g^*) = \mathrm{Grad}(g)$, $k = \mathrm{Grad}(h^*) = \mathrm{Grad}(h)$. Schreibe $g = \sum_{i=0}^{m} b_i X^i$ und $h = \sum_{j=0}^{k} c_j X^j$, wobei $\bar{b}_m \neq \bar{0}$ und $\bar{c}_k \neq \bar{0}$. Es genügt wieder zu zeigen, dass $m = 0$ oder $k = 0$ gilt. Nun, die Voraussetzungen ergeben $f^* = \bar{a}_n X^n \in R/(p)[X]$. Sei $r := \min\{0 \leqslant i \leqslant m \mid \bar{b}_i \neq \bar{0}\}$ und $s := \min\{0 \leqslant j \leqslant k \mid \bar{c}_j \neq \bar{0}\}$. Dann ist $\bar{a}_n X^n = f^* = g^* h^* = \bar{b}_m \bar{c}_k X^{m+k} + \ldots + \bar{b}_r \bar{c}_s X^{r+s}$, wobei $\bar{b}_m \bar{c}_k \neq \bar{0}$ und $\bar{b}_r \bar{c}_s \neq \bar{0}$ (da $R/(p)$ Integritätsring). Also folgt $m = r$ und $k = s$. Wäre $m = r > 0$ und $k = s > 0$, so folgt $\bar{b}_0 = \bar{c}_0 = \bar{0}$, d.h., $p \mid b_0$ und $p \mid c_0$; dann wäre aber $p^2 \mid b_0 c_0 = a_0$, Widerspruch. Also ist $m = 0$ oder $k = 0$. □

Beispiel 13.5. (a) Das obige Polynom $g = X^{24} - 18X^7 + 15 \in \mathbb{Z}[X]$ ist irreduzibel, weil mit $R = \mathbb{Z}$ und $p = 3$ die Voraussetzungen des Eisenstein-Kriteriums erfüllt sind.

(b) Obiges Polynom f ist reduzibel; wir finden die Faktorisierung $f = (X^4 + 1)(X^3 - 5X + 3)$. Untersuchen wir nun den Faktor $X^3 - 5X + 3 \in \mathbb{Z}[X]$. Sei $p = 2$ und betrachte $\pi_2 \colon \mathbb{Z}[X] \to \mathbb{F}_2[X]$ wie in Bemerkung 13.1, wobei $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$. Dann ist $(X^3 - 5X + 3)^* = X^3 + X + \bar{1} \in \mathbb{F}_2[X]$. Wäre $X^3 + X + \bar{1}$ reduzibel, so müsste es einen Faktor vom Grad 1, also eine Nullstelle in $\mathbb{F}_2$ geben; aber die beiden Elemente $\bar{0}, \bar{1}$ von $\mathbb{F}_2$ sind keine Nullstellen dieses Polynoms. Also zeigt Lemma 13.3, dass $X^3 - 5X + 3 \in \mathbb{Z}[X]$ irreduzibel ist.

(c) Was ist mit dem Faktor $h = X^4 + 1 \in \mathbb{Z}[X]$? Wenn wir sonst nicht weiter wissen, können wir folgendes "***Verfahren von Kronecker***" anwenden. Wäre h reduzibel, so müsste es einen Faktor vom Grad 1 oder 2 geben. Da h keine Nullstellen in $\mathbb{Q}$ hat, ist die einzige Möglichkeit $h = h'h''$ mit $\mathrm{Grad}(h') = \mathrm{Grad}(h'') = 2$. Sei $h' = aX^2 + bX + c$ mit $a, b, c \in \mathbb{Z}$. Da h normiert ist, können wir $a = 1$ annehmen. Um die Möglichkeiten für b, c zu bestimmen, wählen wir zwei Werte $x_1, x_2 \in \mathbb{Z}$ und betrachten $y_i = h(x_i)$. Wegen $h' \mid h$ in $\mathbb{Z}[X]$ muss dann auch $h'(x_i) \mid y_i$ in $\mathbb{Z}$ gelten. Für $(x_1, x_2) = (0, 1)$ ist $(y_1, y_2) = (1, 2)$ und damit $c = h'(0) \in \{\pm 1\}$ und $b + c + 1 = h'(1) \in \{\pm 1, \pm 2\}$. Also gibt es nur die Möglichkeiten

$$(b, c) \in \{(-4, 1), (0, 1), (-1, 1), (-3, 1), (-1, -1), (-2, -1), (2, -1), (1, -1)\}.$$

Dann testen wir einfach für jede dieser Möglichkeiten, ob das zugehörige Polynom h' ein Teiler von h ist. Hier stellen wir fest, dass dies niemals der Fall ist. Also ist h irreduzibel. (Dieses Verfahren erfordert einiges an Rechenaufwand, aber es liefert auf jeden Fall in endlich vielen Schritten eine Antwort auf die Frage, ob ein Polynom in $\mathbb{Z}[X]$ irreduzibel ist oder nicht.)

Ist $f \in \mathbb{Z}[X]$ irreduzibel, wie sieht es dann mit der Irreduzibilität in $\mathbb{Q}[X]$ aus? Beispiel:

$$24X^3 + 4X^2 - 6X - 1 = (4X + \tfrac{2}{3})(6X^2 - \tfrac{3}{2}) = (2X + \tfrac{1}{3})(12X^2 - 3) = (6X + 1)(4X^2 - 1).$$

Wir wollen uns nun überlegen, dass dies allgemein richtig ist, d.h., eine Faktorisierung über $\mathbb{Q}[X]$ zieht immer eine Faktorisierung über $\mathbb{Z}[X]$ nach sich. Dazu sind mehrere Schritte nötig.

Lemma 13.6. *Sei* R *faktoriell und* $0 \neq f \in R[X]$*. Dann ist entweder* f *eine Einheit oder lässt sich schreiben als* $f = f_1 f_2 \cdots f_r$ *mit* $r \geqslant 1$ *und* $f_i \in R[X]$ *irreduzibel für alle* i*.*

Beweis. Wir setzen die Längenfunktion $\ell(a)$ aus Bemerkung 11.14 wie folgt auf $R[X]$ fort. Sei $0 \neq g \in R[X]$ mit $\mathrm{Grad}(g) = n \geqslant 0$ und Leitkoeffizient $0 \neq a_n \in R$. Dann setze $L(g) := n + \ell(a_n)$. Wegen $\mathrm{Grad}(gh) = \mathrm{Grad}(g) + \mathrm{Grad}(h)$ für alle $0 \neq g, h \in R[X]$ (da R Integritätsring) und $\ell(ab) = \ell(a) + \ell(b)$ für alle $0 \neq a, b \in R$ (siehe Bemerkung 11.14), gilt:

(1) Ist $L(f) = 0$, so $f \in R^\times$.

(2) $L(gh) = L(g) + L(h)$ für alle $0 \neq g, h \in R[X]$.

Jetzt zeigen wir die gewünschte Aussage mit Induktion nach $L(f)$. Ist $L(f) = 0$, so ist $f \in R^\times$, siehe (1). Sei jetzt $L(f) > 0$. Ist f irreduzibel, so sind wir fertig. Andernfalls ist $f = gh$ mit $g, h \in R[X]$, die beide keine Einheiten sind, also $L(g) > 0$ und $L(h) > 0$, siehe noch einmal (1). Wegen (2) ist dann aber auch $L(g) < L(f)$ und $L(h) < L(f)$. Mit Induktion sind g und h Produkte von irreduziblen Elementen, folglich auch f. □

Lemma 13.7. *Ist* R *Integritätsring und* $p \in R$ *Primelement, so ist* p *Primelement in* $R[X]$*.*

Beweis. Da $p \in R$ Primelement ist, ist $R/(p)$ ein Integritätsring (Lemma 11.4), also auch $R/(p)[X]$ ein Integritätsring. Mit dem Homomorphie-Satz folgt, dass auch $R[X]/\mathrm{Kern}(\pi_p) \cong \mathrm{Bild}(\pi_p) = R/(p)[X]$ ein Integritätsring ist. Nun ist $\mathrm{Kern}(\pi_p)$ das von p erzeugte Hauptideal in $R[X]$ (Bemerkung 13.1), also ist p auch Primelement in $R[X]$ (wiederum Lemma 11.4). □

Lemma 13.8. *Sei* R *faktoriell und* $0 \neq a \in R[X]$ *irreduzibel mit* $\mathrm{Grad}(a) \geqslant 1$*. Ist* $a \mid bc$ *(in* $R[X]$*) mit* $0 \neq b \in R$ *und* $c \in R[X]$*, so gilt* $a \mid c$ *(in* $R[X]$*).*

Beweis. Induktion nach $\ell(b)$ (wie in Bemerkung 11.14 definiert). Ist $\ell(b) = 0$, so $b \in R^\times$ und die Behauptung ist klar. Sei nun $\ell(b) > 0$ und $0 \neq p \in R$ irreduzibel mit $p \mid b$; dann ist $b = pb'$ mit $b' \in R$ und $\ell(b') = \ell(b) - 1$. Nun gilt $bc = ad$ mit $d \in R[X]$, also folgt $p \mid ad$ in $R[X]$. Wegen R faktoriell ist p ein Primelement in R, also auch ein Primelement in $R[X]$ (Lemma 13.7). Damit folgt $p \mid a$ oder $p \mid d$. Wäre $p \mid a$, so $a = pa'$ mit $0 \neq a' \in R[X]$, wobei $\mathrm{Grad}(a') \geqslant 1$ (da $p \in R$ und $\mathrm{Grad}(a) \geqslant 1$); dann ist aber auch a' keine Einheit, Widerspruch zu a irreduzibel. Also muss $p \mid d$ gelten, d.h., $d = pd'$ mit $d' \in R[X]$. Aus $pb'c = bc = ad = pad'$ folgt nun $b'c = ad'$, also $a \mid b'c$ und daher $a \mid c$ mit Induktion. □

Bemerkung 13.9. Sei R ein Integritätsring. Ist R Teilring eines Körpers K und gilt

$$R \subseteq K = \{ab^{-1} \mid a \in R, 0 \neq b \in R\},$$

so heißt K ***Quotientenkörper*** von R. Zum Beispiel ist $\mathbb{Q}$ Quotientenkörper von $\mathbb{Z}$. Genauso, wie $\mathbb{Q}$ aus $\mathbb{Z}$ konstruiert wird (siehe Beispiel 2.1), kann auch zu R ein Quotientenkörper konstruiert werden. Hier die wichtigsten Schritte (Details als Übung):

Definiere eine Relation $\sim$ auf $R \times (R \setminus \{0\})$ durch: $(a,b) \sim (c,d) \overset{\text{def.}}{\iff} ad = bc$. Dies ist eine Äquivalenzrelation. Bezeichne mit a/b die Äquivalenzklasse von (a,b). Definiere dann eine Addition und eine Multiplikation auf $K := \{a/b \mid a \in R, 0 \neq b \in R\}$ durch

$$(a/b) + (c/d) := (ad + bc)/(bd) \qquad \text{und} \qquad (a/b) \cdot (c/d) := (ac)/(bd).$$

Dann wird gezeigt, dass mit diesen Verknüpfungen K ein Körper ist und die Abbildung $R \to K$, $a \mapsto a/1$, injektiv ist. Vermöge dieser Abbildung fassen wir R als Teilring von K auf, d.h., wir identifizieren $a \in R$ mit $a/1 \in K$. Dann ist K ein Quotientenkörper von R.

Satz 13.10 (Satz von Gauß)**.** *Sei R ein faktorieller Ring, mit Quotientenkörper K.*
(a) *Sei $0 \neq f \in R[X]$ mit $\operatorname{Grad}(f) \geqslant 1$. Ist f irreduzibel in $R[X]$, so auch in $K[X]$.*
(b) *Der Polynomring $R[X]$ ist ebenfalls faktoriell.*

Beweis. (a) Sei $f = gh$ mit $g, h \in K[X]$, wobei $\operatorname{Grad}(g) \geqslant 1$ gelte. Dann müssen wir zeigen, dass $h \in K$ gilt. Nun, die Koeffizienten von g und h sind Brüche von Elementen von R. Ist also $0 \neq d \in R$ das Produkt der Nenner in all diesen Brüchen, so gilt $\tilde{g} := dg \in R[X]$ und $\tilde{h} := dh \in R[X]$; hierbei ist natürlich $\operatorname{Grad}(\tilde{g}) = \operatorname{Grad}(g)$ und $\operatorname{Grad}(\tilde{h}) = \operatorname{Grad}(h)$. Also erhalten wir eine Gleichung in $R[X]$, nämlich $d^2 f = \tilde{g}\tilde{h}$. Wegen $\operatorname{Grad}(\tilde{g}) = \operatorname{Grad}(g) \geqslant 1$ ist $\tilde{g}$ keine Einheit. Nach Lemma 13.6 können wir also schreiben: $\tilde{g} = q_1 \cdots q_r$ mit $r \geqslant 1$ und $q_i \in R[X]$ irreduzibel für alle i. Wegen $\operatorname{Grad}(\tilde{g}) \geqslant 1$ können wir die Bezeichnung so wählen, dass $\operatorname{Grad}(q_1) \geqslant 1$ gilt. Nun ist $q_1 \mid d^2 f$, also folgt $q_1 \mid f$ mit Lemma 13.8. Da $f \in R[X]$ irreduzibel und q_1 keine Einheit ist, gilt also $f = u q_1$ mit $u \in R^\times$. Aber dann erhalten wir $d^2 u = q_2 \cdots q_r \tilde{h}$. Die linke Seite hat Grad 0, also ist auch $\operatorname{Grad}(h) = \operatorname{Grad}(\tilde{h}) = 0$.

(b) Nach Lemma 13.6 gilt Bedingung (1) in Definition 11.6. Zu Bedingung (2): Sei $0 \neq f \in R[X]$ irreduzibel. Wir müssen zeigen, dass f ein Primelement in $R[X]$ ist. Ist $\operatorname{Grad}(f) = 0$, d.h., $f \in R$, so ist f offenbar auch irreduzibel in R, also auch Primelement in R (da R faktoriell) und damit auch in $R[X]$ (Lemma 13.7). Sei jetzt $\operatorname{Grad}(f) \geqslant 1$ und seien $g, h \in R[X]$ gegeben mit $f \mid gh$ (in $R[X]$). Dann gilt auch $f \mid gh$ in $K[X]$. Nach (a) ist f irreduzibel in $K[X]$ und damit auch ein Primelement in $K[X]$, da $K[X]$ Euklidischer Ring ist (Folgerung 12.2). Also folgt $f \mid g$ oder $f \mid h$ in $K[X]$. Nehmen wir an, es gelte $f \mid g$ in $K[X]$. (Das Argument für den anderen Fall ist vollkommen analog.) Es gilt also $g = af$ mit $a \in K[X]$. Multiplizieren wir mit dem Produkt der Nenner der Koeffizienten von a, so erhalten wir eine Gleichung $dg = \tilde{a} f$ mit $0 \neq d \in R$ und $\tilde{a} := da \in R[X]$. Mit Lemma 13.8 folgt dann $f \mid g$ (in $R[X]$). □

Folgerung 13.11. *Sei R faktoriell und K Quotientenkörper von R. Erfüllt $0 \neq f \in R[X]$ die Voraussetzungen des Reduktions-Kriteriums (Lemma 13.3) oder des Eisenstein-Kriteriums (Satz 13.4), so ist f nicht nur irreduzibel in $R[X]$, sondern auch in $K[X]$.*

Beweis. Klar nach Satz 13.10(a). □

Beispiel 13.12. Sei $n \geqslant 1$ und R ein kommutativer Ring mit 1. Dann können wir den Polynomring in n Unbestimmten $X_1, \ldots, X_n$ über R rekursiv definieren durch

$$R[X_1, X_2] := (R[X_1])[X_2], \quad \ldots, \quad R[X_1, \ldots, X_n] := (R[X_1, \ldots, X_{n-1}])[X_n].$$

Ein Polynom $f \in R[X_1, \ldots, X_n]$ lässt sich also zunächst schreiben als $f = \sum_{i=0}^{m} f_i X_n^i$ mit $f_i \in R[X_1, \ldots, X_{n-1}]$; analog dann jedes f_i als Polynom in X_{n-1} mit Koeffizienten in $R[X_1, \ldots, X_{n-2}]$ und so fort. Multiplizieren wir all dies aus, so erhalten wir am Ende einen eindeutigen Ausdruck von f als endliche Summe

$$f = \sum_{(i_1,\ldots,i_n)\in\mathbb{N}_0^n} a_{i_1,\ldots,i_n} X_1^{i_1} \cdots X_n^{i_n} \qquad \text{mit} \qquad a_{i_1,\ldots,i_n} \in R,$$

wobei nur endlich viele Koeffizienten $a_{i_1,\ldots,i_n}$ ungleich 0 sind. Ist $f \neq 0$, so definieren wir $\mathrm{Grad}(f) := \max\{i_1 + \ldots + i_n \mid a_{i_1,\ldots,i_n} \neq 0\}$. Beachte, dass es für $n \geqslant 2$ keinen eindeutigen Leitkoeffizienten mehr zu geben braucht; ist etwa $f = 5X_1^3 - X_1X_2^2 + 7X_1^2X_2 \in \mathbb{Z}[X_1, X_2]$, so ist $\mathrm{Grad}(f) = 3$, aber für alle $X_1^{i_1}X_2^{i_2}$, die mit Koeffizient $\neq 0$ in f vorkommen, gilt $i_1 + i_2 = 3$.

Schließlich halten wir fest: Ist R ein Integritätsring, so auch $R[X_1]$, sodann $R[X_1, X_2]$ und so fort, also schließlich auch $R[X_1, \ldots, X_n]$, wobei $(R[X_1, \ldots, X_n])^\times = R^\times$. Ist R faktoriell, so folgt mit wiederholter Anwendung von Satz 13.10, dass auch $R[X_1, \ldots, X_n]$ faktoriell ist. Andererseits beachte: Ist K ein Körper, so ist $K[X_1]$ ein Hauptidealring, aber für $n \geqslant 2$ ist $K[X_1, \ldots, X_n]$ kein Hauptidealring, denn sonst wäre $K[X_1, \ldots, X_{n-1}]$ ein Körper (siehe Bemerkung 12.3), Widerspruch zu $(K[X_1, \ldots, X_{n-1}])^\times = K^\times \subsetneqq K[X_1, \ldots, X_{n-1}] \setminus \{0\}$.

14. *Symmetrische Polynome*

Sei K ein Körper. Gegeben sei ein quadratisches Polynom $f = X^2 + pX + q \in K[X]$. Gibt es Nullstellen $z_1, z_2 \in K$ mit $f = (X - z_1)(X - z_2)$ (hier ist $z_1 = z_2$ erlaubt), so erhalten wir durch Ausmultiplizieren $p = -z_1 - z_2$ und $q = z_1z_2$. Analog ergeben sich für ein Polynom $f = X^3 + a_2X^2 + a_1X + a_0 = (X - z_1)(X - z_2)(X - z_3) \in K[X]$ vom Grad 3 die Formeln

$$a_2 = -z_1 - z_2 - z_3, \qquad a_1 = z_1z_2 + z_2z_3 + z_1z_3, \qquad a_0 = -z_1z_2z_3.$$

Berechnen Sie selbst auch noch analoge Formeln für $n = 4$. Dieser Zusammenhang zwischen den Koeffizienten und den Nullstellen eines Polynoms wird als ***Satz von Vieta*** bezeichnet und geht bis ins 16. Jahrhundert zurück. Allgemein gilt:

Lemma 14.1. *Sei* $n \geqslant 1$ *und* $f = a_0 + a_1X + \ldots + a_{n-1}X^{n-1} + X^n \in K[X]$. *Seien* $z_1, \ldots, z_n \in K$ *mit* $f = (X - z_1) \cdots (X - z_n)$. *(Die* z_i *müssen nicht verschieden sein.) Dann gilt*

$$a_{n-d} = (-1)^d \sum_{1\leqslant i_1<\ldots<i_d\leqslant n} z_{i_1} z_{i_2} \cdots z_{i_d} \qquad \textit{für } 1 \leqslant d \leqslant n.$$

Also $a_{n-1} = -\sum_{1\leqslant i\leqslant n} z_i$, $a_{n-2} = \sum_{1\leqslant i<j\leqslant n} z_iz_j$ und so weiter bis $a_0 = (-1)^n z_1z_2 \cdots z_n$.

Beweis. Für $1 \leqslant d \leqslant n$ setze $I(n,d) := \{(i_1,\dots,i_d) \in \mathbb{N}^d \mid 1 \leqslant i_1 < \dots < i_d \leqslant n\}$ und

$$s_d := \sum_{(i_1,\dots,i_d)\in I(n,d)} z_{i_1} z_{i_2} \cdots z_{i_d};$$

setze auch $s_0 := 1$. Wir wollen also $a_{n-d} = (-1)^d s_d$ zeigen. Dazu benutzen wir Induktion nach n. Ist $n = 1$, so ist $s_1 = z_1$ und $f = a_0 + X$, also $z_1 = -a_0$ und damit $a_0 = -s_1$. Sei nun $n > 1$. Sei $f' := (X - z_1)\cdots(X - z_{n-1}) = b_0 + b_1 X + \cdots + b_{n-2}X^{n-2} + X^{n-1}$ mit $b_j \in K$. Seien $s'_0, s'_1, \dots, s'_{n-1}$ entsprechend für $f' = (X - z_1)\cdots(X - z_{n-1})$ definiert, also $s'_0 := 1$ und

$$s'_d := \sum_{(i_1,\dots,i_d)\in I(n-1,d)} z_{i_1} z_{i_2} \cdots z_{i_d} \qquad \text{für } 1 \leqslant d \leqslant n-1.$$

Behauptung: $(*)$ $\quad s_n = s'_{n-1}X_n$ und $s_d = s'_d + s'_{d-1}X_n$ für $1 \leqslant d \leqslant n-1$.

Dazu: Die Gleichheit $s_n = s'_{n-1}X_n$ ist klar nach Definition von s_n und s'_{n-1}. Sei nun $1 \leqslant d \leqslant n-1$ und $(i_1,\dots,i_d) \in I(n,d)$. Ist $i_d = n$, so ist $(i_1,\dots,i_{d-1}) \in I(n-1,d-1)$; andernfalls ist $(i_1,\dots,i_d) \in I(n-1,d)$. Damit folgt

$$I(n,d) = I(n-1,d) \;\cup\; \{(i_1,\dots,i_{d-1},n) \mid (i_1,\dots,i_{d-1}) \in I(n-1,d-1)\},$$

wobei die Vereinigung disjunkt ist. Damit erhalten wir

$$s_d = \Big(\sum_{(i_1,\dots,i_d)\in I(n-1,d)} z_{i_1} z_{i_2} \cdots z_{i_d}\Big) + \Big(\sum_{(i_1,\dots,i_{d-1})\in I(n-1,d-1)} z_{i_1} z_{i_2} \cdots z_{i_{d-1}}\Big) z_n = s'_d + s'_{d-1} z_n.$$

Also ist $(*)$ gezeigt. Nach Induktion gilt $b_{n-1-d} = (-1)^d s'_d$ für $1 \leqslant d \leqslant n-1$. Nun folgen aus $f = f' \cdot (X - z_n)$ die Relationen $a_0 = -b_0 z_n$ und $a_i = b_{i-1} - b_i z_n$ für $1 \leqslant i \leqslant n-1$, wobei wir $b_{n-1} := 1$ setzen. Also erhalten wir für $1 \leqslant d \leqslant n-1$:

$$a_{n-d} = b_{n-1-d} - b_{n-d} z_n = (-1)^d s'_d - (-1)^{d-1} s'_{d-1} z_n = (-1)^d (s'_d + s'_{d-1} z_n) \overset{(*)}{=} (-1)^d s_d.$$

Für $d = n$ gilt schließlich $a_0 = -b_0 z_n = -(-1)^{n-1} s'_{d-1} z_n = (-1)^n s_n$. □

Die Formeln in Lemma 14.1 zeigen zwei Dinge:

1) Jeder Koeffizient a_{n-d} in f ist ein polynomialer Ausdruck in $z_1,\dots,z_n$.
2) Ist $\pi \in S_n$ eine Permutation, so gilt natürlich auch $f = (X - z_{\pi(1)})\cdots(X - z_{\pi(n)})$; die Ausdrücke in 1) sind also "invariant" unter beliebigen Permutationen der z_i.

Wir betrachten nun den Polynomring $R[X_1,\dots,X_n]$, wobei $n \geqslant 1$ und R ein kommutativer Ring mit 1 ist; siehe Beispiel 13.12. Analog zu Satz 12.6 haben wir folgende universelle Eigenschaft: Ist $\varphi\colon R \to S$ ein Homomorphismus in einen kommutativen Ring S mit 1, und sind $y_1,\dots,y_n \in S$ fest gegeben, so gibt es einen ***Einsetzungs-Homomorphismus***

$$\tilde{\varphi}_{y_1,\dots,y_n}\colon R[X_1,\dots,X_n] \to S$$

mit $\tilde{\varphi}_{y_1,\dots,y_n}(r) = \varphi(r)$ für $r \in R$ und $\tilde{\varphi}_{y_1,\dots,y_n}(X_i) = y_i$ für $1 \leqslant i \leqslant n$. (Dies folgt sofort mit Induktion nach n aus der rekursiven Definition von $R[X_1,\dots,X_n]$ in Beispiel 13.12.) Wir

schreiben dann wieder einfach $f(y_1, \ldots, y_n)$ anstelle von $\tilde{\varphi}_{y_1,\ldots,y_n}(f)$, für alle $f \in R[X_1, \ldots, X_n]$. Wir wenden dies hier in folgender Situation an:

Sei $S = R[X_1 \ldots, X_n]$ und $\varphi = \mathrm{id}$. Ist eine Permutation $\pi \in S_n$ gegeben, so sei $y_i := X_{\pi(i)}$ für $1 \leqslant i \leqslant n$. Wir erhalten also einen Einsetzungs-Homomorphismus

$$\tilde{\pi}\colon R[X_1, \ldots, X_n] \to R[X_1, \ldots, X_n], \qquad f \mapsto f(y_1, \ldots, y_n) = f(X_{\pi(1)}, \ldots, X_{\pi(n)}).$$

Definition 14.2. Sei $f \in R[X_1, \ldots, X_n]$. Dann heißt f ein ***symmetrisches Polynom***, wenn $f = \tilde{\pi}(f) = f(X_{\pi(1)}, \ldots, X_{\pi(n)})$ für alle $\pi \in S_n$ gilt. Sei

$$R[X_1, \ldots, X_n]^{S_n} := \{f \in R[X_1, \ldots, X_n] \mid f \text{ symmetrisch}\}.$$

Weil $\tilde{\pi}\colon R[X_1, \ldots, X_n] \to R[X_1, \ldots, X_n]$ für alle $\pi \in S_n$ ein Ring-Homomorphismus ist, folgt sofort, dass $R[X_1, \ldots, X_n]^{S_n}$ ein Teilring von $R[X_1, \ldots, X_n]$ ist, also Summen, Differenzen und Produkte von symmetrischen Polynomen wieder symmetrisch sind.
Zum Beispiel sind $X_1X_2 \cdots X_n$ und $X_1^k + \ldots + X_n^k$ (mit $k \in \mathbb{N}$) offenbar symmetrisch, oder etwa auch die etwas komplizierter aussehenden Ausdrücke $\prod_{i,j=1}^{n}(X_i + X_j)$ und $\prod_{1 \leqslant i < j \leqslant n}(X_i - X_j)^2$.

Ab hier Woche 10

Beispiel 14.3. Analog zu den Formeln in Lemma 14.1 definieren wir $s_0 := 1$ und

$$s_d := \sum_{1 \leqslant i_1 < \ldots < i_d \leqslant n} X_{i_1} \cdots X_{i_d} \in R[X_1, \ldots, X_n] \qquad \text{für } 1 \leqslant d \leqslant n.$$

Dann ist $\mathrm{Grad}(s_d) = d$ für $1 \leqslant d \leqslant n$. Behauptung: Es gilt $s_d \in R[X_1, \ldots, X_n]^{S_n}$. Um dies zu sehen, betrachten wir den Polynomring $R[X_1, \ldots, X_{n+1}]$ in $n+1$ Unbestimmten und bilden

$$f := (X_{n+1} - X_1)(X_{n+1} - X_2) \cdots (X_{n+1} - X_n) \in R[X_1, \ldots, X_{n+1}].$$

Genauso wie in Lemma 14.1 folgt dann $f = a_0 + a_1X_{n+1} + \ldots + a_{n-1}X_{n+1}^{n-1} + X_{n+1}^n$ mit $a_{n-d} = (-1)^d s_d \in R[X_1, \ldots, X_n]$ für $1 \leqslant d \leqslant n$. Wir fassen nun alle Permutationen $\pi \in S_n$ auch als Permutationen in S_{n+1} auf, indem wir $\pi(n+1) = n+1$ setzen. Dann gilt offensichtlich $\tilde{\pi}(f) = f$ für alle $\pi \in S_n$, also auch $\tilde{\pi}(a_{n-d}) = a_{n-d}$ für $1 \leqslant d \leqslant n$. Also ist $s_d \in R[X_1, \ldots, X_n]^{S_n}$. — Die Polynome s_d heißen ***elementar-symmetrische Polynome***.

Satz 14.4 (Hauptsatz über symmetrische Polynome; Newton, Gauß)**.** *Sei* R *Integritätsring und* $f \in R[X_1, \ldots, X_n]^{S_n}$*. Dann gibt es ein Polynom* $g \in R[X_1, \ldots, X_n]$ *mit* $f = g(s_1 \ldots, s_n)$.

Beweis. Sei $\mathcal{I}_n$ die Menge aller Tupel $\underline{i} = (i_1, \ldots, i_n)$ mit $i_1, \ldots, i_n \in \mathbb{N}_0$. Wir setzen $|\underline{i}| := i_1 + \ldots + i_n$ und definieren wie folgt eine Ordnungrelation $\sqsubseteq$ auf $\mathcal{I}_n$. Seien $\underline{i} \neq \underline{j}$ in $\mathcal{I}_n$ und $k \in \{1, \ldots, n\}$ minimal mit $i_k \neq j_k$. Dann schreibe $\underline{i} \sqsubset \underline{j}$ falls entweder $|\underline{i}| < |\underline{j}|$, oder $|\underline{i}| = |\underline{j}|$ und $i_k < j_k$. Diese Ordnung heißt ***graduierte lexikographische Ordnung***; überzeugen Sie sich selbst, dass die Relation $\sqsubseteq$ transitiv ist. Sind endlich viele Tupel in $\mathcal{I}_n$ gegeben, so

können wir diese stets gemäß $\sqsubseteq$ anordnen. Ist also $0 \neq f \in R[X_1, \ldots, X_n]$ beliebig, so gibt es ein eindeutiges $\underline{i} = (i_1, \ldots, i_n) \in \mathcal{I}_n$ mit

$$f = \underbrace{a_{i_1,\ldots,i_n}}_{\neq 0} X_1^{i_1} \cdots X_n^{i_n} \quad + \quad \text{Summe von Termen } a_{j_1,\ldots,j_n} X_1^{j_1} \cdots X_n^{j_n} \text{ mit } \underline{j} \sqsubset \underline{i}.$$

Wir bezeichnen dann $LT(f) := a_{i_1,\ldots,i_n} X_1^{i_1} \cdots X_n^{i_n}$ als den ***Leitterm*** von f (bezüglich $\sqsubseteq$).

[Zum Beispiel hat $f = 5X_1X_3^2 - X_2^4 + 3X_1^2X_3^2 - 2X_1^3X_2 \in \mathbb{Z}[X_1, X_2, X_3]$ den Leitterm $LT(f) = -2X_1^3X_2$; die Exponenten-Tupel in $\mathcal{I}_3$ zu den vier Termen in f sind $(1,0,2) \sqsubset (0,4,0) \sqsubset (2,0,2) \sqsubset (3,1,0)$.]

In den Übungen werden Sie zeigen:

(1) $\quad LT(f \cdot g) = LT(f) \cdot LT(g) \quad$ für alle $f, g \in R[X_1, \ldots, X_n]$ mit $f \neq 0$, $g \neq 0$.

Schauen wir uns nun die Leitterme von symmetrischen Polynomen an. Behauptung:

(2) Ist $0 \neq f \in R[X_1, \ldots, X_n]^{S_n}$ und $LT(f) = a_{i_1,\ldots,i_n} X_1^{i_1} \cdots X_n^{i_n}$, so gilt $i_1 \geqslant i_2 \geqslant \ldots \geqslant i_n$.

Dazu: Annahme, es gäbe ein k mit $i_k < i_{k+1}$. Dann betrachte die Transposition $\tau = (k\ k{+}1) \in S_n$. Da f symmetrisch ist, gilt $f = \tilde{\tau}(f)$. Wenden wir τ auf $LT(f)$ an, so erhalten wir

$$\tilde{\tau}(LT(f)) = a_{i_1,\ldots,i_n} X_1^{i_1} \cdots X_{k-1}^{i_{k-1}} X_{k+1}^{i_k} X_k^{i_{k+1}} X_{k+2}^{i_{k+2}} \cdots X_n^{i_n},$$

mit Exponenten-Tupel $\underline{j} := (i_1, \ldots, i_{k-1}, i_{k+1}, i_k, i_{k+2}, \ldots, i_n) \in \mathcal{I}_n$. Obiger Term kommt wieder in $f = \tilde{\tau}(f)$ vor, also gilt $\underline{j} \sqsubseteq \underline{i}$ nach Definition des Leitterms. Da $|\underline{j}| = |\underline{i}|$, $j_k = i_{k+1} \neq i_k$ und die ersten $k-1$ Einträge von $\underline{j}$, $\underline{i}$ gleich sind, bedeutet dies gemäß der Definition von $\sqsubseteq$ aber $i_{k+1} = j_k < i_k$, Widerspruch. Also gilt die obige Behauptung (2).

Sei $f = s_d$ das d-te elementar-symmetrische Polynom. Die zugehörigen Exponenten-Tupel sind alle $\underline{i} \in \mathcal{I}_n$ mit $i_1, \ldots, i_n \in \{0, 1\}$, wobei genau d Einsen vorkommen. Das einzige solche Tupel mit $i_1 \geqslant i_2 \geqslant \ldots \geqslant i_n$ ist $(1, \ldots, 1, 0, \ldots, 0)$ (alle Einsen am Anfang). Mit (2) folgt:

(3) $$LT(s_d) = X_1X_2 \cdots X_d \qquad \text{für } 1 \leqslant d \leqslant n.$$

Nach diesen Vorbereitungen beschreiben wir jetzt einen Algorithmus, der zu einem gegebenen $0 \neq f \in R[X_1, \ldots, X_n]^{S_n}$ ein Polynom $g \in R[X_1, \ldots, X_n]$ findet mit $f = g(s_1, \ldots s_n)$. Dazu betrachten wir $LT(f) = a_{i_1,\ldots,i_n} X_1^{i_1} \cdots X_n^{i_n}$. Nach (2) ist $i_1 \geqslant i_2 \geqslant \ldots \geqslant i_n$; damit setzen wir

$$h := a_{i_1,\ldots,i_n} s_1^{i_1-i_2} s_2^{i_2-i_3} s_3^{i_3-i_4} \cdots s_{n-1}^{i_{n-1}-i_n} s_n^{i_n} \in R[X_1, \ldots, X_n]^{S_n}.$$

Durch mehrfache Anwendung von (1) und (3) erhalten wir

$$\begin{aligned} LT(h) &= a_{i_1,\ldots,i_n} LT(s_1)^{i_1-i_2} LT(s_2)^{i_2-i_3} LT(s_3)^{i_3-i_4} \cdots LT(s_{n-1})^{i_{n-1}-i_n} LT(s_n)^{i_n} \\ &= a_{i_1,\ldots,i_n} X_1^{i_1-i_2} (X_1X_2)^{i_2-i_3} (X_1X_2X_3)^{i_3-i_4} \cdots (X_1X_2 \cdots X_{n-1})^{i_{n-1}-i_n} (X_1X_2 \cdots X_n)^{i_n}. \end{aligned}$$

Hier kommt X_1 mit Exponent $(i_1 - i_2) + (i_2 - i_3) + (i_3 - i_4) + \ldots + (i_{n-1} - i_n) + i_n = i_1$ vor, sodann X_2 mit Exponent $(i_2 - i_3) + (i_3 - i_4) + \ldots + (i_{n-1} - i_n) + i_n = i_2$, und so weiter. Also

erhalten wir letztendlich $\mathrm{LT}(h) = a_{i_1,\ldots,i_n} X_1^{i_1} \cdots X_n^{i_n} = \mathrm{LT}(f)$. Da f, h symmetrisch sind, ist auch $f' := f - h$ symmetrisch. Ist $f' = 0$, so ist $f = h$ und wir sind fertig. Ist $f' \neq 0$, so ist das Exponenten-Tupel im Leitterm von f' echt kleiner (bezüglich $\sqsubseteq$) als $(i_1, \ldots, i_n)$, da der Term $a_{i_1,\ldots,i_n} X_1^{i_1} \cdots X_n^{i_n}$ in $f' = f - h$ wegfällt. Also können wir mit f' fortfahren und so fort. Dieses Verfahren muss nach endlich vielen Schritten abbrechen. Dazu beachte: Ist $\underline{i} \in \mathcal{I}_n$ gegeben, so gibt es nur endlich viele $\underline{j} \in \mathcal{I}_n$ mit $\underline{j} \sqsubseteq \underline{i}$. (Denn ist $\underline{j} \sqsubseteq \underline{i}$, so muss insbesondere $|\underline{j}| \leqslant |\underline{i}|$ gelten und damit auch $j_k \leqslant |\underline{i}|$ für $1 \leqslant k \leqslant n$, d.h., die Einträge von $\underline{j}$ sind beschränkt). $\square$

Beispiel 14.5. (a) Sei $f = X_1^2 + X_2^2 + X_3^2 \in \mathbb{Z}[X_1, X_2, X_3]^{S_3}$. Dann ist $\mathrm{LT}(f) = X_1^2$ mit Exponenten-Tupel $(i_1, i_2, i_3) = (2, 0, 0) \in \mathcal{I}_3$. Also betrachten wir $h = s_1^{i_1 - i_2} s_2^{i_2 - i_3} s_3^{i_3} = s_1^2$ und $f' := f - h = X_1^2 + X_2^2 + X_3^2 - s_1^2 = X_2^2 + X_3^2 - (X_1 + X_2 + X_3)^2 = -2X_1X_2 - 2X_1X_3 - 2X_2X_3 = -2s_2$. Hier sind wir also bereits fertig; es gilt $f = h + f' = s_1^2 - 2s_2$.

(b) Sei $f = X_1^4 + X_2^4 \in \mathbb{Z}[X_1, X_2]^{S_2}$. Dann ist $\mathrm{LT}(f) = X_1^4$ mit Exponenten-Tupel $(i_1, i_2) = (4, 0) \in \mathcal{I}_2$. Also betrachten wir $h = s_1^{i_1 - i_2} s_2^{i_2} = s_1^4$ und $f' := f - h = X_1^4 + X_2^4 - (X_1 + X_2)^4 = -4X_1X_2^3 - 6X_1^2X_2^2 - 4X_1^3X_2$. Nun ist $\mathrm{LT}(f') = -4X_1^3X_2$ mit Exponenten-Tupel $(i_1, i_2) = (3, 1) \in \mathcal{I}_2$. Also betrachten wir $h' := -4s_1^{i_1 - i_2} s_2^{i_2} = -4s_1^2 s_2 = -4(X_1 + X_2)^2 X_1X_2$ und $f'' := f' - h' = 2X_1^2X_2^2 = 2s_2^2$. Also sind wir hier fertig und es folgt

$$f = h + f' = h + (h' + f'') = s_1^4 - 4s_1^2s_2 + 2s_2^2.$$

Zum Verständnis dieses Verfahrens mag es wiederum sehr nützlich sein, ein Computer-Programm zu schreiben. Dazu benötigt man nun ein Computer-Algebra-System, in dem man exakt mit Polynomen in mehreren Variablen arbeiten kann, zum Beispiel GAP.

Ein Grund für die Bedeutung von Satz 14.4 liegt darin, dass man damit in bestimmten Situationen Argumente führen kann, die man ansonsten nur mit Galois-Theorie hinbekommen würde; in diesem Sinne ist dieser Satz sozusagen ein Vorläufer der Galois-Theorie.

Beispiel 14.6. Sei $\mathbb{Q} \subseteq K$ und $f = a_0 + a_1X + \ldots + a_{n-1}X^{n-1} + X_n \in K[X]$. Seien $z_1, \ldots, z_n \in K$ mit $f = (X - z_1) \cdots (X - z_n)$. Sei $h \in \mathbb{Z}[X_1, \ldots, X_n]^{S_n}$ beliebig. Was können wir dann über $h(z_1, \ldots, z_n)$ aussagen? Nun, nach dem Hauptsatz gibt es ein $g \in \mathbb{Z}[X_1, \ldots, X_n]$ mit $h = g(s_1, \ldots, s_n)$. Durch Kombination von Lemma 14.1 und Definition 14.2 folgt $s_d(z_1, \ldots, z_n) = (-1)^d a_{n-d}$ für $1 \leqslant d \leqslant n$. Mit Hilfe von g erhalten wir damit die Formel

$$h(z_1, \ldots, z_n) = g\big(s_1(z_1, \ldots, z_n), \ldots, s_n(z_1, \ldots, z_n)\big) = g\big(-a_{n-1}, a_{n-2}, \ldots, (-1)^n a_0\big),$$

d.h., $h(z_1, \ldots, z_n)$ kann man als Polynom in den Koeffizienten a_i von f schreiben.

Verblüffende Konsequenz: *Sind alle Koeffizienten von f bereits in einem Teilring $R \subseteq K$ enthalten (zum Beispiel $R = \mathbb{Z}$ oder $R = \mathbb{Q}$), so ist auch $h(z_1, \ldots, z_n) \in R$.*

Kapitel IV: Körper

Wir kennen bisher vor allem die Körper $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ und $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$, für eine Primzahl p. Ein Ziel wird es sein, Methoden zur Konstruktion weiterer Körper zu beschreiben, zum Beispiel Körper, über denen gegebene Polynome in Linearfaktoren zerfallen, oder endliche Körper.

15. *Algebraische und transzendente Elemente*

Sei R kommutativer Ring mit 1. Für $a \in R$ und $m \in \mathbb{Z}$ können wir $m \cdot a \in R$ definieren, mit den üblichen Konventionen (also etwa $3 \cdot a = a + a + a$, $0 \cdot a = 0_R$, $(-5) \cdot a = -(5 \cdot a)$ usw.). Es folgt sofort, dass die Abbildung $\varphi\colon \mathbb{Z} \to R$, $m \mapsto m \cdot 1_R$, ein Ring-Homomorphismus ist. Da $\mathbb{Z}$ Hauptidealring ist, gilt also $\operatorname{Kern}(\varphi) = p\mathbb{Z} = (p)$ mit einem eindeutigen $p \in \mathbb{N}_0$. Wir setzen $\operatorname{char}(R) := p$ und nennen dies die ***Charakteristik*** von R. Es gibt also 2 Fälle:

(1) $p = 0$. Dies gilt genau dann, wenn $m \cdot 1_R \neq 0$ für alle $m > 0$. In diesem Fall ist φ injektiv und $\mathbb{Z}$ kann als Teilring von R aufgefasst werden vermöge dieser Abbildung.
(2) $p > 0$. Dann ist $p \cdot 1_R = 0$ und p ist minimal mit dieser Eigenschaft. In diesem Fall erhalten wir nach Satz 12.5 einen injektiven Homomorphismus $\mathbb{Z}/p\mathbb{Z} \to R$, $\bar{i} \mapsto i \cdot 1_R$, können also vermöge dieser Abbildung $\mathbb{Z}/p\mathbb{Z}$ als Teilring von R auffassen.

Bemerkung 15.1. Sei R ein Integritätsring und $p := \operatorname{char}(R) \geqslant 0$. Ist $p > 0$, so muss p eine Primzahl sein. Denn: Wäre $p = p_1 p_2$ mit $0 < p_1, p_2 < p$, so $0 = p \cdot 1_R = (p_1 p_2) \cdot 1_R = (p_1 \cdot 1_R)(p_2 \cdot 1_R)$. Da es keine echten Nullteiler gibt, muss also $p_1 \cdot 1_R = 0$ oder $p_2 \cdot 1_R = 0$ gelten, Widerspruch zur Minimalität von p. Insbesondere folgt:

Die Charakteristik eines Körpers ist entweder 0 oder eine Primzahl.

Dies ist die erste wichtige Invariante, um Körper voneinander zu unterscheiden. Eine Besonderheit von Körpern mit positiver Charakteristik ist folgende Aussage.

Lemma 15.2. *Sei R ein Integritätsring (also zum Beispiel ein Körper) mit $\operatorname{char}(R) = p > 0$ wobei p eine Primzahl ist. Dann ist die Abbildung $F\colon R \to R$, $a \mapsto a^p$, ein injektiver Ringhomomorphismus, der **Frobenius-Homomorphismus** genannt wird.*

Beweis. Für $a, b \in R$ gilt $F(ab) = (ab)^p = a^p b^p = F(a)F(b)$ weil R kommutativ ist. Außerdem ist $\operatorname{Kern}(F) = \{a \in R \mid a^p = 0\} = \{0\}$ weil R Integritätsring ist. Es bleibt also zu zeigen: $F(a + b) = F(a) + F(b)$, also $(a + b)^p = a^p + b^p$. Nach der Binomialformel ist

$$(a + b)^p = \sum_{i=0}^{p} \binom{p}{i} \cdot a^i b^{p-i} = a^p + \Big(\sum_{i=1}^{p-1} \binom{p}{i} \cdot a^i b^{p-i}\Big) + b^p.$$

Betrachte nun $\binom{p}{i} = \frac{p(p-1)(p-2)\cdots(p-i+1)}{1\cdot 2\cdot\ldots\cdot i} \in \mathbb{N}$ für $1 \leqslant i \leqslant p-1$. Der Zähler dieses Bruches ist durch p teilbar, aber wegen $1 \leqslant i \leqslant p-1$ ist der Nenner nicht durch p teilbar. Also ist $p \mid \binom{p}{i}$ für $1 \leqslant i \leqslant p-1$ und damit $\binom{p}{i} \cdot a^i b^{p-i} = \left(\binom{p}{i} \cdot 1_R\right) \cdot a^i b^{p-i} = 0$ in R. □

Sei L ein Körper und $K \subseteq L$ ein Teilkörper, d.h., die Addition und Multiplikation in K sind Einschränkungen der analogen Operationen in L. Wir nennen $L \supseteq K$ eine ***Körpererweiterung*** und L einen Erweiterungskörper von K. Standard-Beispiele sind $\mathbb{Q} \subseteq \mathbb{R}$ oder $\mathbb{R} \subseteq \mathbb{C}$. Ist $L \supseteq K$ eine Körpererweiterung, so können wir L auch als einen K-Vektorraum auffassen, wobei die Multiplikation von $v \in L$ mit einem Skalar $s \in K$ einfach durch die Multiplikation sv in L gegeben ist. (Alle Vektorrraum-Axiome gelten automatisch.) Dann heißt

$$[L : K] := \dim_K L$$

der ***Körpergrad*** von L über K. Die Erweiterung $L \supseteq K$ heißt endliche Erweiterung, wenn $[L : K] < \infty$ gilt. Beachte auch: Es gilt $[L : K] = 1$ genau dann, wenn $L = K$ ist.

Beispiel 15.3. (a) $\mathbb{C} \supseteq \mathbb{R}$ ist eine Körpererweiterung mit $[\mathbb{C} : \mathbb{R}] = 2$; eine $\mathbb{R}$-Basis von $\mathbb{C}$ ist gegeben durch $\{1, i\}$. Analog können wir auch $\mathbb{Q}(i) := \{x + iy \mid x, y \in \mathbb{Q}\} \subseteq \mathbb{C}$ definieren. Dann sieht man sofort, dass $\mathbb{Q}(i)$ ein Teilkörper von $\mathbb{C}$ ist und $[\mathbb{Q}(i) : \mathbb{Q}] = 2$. Wiederum analog sieht man leicht, dass zum Beispiel $\mathbb{Q}(\sqrt{2}) := \{x + \sqrt{2}\, y \mid x, y \in \mathbb{Q}\} \subseteq \mathbb{R}$ ein Teilkörper ist. Es gilt $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$, denn $\{1, \sqrt{2}\}$ sind linear unabhängig über $\mathbb{Q}$ (wegen $\sqrt{2} \notin \mathbb{Q}$). Beispiele dieses Typs werden wir im nächsten Abschnitt ausführlicher behandeln.

(b) Es gilt $[\mathbb{R} : \mathbb{Q}] = \infty$. Dies kann einfach gezeigt werden mit Hilfe eines Abzählarguments. Zur Erinnerung: Eine Menge X heißt ***abzählbar***, wenn es eine Bijektion $\mathbb{N}_0 \to X$ gibt. Sind X und Y abzählbar, so auch $X \times Y$. (Zum Beispiel ist $\mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$, $(k, n) \mapsto 2^k(2n+1)-1$, eine Bijektion; also erhalten wir auch eine Bijektion zwischen $X \times Y$ und X.) Folglich ist X^n abzählbar für alle $n \geqslant 1$, wenn X abzählbar ist. Nun ist $|\mathbb{Q}| = \infty$; außerdem haben wir eine surjektive Abbildung $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\}) \to \mathbb{Q}$, $(n, m) \mapsto \frac{n}{m}$. Da $\mathbb{Z}$ abzählbar ist (warum?), ist also $\mathbb{Q}$ abzählbar und damit auch $\mathbb{Q}^n$ abzählbar für alle $n \geqslant 1$. Wäre nun $[\mathbb{R} : \mathbb{Q}] = n < \infty$, so gäbe es eine Bijektion $\mathbb{R} \leftrightarrow \mathbb{Q}^n$, also wäre $\mathbb{R}$ abzählbar, Widerspruch (siehe Analysis I).

Bemerkung 15.4. Sei K beliebiger Körper. Betrachte wie oben den Ring-Homomorphismus $\varphi\colon \mathbb{Z} \to K$, $m \mapsto m \cdot 1_K$; dann ist $\mathrm{Kern}(\varphi) = p\mathbb{Z}$, wobei entweder $p = 0$ oder p eine Primzahl ist. Ist $p > 0$, so haben wir bereits gesehen, dass wir $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ als Teilring (also Teilkörper) von K auffassen können. Ist $p = 0$, so können wir analog $\mathbb{Z}$ als Teilring von K auffassen. Nun sieht man sofort, dass $K_0 := \{(n \cdot 1_K)/(m \cdot 1_K) \mid n \in \mathbb{Z}, 0 \neq m \in \mathbb{Z}\} \subseteq K$ ein Teilkörper ist, den wir mit $\mathbb{Q}$ identifizieren können (indem wir einen Bruch $\frac{n}{m} \in \mathbb{Q}$ mit $(n \cdot 1_K)/(m \cdot 1_K) \in K_0$ gleichsetzen). Mit diesen Identifizierungen haben wir also stets eine Körpererweiterung

$$K \supseteq \mathbb{Q} \qquad \text{oder} \qquad K \supseteq \mathbb{F}_p \quad \text{mit } p \text{ Primzahl.}$$

Hier heißt $\mathbb{Q}$ bzw. $\mathbb{F}_p$ der ***Primkörper*** von K.

Definition 15.5. Sei $L \supseteq K$ eine Körpererweiterung und $\alpha \in L$. Sei $\varphi\colon K \hookrightarrow L$ die Inklusionsabbildung. Betrachte den Einsetzungs-Homomorphismus $\tilde{\varphi}_\alpha\colon K[X] \to L$, $f \mapsto f(\alpha)$.

(a) Ist $\mathrm{Kern}(\tilde{\varphi}_\alpha) = \{0\}$, so heißt α ***transzendent***; sonst heißt α ***algebraisch*** (jeweils über K).
(b) Sei α algebraisch, d.h., es gibt ein $0 \neq f \in K[X]$ mit $f(\alpha) = 0$. Nach Folgerung 12.2 gibt es dann ein eindeutiges normiertes $0 \neq f_0 \in K[X]$ mit $\mathrm{Kern}(\tilde{\varphi}_\alpha) = (f_0)$. Dieses f_0 heißt ***Minimalpolynom*** von α (über K) und wird mit $\mu_\alpha := f_0$ bezeichnet.
(c) Ist jedes $\alpha \in L$ algebraisch über K, so heißt $L \supseteq K$ eine algebraische Körpererweiterung.

Lemma 15.6. *Ist* $[L : K] < \infty$*, so ist* $L \supseteq K$ *eine algebraische Erweiterung.*

Beweis. Sei $\alpha \in L$. Sei $n = [L : K] \geqslant 1$. Dann sind $1, \alpha, \ldots, \alpha^n$ linear abhängig, also gibt es $a_0, a_1, \ldots, a_n \in K$ (die nicht alle gleich Null sind) mit $a_0 + a_1\alpha + \ldots + a_n\alpha^n = 0$. Damit ist $f := a_0 + a_1X + \ldots + a_nX^n \in K[X]$ ein Polynom mit $f \neq 0$ und $f(\alpha) = 0$. □

Lemma 15.7. *Sei* $\alpha \in L$ *algebraisch und* $\mu_\alpha \in K[X]$ *das Minimalpolynom von* α*. Dann ist* μ_α *irreduzibel und es gilt* $\mu_\alpha \mid f$ *für jedes* $f \in K[X]$ *mit* $f(\alpha) = 0$*. Ist umgekehrt* $f \in K[X]$ *irreduzibel mit* $f(\alpha) = 0$*, so gilt* $f = c\mu_\alpha$ *mit einem* $0 \neq c \in K$*.*

Beweis. Zunächst beachte: Da μ_α eine Nullstelle hat (nämlich α), gilt $\mathrm{Grad}(\mu_\alpha) \geqslant 1$ (und μ_α ist keine Einheit in $K[X]$). Sei nun $f \in K[X]$ beliebig mit $f(\alpha) = 0$. Dann ist $f \in \mathrm{Kern}(\tilde{\varphi}_\alpha) = (\mu_\alpha)$, also $\mu_\alpha \mid f$ und $f = g\mu_\alpha$ mit einem $0 \neq g \in K[X]$. Ist f sogar irreduzibel, so muss nun g eine Einheit sein, also $c := g \in K \setminus \{0\}$.
Sei schließlich $\mu_\alpha = f_1f_2$ mit $f_1, f_2 \in K[X]$. Weil Einsetzen ein Homomorphismus ist, folgt $0 = \mu_\alpha(\alpha) = f_1(\alpha)f_2(\alpha)$, also $f_1(\alpha) = 0$ oder $f_2(\alpha) = 0$. Ist $f_1(\alpha) = 0$, so folgt also $\mu_\alpha \mid f_1$ und $f_1 = h\mu_\alpha = hf_1f_2$ mit $0 \neq h \in K[X]$. Durch Kürzen von f_1 folgt $1 = hf_2$, also ist $f_2 \in K$. Analog folgt aus $f_2(\alpha) = 0$, dass $f_1 \in K$ gilt. Also ist μ_α irreduzibel. □

Bemerkung: In der Linearen Algebra wird auch jeder Matrix $A \in M_n(K)$ ein Minimalpolynom zugeordnet, also ein eindeutiges normiertes Polynom kleinsten Grades $\mu_A \in K[X]$ mit $\mu_A(A) = 0_{n\times n}$. Beachte allerdings, dass μ_A nicht irreduzibel sein muss; zum Beispiel hat die Diagonalmatrix $A \in M_2(\mathbb{Q})$ mit Diagonaleinträgen 1 und -1 das Minimalpoynom $\mu_A = (X-1)(X+1)$.

Beispiel 15.8. (a) $\mathbb{C} \supseteq \mathbb{R}$ ist algebraisch, weil $[\mathbb{C} : \mathbb{R}] = 2$ gilt. Das Polynom $f := X^2 + 1 \in \mathbb{R}[X]$ ist irreduzibel (weil es keine Nullstelle in $\mathbb{R}$ hat) und es gilt $f(i) = 0$. Also ist $f = \mu_i$ nach Lemma 15.7. Analog ist $\mathbb{Q}(\sqrt{2}) \supseteq \mathbb{Q}$ algebraisch. Das Polynom $f := X^2 - 2 \in \mathbb{Q}[X]$ ist irreduzibel (Eisenstein-Kriterium mit $p = 2$) und es gilt $f(\sqrt{2}) = 0$. Also ist $\mu_{\sqrt{2}} = f$.
(b) In $\mathbb{R} \supseteq \mathbb{Q}$ sei $\alpha = \sqrt[3]{-1+\sqrt{3}} \in \mathbb{R}$. Dann ist $\alpha^3 = -1 + \sqrt{3}$, also $(\alpha^3 + 1)^2 = 3$ und

$\alpha^6+2\alpha^3-2=0$. Damit ist $f(\alpha)=0$ für $f:=X^6+2X^3-2\in\mathbb{Q}[X]$, also α algebraisch über $\mathbb{Q}$. Mit dem Eisenstein-Kriterium ($p=2$) folgt, dass f sogar irreduzibel ist, also $\mu_\alpha=f$.
(c) In $\mathbb{C}\supseteq\mathbb{Q}$ sei $\alpha=i+\sqrt{2}\in\mathbb{C}$. In diesem Fall gibt es zunächst keinen offensichtlichen Kandidaten für ein $f\in\mathbb{Q}[X]$ mit $f(\alpha)=0$. Wir berechnen dann einige Potenzen $\alpha^2,\alpha^3,\ldots$ und versuchen, eine Linearkombination zwischen diesen zu finden:

$\alpha^2=(i+\sqrt{2})^2=-1+2i\sqrt{2}+2=1+2i\sqrt{2}$.
$\alpha^3=(i+\sqrt{2})(1+2i\sqrt{2})=i-2\sqrt{2}+\sqrt{2}+4i=5i-\sqrt{2}$.
$\alpha^4=(1+2i\sqrt{2})^2=1+4i\sqrt{2}-8=4i\sqrt{2}-7=2(\alpha^2-1)-7=2\alpha^2-9$.

Also ist $f(\alpha)=0$ mit $f:=X^4-2X^2+9\in\mathbb{Q}[X]$. Hier kann man zwar das Eisenstein-Kriterium nicht anwenden, aber man zeigt auch leicht direkt (zum Beispiel mit dem Kronecker-Verfahren in Beispiel 13.5(c)), dass f irreduzibel ist, also $f=\mu_\alpha$. Wir werden dies auch noch auf etwas andere Weise in Beispiel 16.8 im nächsten Abschnitt sehen.

Im nächsten Beispiel wird mit einem abstrakten Argument gezeigt, dass $\mathbb{R}\supseteq\mathbb{Q}$ keine algebraische Erweiterung ist, also ist auch $\mathbb{C}\supseteq\mathbb{Q}$ keine algebraische Erweiterung. Für ein konkret gegebenes $\alpha\in\mathbb{C}$ kann es im Einzelfall extrem schwierig sein, zu entscheiden, ob α algebraisch ist oder nicht. Wir werden aber zeigen könnnen, dass durch algebraische Operationen (also Summen, Produkte, Wurzelziehen, ...) gebildete Ausdrücke wie oben in (b) oder (c) immer algebraisch sind, ohne dies explizit nachrechnen zu müssen.

Ab hier Woche 11

Beispiel 15.9. Es gibt unendlich viele $\alpha\in\mathbb{R}$, die transzendent über $\mathbb{Q}$ sind. Dazu benutzen wir ein Zählargument, das auf ***Cantor*** zurückgeht. Sei $\mathbb{Q}[X]^\#$ die Menge der nicht-konstanten Polynome in $\mathbb{Q}[X]$. Dann ist $\mathbb{Q}[X]^\#=\bigcup_{n\geqslant 1}P_n$, wobei $P_n=\{f\in\mathbb{Q}[X]^\#\mid \mathrm{Grad}(f)=n\}$. Nun ist P_n in Bijektion zu $\mathbb{Q}^n\times\mathbb{Q}^\times$ (die Koeffizienten $a_0,\ldots,a_n$ eines Polynoms $\neq 0$ und vom Grad n bilden ein $(n+1)$-Tupel mit $a_n\neq 0$), also ist P_n abzählbar. Für jedes $n\geqslant 1$ sei $a_n\colon\mathbb{N}_0\to P_n$ eine Bijektion. Dann ist $f\colon\mathbb{N}\times\mathbb{N}_0\to\mathbb{Q}[X]^\#$, $(n,k)\mapsto a_n(k)$, bijektiv, also $\mathbb{Q}[X]^\#$ abzählbar. D.h., es gibt eine Aufzählung $\mathbb{Q}[X]^\#=\{f_m\mid m\in\mathbb{N}_0\}$. Für jedes $m\geqslant 0$ sei $N_m\subseteq\mathbb{R}$ die Menge der Nullstellen von f_m; dann ist $|N_m|\leqslant\mathrm{Grad}(f_m)<\infty$. Nun folgt, dass

$$\{\alpha\in\mathbb{R}\mid \alpha\text{ algebraisch über }\mathbb{Q}\}=\bigcup_{m\in\mathbb{N}_0}N_m$$

abzählbar ist (weil eine abzählbare Vereinigung von endlichen Mengen wieder abzählbar ist). Da $\mathbb{R}$ überabzählbar ist, sind also tatsächlich die meisten reellen Zahlen transzendent über $\mathbb{Q}$. — Allerdings kennen wir damit noch kein einziges konkretes Beispiel! Der Beweis, dass eine gegebene Zahl wirklich transzendent ist, ist erheblich schwieriger. Beispiele:

Hermite (1873): e ist transzendent; **Lindemann** (1882): π ist transzendent.
Liouville (1844): Schnell konvergierende Reihen sind transzendent, z.B. $\sum_{n=1}^{\infty}10^{-n!}$.
Literatur dazu: Siehe die Bücher von Fischer–Sacher [FS, Anhang 2] und Lorenz [Lo, §17].

16. *Der Gradsatz*

Sei $L \supseteq K$ eine Körpererweiterung. Sei $M \subseteq L$ ein Teilkörper mit $K \subseteq M$; dann bezeichnen wir M auch als ***Zwischenkörper*** der Erweiterung $L \supseteq K$. Ist $[L : K] < \infty$, so gilt auch $[L : M] < \infty$ und $[M : K] < \infty$. Die folgende Umkehrung wird sich als sehr nützlich erweisen.

Satz 16.1 (Gradsatz)**.** *Sei $M \subseteq L$ Teilkörper mit $K \subseteq M$. Gilt $[L : M] < \infty$ und $[M : K] < \infty$, so ist auch $[L : K] < \infty$ und $[L : K] = [L : M][M : K]$.*

Beweis. Sei $n := [M : K] < \infty$ und $m := [L : M] < \infty$. Sei $\{x_1, \ldots, x_n\}$ eine K-Basis von M und $\{y_1, \ldots, y_m\}$ eine M-Basis von L. Wir zeigen:

$(*)$ $\qquad \{x_i y_j \mid i \in 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m\}$ ist eine K-Basis von L.

Dazu: Sei $z \in L$ beliebig. Dann ist $z = \sum_{1 \leqslant j \leqslant m}^{m} b_j y_j$ mit $b_j \in M$. Außerdem $b_j = \sum_{1 \leqslant i \leqslant n} a_{ij} x_i$ mit $a_{ij} \in K$. Also erhalten wir eine endliche Summe $z = \sum_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m} a_{ij} x_i y_j$, d.h., obige Menge ist ein Erzeugendensystem von L über K. Nun zur linearen Unabhängigkeit. Seien $a_{ij} \in K$ und $0 = \sum_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m} a_{ij} x_i y_j$. Dann ist

$$0 = \sum_{1 \leqslant j \leqslant m} \Big(\sum_{1 \leqslant i \leqslant n} a_{ij} x_i \Big) y_j = \sum_{1 \leqslant j \leqslant m} c_j y_j \qquad \text{mit} \qquad c_j := \sum_{1 \leqslant i \leqslant n} a_{ij} x_i \in M.$$

Da $\{y_j\}$ linear unabhängig über M sind, folgt also $c_j = 0$ für alle j. Weil $\{x_i\}$ linear unabhängig über K sind, folgt dann auch $a_{ij} = 0$ für alle i, j. Aus $(*)$ folgt schließlich $[L : K] = mn$. $\square$

Sei $L \supseteq K$ Körpererweiterung und $S \subseteq L$ eine Teilmenge. Analog zur Definition in §6 bei Gruppen definieren wir $K(S) \subseteq L$ als den Durchschnitt aller Teilkörper $K' \subseteq L$ mit $K \subseteq K'$ und $S \subseteq K'$. Da beliebige Schnitte von Teilkörpern wieder Teilkörper sind, ist $K(S)$ ein Zwischenkörper der Erweiterung $L \supseteq K$. Ist $S = \{\alpha_1, \ldots, \alpha_r\}$, so schreiben wir kurz $K(\alpha_1, \ldots, \alpha_r)$ anstelle von $K(S)$ und nennen dies die ***Adjunktion*** von $\alpha_1, \ldots, \alpha_r \in L$ zu K. Wir befassen uns nun als Erstes mit dem Spezialfall $K(\alpha)$ für ein $\alpha \in L$; derartige Zwischenkörper heißen auch ***einfache Körpererweiterungen***.

Beispiel 16.2. Sei $\alpha \in L$ transzendent. Dann ist $K(\alpha) = \{f(\alpha)/g(\alpha) \mid f, g \in K[X], g \neq 0\}$, denn die Menge auf der rechten Seite ist ein Teilkörper von L, und diese Menge ist sicherlich in jedem Teilkörper enthalten, der K und α enthält.

Satz 16.3. *Sei $\alpha \in L$ algebraisch über K, mit Minimalpolynom $\mu_\alpha \in K[X]$. Dann gilt $K(\alpha) = \{f(\alpha) \mid f \in K[X]\} \subseteq L$. Jedes Element von $K(\alpha)$ lässt sich eindeutig schreiben als $\sum_{i=0}^{d-1} a_i \alpha^i$ mit $a_i \in K$, wobei $d := \mathrm{Grad}(\mu_\alpha) \geqslant 1$. Insbesondere ist $[K(\alpha) : K] = d = \mathrm{Grad}(\mu_\alpha)$.*

Beweis. Betrachte den Einsetzungs-Homomorphismus $\tilde{\varphi}_\alpha \colon K[X] \to L$, $f \mapsto f(\alpha)$; es gilt also $\mathrm{Kern}(\tilde{\varphi}_\alpha) = (\mu_\alpha)$. Sei $R := \{f(\alpha) \mid f \in K[X]\} \subseteq L$ das Bild von $\tilde{\varphi}_\alpha$. Dies ist ein Teilring von L

mit $K \subseteq R$; insbesondere ist R ein Integritätsring. Nach Satz 12.5 ist $K[X]/(\mu_\alpha) \cong R \subseteq L$. Da $K[X]$ ein Hauptidealring ist und μ_α irreduzibel, ist $K[X]/(\mu_\alpha)$ ein Körper; siehe Satz 11.9(a). Also ist $R \subseteq L$ ein Teilkörper mit $K \subseteq R$ und $\alpha \in R$; insbesondere also $K(\alpha) \subseteq R$ (nach Definition von $K(\alpha)$). Andererseits ist klar, dass R in jedem Teilkörper $K' \subseteq L$ mit $K \subseteq K'$ und $\alpha \in K'$ enthalten ist. Also folgt $K(\alpha) = R$.
Sei nun $\beta \in K(\alpha)$ beliebig, also $\beta = f(\alpha)$ mit einem $f \in K[X]$. Division mit Rest ergibt $f = g\mu_\alpha + r$ mit $g, r \in K[X]$ und entweder $r = 0$, oder $r \neq 0$ und $\mathrm{Grad}(r) < \mathrm{Grad}(\mu_\alpha) = d$. In jedem Fall ist $r = \sum_{i=0}^{d-1} a_i X^i$ mit $a_i \in K$. Dann folgt

$$\beta = f(\alpha) = g(\alpha)\underbrace{\mu_\alpha(\alpha)}_{=0} + r(\alpha) = r(\alpha) = \sum_{i=0}^{d-1} a_i \alpha^i.$$

Dies zeigt, dass $K(\alpha)$ von $1, \alpha, \ldots, \alpha^{d-1}$ als K-Vektorraum erzeugt wird. Seien nun $a_i \in K$ mit $\sum_{i=0}^{d-1} a_i \alpha^i = 0$. Dann gilt $g(\alpha) = 0$ für $g := \sum_{i=0}^{d-1} a_i X^i \in K[x]$, also $\mu_\alpha \mid g$. Aus $g \neq 0$ würde $d = \mathrm{Grad}(\mu_\alpha) \leqslant \mathrm{Grad}(g) \leqslant d - 1$ folgen, Widerspruch. Also ist $g = 0$, d.h., $a_i = 0$ für alle i. Damit ist gezeigt, dass $1, \alpha, \ldots, \alpha^{d-1}$ auch linear unabhängig über K sind. Also ist $\{1, \alpha, \ldots, \alpha^{d-1}\}$ eine K-Basis von $K(\alpha)$ und damit $[K(\alpha) : K] = d$. □

Beispiel 16.4. Sei K ein endlicher Körper. Die Charakteristik von K kann nicht 0 sein, denn sonst wäre $\mathbb{Q} \subseteq K$, also $|K| = \infty$, Widerspruch. Also ist $\mathrm{char}(K) = p$ und $\mathbb{F}_p \subseteq K$ für eine Primzahl p. Wegen $|K| < \infty$ folgt auch $n := [K : \mathbb{F}_p] = \dim_{\mathbb{F}_p} K < \infty$. Nach Satz 12.8 gibt es ein $0 \neq \alpha \in K$ mit $K^\times = \langle \alpha \rangle$. Dann ist $K = \{0\} \cup \{\alpha^i \mid 1 \leqslant i \leqslant |K| - 1\}$; insbesondere ist $K = \mathbb{F}_p(\alpha)$ eine einfache Erweiterung. Sei $0 \neq \mu_\alpha \in \mathbb{F}_p[X]$ das Minimalpolynom von α. Nach Satz 16.3 ist $\mathrm{Grad}(\mu_\alpha) = [K : \mathbb{F}_p] = n$. Schließlich sei $\{z_1, \ldots, z_n\}$ eine $\mathbb{F}_p$-Basis von K. Dann lässt sich jedes $z \in K$ eindeutig schreiben als $z = \sum_{i=1}^{n} s_i z_i$ mit $s_i \in \mathbb{F}_p$, also folgt $|K| = p^n$.

Beispiel 16.5. Seien $\alpha_1, \ldots, \alpha_r \in L$ algebraisch über K. Ist $r > 1$, so kann der Teilkörper $L' := K(\alpha_1, \ldots, \alpha_r) \subseteq L$ sukzessive aus einfachen Erweiterungen aufgebaut werden:

$$K(\alpha_1, \alpha_2) = (K(\alpha_1))(\alpha_2), \quad \ldots, \quad L' = K(\alpha_1, \ldots, \alpha_r) = (K(\alpha_1, \ldots, \alpha_{r-1}))(\alpha_r).$$

Jedes α_i ist dabei natürlich weiterhin auch algebraisch über $K(\alpha_1, \ldots, \alpha_{i-1})$. Behauptung:

$$[L' : K] < \infty \qquad \text{und} \qquad L' = \{f(\alpha_1, \ldots, \alpha_r) \mid f \in K[X_1, \ldots, X_r]\}.$$

Beweis mit Induktion nach r. Für $r = 1$ gelten die Aussagen nach Satz 16.3. Sei nun $r > 1$; dann ist $M := K(\alpha_1, \ldots, \alpha_{r-1}) \subseteq L'$ mit $K \subseteq M$; außerdem ist $L' = M(\alpha_r)$ und α_r algebraisch über M. Nach Induktion gilt $[M : K] < \infty$, und nach Satz 16.3 auch $[L' : M] < \infty$. Mit dem Gradsatz folgt also $[L' : K] = [L' : M][M : K] < \infty$. Sei nun $R = K[X_1, \ldots, X_r]$ und $R' = K[X_1, \ldots, X_{r-1}]$. Nach der rekursiven Definition in Beispiel 13.12 ist $R = R'[X_n]$. Nach Satz 16.3 ist jedes $\beta \in L'$ von der Form $\beta = g(\alpha_r)$ mit $g \in M[X_r]$. Schreiben wir $g = \sum_j b_j X_r^j$

mit $b_j \in M$, so gibt es $g_j \in R'$ mit $b_j = g_j(\alpha_1, \ldots, \alpha_{r-1})$. Also ist $\beta = f(\alpha_1, \ldots, \alpha_r)$, wobei $f := \sum_j g_j X_r^j \in R'[X_r] = R$. Umgekehrt ist natürlich auch $f(\alpha_1, \ldots, \alpha_r) \in L'$ für alle $f \in R$. □

Folgerung 16.6. *Sei $M \subseteq L$ Teilkörper mit $K \subseteq M$. Sind $L \supseteq M$ und $M \supseteq K$ jeweils algebraische Erweiterungen, so ist auch die Erweiterung $L \supseteq K$ algebraisch.*

Beweis. Sei $\alpha \in L$ und $\mu_\alpha = \beta_0 + \beta_1 X + \ldots + \beta_{m-1}X^{m-1} + X^m \in M[X]$ (mit $m \geqslant 1$) das Minimalpolynom von α über M. Jedes $\beta_i \in M$ ist algebraisch über K. Nach Beispiel 16.5 ist $K_1 := K(\beta_0, \beta_1 \ldots, \beta_{m-1}) \subseteq L$ Teilkörper mit $[K_1 : K] < \infty$. Wegen $\mu_\alpha \in K_1[X]$ ist außerdem α algebraisch über K_1, also $[K_1(\alpha) : K_1] < \infty$ mach Satz 16.3. Mit dem Gradsatz folgt $[K_1(\alpha) : K] = [K_1(\alpha) : K_1][K_1 : K] < \infty$, also ist α algebraisch über K. □

Folgerung 16.7. *Sei $L \supseteq K$ eine Körpererweiterung und M die Menge aller $\alpha \in L$, die über K algebraisch sind. Dann ist M ein Teilkörper von L.*

Beweis. Seien $0 \neq \alpha, \beta \in M$. Sei $\gamma \in L$ eines der Elemente $\alpha \pm \beta$, $\alpha \cdot \beta$ und $\beta \cdot \alpha^{-1}$. Dann gilt $\gamma \in K(\alpha, \beta)$. Nach Beispiel 16.5 ist $[K(\alpha, \beta) : K] < \infty$ und damit γ algebraisch über K. □

Beispiel 16.8. Betrachte noch einmal $\alpha = i + \sqrt{2} \in \mathbb{C}$. Hier sind i, $\sqrt{2}$ jeweils algebraisch über $\mathbb{Q}$, mit Minimalpolynomen $\mu_i = X^2 + 1$ und $\mu_{\sqrt{2}} = X^2 - 2$. Also ist bereits nach Folgerung 16.7 klar (ohne weitere Rechnungen), dass α algebraisch über $\mathbb{Q}$ sein muss.

Sei $\mu_\alpha \in \mathbb{Q}[X]$ das Minimalpolynom. In Beispiel 15.8(c) haben wir bereits gesehen, dass $f(\alpha) = 0$ gilt für $f = X^4 - 2X^2 + 9$; also folgt $\mu_\alpha \mid f$. Um zu zeigen, dass $\mu_\alpha = f$ gilt, betrachten wir den Körper $L := \mathbb{Q}(\sqrt{2}, i)$. Wegen $i \notin \mathbb{Q}(\sqrt{2})$ ist μ_i weiterhin irreduzibel in $(\mathbb{Q}(\sqrt{2})[X]$; also ist μ_i auch das Minimalpolynom von i über $\mathbb{Q}(\sqrt{2})$. Also folgt $[L : \mathbb{Q}] = [L : \mathbb{Q}(\sqrt{2})][\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2 \cdot 2 = 4$ mit dem Gradsatz. Behauptung: Es gilt $\mathbb{Q}(i+\sqrt{2}) = \mathbb{Q}(\sqrt{2}, i)$.

Dazu: Offensichtlich ist $\alpha \notin \mathbb{Q}$ und damit $\mathbb{Q} \subsetneqq \mathbb{Q}(\alpha)$. Wäre auch $\mathbb{Q}(\alpha) \subsetneqq L$, so müsste wegen $[L : \mathbb{Q}] = 4$ nach dem Gradatz $[\mathbb{Q}(\alpha) : \mathbb{Q}] = 2$ gelten, d.h., das Minimalpolynom $\mu_\alpha \in \mathbb{Q}[X]$ hätte Grad 2. Sei $\mu_\alpha = X^2 + aX + b$ mit $a, b \in \mathbb{Q}$. Dann $\alpha^2 = -a\alpha - b$. Mit den Rechnungen in Beispiel 15.8(c) folgt $1 + 2i\sqrt{2} = \alpha^2 = (-a\sqrt{2} - b) - ai$, also $-a\sqrt{2} - b = 1$ und $a = -2\sqrt{2} \in \mathbb{Q}$, Widerspruch. Also ist $\mathbb{Q}(\alpha) = L$, wie behauptet. Wegen $[L : \mathbb{Q}] = 4$ folgt nun auch, dass $\mathrm{Grad}(\mu_\alpha) = 4$ ist, und damit $\mu_\alpha = f$.

Definition 16.9. Sei $L \supseteq K$ eine Körpererweiterung und $0 \neq f \in K[X]$ mit $n = \mathrm{Grad}(f) \geqslant 1$. Wir sagen, dass f über L in ***Linearfaktoren*** zerfällt, wenn gilt

$$f = c(X - \alpha_1) \cdots (X - \alpha_n) \qquad \text{mit} \qquad c \in K \quad \text{und} \quad \alpha_1, \ldots, \alpha_n \in L.$$

(Die α_i müssen hier nicht verschieden sein.) Beachte, dass jedes α_i algebraisch über K ist. Gilt außerdem $L = K(\alpha_1, \ldots, \alpha_n)$, so nennen wir L einen ***Zerfällungskörper*** von f.

Gibt es Indizes $i \neq j$ mit $\alpha_i = \alpha_j$, so heißt α_i eine ***mehrfache Nullstelle*** von f; dann ist also $(X - \alpha_i)^2$ ein Teiler von f in $L[X]$.

Beispiel 16.10. (a) Sei $f = X^2 + pX + q \in \mathbb{Q}[X]$. Durch quadratische Ergänzung erhalten wir $f = (X + p/2)^2 - \Delta/4$ mit $\Delta := p^2 - 4q$. In $\mathbb{C}$ können wir $\sqrt{\Delta}$ bilden. Es gilt sogar: Zu jedem beliebigen $z = a + bi \in \mathbb{C}$ (mit $i = \sqrt{-1}$ und $a, b \in \mathbb{R}$) gibt es eine explizite Formel für eine Quadratwurzel in $\mathbb{C}$, nämlich

$$z = \left(\sqrt{\tfrac{1}{2}\left(\sqrt{a^2 + b^2} + a\right)} \pm i\sqrt{\tfrac{1}{2}\left(\sqrt{a^2 + b^2} - a\right)} \right)^2,$$

wobei das Vorzeichen gleich "+" ist, falls $b \geqslant 0$, und gleich "−", falls $b < 0$. (Einfaches Nachrechnen.) Also ist $f = (X - \alpha_1)(X - \alpha_2)$ mit $\alpha_i = \frac{1}{2}(-p \pm \sqrt{\Delta}) \in \mathbb{C}$. Wegen $\alpha_1, \alpha_2 \in \mathbb{Q}(\sqrt{\Delta})$, $\alpha_1 - \alpha_2 = \sqrt{\Delta}$ ist $L = \mathbb{Q}(\alpha_1, \alpha_2) = \mathbb{Q}(\sqrt{\Delta})$ Zerfällungskörper von f, mit $[L : \mathbb{Q}] \leqslant 2$.

(b) Sei $f = X^4 - 2 \in \mathbb{Q}[X]$. Dann ist f irreduzibel (Eisenstein mit $p = 2$); wegen

$$f = X^4 - 2 = (X^2 - \sqrt{2})(X^2 + \sqrt{2}) = (X - \sqrt[4]{2})(X + \sqrt[4]{2})(X - i\sqrt[4]{2})(X + i\sqrt[4]{2})$$

ist $L = \mathbb{Q}(\pm\sqrt[4]{2}, \pm i\sqrt[4]{2}) \subseteq \mathbb{C}$ ein Zerfällungskörper von f. Was ist $[L : \mathbb{Q}]$? Dazu: Es gilt auch $L = \mathbb{Q}(\sqrt[4]{2}, i)$. Nun ist $[\mathbb{Q}(\sqrt[4]{2}) : \mathbb{Q}] = \mathrm{Grad}(f) = 4$ (da $f = \mu_{\sqrt[4]{2}}$, siehe Satz 16.3) und $i \notin \mathbb{Q}(\sqrt[4]{2})$. Also ist $X^2 + 1$ auch das Minimalpolynom von i über $\mathbb{Q}(\sqrt[4]{2})$. Mit dem Gradsatz folgt $[L : \mathbb{Q}] = [L : \mathbb{Q}(\sqrt[4]{2})][\mathbb{Q}(\sqrt[4]{2}) : \mathbb{Q}] = 2 \cdot 4 = 8$.

(c) Etwas allgemeiner: Sei $a \in \mathbb{Q}$, $a > 0$, fest und $f = X^n - a \in \mathbb{Q}[X]$. Sei $\alpha = \sqrt[n]{a} \in \mathbb{R}$ die positive n-te Wurzel aus a und $\zeta_n \in \mathbb{C}$ eine Einheitswurzel der Ordnung n. Dann ist $(\zeta_n^i \alpha)^n = \alpha^n = a$ für $1 \leqslant i \leqslant n$, also gilt $f = \prod_{i=1}^{n}(X - \zeta_n^i \alpha)$ und $L = \mathbb{Q}(\alpha, \zeta_n) \subseteq \mathbb{C}$ ist ein Zerfällungskörper von f. Nun ist $\Phi_n(\zeta_n) = 0$, also $[\mathbb{Q}(\zeta_n) : \mathbb{Q}] \leqslant \phi(n)$; außerdem $f(\alpha) = 0$, also $[L : \mathbb{Q}(\zeta_n)] \leqslant n$. Mit dem Gradsatz folgt $[L : \mathbb{Q}] = [L : \mathbb{Q}(\zeta_n)] \cdot [\mathbb{Q}(\zeta_n) : \mathbb{Q}] \leqslant n\phi(n)$.

Satz 16.11. *Sei* $0 \neq f \in K[X]$ *mit* $n = \mathrm{Grad}(f) \geqslant 1$. *Ist* $L \supseteq K$ *ein Zerfällungskörper von* f, *so gilt* $[L : K] \leqslant n!$. *Insbesondere ist* $L \supseteq K$ *eine algebraische Erweiterung.*

Beweis. Induktion nach n. Ist $n = 1$, so ist $f = X - a$ mit $a \in K$, also $L = K$ und damit $[L : K] = 1$. Sei nun $n > 1$ und $f = c(X - \alpha_1) \cdots (X - \alpha_n)$ mit $0 \neq c \in K$ und $\alpha_i \in L$. Sei $f_1 \in K[X]$ das Minimalpolynom von α_1. Wegen $f(\alpha_1) = 0$ folgt dann $f_1 \mid f$, also $\mathrm{Grad}(f_1) \leqslant n$. Nach Satz 16.3 ist $M := K(\alpha_1) \subseteq L$ ein Teilkörper mit $K \subseteq M$ und $[M : K] = \mathrm{Grad}(f_1) \leqslant n$.

Wir fassen nun f als Polynom in $M[X]$ auf. Wegen $\alpha_1 \in M$ und $f(\alpha_1) = 0$ folgt $f = (X - \alpha_1)g$ mit einem Polynom $g \in M[X]$, wobei $\mathrm{Grad}(g) = n - 1 \geqslant 1$ (siehe $(*)$ im Beweis von Lemma 12.7). Andererseits ist $g = c\prod_{i=2}^{n}(X - \alpha_i)$, also ist $L = M(\alpha_2, \ldots, \alpha_n) \supseteq M$ auch ein Zerfällungskörper von g. Mit Induktion folgt $[L : M] \leqslant (n-1)!$. Mit dem Gradsatz ergibt sich schließlich $[L : K] = [L : M][M : K] \leqslant (n-1)!n = n!$. □

17. *Konstruktion von Körpererweiterungen*

Wir stellen nun eine allgemeine Methode vor, um algebraische Erweiterungen zu konstruieren. (Dies ist ein weiteres Beispiel für die Verwendung von Faktorstrukturen wie in §2.) Sei zunächst R Integritätsring und $R[X]$ der zugehörige Polynomring. Sei $0 \neq f \in R[X]$ ein normiertes Polynom mit $\mathrm{Grad}(f) = n \geqslant 1$. Wir betrachten das Hauptideal (f) und bilden

$$R_f := R[X]/(f) = \{\bar{g} \mid g \in R[X]\} \qquad \text{wobei} \qquad \bar{g} := g + (f).$$

Sei $\pi\colon R[X] \to R_f$, $g \mapsto \bar{g}$, der kanonische Homomorphismus. Dann ist $\mathrm{Kern}(\pi|_R) = \{a \in R \mid a \in (f)\} = \{0\}$ weil $\mathrm{Grad}(f) = n \geqslant 1$ und R ein Integritätsring ist. Also ist $\pi|_R$ injektiv und wir können R vermöge dieses Homomorphismus als Teilring von R_f auffassen. D.h., für $a \in R$ schreiben wir einfach wieder $a \in R_f$ anstelle von $\bar{a}$. Sei $\alpha := \bar{X} \in R_f$. Dann haben wir den Einsetzungs-Homomorphismus, welcher $g \in R[X]$ auf $g(\alpha) \in R_f$ abbildet. Ist $g = \sum_{j=0}^{m} b_j X^j \in R[X]$, so gilt $g(\alpha) = \sum_{j=0}^{m} b_j \alpha^j = \sum_{j=0}^{m} \bar{b}_j \bar{X}^j = \bar{g}$. Dies zeigt $R_f = \{g(\alpha) \mid g \in R[X]\}$.

Lemma 17.1. *Mit obigen Bezeichnungen gilt* $R \subseteq R_f = \left\{\sum_{i=0}^{n-1} a_i \alpha^i \;\middle|\; a_i \in R\right\}$, *und die Darstellung der Elemente auf der rechten Seite ist eindeutig.*

Beweis. Wir haben bereits gesehen, dass $R_f = \{g(\alpha) \mid g \in R[X]\}$ gilt. Sei nun $g \in R[X]$. Wir schreiben $g = fq + r$ mit $q, r \in R[X]$, wobei entweder $r = 0$, oder $r \neq 0$ und $\mathrm{Grad}(r) < \mathrm{Grad}(f)$. Wegen $f(\alpha) = 0$ ist also $g(\alpha) = r(\alpha)$. Genauso wie im Beweis von Satz 16.3 folgt, dass sich jedes Element in R_f wie oben auf der rechten Seite schreiben lässt. Seien nun $a_i, b_i \in R$ mit $\sum_{i=0}^{n-1} a_i \alpha^i = \sum_{i=0}^{n-1} b_i \alpha^i$. Mit $c_i := a_i - b_i$ ist dann $0 = \sum_{i=0}^{n-1} c_i \alpha^i = \sum_{i=0}^{n-1} c_i \bar{X}^i = \overline{\sum_{i=0}^{n-1} c_i X^i}$, also $g := \sum_{i=0}^{n-1} c_i X^i \in \mathrm{Kern}(\pi) = (f)$ und damit $f \mid g$. Aus $g \neq 0$ würde dann $n = \mathrm{Grad}(f) \leqslant \mathrm{Grad}(g) \leqslant n - 1$ folgen, Widerspruch. Also ist $g = 0$, d.h., $c_i = 0$ und damit $a_i = b_i$ für alle i. □

Satz 17.2 (Kronecker)**.** *Sei* $R = K$ *ein Körper. Dann ist* K_f *ein* K*-Vektorraum mit* $\dim K_f = \mathrm{Grad}(f) = n$ *und Basis* $\{1, \alpha, \ldots, \alpha^{n-1}\}$. *Ist außerdem* f *irreduzibel, so ist* $K_f = K(\alpha)$ *selbst ein Körper und* $K_f \supseteq K$ *eine Körpererweiterung mit* $[K_f : K] = n$ *und* $f(\alpha) = 0$.

Beweis. Nach obiger Diskussion ist K ein Teilring von K_f. Da K ein Körper ist, ist damit K_f ein K-Vektorraum. Jedes Element von K_f lässt sich eindeutig schreiben als $\sum_{i=0}^{n-1} a_i \alpha^i$ mit $a_i \in K$. Dies zeigt, dass $\{1, \alpha, \ldots, \alpha^{n-1}\}$ eine K-Basis von K_f ist.
Nun ist $K[X]$ ein Hauptidealring. Ist also f irreduzibel, so ist K_f ein Körper nach Satz 11.9(a). Es gilt $f(\alpha) = 0$, weil $f(\alpha) = \bar{f} = 0$. Also ist α algebraisch über K, mit $\mu_\alpha = f$. Wegen $K_f = \{g(\alpha) \mid g \in K[X]\}$ (siehe oben) folgt $K_f = K(\alpha)$. □

Beispiel 17.3. (a) Sei $K = \mathbb{R}$ und $f = X^2 + 1 \in \mathbb{R}[X]$. Dann hat f keine Nullstelle in $\mathbb{R}$, also ist f irreduzibel. Nach obigem Verfahren ist also $\mathbb{R}_f = \mathbb{R}[X]/(f)$ ein Körper mit $\mathbb{R} \subseteq \mathbb{R}_f$ und

$f(\alpha) = 0$, d.h., $\alpha^2 = -1$. Dies ist die formale algebraische Konstruktion des Körpers $\mathbb{C}$. Wir haben nämlich den Einsetzungs-Homomorphismus

$$\mathbb{R}[X] \to \mathbb{C}, \qquad g \mapsto g(i) \qquad (\text{wobei } i = \sqrt{-1} \in \mathbb{C}).$$

Der Kern ist gerade $(X^2 + 1)$ also gilt $\mathbb{R}_f = \mathbb{R}[X]/(f) \cong \mathbb{C}$ nach dem Homomorphiesatz 12.5.

(b) Sei $p = 2$ und $f = X^2+X+1 \in \mathbb{F}_2[X]$. Dann ist f irreduzibel (weil es keine Nullstelle in $\mathbb{F}_2$ gibt), also ist $K := (\mathbb{F}_2)_f$ ein Körper. Sei $\alpha := X+(f)$. Dann gilt also $K = \{a+b\alpha \mid a, b \in \mathbb{F}_2\}$, d.h., K ist ein Körper mit 4 Elementen. Es gilt $\alpha^2+\alpha+1 = 0$ und mit dieser Relation können wir in K rechnen, also zum Beispiel ein Produkt $(a + b\alpha)(a' + b'\alpha)$ wieder in der Form $c + d\alpha$ mit $c, d \in \mathbb{F}_2$ darstellen. Etwas effizienter ist folgende Darstellungsweise. Betrachte die multiplikative Gruppe $K^\times$; es gilt $|K^\times| = 3$. Wegen $\alpha \neq 1$ und $\alpha^2 = \alpha+1 \neq 1$ ist $o(\alpha) > 2$, also $K^\times = \langle\alpha\rangle$, d.h., $K = \{0, 1, \alpha, \alpha^2\}$. Damit erhalten wir folgende Verknüpfungstabellen:

$+$	0	1	α	α^2
0	0	1	α	α^2
1	1	0	α^2	α
α	α	α^2	0	1
α^2	α^2	α	1	0

$\cdot$	0	1	α	α^2
0	0	0	0	0
1	0	1	α	α^2
α	0	α	α^2	1
α^2	0	α^2	1	α

Ab hier Woche 12

(c) Allgemeiner: Sei p eine Primzahl und $n \geqslant 1$ beliebig. Nehmen wir an, $0 \neq f \in \mathbb{F}_p[X]$ ist ein normiertes, irreduzibles Polynom mit $\mathrm{Grad}(f) = n$. Dann ist $K := (\mathbb{F}_p)_f \supseteq \mathbb{F}_p$ nach Satz 17.2 ein Körper mit $[K : \mathbb{F}_p] = n$. Wie in Beispiel 16.4 folgt $|K| = p^n$. Problem: Gibt es ein solches f ? (Dies ist nicht offensichtlich!)

Satz 17.4. *Sei K ein Körper und $0 \neq f \in K[X]$ nicht-konstant; sei $n = \mathrm{Grad}(f) \geqslant 1$. Dann gibt es einen Zerfällungskörper $L \supseteq K$ von f und es gilt $[L : K] \leqslant n!$.*

Beweis. Induktion nach n. Für $n = 1$ ist $f = X - a$ mit $a \in K$, also ist $L = K$ Zerfällungskörper. Sei nun $n > 1$ und $f_1 \in K[X]$ normiert, irreduzibel mit $f_1 \mid f$. Nach Satz 17.2 ist $K_1 := K_{f_1} \supseteq K$ eine Erweiterung, so dass f_1 (und damit f) zumindest eine Nullstelle $\alpha_1 \in K_1$ hat. Wir fassen f als Polynom in $K_1[X]$ auf und schreiben $f = (X-\alpha_1)g$ mit einem $g \in K_1[X]$. Wegen $\mathrm{Grad}(g) = n - 1 \geqslant 1$ gibt es nach Induktion einen Zerfällungskörper $L_1 \supseteq K_1$ von g. Offenbar zerfällt f auch in L_1 in Linearfaktoren. Betrachte nun den Teilkörper $L \subseteq L_1$, der durch Adjunktion der Nullstellen von f an K entsteht. Dann ist L ein Zerfällungskörper, und nach Satz 16.11 gilt $[L : K] \leqslant n!$. □

Es gibt einen einfachen Trick, um einem Polynom in $K[X]$ direkt anzusehen, ob es mehrfache Nullstellen in einem Erweiterungskörper $L \supseteq K$ gibt oder nicht. In Analogie zur Analysis definieren wir dazu wie folgt eine ***formale Ableitung*** für Polynome in $K[X]$:

$$D\colon K[X] \to K[X], \qquad f = \sum_{i=0}^{n} a_i X^i \quad \mapsto \quad D(f) := \sum_{i=1}^{n} i a_i X^{i-1}.$$

(Hierbei ist ia für ein $a \in K$ als $a + a + \cdots + a$ mit i Summanden zu verstehen.) Folgende Eigenschaften lassen sich leicht durch direktes Nachrechnen zeigen:

(1) D ist K-linear, also $D(af + bg) = aD(f) + bD(g)$ für alle $a, b \in K$ und $f, g \in K[X]$.
(2) Es gilt die Produkt-Regel $D(fg) = fD(g) + D(f)g$ für alle $f, g \in K[X]$.
(3) Es gilt $D(f^n) = nD(f)f^{n-1}$ für alle $f \in K[X]$. (Wiederholte Anwendung von (2).)

Anwendung: Seien $0 \neq f, g \in K[X]$ mit $g \mid f$. Wir sagen, dass g ein ***mehrfacher Faktor*** von f ist, wenn $g^2 \mid f$ gilt. Solche mehrfachen Faktoren können wie folgt gefunden werden. Ist $f = g^2 h$ mit $h \in K[X]$, so folgt $D(f) = 2gD(g)h + g^2 D(h) = g(2D(g)h + gD(h))$. Also ist $g \mid f$ und $g \mid D(f)$. Zum Beispiel kann $f = X^4 + 1 \in \mathbb{Z}[X]$ keine echten mehrfachen Faktoren enthalten, weil f und $D(f) = 4X^3$ überhaupt keine gemeinsamen Faktoren außer ± 1 haben.

Lemma 17.5. *Sei $0 \neq f \in K[X]$ nicht-konstant und $L \supseteq K$ eine Körpererweiterung.*
(a) *Haben f und $D(f)$ keinen gemeinsamen nicht-konstanten Teiler in $K[X]$, so hat f keine mehrfachen Nullstellen in L.*
(b) *Ist f irreduzibel und $D(f) \neq 0$, so hat f keine mehrfachen Nullstellen in L.*

Beweis. (a) Weil $K[X]$ ein Hauptidealring ist, gilt das Lemma von Bézout; siehe Bemerkung 11.8. Also gibt es ein $0 \neq d \in K[X]$, welches f und $D(f)$ teilt, sowie $g, h \in K[X]$ mit $d = gf + hD(f)$. Nach Voraussetzung in (a) ist $\mathrm{Grad}(d) = 0$. Die Gleichung $d = gf + hD(f)$ können wir natürlich auch als Gleichung von Polynomen in $L[X]$ auffassen. Angenommen, es gäbe ein $\alpha \in L$ mit $(X - \alpha)^2 \mid f$ in $L[X]$, d.h., $X - \alpha$ ist ein mehrfacher Faktor von f in $L[X]$. Wie oben gesehen, ist dann aber $X - \alpha$ ein gemeinsamer Teiler von f und $D(f)$ in $L[X]$. Wegen $d = gf + hD(f)$ folgt dann auch $(X - \alpha) \mid d$, Widerspruch zu $\mathrm{Grad}(d) = 0$.

(b) Ist $D(f) \neq 0$, so ist $\mathrm{Grad}(D(f)) < \mathrm{Grad}(f)$. Ist f irreduzibel, so kann es keinen nicht-konstanten gemeinsamen Teiler von f und $D(f)$ geben, also folgt die Behauptung mit (a). □

Bemerkung 17.6. Sei $f \in K[X]$ nicht-konstant, irreduzibel. Ist $\mathrm{char}(K) = 0$, so ist stets $D(f) \neq 0$; also hat f keine mehrfachen Nullstellen. Ist dagegen $\mathrm{char}(K) = p > 0$, so kann es passieren, dass f mehrfache Nullstellen hat. Beispiel: Sei $K = \mathbb{F}_p(t)$ mit einer Unbestimmten t über $\mathbb{F}_p$. Dann ist $f = X^p - t \in K[X]$ irreduzibel (Eisenstein-Kriterium bezüglich des Primelementes $t \in \mathbb{F}_p[t]$). Sei $L \supseteq K$ ein Zerfällungskörper von f; es gibt also ein $\alpha \in L$ mit $f(\alpha) = 0$. Dann ist $\alpha^p = t$ und $f = X^p - \alpha^p = (X - \alpha)^p$, wobei wir Lemma 15.2 benutzen. Also ist f zwar irreduzibel, hat aber eine mehrfache Nullstelle in L. Beachte: Dies ist kein Widerspruch zu Lemma 17.5(b), denn hier haben wir $D(f) = pX^{p-1} = 0$.

Mit Hilfe der obigen Konstruktionen, und kombiniert mit Beispiel 16.4, erhalten wir nun einen vollständigen Überblick über alle endlichen Körper.

Satz 17.7 (Hauptsatz über endliche Körper)**.** *Sei* $n \geqslant 1$ *und* p *eine Primzahl. Dann gibt es einen endlichen Körper* K *mit* $|K| = p^n$*, und alle diese Körper sind zueinander isomorph. Konkret erhält man* K *als* $\mathbb{F}_p[X]/(f)$*, mit* $f \in \mathbb{F}_p[X]$ *normiert, irreduzibel und* $\mathrm{Grad}(f) = n$*.*

Beweis. Nach Satz 17.4 gibt es einen Zerfällungskörper $L \supseteq \mathbb{F}_p$ des Polynoms $\Phi := X^{p^n} - X \in \mathbb{F}_p[X]$. Sei $K := \{z \in L \mid \Phi(z) = 0\}$ die Menge der Nullstellen in L. Wegen $D(\Phi) = p^n X^{p^n-1} - 1 = -1$ haben Φ und $D(\Phi)$ keine nicht-konstanten gemeinsamen Teiler. Nach Lemma 17.5 hat Φ in L also keine mehrfachen Nullstellen. Wegen $\mathrm{Grad}(\Phi) = p^n$ folgt $|K| = p^n$. Behauptung: $K \subseteq L$ ist ein Teilkörper. Dazu: Es ist $\Phi(0) = 0$, also $0 \in K$. Sind $\alpha, \beta \in K$, so folgt $\alpha^{p^n} = \alpha$ und $\beta^{p^n} = \beta$, also auch $(\alpha\beta)^{p^n} = \alpha^{p^n}\beta^{p^n} = \alpha\beta$ und damit $\alpha\beta \in K$. Weiterhin gilt $(-\alpha)^{p^n} = (-1)^{p^n}\alpha^{p^n} = (-1)^{p^n}\alpha$. Ist $p > 2$, so ist $(-1)^{p^n} = -1$, also folgt $-\alpha \in K$. Ist $p = 2$, so spielt das Vorzeichen keine Rolle. Für $\alpha \neq 0$ folgt außerdem $(\alpha^{-1})^{p^n} = \alpha^{-1}$ und damit $\alpha^{-1} \in K$. Es bleibt noch zu zeigen: $\alpha + \beta \in K$ für alle $\alpha, \beta \in K$. Dazu verwenden wir den Homomorphismus $F\colon L \to L$, $z \mapsto z^p$; siehe Lemma 15.2. Es gilt $(\alpha + \beta)^p = F(\alpha + \beta) = F(\alpha) + F(\beta) = \alpha^p + \beta^p$. Durch wiederholte Anwendung folgt $(\alpha+\beta)^{p^n} = \alpha^{p^n} + \beta^{p^n} = \alpha + \beta \in K$. Also ist in der Tat $K \subseteq L$ ein Teilkörper mit $|K| = p^n$. (Weil K alle Nullstellen von Φ enthält, gilt sogar $K = L$.) Mit diesem abstrakten Argument ist jedenfalls gezeigt, dass es einen Körper K mit p^n Elementen gibt. Nach Beispiel 16.4 gibt es ein $0 \neq \alpha \in K$ mit $K = \mathbb{F}_p(\alpha)$, wobei das Minimalpolynom $f_0 := \mu_\alpha \in \mathbb{F}_p[X]$ den Grad n hat. Wegen $\Phi(\alpha) = 0$ folgt dann $f_0 = \mu_\alpha \mid \Phi$.

Sei K' ein beliebiger Körper mit $|K'| = p^n$; dann ist wieder $\mathbb{F}_p \subseteq K'$ und $[K' : \mathbb{F}_p] = n$. Nach Lagrange (genauer: Folgerung 3.6) gilt $\beta^{p^n-1} = 1_{K'}$ für alle $\beta \in K'^\times$, und damit $\Phi(\beta) = \beta^{p^n} - \beta = 0$; dies gilt auch für $\beta = 0$. Also sind alle Elemente von K' Nullstellen von Φ; wegen $\mathrm{Grad}(\Phi) = p^n$ folgt $\Phi = \prod_{\beta \in K'}(X - \beta)$. Nun ist $f_0 \mid \Phi$, also gibt es ein $\beta \in K'$ mit $f_0(\beta) = 0$; damit ist f_0 auch das Minimalpolynom von β über $\mathbb{F}_p$. Nach Satz 16.3 ist $\mathbb{F}_p(\beta) \subseteq K'$ ein Teilkörper mit $[\mathbb{F}_p(\beta) : \mathbb{F}_p] = \mathrm{Grad}(\mu_\alpha) = n$. Wie in Beispiel 16.4 folgt $|\mathbb{F}_p(\beta)| = p^n$, also $K' = \mathbb{F}_p(\beta)$. Betrachte den Einsetzungs-Homomorphismus $\varphi\colon \mathbb{F}_p[X] \to K'$, $g \mapsto g(\beta)$. Wegen $K' = \mathbb{F}_p(\beta)$ ist dieser surjektiv; da f_0 das Minimalpolynom von β ist, folgt $\mathrm{Kern}(\varphi) = (f_0)$; mit Satz 12.5 dann auch $\mathbb{F}_p[X]/(f_0) \cong K'$. Damit ist gezeigt, dass jeder endliche Körper K' mit $|K'| = p^n$ isomorph zu dem festen Körper $\mathbb{F}_p[X]/(f_0)$ ist. □

Beispiel 17.8. Sei $n \geqslant 1$ und p eine Primzahl. Um einen endlichen Körper mit p^n Elementen zu konstruieren, müssen wir also "nur" ein normiertes irreduzibles Polynom $f \in \mathbb{F}_p[X]$ mit $\mathrm{Grad}(f) = n$ finden. Der obige Beweis zeigt, dass ein solches f unter den Teilern von $X^{p^n} - X \in \mathbb{F}_p[X]$ zu finden ist. In Beispiel 17.3(b) haben wir bereits den Fall $n = p = 2$ betrachtet: Dort

gilt $X^{2^2} - X = X(X^3 - \bar{1}) = X(X - \bar{1})(X^2 + X + \bar{1})$, und $f = X^2 + X + \bar{1} \in \mathbb{F}_2[X]$ ist irreduzibel. Sei nun $n = 2$ und $p = 3$. Dann ist

$$X^{3^2} - X = X(X^8 - \bar{1}) = X(X^4 - \bar{1})(X^4 + \bar{1}) = X(X - \bar{1})(X + \bar{1})(X^2 + \bar{1})(X^4 + \bar{1}) \in \mathbb{F}_3[X].$$

Hier ist $X^2 + 1$ irreduzibel (keine Nullstellen in $\mathbb{F}_3$). Außerdem stellt man fest, dass $X^4 + \bar{1} = (X^2 + X - \bar{1})(X^2 - X - \bar{1})$ gilt, wobei $X^2 + X - \bar{1}$ und $X^2 - X - \bar{1}$ irreduzibel sind (wiederum keine Nullstellen in $\mathbb{F}_3$). Wir können also einen endlichen Körper mit 9 Elementen bilden als $K = \mathbb{F}_3[X]/(f)$, wobei $f \in \{X^2 + \bar{1}, X^2 + X - \bar{1}, X^2 - X - \bar{1}\}$. Gibt es Wahlen, die sinnvoller sind als andere? Um übersichtliche Verknüpfungstabellen wie in Beispiel 17.3(b) zu erhalten, erscheint es sinnvoll zu sein, f so zu wählen, dass $\alpha = X + (f) \in K = \mathbb{F}_3[X]/(f)$ die multiplikative Gruppe von $K^\times$ erzeugt[8]. Für $f = X^2 + \bar{1}$ folgt aber $\alpha^2 = -\bar{1}$, also $\alpha^4 = \bar{1}$ und damit $\langle\alpha\rangle \subsetneq K^\times$. Für $f = X^2 \pm X - \bar{1}$ gilt dagegen $K^\times = \langle\alpha\rangle$. Mehr dazu in den Übungen.

Sei $K[X]^\#$ die Menge der nicht konstanten Polynome in $K[X]$. In Satz 17.4 haben wir gesehen, dass es zu jedem $f \in K[X]^\#$ eine algebraische Erweiterung $L \supseteq K$ gibt, so dass f über L in Linearfaktoren zerfällt. Dann gibt es auch zu jeder endlichen Teilmenge $S \subseteq K[X]^\#$ eine algebraische Erweiterung $L \supseteq K$, so dass alle $f \in S$ über L in Linearfaktoren zerfallen. (Nehme den Zerfällungskörper des Produkts der Polynome in S.) Es stellt sich die Frage, ob dies auch für größere Teilmengen S, und schließlich für $S = K[X]^\#$ selbst möglich ist.

Definition 17.9. Ein Körper K heißt ***algebraisch abgeschlossen***, wenn jedes nicht-konstante Polynom in $K[X]$ über K in Linearfaktoren zerfällt. Ist K selbst nicht algebraisch abgeschlossen, so heißt ein Körper L mit $L \supseteq K$ ein ***algebraischer Abschluss*** von K, wenn L algebraisch abgeschlossen und die Erweiterung $L \supseteq K$ algebraisch ist.

Allgemein kann mit Hilfe des ***Lemmas von Zorn*** gezeigt werden, dass es zu jedem beliebigen K einen algebraischen Abschluss $L \supseteq K$ gibt (***Satz von Steinitz***). Wir werden dies am Ende dieses Skripts in einem Anhang B weiter diskutieren; insbesondere werden wir dort Fälle behandeln, wo man das Lemma von Zorn nicht benötigt.

18. *Eine Anwendung: Konstruktion mit Zirkel und Lineal*

Die folgenden Fragestellungen sind klassische Probleme der antiken Mathematik:

- Ist die Quadratur des Kreises möglich? (Also: Kann man aus einem gegebenen Kreis ein Quadrat mit gleichem Flächeninhalt konstruieren?).
- Kann man jeden gegebenen Winkel in drei gleich große Teile unterteilen?

[8] Auf diese Weise werden zum Beispiel endliche Körper in Computer-Algebra-Systemen, etwa GAP [GAP], dargestellt: Jedes Element ist entweder 0 oder Potenz eines Erzeugers der multiplikativen Gruppe.

- Würfelverdoppelung? (Also: Kann man aus einem gegebenen Würfel einen Würfel mit doppeltem Volumen konstruieren?)

Erlaubte Hilfsmittel sind hierbei nur die "Euklidischen Werkzeuge" Zirkel und Lineal; siehe auch `https://de.wikipedia.org/wiki/Konstruktion_mit_Zirkel_und_Lineal`.
Eine Lösung dieser Probleme erfolgte erst im 19. Jahrhundert, durch Gauß, Galois, Wantzel sowie Lindemanns Beweis der Transzendenz von π. — Dies ist ein Musterbeispiel für die Anwendung algebraischer Methoden auf geometrische Fragestellungen.

Wir präzisieren nun, was genau "Konstruktion mit Zirkel und Lineal" bedeutet. Alles spielt sich in der reellen Ebene $\mathbb{R}^2$ ab. Gegeben sei eine Teilmenge $P_0 \subseteq \mathbb{R}^2$; wir nehmen stets an, dass $(0,0)$ und $(1,0)$ zu P_0 gehören (als Startpunkte). Dann gibt es die beiden folgenden elementaren Konstruktionsschritte, wobei nur Zirkel und Lineal benutzt werden:

(L) Durch zwei verschiedene Punkte von P_0 ziehe eine Gerade.

(Z) Schlage einen Kreis um einen Punkt aus P_0 als Mittelpunkt, wobei der Radius der Abstand zweier verschiedener Punkte aus P_0 ist.

(Das Lineal enthält dabei keine Markierungen von Maßeinheiten; es wird nur benutzt, um eine gerade Linie zu ziehen. Abstände wie in (Z) werden also nicht gemessen, sondern mit dem Zirkel abgegri en.)

Definition 18.1. Sei $P_0 \subseteq \mathbb{R}^2$ gegeben, wobei $(0,0) \in P_0$ und $(1,0) \in P_0$. Sei $(x,y) \in \mathbb{R}^2$.
(a) (x,y) heißt elementar aus P_0 konstruierbar, wenn (x,y) der Schnittpunkt von zwei verschiedenen Graden wie in (L) ist, oder ein Schnittpunkt einer Geraden wie in (L) und eines Kreises wie in (Z), oder ein Schnittpunkt von zwei verschiedenen Kreisen wie in (Z).

(b) (x,y) heißt mit Zirkel und Lineal aus P_0 konstruierbar, wenn es ein $n \geqslant 1$ und Punkte $(x_1,y_1),\dots,(x_n,y_n) = (x,y)$ in $\mathbb{R}^2$ gibt, so dass jedes (x_i,y_i) für $1 \leqslant i \leqslant n$ rekursiv aus $P_0 \cup \{(x_1,y_1),\dots,(x_{i-1},y_{i-1})\}$ elementar konstruierbar ist.

Bemerkung: Man sieht leicht, dass $(x,y) \in \mathbb{R}^2$ genau dann mit Zirkel und Lineal aus P_0 konstruierbar ist, wenn $(x,0)$ und $(0,y)$ aus P_0 konstruierbar sind.

Lemma 18.2. *Sei* $P_0 \subseteq \mathbb{R}^2$ *wie oben und* $(x,y) \in \mathbb{R}^2$ *elementar aus* P_0 *konstruierbar. Sei* $K_0 = \mathbb{Q}(u,v \mid (u,v) \in P_0) \subseteq \mathbb{R}$. *Dann gilt* $[K_0(x) : K_0] \leqslant 2$ *und* $[K_0(y) : K_0] \leqslant 2$.

Beweis. Vorbetrachtung: Sind zwei verschiedene Punkte $(x_1,y_1) \in P_0$ und $(x_2,y_2) \in P_0$ gegeben, so ist die Gerade durch diese beiden Punkte gegeben durch

$$G = \{(u,v) \in \mathbb{R}^2 \mid au + bv = c\}, \quad \text{wobei } a = y_1 - y_2,\ b = x_2 - x_1,\ c = x_2y_1 - x_1y_2;$$

beachte $a,b,c \in K_0$, sowie $a \neq 0$ oder $b \neq 0$. Ist $(x_0,y_0) \in P_0$ und r der Abstand zwischen (x_1,y_1) und (x_2,y_2), so ist der Kreis mit Mittelpunkt (x_0,y_0) und Radius r gegeben durch

$$\{(u,v) \in \mathbb{R}^2 \mid (u-x_0)^2 + (v-y_0)^2 = r^2\}, \quad \text{wobei } r^2 = (x_2-x_1)^2 + (y_2-y_1)^2 \in K_0.$$

Nun betrachten wir die drei möglichen Konstruktionsschritte für (x, y).

1. Fall: (x, y) ist der Schnittpunkt von zwei verschiedenen Geraden wie in (L). Diese beiden Geraden sind also gegeben durch $G_i = \{(u, v) \in \mathbb{R}^2 \mid a_i u + b_i v = c_i\}$ für $i = 1, 2$, wobei $a_i, b_i, c_i \in K_0$. Wegen $G_1 \neq G_2$ gibt es genau einen Schnittpunkt, den man als Lösung des linearen Gleichungssystems $a_1 x + b_1 y = c_1$, $a_2 x + b_2 y = c_2$ erhält. Hier ist also $x, y \in K_0$.

2. Fall: (x, y) ist ein Schnittpunkt einer Geraden wie in (L) und eines Kreises wie in (Z). Dann erfüllt (x, y) also zwei Gleichungen der Form

$$ax + by = c \qquad \text{und} \qquad (x - x_0)^2 + (y - y_0)^2 = r^2,$$

wobei $a, b, c \in K_0$, $x_0, y_0 \in K_0$ und $r^2 \in K_0$. Sei zuerst $b \neq 0$. Dann ist $y = b^{-1}c - b^{-1}ax$ und wir erhalten $(x - x_0)^2 + (b^{-1}c - b^{-1}ax - y_0)^2 = r^2 \in K_0$. Multiplizieren wir die linke Seite aus, so erhalten wir $\alpha x^2 + \beta x + \gamma = r^2$ mit $\alpha, \beta, \gamma \in K_0$; dabei ist $\alpha = 1 + b^{-2}a^2 \neq 0$. Also ist x Lösung einer quadratischen Gleichung mit Koeffizienten in K_0 und damit $[K_0(x) : K] \leqslant 2$. Aus $y = b^{-1}c - b^{-1}ax$ folgt dann auch $y \in K_0(x)$, also $[K_0(y) : K_0] \leqslant 2$. Ist $b = 0$, so muss $a \neq 0$ gelten und man kann völlig analog argumentieren.

3. Fall: (x, y) ist ein Schnittpunkt von zwei verschiedenen Kreisen wie in (Z). Dann erfüllt (x, y) also zwei Gleichungen der Form

$$(x - x_1)^2 + (y - y_1)^2 = r_1^2 \qquad \text{und} \qquad (x - x_2)^2 + (y - y_2)^2 = r_2^2,$$

wobei $x_1, x_2, y_2, y_2 \in K_0$ und $r_1^2, r_2^2 \in K_0$; dabei ist $(x_1, y_1) \neq (x_2, y_2)$. Multiplizieren wir die obigen Gleichungen aus und subtrahieren sie voneinander, so erhalten wir

$$2(x_2 - x_1)x + 2(y_2 - y_1)y = r_1^2 - r_2^2 + y_2^2 - y_1^2 + x_2^2 - x_1^2,$$

also die Gleichung einer Geraden über K_0. Damit ist das Problem auf den 2. Fall zurückgeführt; es folgt also wieder $[K_0(x) : K_0] \leqslant 2$ und $[K_0(y) : K_0] \leqslant 2$. □

Satz 18.3. *Seien* $P_0 \subseteq \mathbb{R}^2$ *und* $K_0 = \mathbb{Q}(u, v \mid (u, v) \in P_0) \subseteq \mathbb{R}$ *wie oben. Ist* $(x, y) \in \mathbb{R}^2$ *aus* P_0 *konstruierbar, so sind die Körpergrade* $[K_0(x) : K_0]$ *und* $[K_0(y) : K_0]$ *Potenzen von* 2.

Beweis. Nach Definition 18.1 gibt es ein $n \geqslant 1$ und $(x_1, y_1), \ldots, (x_n, y_n) = (x, y)$ in $\mathbb{R}^2$, so dass jedes (x_i, y_i) für $1 \leqslant i \leqslant n$ aus $P_{i-1} := P_0 \cup \{(x_1, y_1), \ldots, (x_{i-1}, y_{i-1})\}$ elementar konstruierbar ist. Setze $K_1 := K_0(x_1, y_1)$, $K_2 := K_1(x_2, y_2)$, …, $K_n := K_{n-1}(x_n, y_n)$. Dann ist

$$K_{i-1} = \mathbb{Q}(u, v \mid (u, v) \in P_{i-1}) \qquad \text{für } 1 \leqslant i \leqslant n.$$

Nach Lemma 18.2 folgt also $[K_{i-1}(x_i) : K_{i-1}] \leqslant 2$ und $[K_{i-1}(y_i) : K_{i-1}] \leqslant 2$. Insbesondere ist y_i Nullstelle eines Polynoms vom Grad $\leqslant 2$ mit Koeffizienten in K_{i-1}. Fassen wir dieses Polynom

als Polynom mit Koeffizienten in $K_{i-1}(x_i)$ auf, so folgt also auch $[K_{i-1}(x_i, y_i) : K_{i-1}(x_i)] \leqslant 2$. Mit dem Gradsatz erhalten wir dann

$$[K_i : K_{i-1}] = [K_{i-1}(x_i, y_i) : K_{i-1}] = [K_{i-1}(x_i, y_i) : K_{i-1}(x_i)] \cdot [K_{i-1}(x_i) : K_{i-1}],$$

also $[K_i : K_{i-1}] \in \{1, 2, 4\}$. Wieder mit dem Gradsatz folgt $[K_n : K_0] = [K_n : K_{n-1}] \cdots [K_1 : K_0] =$ Potenz von 2. Schließlich ist $[K_n : K_0] = [K_n : K_0(x)][K_0(x) : K_0]$, also auch $[K_0(x) : K_0]$ eine Potenz von 2; analog ebenso $[K_0(y) : K_0]$. □

Beispiel 18.4. Sei jeweils $P_0 = \{(0,0), (1,0)\}$.
(a) Die ***Quadratur des Kreises*** ist unmöglich. Genauer: Gegeben sei ein Kreis mit Mittelpunkt $(0,0)$ und Radius 1, also genau dem Abstand der beiden Punkte in P_0. Ist es dann möglich, ausgehend von P_0 die Eckpunkte $(0,0), (a,0), (0,a), (a,a)$ (mit $a > 0$) eines Quadrats mit Zirkel und Lineal zu konstruieren, dessen Flächeninhalt gleich dem Flächeninhalt des Kreises, also π, ist? Wenn ja, dann wäre $a = \sqrt{\pi}$ und mit Satz 18.3 folgt, dass $[\mathbb{Q}(\sqrt{\pi}) : \mathbb{Q}]$ eine 2-Potenz ist. Insbesondere müsste $\sqrt{\pi}$ algebraisch sein, und damit auch $\pi = \sqrt{\pi}^2$, Widerspruch zum Satz von Lindemann (Transzendenz von π).

(b) Die ***Würfelverdoppelung*** (auch Delisches Problem genannt) ist unmöglich. Genauer: Gegeben sei ein Würfel mit Kantenlänge 1, also genau dem Abstand der beiden Punkte in P_0. Ist es dann möglich, die Kantenlänge eines Würfels mit doppeltem Volumen zu konstruieren, also ausgehend von P_0 den Punkt $(\sqrt[3]{2}, 0) \in \mathbb{R}^2$? Wenn ja, so müsste $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}]$ eine 2-Potenz sein; aber mit dem Eisenstein-Kriterium (für $p = 2$) folgt, dass $X^3 - 2 \in \mathbb{Q}[X]$ das Minimalpolynom von $\sqrt[3]{2}$ ist, also $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$, Widerspruch.

(c) Die ***Winkeldreiteilung*** wird in den Übungen behandelt.

Betrachte für $n \geqslant 3$ die Punkte $z_{n,k} := \big(\cos(2\pi k/n), \sin(2\pi k/n)\big) \in \mathbb{R}^2$ für $0 \leqslant k \leqslant n-1$. Diese Punkte liegen auf dem Kreis mit $(0,0)$ als Mittelpunkt und Radius 1, und teilen diesen Kreis in n gleiche Abschnitte ab. Wir erhalten also ein ***regelmäßiges n-Eck*** mit Eckpunkten $z_{n,0}, z_{n,1}, \ldots, z_{n,n-1}$. Lassen sich diese mit Zirkel und Lineal konstruieren? Dazu zeigen wir zuerst den folgenden Satz, der unabhängig von Zirkel und Lineal interessant ist.

Satz 18.5 (Gauß, Kronecker)**.** *Sei $n \geqslant 1$. Dann ist das n-te Kreisteilungspolynom $\Phi_n \in \mathbb{Q}[X]$ irreduzibel. Also ist Φ_n das Minimalpolynom (über $\mathbb{Q}$) der Einheitswurzel $\zeta_n \in \mathbb{C}$.*

Ab hier Woche 13

Beweis. Nach Satz 13.10(a) genügt es zu zeigen, dass Φ_n irreduzibel in $\mathbb{Z}[X]$ ist. Sei $\Phi_n = fg$ mit $f, g \in \mathbb{Z}[X]$ wobei f irreduzibel, $\mathrm{Grad}(f) \geqslant 1$ und $f(\zeta_n) = 0$ gilt. Weil Φ_n normiert ist, können wir auch annehmen, dass f, g normiert sind. Wiederum nach Satz 13.10(a) ist f auch irreduzibel in $\mathbb{Q}[X]$. Wir behaupten nun, mit Bezeichnungen wie in Definition 12.10:

(∗) Ist $\alpha \in E_n$ mit $f(\alpha) = 0$ und p Primzahl mit $p \nmid n$, so gilt auch $f(\alpha^p) = 0$.

Zunächst beachte: Wegen $f \mid \Phi_n$ ist $\Phi_n(\alpha) = 0$, also $\langle\alpha\rangle = E_n$. Und wegen $p \nmid n$ gilt auch $\langle\alpha^p\rangle = E_n$, d.h., $\Phi_n(\alpha^p) = 0$. Wäre also $f(\alpha^p) \neq 0$, dann müsste $g(\alpha^p) = 0$ gelten, d.h., α ist eine Nullstelle des Polynoms $\tilde{g} := g(X^p) \in \mathbb{Z}[X]$. Wir müssen dies nun zu einem Widerspruch führen. Dazu: Da f normiert und irreduzibel ist, gilt $\mu_\alpha = f$; also folgt $f \mid \tilde{g}$ (in $\mathbb{Q}[X]$) mit Lemma 15.7. Es gibt also ein $h \in \mathbb{Q}[X]$ mit $\tilde{g} = fh$. Sei $0 \neq c \in \mathbb{Z}$ so, dass $\tilde{h} := ch \in \mathbb{Z}[X]$ gilt. Dann ist $c\tilde{g} = (ch)f = \tilde{h}f$. Mit Lemma 13.8 gilt dann auch $f \mid \tilde{g}$ in $\mathbb{Z}[X]$.
Wir reduzieren modulo p und rechnen in $\mathbb{F}_p[X]$ wobei $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ Körper mit p Elementen ist. Für ein beliebiges $h \in \mathbb{Z}[X]$ sei dann wie in Bemerkung 13.1 $h^* \in \mathbb{F}_p[X]$ das Polynom, das wir durch Reduktion der Koeffizienten modulo p erhalten. Damit ist $\Phi_n^* = f^*g^*$ und $f^* \mid g(X^p)^*$ in $\mathbb{F}_p[X]$. Sei $g = \sum_j b_j X^j$. Dann ist

$$g(X^p)^* = \sum_j \bar{b}_j X^{jp} = \sum_j \bar{b}_j^p X^{jp} = \sum_j \big(\bar{b}_j X^j\big)^p = \big(\sum_j \bar{b}_j X^j\big)^p = (g^*)^p,$$

wobei wir den kleinen Satz von Fermat und Lemma 15.2 benutzt haben. Also folgt auch $f^* \mid (g^*)^p$ in $\mathbb{F}_p[X]$. Sei $f_1 \in \mathbb{F}_p[X]$ ein irreduzibler Faktor von f^* (also auch $\mathrm{Grad}(f_1) \geqslant 1$). Dann ist $f_1 \mid (g^*)^p$ also auch $f_1 \mid g^*$ und es folgt $f_1^2 \mid \Phi_n^*$ in $\mathbb{F}_p[X]$. Dann ist aber auch $f_1^2 \mid X^n - \bar{1}$ und damit $f_1 \mid D(X^n - \bar{1}) = \bar{n}X^{n-1}$ in $\mathbb{F}_p[X]$ (siehe die Bemerkungen vor Lemma 17.5). Wegen $\bar{n} \neq 0$ müsste dann $f_1 = cX$ mit $0 \neq c \in \mathbb{F}_p$ sein, Widerspruch zu $f_1 \mid X^n - 1$.

Damit ist (∗) gezeigt. Sei nun $d \in \{1, \ldots, n\}$ beliebig mit $\mathrm{ggT}(d, n) = 1$. Wir schreiben $d = p_1 p_2 \cdots p_k$ mit Primzahlen p_i. Es kann hier $p_i = p_j$ für $i \neq j$ gelten, aber auf jeden Fall ist $p_i \nmid n$ für alle i. Da nun $f(\zeta_n) = 0$ gilt, folgt mit (∗) auch $f(\zeta_n^{p_1}) = 0$. Anwenden von (∗) auf $\zeta_n^{p_1}$ zeigt dann auch $f(\zeta_n^{p_1 p_2}) = 0$ und so fort, bis zu $f(\zeta_n^d) = 0$. Also hat f alle ζ_n^d mit $\mathrm{ggT}(d, n) = 1$ als Nullstellen. Damit $\mathrm{Grad}(f) \geqslant \phi(n)$, also schließlich $\Phi_n = f$. □

Satz 18.6 (Gauß). *Sei $n \geqslant 3$. Sind die Eckpunkte des regelmäßigen n-Ecks mit Zirkel und Lineal konstruierbar (aus $P_0 = \{(0,0), (1,0)\}$), so muss $\phi(n)$ eine 2-Potenz sein. (Hier ist wieder $\phi(n)$ die **Eulersche Phi-Funktion**.)*

Beweis. Nach Annahme ist insbesondere $z_{n,1} \in \mathbb{R}^2$ konstruierbar, also ist $[\mathbb{Q}(\cos(2\pi/n)) : \mathbb{Q}]$ nach Satz 18.3 eine 2-Potenz. Wegen $\cos(2\pi/n)^2 + \sin(2\pi/n)^2 = 1$ ist $\sin(2\pi/n)$ Nullstelle eines Polynoms vom Grad 2 über $\mathbb{Q}(\cos(2\pi/n))$, also gilt $[\mathbb{Q}(\cos(2\pi/n), \sin(2\pi/n)) : \mathbb{Q}(\cos(2\pi/n))] \leqslant 2$. Mit dem Gradsatz folgt, dass $[\mathbb{Q}(\cos(2\pi/n), \sin(2\pi/n)) : \mathbb{Q}]$ eine 2-Potenz ist. Wir gehen nun zu $\mathbb{C}$ über und betrachten $K := \mathbb{Q}(\cos(2\pi/n), \sin(2\pi/n), i) \subseteq \mathbb{C}$. Wegen $i^2 + 1 = 0$ ist $[K : \mathbb{Q}(\cos(2\pi/n), \sin(2\pi/n))] \leqslant 2$. Mit dem Gradsatz folgt, dass $[K : \mathbb{Q}]$ eine 2-Potenz ist. Nun ist $\zeta_n = \cos(2\pi/n) + i\sin(2\pi/n) \in K$, also folgt auch $[\mathbb{Q}(\zeta_n) : \mathbb{Q}] = 2^r$

für ein $r \geqslant 0$ (wiederum mit dem Gradsatz). Aber nach Satz 18.5 ist Φ_n das Minimalpoynom von ζ_n über $\mathbb{Q}$; also schließlich $\phi(n) = \mathrm{Grad}(\Phi_n) = [\mathbb{Q}(\zeta_n) : \mathbb{Q}] = 2^r$. □

Gauß hat auch die Umkehrung des obigen Satzes gezeigt, d.h., ist $\phi(n)$ eine 2-Potenz, so ist das regelmässige n-Eck mit Zirkel und Lineal konstruierbar. Damit konnte er zum Beispiel zeigen, dass das regelmäßige 17-Eck konstruierbar ist, was der erste echte Fortschritt zu diesem Problem seit der Antike war. (Gauß war gerade 19 Jahre alt, als er dies löste.) Für weitere Informationen siehe `https://de.wikipedia.org/wiki/Konstruierbares_Polygon`.

Beispiel 18.7. Sei $n = p \geqslant 3$ eine Primzahl. Dann ist $\phi(p) = p - 1$. Ist das regelmäßige p-Eck mit Zirkel und Lineal konstruierbar, so muss $p-1$ eine 2-Potenz sein. Solche Primzahlen heißen ***Fermat'sche Primzahlen***; Beispiele sind $3 = 2 + 1$, $5 = 2^2 + 1$, $17 = 2^4 + 1$, $257 = 2^8 + 1$, $65537 = 2^{16} + 1$. Es ist nicht bekannt, ob es noch weitere gibt!

19. *Lösbarkeit algebraischer Gleichungen*

In einer Vorlesung zur Analysis (oder komplexen Analysis) haben Sie vielleicht schon einen Beweis des folgenden Satzes gesehen.

Satz 19.1 (Fundamentalsatz der Algebra).[9] *Jedes nicht-konstante Polynom in* $\mathbb{C}[X]$ *zerfällt in Linearfaktoren über* $\mathbb{C}$*, d.h.,* $\mathbb{C}$ *ist algebraisch abgeschlossen.*

Der folgende Beweis des Fundamentalsatzes ist "so algebraisch wie möglich". Er geht auf Laplace zurück (um 1795, siehe [Eb, Kap. 4, Anhang]) und benutzt den ***Hauptsatz über symmetrische Polynome***. Wegen $\mathbb{C} \supseteq \mathbb{R}$ werden wir nicht ganz ohne Analysis auskommen; wir benutzen allerdings nur die folgenden beiden "analytischen" Tatsachen:

(1) Ist $f \in \mathbb{R}[X]$ nicht-konstant mit $\mathrm{Grad}(f)$ ungerade, so hat f eine Nullstelle in $\mathbb{R}$. [Denn: Sei f normiert. Da $\mathrm{Grad}(f)$ ungerade, ist $\lim_{x\to+\infty} f(x) = +\infty$, $\lim_{x\to-\infty} f(x) = -\infty$; nach dem Zwischenwertsatz der Analysis muss es also ein $x_0 \in \mathbb{R}$ geben mit $f(x_0) = 0$.]

(2) Wie bereits in Beispiel 16.10 bemerkt, hat jedes $z = a + ib \in \mathbb{C}$ (mit $i = \sqrt{-1}$) eine Quadratwurzel in $\mathbb{C}$. Also kann man jedes $0 \neq f \in \mathbb{C}[X]$ vom Grad 2 nach dem üblichen Verfahren (quadratische Ergänzung) in Linearfaktoren über $\mathbb{C}$ zerlegen.

Nun zum eigentlichen Beweis. Es genügt zu zeigen, dass jedes nicht-konstante $f \in \mathbb{C}[X]$ eine Nullstelle in $\mathbb{C}$ hat. Denn ist $f(z) = 0$ für ein $z \in \mathbb{C}$, so ist $f = (X - z)g$ mit $\mathrm{Grad}(g) = \mathrm{Grad}(f) - 1$ und wir können mit g fortfahren. Es genügt sogar zu zeigen:

(∗) Jedes nicht-konstante $g \in \mathbb{R}[X]$ hat eine Nullstelle in $\mathbb{C}$.

[9] Siehe [Eb, Kap. 4] und die "Note historique" [Bo, VII.69] für einen Überblick zur Geschichte dieses Satzes.

Denn ist $f \in \mathbb{C}[X]$ beliebig, so schreibe $f = \sum_{i=0}^{n} a_i X^i$ mit $a_i \in \mathbb{C}$ und setze $\bar{f} := \sum_{i=0}^{n} \bar{a}_i X^i$. Dann ist $g := f\bar{f} \in \mathbb{R}[X]$. Gilt $(*)$, so gibt es ein $z \in \mathbb{C}$ mit $g(z) = (f\bar{f})(z) = f(z)\bar{f}(z) = 0$. Also ist $f(z) = 0$ oder $\bar{f}(z) = 0$; im letzteren Fall ist auch $f(\bar{z}) = 0$. Also hat f auf jeden Fall eine Nullstelle in $\mathbb{C}$.

Sei also nun $0 \neq g \in \mathbb{R}[X]$ mit $n = \mathrm{Grad}(g) \geqslant 1$; um zu zeigen, dass $(*)$ gilt, können wir g als normiert annehmen und schreiben $n = 2^l m$ mit $l \geqslant 0$ und $m \geqslant 1$ ungerade. Der Beweis erfolgt mit Induktion nach l. Ist $l = 0$, so ist n ungerade, also hat g eine Nullstelle, sogar in $\mathbb{R}$, nach (1). Sei nun $l > 0$ und $L \supseteq \mathbb{C}$ ein Zerfällungskörper von g. Da g normiert ist, können wir schreiben $g = (X - \alpha_1) \cdots (X - \alpha_n)$ mit $\alpha_i \in L$. Der geniale Trick des Beweises besteht darin, für ein festes $\lambda \in \mathbb{R}$ das folgende Polynom zu betrachten:

$$g_\lambda := \prod_{1 \leqslant r < s \leqslant n} \left(X - \alpha_r - \alpha_s - \lambda \alpha_r \alpha_s\right) \in L[X].$$

Es gilt $\mathrm{Grad}(g_\lambda) = \frac{1}{2} n(n-1) = \frac{1}{2} 2^l m(2^l m - 1) = 2^{l-1} m(2^l m - 1) \geqslant 1$. Also ist der 2-Anteil im Grad von g_λ nur noch 2^{l-1}, d.h., wir können versuchen, Induktion anzuwenden. Dazu müssen wir noch zeigen, dass $g_\lambda \in \mathbb{R}[X]$ gilt. Bei diesem entscheidenden Punkt benutzen wir nun den Hauptsatz über symmetrische Polynome. Betrachte dazu

$$G_\lambda := \prod_{1 \leqslant r < s \leqslant n} \left(X_{n+1} - X_r - X_s - \lambda X_r X_s\right) \in \mathbb{R}[X_1, \ldots, X_n, X_{n+1}].$$

Wenden wir irgendeine Permutation $\pi \in S_n$ (aufgefasst also Permutation in S_{n+1} wobei $\pi(n+1) = n+1$) auf obiges Polynom an, so werden lediglich die Faktoren umgeordnet, man erhält also das gleiche Polynom. Schreiben wir also $G_\lambda = \sum_j g_j X_{n+1}^j$ mit $g_j \in \mathbb{R}[X_1, \ldots, X_n]$, so folgt $g_j \in \mathbb{R}[X_1, \ldots, X_n]^{S_n}$ für alle j. Durch Einsetzen erhalten wir

$$g_\lambda = G_\lambda(\alpha_1 \ldots, \alpha_n, X) = \sum_j g_j(\alpha_1, \ldots, \alpha_n) X^j \in L[X].$$

Nun war $g = (X - \alpha_1) \cdots (X - \alpha_n) \in \mathbb{R}[X]$. Mit der "verblüffenden Konsequenz" in Beispiel 14.6 gilt also $h(\alpha_1, \ldots, \alpha_n) \in \mathbb{R}$ für *jedes* $h \in \mathbb{R}[X_1, \ldots, X_n]^{S_n}$. Insbesondere damit auch $g_j(\alpha_1, \ldots, \alpha_n) \in \mathbb{R}$ für alle j, also $g_\lambda \in \mathbb{R}[X]$. Wir können also tatsächlich Induktion auf g_λ anwenden; damit gibt es ein $z \in \mathbb{C}$ mit $g_\lambda(z) = 0$. Die Definition von g_λ zeigt, dass es Indizes $1 \leqslant r < s \leqslant n$ gibt mit $\alpha_r + \alpha_s + \lambda \alpha_r \alpha_s = z \in \mathbb{C}$. Wählen wir nun ein weiteres $\lambda' \in \mathbb{R}$ und verfahren analog, so muss es wieder (möglicherweise andere) Indizes $1 \leqslant r' < s' \leqslant n$ geben mit $\alpha_{r'} + \alpha_{s'} + \lambda' \alpha_{r'} \alpha_{s'} \in \mathbb{C}$. Da es unendlich viele Möglichkeiten für $\lambda \in \mathbb{R}$ gibt, aber nur endlich viele Paare von Indizes in $\{1, \ldots, n\}$, muss es verschiedene $\lambda, \lambda' \in \mathbb{R}$ geben, die zu dem gleichen Paar von Indizes führen, d.h., es gibt $1 \leqslant r < s \leqslant n$ mit

$$\alpha_r + \alpha_s + \lambda \alpha_r \alpha_s \in \mathbb{C}, \qquad \alpha_r + \alpha_s + \lambda' \alpha_r \alpha_s \in \mathbb{C} \qquad \text{und} \qquad \lambda \neq \lambda'.$$

Daraus folgt dann aber $p := \alpha_r + \alpha_s \in \mathbb{C}$ und $q := \alpha_r\alpha_s \in \mathbb{C}$, also ist

$$(X - \alpha_r)(X - \alpha_s) = X^2 - (\alpha_r + \alpha_s)X + \alpha_r\alpha_s = X^2 - pX + q \in \mathbb{C}[X].$$

Nach (2) sind α_r, α_s in $\mathbb{C}$, also haben wir tatsächlich Nullstellen von g in $\mathbb{C}$ gefunden. □

Beispiel 19.2. Sei $f \in \mathbb{C}[X]$ normiert mit $\mathrm{Grad}(f) = 3$. Durch eine Substitution der Form $X \mapsto X + a$ kann man stets erreichen, dass $f = X^3 + pX + q$ mit $p, q \in \mathbb{C}$ gilt. Da $\mathbb{C}$ algebraisch abgeschlossen ist, gilt dann $f = (X - \alpha_1)(X - \alpha_2)(X - \alpha_3)$ mit $\alpha_i \in \mathbb{C}$. Die ***Cardano'schen Formeln*** (16. Jahrhundert) zeigen, dass sich die α_i wie folgt explizit aus p, q berechnen lassen: Da wir in $\mathbb{C}$ beliebige Wurzeln ziehen können, bestimmen wir zuerst $u, v \in \mathbb{C}$ mit

$$u^3 = -\tfrac{q}{2} + \sqrt{\Delta} \qquad \text{und} \qquad v^3 = -\tfrac{q}{2} - \sqrt{\Delta}, \qquad \text{wobei} \qquad \Delta := \tfrac{q^2}{4} + \tfrac{p^3}{27} \in \mathbb{C}$$

und $\sqrt{\Delta} \in \mathbb{C}$ wie in Beispiel 16.10(a) gebildet werden kann. Dann folgt $u^3 + v^3 = -q$ und $(uv)^3 = u^3v^3 = (-\frac{q}{2} + \sqrt{\Delta})(-\frac{q}{2} - \sqrt{\Delta}) = \frac{q^2}{4} - \Delta = -\frac{p^3}{27} = (-\frac{p}{3})^3$; also können wir u, v so wählen, dass $uv = -\frac{p}{3}$ gilt. Mit dieser Nebenbedingung sind die Lösungen gegeben durch

$$\alpha_1 := u + v, \qquad \alpha_2 := \zeta_3 u + \zeta_3^2 v, \qquad \alpha_3 := \zeta_3^2 u + \zeta_3 v,$$

wobei $\zeta_3 \in \mathbb{C}$ eine primitive dritte Einheitswurzel ist. (Beweis durch einfaches Nachrechnen: $\alpha_1^3 + p\alpha_1 + q = (u + v)^3 + p(u + v) + q = u^3 + 3u^2v + 3uv^2 + v^3 + p(u + v) + q = u^3 + v^3 + (3uv + p)(u + v) + q = 0$; analog für α_2 und α_3.) Beispiele: $f = X^3 + 6X - 20$ mit genau einer reellen Lösung, sowie $f = X^3 - 15X - 4$ mit drei reellen Lösungen, wie man leicht mit einer Kurvendiskussion sieht; was ergibt sich jeweils für $\alpha_1, \alpha_2, \alpha_3$ aus obigen Formeln?

Wir erhalten hier die α_i also durch wiederholtes Wurzelziehen und die üblichen Körper-Operationen $+, -, \cdot, /$ aus den Koeffizienten von f, ähnlich (nur etwas komplizierter) zum Fall von Polynomen vom Grad 2. — Dies führt zur folgenden allgemeinen Definition.

Definition 19.3. Sei $f \in K[X]$ nicht-konstant und $L \supseteq K$ ein Zerfällungskörper von f. Wir sagen, dass f durch ***Radikale*** auflösbar ist, wenn es eine endliche Folge von einfachen algebraischen Körpererweiterungen $K = K_0 \subseteq K_1 = K(\beta_1) \subseteq K_2 = K_1(\beta_2) \subseteq \ldots \subseteq K_r = K_{r-1}(\beta_r)$ gibt, so dass $L \subseteq K_r$ gilt und $\beta_i^{n_i} \in K_{i-1}$ mit $n_i \geqslant 1$ für alle i.

Nach den Beispielen 16.10(a) und 19.2 sind also Polynome vom Grad 2 und 3 durch Radikale auflösbar, und es ist bekannt, dass dies auch für Polynome vom Grad 4 geht (Ferrari-Formeln, ebenfalls 16. Jahrhundert). Ein berühmtes Problem der Algebra ist die Frage, ob dies für beliebige Polynome möglich ist. Die systematische Lösung dieser Frage (durch Abel, Galois) ist ein Musterbeispiel für das Zusammenspiel von Gruppen, Ringen und Körpern.

Sei $L \supseteq K$ eine Körpererweiterung und $G := \mathrm{Aut}(L, K)$ die Menge der Körper-Isomorphismen $\sigma\colon L \to L$ mit $\sigma(x) = x$ für alle $x \in K$. Dann ist G eine Gruppe (mit der Hintereinanderausführung als Verknüpfung), die auf L operiert durch $G \times L \to L$, $(\sigma, y) \mapsto \sigma(y)$.

Lemma 19.4. *Sei* $L \supseteq K$ *Zerfällungskörper von* $f \in K[X]$*, wobei* $f \neq 0$ *und* $n := \mathrm{Grad}(f) \geqslant 1$*. Dann operiert* $G = \mathrm{Aut}(L, K)$ *auf* $\mathcal{O} := \{z \in L \mid f(z) = 0\}$ *und der zugehörige Homomorphismus* $\rho\colon G \to S_{\mathcal{O}}$ *ist injektiv; die Gruppe* G *ist isomorph zu einer Untergruppe von* S_n*.*

Beweis. Sei $|\mathcal{O}| = m$ und $\mathcal{O} = \{z_1, \ldots, z_m\}$ mit $m \leqslant n$. (Es könnte ja mehrfache Nullstellen geben.) Sei $\sigma \in \mathrm{Aut}(L, K)$ und $z \in \mathcal{O}$. Ist $f = \sum_{i=0}^{n} a_i X^i$, so folgt

$$0 = \sigma(0) = \sigma(f(z)) = \sigma\Big(\sum_{i=0}^{n} a_i z^i\Big) = \sum_{i=0}^{n} \sigma(a_i)\sigma(z)^i = \sum_{i=0}^{n} a_i \sigma(z)^i = f(\sigma(z)),$$

wobei wir $\sigma(x) = x$ für alle $x \in K$ benutzt haben; also ist auch $\sigma(z) \in \mathcal{O}$. Damit schränkt die Operation von G auf L auf eine Operation von G auf $\mathcal{O}$ ein. Nach Beispiel 8.11 erhalten wir einen Gruppen-Homomorphismus $\rho\colon G \to S_{\mathcal{O}} \cong S_m \leqslant S_n$ mit $\mathrm{Kern}(\rho) = \{\sigma \in G \mid \sigma(z_i) = z_i \text{ für } 1 \leqslant i \leqslant m\}$. Nun ist $L = K(z_1, \ldots, z_m)$; nach Beispiel 16.5 ist ein beliebiges $y \in L$ von der Form $y = g(z_1, \ldots, z_m)$ mit $g \in K[X_1, \ldots, X_m]$. Ist also $\sigma(z_i) = z_i$ für alle i, so folgt $\sigma(y) = y$ für alle $y \in L$, d.h., $\sigma = \mathrm{id}_L$. Also ist ρ injektiv. □

Uns stehen an dieser Stelle eigentlich überhaupt keine Methoden zur Verfügung, um Elemente von $\mathrm{Aut}(L, K)$ zu konstruieren; wir können nicht einmal ausschließen, dass $\mathrm{Aut}(L, K)$ nur aus der Identität id_L besteht. Um dennoch ein substanzielles Beispiel behandeln zu können, kommen uns wiederum die symmetrischen Polynome aus §14 zu Hilfe.

Beispiel 19.5. Sei $n \geqslant 1$ und L_n Quotientenkörper des Polynomrings $\mathbb{Q}[X_1, \ldots, X_n]$. Es gilt $L_n = \{f/g \mid f, g \in \mathbb{Q}[X_1, \ldots, X_n], g \neq 0\}$, und es wird wie üblich mit solchen Brüchen gerechnet. Sei $R_n = \mathbb{Q}[X_1, \ldots, X_n]^{S_n} \subseteq \mathbb{Q}[X_1, \ldots, X_n] \subseteq L_n$ der Teilring der symmetrischen Polynome. Dann ist $K_n := \{f/g \in L_n \mid f, g \in R_n, g \neq 0\}$ ein Teilkörper von L_n. Wir definieren das ***allgemeine Polynom n-ten Grades*** als

$$f_n := \sum_{d=0}^{n} (-1)^{n-d} s_{n-d} X^d \in K_n[X] \qquad \text{(wobei } s_{n-d} = \text{elementar-symmetrisches Polynom)}.$$

Analog zu Beispiel 14.3 gilt die Formel $f_n = (X - X_1) \cdots (X - X_n) \in L_n[X]$, d.h., f_n zerfällt in $L_n \supseteq K_n$ in Linearfaktoren. Wegen $L_n = K_n(X_1, \ldots, X_n)$ folgt, dass L_n sogar ein Zerfällungskörper von f_n ist, und damit $[L_n : K_n] \leqslant n!$ (siehe Satz 16.11). Nach Lemma 19.4 ist $G_n := \mathrm{Aut}(L_n, K_n)$ isomorph zu einer Untergruppe von S_n, also auch $|G_n| \leqslant n!$.

Satz 19.6. *Mit obigen Bezeichnungen gilt* $G_n \cong S_n$ *und* $[L_n : K_n] = n!$*.*

Beweis. Um $G_n \cong S_n$ zu zeigen, genügt es noch zu zeigen, dass $|G_n| \geqslant n!$ gilt. Dazu: Für jedes $\pi \in S_n$ haben wir den Einsetzungs-Homomorphismus $\tilde{\pi}\colon \mathbb{Q}[X_1, \ldots, X_n] \to \mathbb{Q}[X_1, \ldots, X_n]$ mit $\tilde{\pi}(f) = f(X_{\pi(1)}, \ldots, X_{\pi(n)})$, wie in §14. Man sieht leicht, dass $\tilde{\pi}$ bijektiv ist, mit inverser Abbildung $\tilde{\sigma}\colon \mathbb{Q}[X_1, \ldots, X_n] \to \mathbb{Q}[X_1, \ldots, X_n]$ wobei $\sigma = \pi^{-1}$. Dann können wir $\tilde{\pi}$ auch

zu einem Automorphismus von L_n fortsetzen durch $\tilde{\pi}(f/g) := \tilde{\pi}(f)/\tilde{\pi}(g)$ für $f/g \in L_n$. (Man prüft leicht nach, dass dies wohl-definiert ist.) Wegen $\tilde{\pi}(f) = f$ für alle $f \in R_n$ folgt $\tilde{\pi}|_{K_n} = \mathrm{id}_{K_n}$, also ist $\tilde{\pi} \in G_n$. Sind $\pi, \pi' \in S_n$ mit $\pi \neq \pi'$, so gibt es ein $i \in \{1, \ldots, n\}$ mit $\pi(i) \neq \pi'(i)$; dann ist aber auch $\tilde{\pi}(X_i) = X_{\pi(i)} \neq X_{\pi'(i)} = \tilde{\pi}'(X_i)$, also $\tilde{\pi} \neq \tilde{\pi}'$. Damit ist $S_n \to G_n$, $\pi \mapsto \tilde{\pi}$, injektiv, also gilt auch $|G_n| \geqslant n!$.

Nun zu $[L_n : K_n]$. Wir wissen bereits, dass $[L_n : K_n] \leqslant n!$ gilt. Um Gleichheit zu zeigen, betrachten wir für $0 \leqslant i \leqslant n$ den Teilkörper $M_i := K_n(X_1, \ldots, X_i) \subseteq L_n$. Es gilt also

$$M_0 = K_n \subseteq M_1 = K_n(X_1) \subseteq M_2 = K_n(X_1, X_2) \subseteq \ldots \subseteq M_n = K_n(X_1, \ldots, X_n) = L_n;$$

analog wie in Beispiel 16.5. Beachte außerdem $M_i = M_{i-1}(X_i)$ für $1 \leqslant i \leqslant n$. Mit dem Gradsatz folgt $[L_n : K_n] = \prod_{i=1}^{n} [M_i : M_{i-1}]$. Wenn wir also zeigen können, dass $[M_i : M_{i-1}] \geqslant n-i+1$ für all i gilt, so folgt $[L_n : K_n] \geqslant n!$ und wir sind fertig. Sei nun $i \in \{1, \ldots, n\}$ fest und $\mu_i \in M_{i-1}[X]$ das Minimalpolynom von $X_i \in M_i$. Genauso wie im Beweis von Lemma 19.4 sieht man, dass dann auch $\mu_i(\sigma(X_i)) = 0$ gilt für alle $\sigma \in G_n$ mit der Eigenschaft, dass σ die Koeffizienten von μ_i festlässt. Für $i < j \leqslant n$ sei nun $\tau_{ij} \in S_n$ die Transposition, die i und j vertauscht. Dann ist $\tilde{\tau}_{ij} \in G_n$ (siehe oben). Wegen $\tau_{ij}(l) = l$ für $1 \leqslant l \leqslant i-1$ folgt $\tilde{\tau}_{ij}(y) = y$ für alle $y \in M_{i-1}$ und damit $\mu_i(X_j) = \mu_i(\tilde{\tau}_{ij}(X_i)) = 0$. Also sind $X_i, \ldots, X_n$ Nullstellen von μ_i; es folgt mit Satz 16.3 in der Tat $[M_n : M_{n-1}] = \mathrm{Grad}(\mu_i) \geqslant n-i+1$. □

Wir zitieren nun das folgende bedeutende Ergebnis (Beweis siehe Anhang A).

Satz 19.7 (Galois)**.** *Sei* $f \in K[X]$ *nicht-konstant und* $L \supseteq K$ *ein Zerfällungskörper von* f*. Ist* f *durch Radikale auflösbar, so ist* $\mathrm{Aut}(L, K)$ *eine auflösbare Gruppe.*

Nun ist S_n für $n \geqslant 5$ nach Satz 9.11 nicht auflösbar. Also folgt, dass das obige allgemeine Polynom $f_n \in K_n[X]$ für $n = \mathrm{Grad}(f_n) \geqslant 5$ nicht durch Radikale auflösbar ist.

Insbesondere ist damit der zuerst von Abel um 1824 vollständig bewiesene Satz[10] *gezeigt, dass sich Polynome vom Grad* $\geqslant 5$ *im Allgemeinen nicht durch Radikale auflösen lassen. Und letztlich ist der Grund dafür, dass die Gruppen* S_n *mit* $n \geqslant 5$ *nicht auflösbar sind !*

Natürlich würde man gern auch ein konkretes Polynom sehen, welches nicht durch Radikale auflösbar ist. Man kann etwa für $f = X^5 - 4X + 2 \in \mathbb{Q}[X]$ zeigen, dass $\mathrm{Aut}(L, \mathbb{Q}) \cong S_5$ gilt (siehe Anhang A, Beispiel 20.8 mit $n = 2$); *also ist* $f = X^5 - 4X + 2$ *definitiv nicht durch Radikale über* $\mathbb{Q}$ *auflösbar!* (Und Beispiel 20.8 liefert sogar unendlich viele solche Beispiele[11].)

---

[10] Siehe den Artikel von Rosen [Ro] für eine moderne Fassung von Abels ursprünglichem Beweis.

[11] Weitere Literatur dazu: H. OSADA, The Galois groups of the polynomials $X^n + aX^l + b$, J. Number Theory **25** (1987), 230–238. Zum Beispiel ist $\mathrm{Aut}(L, \mathbb{Q}) \cong S_n$ für $f = X^n - X - 1 \in \mathbb{Q}[X]$, für alle $n \geqslant 5$.

Anhänge (nicht in der Vorlesung behandelt)

20. *Anhang A: Beweis des Satzes von Galois*

Der Satz 19.7 von Galois wird oft in Lehrbüchern (und Vorlesungen) im Anschluss an ein Kapitel zur "Galois-Theorie" behandelt. Eine Untersuchung der tatsächlich benötigten Voraussetzungen zeigt allerdings (wie z.B. im Artikel von Rosen [Ro] oder im Buch von Stillwell [St]), dass dies auch wesentlich direkter geht — und genau das ist das Ziel dieses Abschnittes.

Wir beginnen mit einigen allgemeinen Aussagen zu Körper-Automorphismen, die auch von unabhängigem Interesse sind. Zur Erinnerung: Ist $L \supseteq K$ eine Körpererweiterung, so bezeichnen wir mit $\mathrm{Aut}(L, K)$ die Gruppe aller Körper-Isomorphismen $\sigma\colon L \to L$ (auch Automorphismen genannt) mit $\sigma(x) = x$ für alle $x \in K$. Ein grundlegendes Problem, bereits in §19 erwähnt, ist: Wie können wir überhaupt Elemente in $\mathrm{Aut}(L, K)$ konstruieren?

Seien K, K' Körper und $\sigma\colon K \to K'$ ein Isomorphismus. Dann können wir σ zu einem Isomorphismus der zugehörigen Polynomringe fortsetzen, durch die Definition:

$$\tilde{\sigma}(f) := \sum_{i=0}^{n} \sigma(a_i) X^i \in K'[X] \qquad \text{wobei} \qquad f = \sum_{i=0}^{n} a_i X^i \in K[X].$$

Dies ist ein Einsetzungs-Homomorphismus im Sinne von Satz 12.6 (wobei σ auf die Koeffizienten eines Polynoms angewandt wird, und X wieder auf X abgebildet wird).

Lemma 20.1. *Seien K, K' Körper und $\sigma\colon K \to K'$ ein Isomorphismus. Sei $f \in K[X]$ normiert, irreduzibel und $f' := \tilde{\sigma}(f) \in K'[X]$. Seien $L \supseteq K$ und $L' \supseteq K'$ Erweiterungen, die Elemente $\alpha \in L$ mit $f(\alpha) = 0$ und $\beta \in L'$ mit $f'(\beta) = 0$ enthalten. Dann gibt es genau einen Isomorphismus $\sigma_1\colon K(\alpha) \to K'(\beta)$ mit $\sigma_1(\alpha) = \beta$ und $\sigma_1(x) = \sigma(x)$ für alle $x \in K$.*

Beweis. Es ist $K(\alpha) = \{g(\alpha) \mid g \in K[X]\}$. Nehmen wir zuerst an, es existert ein Isomorphismus $\sigma_1\colon K(\alpha) \to K'(\beta)$ mit den gewünschten Eigenschaften. Ist $g = \sum_{i=0}^{m} b_i X^i \in K[X]$, so folgt

$$\sigma_1\big(g(\alpha)\big) = \sigma_1\Big(\sum_{i=0}^{m} b_i \alpha^i\Big) = \sum_{i=0}^{m} \sigma_1(b_i)\sigma_1(\alpha)^i = \sum_{i=0}^{m} \sigma(b_i)\beta^i = \tilde{\sigma}(g)(\beta).$$

Dies zeigt bereits, dass σ_1 dann eindeutig bestimmt ist. Umgekehrt zeigt die obige Formel auch, wie wir die gewünschte Abbildung $\sigma_1\colon K(\alpha) \to L'$ definieren müssen: Für $g \in K[X]$ setze $\sigma_1(g(\alpha)) := \tilde{\sigma}(g)(\beta) \in L'$. Dies ist wohl-definiert. Denn sind $g_1, g_2 \in K[X]$ so, dass $g_1(\alpha) = g_2(\alpha)$ gilt, dann ist $(g_1 - g_2)(\alpha) = 0$, also $f \mid g_1 - g_2$ in $K[X]$ (wegen $f = \mu_\alpha$). Weil $\tilde{\sigma}$ ein Homomorphismus ist, gilt auch $f' = \tilde{\sigma}(f) \mid \tilde{\sigma}(g_1) - \tilde{\sigma}(g_2)$ in $K'[X]$. Wegen $f'(\beta) = 0$ folgt $\tilde{\sigma}(g_1)(\beta) = \tilde{\sigma}(g_2)(\beta)$, also ist σ_1 in der Tat wohl-definiert. Weil die Abbildung $K[X] \to L'$, $g \mapsto \tilde{\sigma}(g)(\beta)$, ebenfalls ein Einsetzungs-Homomorphismus ist (bei dem σ auf die Koeffizienten eines Polynoms angewandt wird, und β für X eingesetzt wird), folgt dann sofort,

dass σ_1 ein Homomorphismus ist. Es gilt $\mathrm{Bild}(\sigma_1) = \{h(\beta) \mid h \in K'[X]\} = K'(\beta)$, da σ ein Isomorphismus ist. Da $K(\alpha)$ ein Körper ist, ist σ_1 automatisch injektiv. □

Beispiel 20.2. Sei $m \in \mathbb{N}$ und $\zeta_m = \cos(2\pi/m) + i\sin(2\pi/m) \in \mathbb{C}$. Dann ist $L := \mathbb{Q}(\zeta_m)$ Zerfällungskörper von $X^m - 1 \in \mathbb{Q}[X]$. Nach Satz 18.5 ist das m-te Kreisteilungspolynom $\Phi_m \in \mathbb{Q}[X]$ das Minimalpolynom von ζ_m und damit $[L : \mathbb{Q}] = \phi(m)$; wegen $\Phi_m \mid X^m - 1$ ist $L \supseteq \mathbb{Q}$ auch Zerfällungskörper von Φ_m. Sei nun $1 \leqslant l \leqslant m$ gegeben mit $\mathrm{ggT}(l, m) = 1$. Dann ist auch ζ_m^l eine Nullstelle von Φ_m (siehe Beispiel 12.11) und nach Lemma 20.1 gibt es einen Automorphismus $\sigma_l \in \mathrm{Aut}(L, \mathbb{Q})$ mit $\sigma_l(\zeta_m) = \zeta_m^l$. Ist umgekehrt $\sigma \in \mathrm{Aut}(L, \mathbb{Q})$ beliebig, so ist $\sigma(\zeta_m)$ wieder eine Nullstelle von Φ_m (siehe Lemma 19.4), also $\sigma(\zeta_m) = \zeta_m^l$ mit $1 \leqslant l \leqslant m$ und $\mathrm{ggT}(l, m) = 1$. Wegen $L = \mathbb{Q}(\zeta_m)$ folgt $\sigma = \sigma_l$. Damit haben wir $\mathrm{Aut}(L, \mathbb{Q})$ vollständig bestimmt: Es gilt $\mathrm{Aut}(L, \mathbb{Q}) = \{\sigma_l \mid 1 \leqslant l \leqslant m,\ \mathrm{ggT}(l, m) = 1\}$.

Beispiel 20.3. Sei K beliebiger Körper, $m \geqslant 1$ und $\tilde{K} \supseteq K$ ein Zerfällungskörper von $X^m - 1 \in K[X]$. Sei $\zeta \in \tilde{K}$ eine primitive m-te Einheitswurzel. Dann sind alle Nullstellen von $X^m - 1$ durch Potenzen von ζ gegeben, also ist $\tilde{K} = K(\zeta)$; siehe auch noch einmal Definition 12.10. Behauptung: $\mathrm{Aut}(\tilde{K}, K)$ ist abelsch.

(Dies ist also ein Spezialfall von Satz 19.7, denn wegen $\zeta^m = 1 \in K$ ist $X^m - 1$ durch Radikale auflösbar.)

Dazu: Sei $\sigma \in \mathrm{Aut}(\tilde{K}, K)$. Wegen $\tilde{K} = \{g(\zeta) \mid g \in K[X]\}$ ist dann σ eindeutig festgelegt durch $\sigma(\zeta)$. Wegen $\sigma(\zeta)^m = \sigma(\zeta^m) = \sigma(1) = 1$ ist $\sigma(\zeta) = \zeta^l$ mit einem $l \in \mathbb{N}_0$. Sei auch $\sigma' \in \mathrm{Aut}(\tilde{K}, K)$; dann ist analog $\sigma'(\zeta) = \zeta^{l'}$ mit $l' \in \mathbb{N}_0$. Es folgt $(\sigma \circ \sigma')(\zeta) = \sigma(\sigma'(\zeta)) = \sigma(\zeta^{l'}) = \sigma(\zeta)^{l'} = (\zeta^l)^{l'} = \zeta^{l'l}$. Analog ergibt sich $(\sigma' \circ \sigma)(\zeta) = \zeta^{ll'} = (\sigma \circ \sigma')(\zeta)$. Wegen $\tilde{K} = \{g(\zeta) \mid g \in K[X]\}$ ist dann $\sigma \circ \sigma' = \sigma' \circ \sigma$. Also ist $\mathrm{Aut}(\tilde{K}, K)$ abelsch.

Satz 20.4 (Fortsetzungssatz)**.** *Seien K, K' Körper und $\sigma\colon K \to K'$ ein Isomorphismus. Sei $f \in K[X]$ nicht-konstant und $L \supseteq K$ ein Zerfällungskörper von f. Sei $L' \supseteq K'$ ein Zerfällungskörper von $f' := \tilde{\sigma}(f) \in K'[X]$. Dann gibt es einen Isomorphismus $\tau\colon L \to L'$ mit $\tau|_K = \sigma$.*

Beweis. Induktion nach $n = [L : K]$. Ist $n = 1$, so ist $L = K$, also nichts zu zeigen. Sei nun $n > 1$. Wegen $K \subsetneqq L$ gibt es ein normiertes, irreduzibles $g \in K[X]$ mit $\mathrm{Grad}(g) \geqslant 2$ und $g \mid f$. Sei $g' := \tilde{\sigma}(g) \in K'[X]$; dann ist $g' \mid f'$ in $K'[X]$. Da f in L in Linearfaktoren zerfällt, gilt das Gleiche auch für g; analog für f' und g'. Sei $\alpha \in L$ eine Nullstelle von g und $\beta \in L'$ eine Nullstelle von g'. Nach Lemma 20.1 gibt es einen Isomorphismus $\sigma_1\colon K(\alpha) \to K'(\beta)$ mit $\sigma_1(\alpha) = \beta$ und $\sigma_1(x) = \sigma(x)$ für alle $x \in K$. Nun setze $K_1 := K(\alpha)$ und $K_1' := K'(\beta)$; dann ist $[K_1 : K] = \mathrm{Grad}(g) \geqslant 2$. Wir fassen f als Polynom in $K_1[X]$ und f' als Polynom in $K_1'[X]$ auf. Trivialerweise sind dann auch $L \supseteq K_1$ und $L' \supseteq K_1'$ Zerfällungskörper von $f \in K_1[X]$ bzw. $f' \in K_1'[X]$. Wegen $[L : K] = [L : K_1][K_1 : K] > [L : K_1]$ (Gradsatz) können wir Induktion auf den Isomorphismus $\sigma_1\colon K_1 \to K_1'$ und $L \supseteq K_1$, $L' \supseteq K_1'$ anwenden. Also gibt es einen Isomorphismus $\tau\colon L \to L'$ mit $\tau|_{K_1} = \sigma_1$, und damit auch $\tau|_K = (\tau|_{K_1})|_K = \sigma_1|_K = \sigma$. □

Folgerung 20.5. *Sei* $f \in K[X]$ *nicht-konstant. Sind* $L \supseteq K$ *und* $L' \supseteq K$ *Zerfällungskörper von* f*, so gibt es einen Isomorphismus* $\tau\colon L \to L'$ *mit* $\tau(x) = x$ *für alle* $x \in K$*.*

Beweis. Wende Satz 20.4 an mit $\sigma = \mathrm{id}_K$; beachte, dass $f' := \tilde{\sigma}(f) = f$ gilt. □

Folgerung 20.6. *Sei* $L \supseteq K$ *Zerfällungskörper eines nicht-konstanten* $f \in K[X]$*. Gegeben seien Zwischenkörper* $K \subseteq K_1 \subseteq L$ *und* $K \subseteq K_1' \subseteq L$*. Ist* $\sigma\colon K_1 \to K_1'$ *ein Isomorphismus mit* $\sigma(x) = x$ *für alle* $x \in K$*, so gibt es ein* $\tau \in \mathrm{Aut}(L, K)$ *mit* $\tau|_{K_1} = \sigma$*.*

Beweis. Fassen wir f als Polynom in $K_1[X]$ auf, so ist L auch Zerfällungskörper von f über K_1; analog ist $L' := L$ auch Zerfällungskörper von f über K_1'. Wegen $\sigma(x) = x$ für alle $x \in K$ ist $f' := \tilde{\sigma}(f) = f$. Wende nun Satz 20.4 auf $L \supseteq K_1$, $L' \supseteq K_1'$ und σ an. □

Folgerung 20.7. *Sei* $L \supseteq K$ *Zerfällungskörper eines nicht-konstanten* $f \in K[X]$*. Ist* f *irreduzibel, so ist die Operation der Gruppe* $\mathrm{Aut}(L, K)$ *auf der Menge* $\mathcal{O} := \{\alpha \in L \mid f(\alpha) = 0\}$ *der Nullstellen von* f *(siehe Lemma 19.4) transitiv.*

Beweis. Seien $\alpha \neq \beta$ in $\mathcal{O}$ beliebig. Wir betrachten $K \subseteq K_1 := K(\alpha) \subseteq L$ und $K \subseteq K_1' := K(\beta) \subseteq L$. Weil α und β Nullstellen des irreduziblen Polynoms $f \in K[X]$ sind, erhalten wir mit Lemma 20.1 einen Isomorphismus $\sigma\colon K_1 \to K_1'$ mit $\sigma(\alpha) = \beta$ und $\sigma(x) = x$ für alle $x \in K$. Nach Folgerung 20.6 gibt es ein $\tau \in \mathrm{Aut}(L, K)$ mit $\tau|_{K_1} = \sigma_1$ und damit $\tau(\alpha) = \beta$. □

Beispiel 20.8. Wir betrachten ein Polynom der Form $f := X^5 - 2nX + 2 \in \mathbb{Q}[X]$, wobei $n \in \mathbb{N}$, $n \geqslant 2$. Sei $L \supseteq \mathbb{Q}$ ein Zerfällungskörper von f. Da $\mathbb{C}$ algebraisch abgeschlossen ist, können wir $L \subseteq \mathbb{C}$ annehmen. Wir behaupten, dass $\mathrm{Aut}(L, \mathbb{Q}) \cong S_5$ gilt.
Dazu: Nach Lemma 19.4 operiert $\mathrm{Aut}(L, \mathbb{Q})$ auf der Menge $\mathcal{O} \subseteq \mathbb{C}$ der Nullstellen von f und es gibt einen injektiven Gruppen-Homomorphismus $\mathrm{Aut}(L, \mathbb{Q}) \to S_5$; sei $U \leqslant S_5$ das Bild dieses Homomorphismus. Wir müssen zeigen, dass $U = S_5$ gilt. Mit einer Kurvendiskussion sieht man leicht, dass f genau 3 verschiedene reelle und 2 konjugiert-komplexe Nullstellen hat; seien diese $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ und $\alpha_4 = \bar{\alpha}_5 \in \mathbb{C} \setminus \mathbb{R}$. Schauen wir uns nun die Operation von $\mathrm{Aut}(L, \mathbb{Q})$ auf $\mathcal{O} = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5\}$, und damit $U \leqslant S_5$, genauer an.
Sei $\tau\colon \mathbb{C} \to \mathbb{C}$, $z \mapsto \bar{z}$, die komplexe Konjugation. Dann ist $\tau \in \mathrm{Aut}(\mathbb{C}, \mathbb{Q})$; es gilt $\tau(\alpha_i) = \alpha_i$ für $i = 1, 2, 3$ und $\tau(\alpha_4) = \alpha_5$, $\tau(\alpha_5) = \alpha_4$. Damit folgt $\tau(L) = L$. Also ist $\tau' := \tau|_L \in \mathrm{Aut}(L, \mathbb{Q})$; unter dem Homomorphismus $\mathrm{Aut}(L, \mathbb{Q}) \to S_5$ wird τ' auf die Transposition $(4, 5) \in S_5$ abgebildet. Also ist $(4, 5) \in U$. Weiterhin folgt mit Eisenstein ($p = 2$) und dem Satz von Gauss, dass f irreduzibel ist. Nach Folgerung 20.7 operiert $\mathrm{Aut}(L, \mathbb{Q})$ also transitiv auf $\mathcal{O}$. Mit dem Bahnensatz folgt, dass $|\mathcal{O}| = 5$ ein Teiler von $|\mathrm{Aut}(L, \mathbb{Q})|$ ist. Nach dem Satz von Cauchy gibt es in $\mathrm{Aut}(L, \mathbb{Q})$ ein Element der Ordnung 5; dieses wird also auf ein $\pi \in U$ mit $o(\pi) = 5$ abgebildet. In S_5 sind aber die einzigen Permutationen der Ordnung 5 genau

die 5-Zykel. Also enthält die Untergruppe $U \leqslant S_5$ eine Transposition sowie einen 5-Zykel. Es ist dann leicht zu sehen (zum Beispiel mit GAP), dass $U = S_5$ sein muss.

Satz 20.9. *Sei* $L \supseteq K$ *eine Erweiterung mit* $[L : K] < \infty$. *Seien* $\alpha_1, \ldots, \alpha_r \in L$ *mit* $L = K(\alpha_1, \ldots, \alpha_r)$. *Für* $1 \leqslant i \leqslant r$ *sei* $f_i := \mu_{\alpha_i} \in K[X]$ *das Minimalpolynom von* α_i. *Sei* $L' \supseteq K$ *ein Zerfällungskörper von* $g := f_1 \cdots f_r \in K[X]$*; wegen* $\alpha_1, \ldots, \alpha_r \in L$ *können wir dabei* $L \subseteq L'$ *annehmen. Dann gilt* $L' = K\big(\tau(\alpha_i) \mid 1 \leqslant i \leqslant r, \tau \in \mathrm{Aut}(L', K)\big)$.

Beweis. Sei $\mathcal{O} := \{\beta \in L' \mid g(\beta) = 0\}$. Weil $\mathrm{Aut}(L', K)$ auf $\mathcal{O}$ operiert, gilt $\tau(\alpha_i) \in \mathcal{O}$ für alle $\tau \in \mathrm{Aut}(L', K)$ und $1 \leqslant i \leqslant r$. Sei umgekehrt $\beta \in \mathcal{O}$ beliebig. Dann ist $f_i(\beta) = 0$ für ein $i \in \{1, \ldots, r\}$. Wir betrachten $K \subseteq K_1 := K(\alpha_i) \subseteq L$ und $K \subseteq K_1' := K(\beta) \subseteq L'$. Mit Lemma 20.1 erhalten wir einen Isomorphismus $\sigma_1 \colon K_1 \to K_1'$ mit $\sigma_1(\alpha_i) = \beta$ und $\sigma_1(x) = x$ für alle $x \in K$. Nach Folgerung 20.6 gibt es dann auch wieder ein $\tau \in \mathrm{Aut}(L', K)$ mit $\tau|_{K_1} = \sigma_1$ und damit $\tau(\alpha_i) = \beta$. Also ist $\mathcal{O} = \{\tau(\alpha_i) \mid 1 \leqslant i \leqslant r, \tau \in \mathrm{Aut}(L', K)\}$ und damit $K(\mathcal{O}) \subseteq L'$ bereits Zerfällungskörper von g. Also folgt $L' = K(\mathcal{O})$. □

Nach den obigen Vorbereitungen steuern wir nun auf den Beweis des Satzes von Galois zu.

Bemerkung 20.10. Wir betrachten Körpererweiterungen $K \subseteq L \subseteq M$, wobei $L \supseteq K$ ein Zerfällungskörper eines nicht-konstanten $f \in K[X]$ sei. Sei $\mathcal{O} := \{\alpha_1, \ldots, \alpha_r\}$ die Menge der Nullstellen von f in L. Mit dem gleichen Argument wie im Beweis von Lemma 19.4 folgt, dass $\mathrm{Aut}(M, K)$ auf $\mathcal{O}$ operiert; es gilt also $\tau(\mathcal{O}) = \mathcal{O}$ für jedes $\tau \in \mathrm{Aut}(M, K)$. Damit folgt $\tau(L) = L$, denn jedes Element $z \in L$ ist von der Form $g(\alpha_1, \ldots, \alpha_r)$ mit $g \in K[X_1, \ldots, X_r]$. Also erhalten wir durch Einschränkung eine Abbildung $\tau|_L \colon L \to L$, die offenbar in $\mathrm{Aut}(L, K)$ ist; wir erhalten also eine Abbildung

$$\rho \colon \mathrm{Aut}(M, K) \to \mathrm{Aut}(L, K), \qquad \tau \mapsto \tau|_L. \tag{a}$$

Man sieht sofort, dass dies ein Gruppen-Homomorphismus ist, mit

$$\mathrm{Kern}(\rho) = \{\tau \in \mathrm{Aut}(M, K) \mid \tau(z) = z \text{ für alle } z \in L\} = \mathrm{Aut}(M, L). \tag{b}$$

Insbesondere zeigt dies, dass $\mathrm{Aut}(M, L) = \mathrm{Kern}(\rho) \trianglelefteq \mathrm{Aut}(M, K)$ ein Normalteiler ist. Mit dem Homomorphiesatz folgt außerdem, dass die Faktorgruppe $\mathrm{Aut}(M, K)/\mathrm{Aut}(M, L)$ isomorph zu $\mathrm{Bild}(\rho)$, also zu einer Untergruppe von $\mathrm{Aut}(L, K)$ ist.

Lemma 20.11. *Seien* $f, g \in K[X]$ *nicht-konstante Polynome. Sei* $L \supseteq K$ *ein Zerfällungskörper von* f *und* $M \supseteq L$ *ein Zerfällungskörper von* g *(aufgefasst als Polynom in* $L[X]$*). Dann gilt: Ist* $\mathrm{Aut}(M, K)$ *auflösbar, so ist auch* $\mathrm{Aut}(L, K)$ *auflösbar.*

Beweis. Nach Bemerkung 20.10 erhalten wir durch Einschränkung einen Homomorphismus $\rho\colon \mathrm{Aut}(M,K) \to \mathrm{Aut}(L,K)$, $\tau \mapsto \tau|_L$. Nun beachte, dass $M \supseteq K$ auch Zerfällungskörper von $fg \in K[X]$ ist. Also können wir Folgerung 20.6 anwenden: Also lässt sich jeder Automomorphismus in $\mathrm{Aut}(L,K)$ zu einem Automorphismus in $\mathrm{Aut}(M,K)$ fortsetzen. Dies bedeutet, dass ρ surjektiv ist. Sei nun $\mathrm{Aut}(M,K)$ auflösbar. Mit den Bezeichnungen wie in §9 gibt es also ein $k \in \mathbb{N}$ mit $\mathrm{Aut}(M,K)^{(k)} = \{\mathrm{id}_M\}$. Wie im Beweis von Lemma 9.6 folgt dann auch $\mathrm{Aut}(L,K)^{(k)} = \rho\big(\mathrm{Aut}(M,K)^{(k)}\big) = \rho(\{\mathrm{id}_M\}) = \{\mathrm{id}_L\}$, also ist $\mathrm{Aut}(L,K)$ auflösbar. □

Definition 20.12. Sei $M \supseteq K$ eine Körpererweiterung. Wir sagen, dass dies eine ***Radikalerweiterung*** ist, wenn es eine endliche Folge von Elementen $\beta_1, \ldots, \beta_r \in M \setminus \{0\}$ und natürliche Zahlen $n_1, \ldots, n_r \in \mathbb{N}$ gibt mit

$$M = K(\beta_1, \ldots, \beta_r) \qquad \text{und} \qquad \beta_i^{n_i} \in K(\beta_1, \ldots, \beta_{i-1}) \quad \text{für } 1 \leqslant i \leqslant r.$$

Setzen wir $K_0 := K$ und $K_i := K(\beta_1, \ldots, \beta_i)$ für $1 \leqslant i \leqslant r$, so gilt $K_i = K_{i-1}(\beta_i)$, $\beta_i^{n_i} \in K_{i-1}$ und $K_0 \subseteq K_1 \subseteq \ldots \subseteq K_r = M$. Dies sind also genau die Erweiterungen, die in Definition 19.3 vorkommen. Es gilt $[K_i : K_{i-1}] \leqslant n_i$ für $1 \leqslant i \leqslant r$ und damit $[M : K] < \infty$ (Gradsatz).

Mit den obigen Bezeichnungen sei nun $G := \mathrm{Aut}(M,K)$. Wir setzen

$$G_r := \{\mathrm{id}_M\} \qquad \text{und} \qquad G_i := \{\varphi \in G \mid \varphi(z) = z \text{ für alle } z \in K_i\} \quad \text{für } 0 \leqslant i \leqslant r-1.$$

Man sieht sofort, dass jedes G_i eine Untergruppe von G ist; außerdem gilt offenbar $G_r = \{\mathrm{id}_M\} \subseteq G_{r-1} \subseteq \ldots \subseteq G_1 \subseteq G_0 = G$. Der grundlegende Zusammenhang zwischen Radikalerweiterungen und der Auflösbarkeit von Gruppen ist im folgenden Hilfssatz enthalten.

Lemma 20.13. *Sei $m \in \mathbb{N}$ eine natürliche Zahl mit $n_i \mid m$ für $1 \leqslant i \leqslant r$. Wir nehmen an, dass das Polynom $X^m - 1 \in K[X]$ bereits über K in Linearfaktoren zerfällt. Für $1 \leqslant i \leqslant r$ ist dann $G_i \trianglelefteq G_{i-1}$ ein Normalteiler und die Faktorgruppe G_{i-1}/G_i ist abelsch. Folglich ist G eine auflösbare Gruppe (siehe Lemma 9.8).*

Beweis. Sei $i \in \{1, \ldots, r\}$ fest und $c_i := \beta_i^{n_i} \in K_{n-1}$. Sei $f_i := X^{n_i} - c_i \in K_{i-1}[X]$.

Behauptung (1): $K_i = K_{i-1}(\beta_i)$ ist ein Zerfällungskörper von f_i.
Dazu: Sei $L \supseteq K_{i-1}$ ein Zerfällungskörper von f_i. Wegen $f_i(\beta_i) = 0$ können wir $K_i = K_{i-1}(\beta_i) \subseteq L$ annehmen. Sei $\gamma \in L$ eine beliebige Nullstelle von f_i. Dann ist $\gamma^{n_i} = c_i = \beta_i^{n_i}$. Mit $\zeta := \gamma\beta_i^{-1} \in L$ folgt also $\zeta^{n_i} = (\gamma\beta_i^{-1})^{n_i} = 1$ und dann auch $\zeta^m = 1$, d.h., ζ ist eine Nullstelle von $X^m - 1$. Nach Voraussetzung ist folglich $\zeta \in K \subseteq K_i$ und damit $\gamma = \zeta\beta_i \in K_i$. Wegen $L = K_{i-1}(\gamma \in L \mid f_i(\gamma) = 0)$ folgt $L \subseteq K_i$ und damit $K_i = L$.

Behauptung (2): Die Gruppe $\mathrm{Aut}(K_i, K_{i-1})$ ist abelsch.
Dazu: Seien $\varphi, \varphi' \in \mathrm{Aut}(K_i, K_{i-1})$. Es gilt $\varphi(\beta_i)^{n_i} = \varphi(\beta_i^{n_i}) = \varphi(c_i) = c_i$, also ist $\varphi(\beta_i)$

Nullstelle von f_i und damit, wie oben, $\varphi(\beta_i) = \zeta\beta_i$ mit $\zeta \in K \subseteq K_{i-1}$ und $\zeta^m = 1$. Analog ist $\varphi'(\beta_i)^{n_i} = \zeta'\beta_i$ mit $\zeta' \in K \subseteq K_{i-1}$ und $(\zeta')^m = 1$. Es folgt $(\varphi \circ \varphi')(\beta_i) = \varphi\big(\varphi'(\beta_i)\big) = \varphi\big(\zeta'\beta_i\big) = \zeta'\varphi(\beta_i) = \zeta\zeta'\beta_i$ und analog $(\varphi' \circ \varphi)(\beta_i) = \zeta'\zeta\beta_i$. Also ist $(\varphi \circ \varphi')(\beta_i) = (\varphi' \circ \varphi)(\beta_i)$. Wegen $K_i = K_{i-1}(\beta_i)$ ist jeder Automorphismus in $\mathrm{Aut}(K_i, K_{i-1})$ eindeutig festgelegt durch seinen Wert auf β_i. Also folgt $\varphi \circ \varphi' = \varphi' \circ \varphi$. Damit ist auch (2) gezeigt.

Wegen (1) können wir nun Bemerkung 20.10 auf die Erweiterungen $K_{i-1} \subseteq K_i \subseteq M$ anwenden. Es folgt, dass $\mathrm{Aut}(M, K_i) \trianglelefteq \mathrm{Aut}(M, K_{i-1})$ ein Normalteiler ist, und die Faktorgruppe $\mathrm{Aut}(M, K_{i-1})/\mathrm{Aut}(M, K_i)$ isomorph zu einer Untergruppe von $\mathrm{Aut}(K_i, K_{i-1})$ ist, also abelsch wegen (2). Schließlich beachte, dass $G_i = \mathrm{Aut}(M, K_i)$ und $G_{i-1} = \mathrm{Aut}(M, K_{i-1})$ gilt. □

Beweis von Satz 19.7. Sei $f \in K[X]$ nicht-konstant und $L \supseteq K$ ein Zerfällungskörper von f. Wir nehmen an, dass f durch Radikale auflösbar ist, d.h., es gibt eine Radikalerweiterung $M \supseteq K$ mit $L \subseteq M$; wie in Definition 20.12 sei $M = K(\beta_1, \ldots, \beta_r)$ wobei $\beta_i^{n_i} \in K(\beta_1, \ldots, \beta_{i-1})$ für $1 \leqslant i \leqslant r$. Wir wählen ein $m \in \mathbb{N}$ mit $n_i \mid m$ für alle i. Um Lemma 20.13 anwenden zu können, gehen wir wie folgt vor. Sei $\tilde{M} \supseteq M$ ein Zerfällungskörper von $X^m - 1 \in M[X]$. Sei $\zeta \in \tilde{M}$ eine primitive m-te Einheitswurzel. Dann ist $\tilde{M} = M(\zeta)$ (siehe Beispiel 20.3). Setzen wir $\tilde{K} := K(\zeta) \subseteq \tilde{M}$, so gilt dann auch (weil alle Nullstellen von $X^m - 1$ Potenzen von ζ sind):

$$(*) \qquad \tilde{K} = K(\zeta) \supseteq K \text{ ist ein Zerfällungskörper von } X^m - 1 \in K[X].$$

Nun gilt weiterhin $\beta_i^{n_i} \in K(\beta_1, \ldots, \beta_{i-1}) \subseteq \tilde{K}(\beta_1, \ldots, \beta_{i-1})$ für $1 \leqslant i \leqslant r$. Wegen $\tilde{M} = K(\beta_1, \ldots, \beta_r, \zeta) = \tilde{K}(\beta_1, \ldots, \beta_r)$ ist also auch $\tilde{M} \supseteq \tilde{K}$ eine Radikalerweiterung. Um auch Lemma 20.11 benutzen zu können, vergrößern wir $\tilde{M}$ noch weiter. Dazu wenden wir Satz 20.9 auf $\tilde{M} \supseteq \tilde{K}$ an. Sei $g \in \tilde{K}[X]$ das Produkt der Minimalpolynome von $\beta_1, \ldots, \beta_r$ über $\tilde{K}$, und $\tilde{M}' \supseteq \tilde{K}$ ein Zerfällungskörper von $g \in \tilde{K}[X]$ mit $\tilde{M}' \supseteq \tilde{M}$; wir erhalten also

$$\tilde{M}' = \tilde{K}(\tau_j(\beta_i) \mid 1 \leqslant i \leqslant r, 1 \leqslant j \leqslant n) \quad \text{wobei} \quad \mathrm{Aut}(\tilde{M}', \tilde{K}) = \{\tau_1 = \mathrm{id}_{\tilde{M}'}, \tau_2, \ldots, \tau_n\}.$$

Behauptung: $\tilde{M}' \supseteq \tilde{K}$ ist immer noch eine Radikalerweiterung. Dazu betrachte die Folge der erzeugenden Elemente von $\tilde{M}'$ über $\tilde{K}$, wie folgt angeordnet:

$$\beta_1,\ \beta_2,\ \ldots,\ \beta_r,\ \tau_2(\beta_1),\ \tau_2(\beta_2),\ \ldots,\ \tau_2(\beta_r),\quad \ldots,\quad \tau_n(\beta_1),\ \tau_n(\beta_2),\ \ldots,\ \tau_n(\beta_r).$$

Für ein Element $\tau_j(\beta_i)$ in dieser Folge gilt $\tau_j(\beta_i)^{n_i} = \tau_j\big(\beta_i^{n_i}\big) \in \tau_j(\tilde{K}(\beta_1, \ldots, \beta_{i-1})) \subseteq \tilde{K}(\tau_j(\beta_1), \ldots, \tau_j(\beta_{i-1}))$. Also ist $\tau_j(\beta_i)^{n_i}$ im von den strikt vorherigen Elementen der obigen Folge erzeugten Teilkörper enthalten — genau wie in Definition 20.12 verlangt. Außerdem sehen wir, dass die gewählte natürliche Zahl m auch weiter die Annahme in Lemma 20.13 bezüglich der Radikalerweiterung $\tilde{M}' \supseteq \tilde{K}$ erfüllt. Also folgt:

$$\mathrm{Aut}(\tilde{M}', \tilde{K}) \quad \text{ist eine auflösbare Gruppe.}$$

Nachdem dies gezeigt ist, gehen wir in drei Schritten von $\tilde{M}' \supseteq \tilde{K}$ auf $L \supseteq K$ zurück. Sei zunächst $\tilde{L} := L(\zeta) \subseteq \tilde{M}'$. Wegen $(*)$ ist $\tilde{L} \supseteq K$ Zerfällungskörper von $(X^m - 1)f \in K[X]$. Trivialerweise ist dann $\tilde{L} \supseteq \tilde{K}$ auch Zerfällungskörper von $(X^m - 1)f \in \tilde{K}[X]$. Mit Lemma 20.11 (angewandt auf $\tilde{K} \subseteq \tilde{L} \subseteq \tilde{M}'$) folgt, dass $\mathrm{Aut}(\tilde{L}, \tilde{K})$ auflösbar ist. Sodann können wir Bemerkung 20.10 wegen $(*)$ auf $K \subseteq \tilde{K} \subseteq \tilde{L}$ anwenden. Es folgt, dass $\mathrm{Aut}(\tilde{L}, \tilde{K}) \trianglelefteq \mathrm{Aut}(\tilde{L}, K)$ ein Normalteiler ist, mit Faktorgruppe isomorph zu einer Untergruppe von $\mathrm{Aut}(\tilde{K}, K)$. Letztere ist abelsch nach Bemerkung 20.3; da auch der Normalteiler $\mathrm{Aut}(\tilde{L}, \tilde{K})$ auflösbar ist, folgt mit Lemma 9.6, dass $\mathrm{Aut}(\tilde{L}, K)$ auflösbar ist. Schließlich können wir Lemma 20.11 auch auf $K \subseteq L \subseteq \tilde{L}$ anwenden und es folgt, dass $\mathrm{Aut}(L, K)$ auflösbar ist. □

Bemerkung 20.14. Sei $f \in K[X]$ nicht-konstant und $L \supseteq K$ Zerfällungskörper von f.

(a) Die Umkehrung von Satz 19.7 gilt auch, d.h., ist $\mathrm{Aut}(L, K)$ auflösbar, so ist f durch Radikale auflösbar. Aber dazu muss vorausgesetzt werden, dass $\mathrm{char}(K) = 0$ gilt (was in obiger Diskussion nicht nötig war); siehe §30.3 im Buch von Karpfinger–Meyberg [KM].

(b) Sei f durch Radikale auflösbar und $M \supseteq K$ eine Radikalerweiterung mit $L \subseteq M$, wie in Definition 20.12. Ist die Annahme (bezüglich m-ter Einheitswurzeln in K) von Lemma 20.13 erfüllt, so gilt sogar, dass $L \supseteq K$ selbst bereits eine Radikalerweiterung ist; siehe §6 im Artikel von Rosen [Ro]. Aber der Beweis ist nicht unbedingt kürzer als das obige Argument, in dem von $\tilde{M}$ zu einer noch größeren Radikalerweiterung $\tilde{M}' \supseteq \tilde{K}$ übergegangen wird.

(c) Im Allgemeinen ist $L \supseteq K$ selbst keine Radikalerweiterung. Ein Beispiel ist gegeben durch den Zerfällungskörper $L \supseteq \mathbb{Q}$ von $f = X^3 - 3X + 1 \in \mathbb{Q}[X]$. Für den Beweis (Selbststudium) wird die folgende, auch an sich nützliche Charakterisierung von Zerfällungskörpern benötigt.

Satz 20.15. *Sei $L \supseteq K$ eine Körpererweiterung mit $[L : K] < \infty$. Genau dann ist L ein Zerfällungskörper eines nicht-konstanten Polynoms $f \in K[X]$, wenn für jedes $\alpha \in L$ das Minimalpolynom $\mu_\alpha \in K[X]$ in Linearfaktoren über L zerfällt.*

Beweis. Sei $L = K(\alpha_1, \ldots, \alpha_r)$ mit $\alpha_1, \ldots, \alpha_r \in L$. Für $1 \leqslant i \leqslant r$ sei $f_i := \mu_{\alpha_i} \in K[X]$ das Minimalpolynom von α_i. Wenn wir annehmen, dass jedes f_i bereits über L in Linearfaktoren zerfällt, so ist $L \supseteq K$ Zerfällungskörper des Polynoms $f := f_1 \cdots f_r \in K[X]$.
Umgekehrt sei $L \subseteq K$ Zerfällungskörper eines nicht-konstanten Polynoms $f \in K[X]$. Sei $\alpha \in L$ beliebig und $M \supseteq L$ Zerfällungskörper von μ_α (aufgefasst als Polynom in $L[X]$). Dann ist $M \supseteq K$ auch Zerfällungskörper des Polynoms $f\mu_\alpha \in K[X]$. Sei $\beta \in M$ eine beliebige Nullstelle von μ_α. Nach Lemma 20.1 gibt es einen Isomorphismus $\sigma\colon K(\alpha) \to K(\beta)$. Nach Folgerung 20.6 gibt es ein $\tau \in \mathrm{Aut}(M, K)$ mit $\tau|_L = \sigma$, also $\tau(\alpha) = \beta$. Nach Bemerkung 20.10 ist $\tau(L) \subseteq L$, also $\beta = \tau(\alpha) \in L$. Also zerfällt μ_α bereits über L in Linearfaktoren. □

21. *Anhang B: Existenz eines algebraischen Abschlusses*

In diesem Abschnitt beschäftigen wir uns mit der Frage, wie man zu einem gegebenem Körper K einen algebraischen Abschluss konstruieren kann, also eine algebraische Erweiterung $L \supseteq K$, so dass *jedes* nicht-konstante Polynom in $L[X]$ über L in Linearfaktoren zerfällt. Um dies in aller Allgemeinheit zu zeigen, wird das ***Lemma von Zorn*** benötigt, was die ganze Angelegenheit mengentheoretisch anspruchsvoll macht[12]. Wir wollen hier diejenigen Teile des Beweises ausführlich behandeln, die nichts mit dem Lemma von Zorn zu tun haben. Als Nebenprodukt erhalten wir das Ergebnis, dass man auf das Lemma von Zorn verzichten kann, wenn man nur Körper K mit höchstens abzählbar vielen Elementen betrachtet. Zusammen mit dem Fundamentalsatz der Algebra ist dies für viele Zwecke völlig ausreichend. Wir beginnen mit einer allgemeinen Vorbetrachtung über Polynome.

In Beispiel 13.12 haben wir gesehen, wie man den Polynomring in Unbestimmten $X_1, \ldots, X_n$ (mit $n \geqslant 2$) rekursiv aus Polynomringen mit jeweils einer Unbestimmten konstruieren kann. Wir stellen nun eine allgemeine Konstruktion von Polynomringen in beliebig vielen Unbestimmten vor. Sei dazu I eine nicht-leere Menge. (Dies wird als Indexmenge für die zu konstruierenden Unbestimmten dienen.) Sei $\mathfrak{X}(I)$ die Menge aller Funktionen $\xi\colon I \to \mathbb{N}_0$ mit $|\{i \in I \mid \xi(i) \neq 0\}| < \infty$. Für $\xi, \xi' \in \mathfrak{X}(I)$ definieren wir ein Produkt durch

$$(\xi \cdot \xi')(i) := \xi(i) + \xi'(i) \qquad \text{für alle } i \in I.$$

Dann ist auch $\xi \cdot \xi' \in \mathfrak{X}(I)$ und man sieht sofort, dass dieses Produkt assoziativ ist; die Abbildung $1_I\colon I \to \mathbb{N}_0$, $i \mapsto 0$, ist neutrales Element bezüglich dieser Verknüpfung. Für festes $i \in I$ sei $\xi_i \in \mathfrak{X}$ definiert durch $\xi_i(i) := 1$ und $\xi_i(j) := 0$ für $j \in I \setminus \{i\}$. Dann lässt sich jedes $1_I \neq \xi \in \mathfrak{X}(I)$ auf eindeutige Weise schreiben als

$$\xi = \xi_{i_1}^{n_1} \cdots \xi_{i_r}^{n_r} \qquad \text{mit } r \geqslant 1,\ i_1, \ldots, i_r \in I \text{ und } n_1, \ldots, n_r \in \mathbb{N}. \tag{1}$$

[Denn sei $\{j \in I \mid \xi(j) \neq 0\} = \{i_1, \ldots, i_r\}$ mit $r \geqslant 1$ und verschiedenen $i_1, \ldots, i_r \in I$. Mit $n_j := \xi(j) \in \mathbb{N}$ für $1 \leqslant j \leqslant r$ folgt $\xi = \xi_{i_1}^{n_1} \cdots \xi_{i_r}^{n_r}$. Umgekehrt sind $i_1, \ldots, i_r$ und $n_1, \ldots, n_r$ hierdurch eindeutig bestimmt.]

Die Elemente in $\mathfrak{X}(I)$ werden als ***Monome über*** I bezeichnet.

Sei nun R ein kommutativer Ring mit 1. Für eine beliebige Funktion $f\colon \mathfrak{X}(I) \to R$ setzen wir $\operatorname{supp}(f) := \{\xi \in \mathfrak{X}(I) \mid f(\xi) \neq 0\}$. Wir definieren $R[I]$ als die Menge aller $f\colon \mathfrak{X}(I) \to R$ mit $|\operatorname{supp}(f)| < \infty$. Für $f, g \in R[I]$ definieren wir eine Summe und ein Produkt durch

$$(f+g)(\xi) := f(\xi) + g(\xi) \qquad \text{und} \qquad (f \cdot g)(\xi) := \sum_{\xi_1, \xi_2 \in \mathfrak{X}(I) \text{ mit } \xi_1 \cdot \xi_2 = \xi} f(\xi_1) g(\xi_2).$$

[12] Für eine Herleitung des Lemmas von Zorn aus den üblichen Axiomen der Zermelo–Frenkel Mengentheorie sowie dem Auswahlaxiom siehe etwa §16 im Buch von Halmos [Ha].

Beachte: Zu gegebenem $\xi \in \mathfrak{X}(I)$ gibt es wegen der Eindeutigkeit in (1) nur endlich viele $\xi_1, \xi_2 \in \mathfrak{X}(I)$ mit $\xi = \xi_1 \cdot \xi_2$; also ist die Summe in der Definition von $f \cdot g$ endlich. Man sieht dann leicht, dass auch $\mathrm{supp}(f+g)$ und $\mathrm{supp}(f \cdot g)$ endliche Mengen sind, also gilt $f + g \in R[I]$ und $f \cdot g \in R[I]$. Man muss dann nachrechnen, dass mit diesen Verknüpfungen $R[I]$ ein kommutativer Ring ist. (Dies ist etwas mühsam, aber der Nachweis der einzelnen Ring-Axiome ist jeweils eine Routine-Angelegenheit.) Dieser Ring besitzt ein Eins-Element, nämlich $1_{R[I]} \in R[I]$, definiert durch $1_{R[I]}(1_I) := 1_R$ und $1_{R[i]}(\xi) := 0$ für $\xi \in \mathfrak{X}(I) \setminus \{1_I\}$.

Wir definieren außerdem eine skalare Multiplikation von $f \in R[I]$ mit $a \in R$ durch $(a \cdot f)(\xi) := af(\xi)$ für alle $\xi \in \mathfrak{X}(I)$. Dann gelten die üblichen Distributivregeln; außerdem $0_R \cdot f = 0_{R[I]}$ und $1_R \cdot f = f$ für alle $f \in R[I]$. Damit können wir R als Teilring von $R[I]$ auffassen, indem wir $a \in R$ mit $a \cdot 1_{R[I]}$ identifizieren. Für $\xi \in \mathfrak{X}(I)$ definiere $X_\xi \in R[I]$ durch $X_\xi(\xi) := 1_R$ und $X_\xi(\xi') := 0$ für $\xi' \in \mathfrak{X}(I) \setminus \{\xi\}$. Damit können wir jedes $0_{R[I]} \neq f \in R[I]$ auf eindeutige Weise schreiben als endliche Summe

$$f = \sum_{\xi \in \mathrm{supp}(f)} a_\xi \cdot X_\xi \qquad \text{mit} \qquad 0 \neq a_\xi \in R \text{ für alle } \xi \in \mathrm{supp}(f). \tag{2}$$

[Denn sei $a_\xi := f(\xi) \in R$ für alle $\xi \in \mathrm{supp}(f)$. Setze dann $f' := \sum_{\xi \in \mathrm{supp}(f)} a_\xi \cdot X_\xi \in R[I]$. Für $\xi' \in \mathfrak{X}(I)$ beliebig folgt $f'(\xi') = 0$ falls $\xi' \notin \mathrm{supp}(f)$; für $\xi' \in \mathrm{supp}(f)$ erhält man genau $a_{\xi'} = f(\xi')$. Also ist $f = f'$.]

Speziell für $\xi = \xi_i$ mit $i \in I$ (wie oben definiert) schreiben wir $X_i := X_{\xi_i}$. Stellen wir ein beliebiges $\xi \in \mathfrak{X}(I)$ wie oben in (1) dar, so folgt aus der Definition der Multiplikation in $R[I]$:

$$X_\xi = X_{i_1}^{n_1} \cdots X_{i_r}^{n_r}. \tag{3}$$

[Denn: Sei $n_r > 0$. Dann gilt $\xi = \xi' \cdot \xi_{i_r}$ wobei $\xi' := \xi_{i_1}^{n_1} \cdots \xi_{i_{r-1}}^{i_{r-1}} \cdot \xi_{i_r}^{i_r - 1} \in \mathfrak{X}(I)$. Damit folgt

$$(X_{\xi'} \cdot X_{i_r})(\zeta) = \sum_{\zeta_1, \zeta_2 \in \mathfrak{X}(I) \text{ mit } \zeta_1 \cdot \zeta_2 = \zeta} X_{\xi'}(\zeta_1) X_{i_r}(\zeta_2).$$

Ist $\zeta_1 = \xi'$ und $\zeta_2 = \xi_{i_r}$, so gilt $\zeta = \zeta_1 \cdot \zeta_2 = \xi' \cdot \xi_{i_r} = \xi$ und $X_{\xi'}(\zeta_1) = X_{\xi_{i_r}}(\zeta_2) = 1$. Dieses Paar (ζ_1, ζ_2) liefert also den Beitrag 1 zu obiger Summe. Seien nun $\zeta_1, \zeta_2 \in \mathfrak{X}(I)$ beliebig, so dass der zugehörige Term in der obigen Summe $\neq 0$ ist. Dann folgt $X_{\xi'}(\zeta_1) \neq 0$, also $\zeta_1 = \xi'$; analog folgt $X_{i_r}(\zeta_2) \neq 0$, also $\zeta_2 = \xi_{i_r}$. Also gilt $(X_{\xi'} \cdot X_{i_r})(\xi) = 1$ und $(X_{\xi'} \cdot X_{i_r})(\zeta) = 0$ für $\zeta \in \mathfrak{X}(I) \setminus \{\xi\}$. Es folgt $X_{\xi'} \cdot X_{i_r} = X_\xi$. Ist $n_r = 0$, so fahren wir entsprechend mit n_{r-1} fort, dann mit n_{r-2} und so weiter, bis alle Terme abgearbeitet sind.]

Durch Kombination von (2) und (3) können wir jedes $f \in R[I]$ als endliche Summe der Form

$$f = \sum a_{i_1, \ldots, i_r} X_{i_1}^{n_1} \cdots X_{i_r}^{n_r} \qquad (a_{i_1, \ldots, i_r} \in R)$$

schreiben, wobei $r \geqslant 0$ beliebig, $i_1, \ldots, i_r \in I$ und $n_1, \ldots, n_r \in \mathbb{N}_0$. Für eine Teilmenge $I' \subseteq I$ erhalten wir einen Teilring von $R[I]$, der aus allen f wie in (3) besteht, wobei $a_{i_1, \ldots, i_r} \neq 0$ nur für $i_1, \ldots, i_r \in I'$ gilt; diesen Teilring können wir einfach mit $R[I']$ identifizieren.

Wir bezeichnen $R[I]$ als den Ring der Polynome in den Unbestimmten X_i ($i \in I$) und mit Koeffizienten in R. Beachte: Ist $I = \{i\}$ eine 1-elementige Menge, so ist $R[I] = R[X_i]$ der Polynomring in einer Unbestimmten $X := X_i$; für $I = \{1, \ldots, n\}$ mit $n \geqslant 1$ ist $R[I] = R[X_1, \ldots, X_n]$, wie in Beispiel 13.12. Wie in diesen beiden Fällen gilt auch für $R[I]$, mit I beliebig, eine "universelle Eigenschaft": Sei S ein kommutativer Ring mit 1 und $\varphi\colon R \to S$ ein Ring-Homomorphismus. Sei $\underline{y} = \{y_i \mid i \in I\} \subseteq S$ eine feste Familie von Elementen in S. Dann gibt es, völlig analog zu Satz 12.6, einen ***Einsetzungs-Homomorphismus***

$$\tilde{\varphi}_{\underline{y}}\colon R[I] \to S \qquad \text{mit} \qquad \tilde{\varphi}_{\underline{y}}(a) = \varphi(a) \ \text{(für } a \in R) \quad \text{und} \quad \tilde{\varphi}_{\underline{y}}(X_i) = y_i \ \text{(für } i \in I).$$

Es wird also $f = \sum a_{i_1,\ldots,i_r} X_{i_1}^{n_1} \cdots X_{i_r}^{n_r}$ (wie oben) abgebildet auf $\sum \varphi(a_{i_1,\ldots,i_r}) y_{i_1}^{n_1} \cdots y_{i_r}^{n_r} \in S$.

Mit Hilfe der obigen allgemeinen Konstruktion können wir nun zeigen:

Lemma 21.1 (Artin)**.** *Es gibt einen kommutativen Ring R mit 1, so dass $K \subseteq R$ ein Teilring ist und jedes nicht-konstante Polynom in $K[X]$ eine Nullstelle in R hat.*

Beweis. Sei $I := K[X]^{\#}$ die Menge der nicht-konstanten Polynome in $K[X]$. Wir bilden den Polynomring $K[I]$. Für $f \in I$ setze X_f in f ein, erhalte also $f(X_f) \in K[I]$. Dann sei $J \trianglelefteq K[I]$ das von allen $f(X_f)$ mit $f \in I$ erzeugte Ideal, also die Menge aller endlichen Summen $\sum_{i=1}^{r} g_i f_i(X_{f_i})$ mit $r \geqslant 1$, $g_i \in K[I]$ und $f_i \in I$ (siehe auch Ü7A4). Behauptung: $J \neq K[I]$. Dazu: Wäre $J = K[I]$, so $1_K \in J$, also gäbe es $g_1, \ldots, g_r \in K[I]$ (mit $r \geqslant 1$) und $f_1, \ldots, f_r \in I$ mit

$$1 = g_1 f_1(X_{f_1}) + \ldots + g_r f_r(X_{f_r}).$$

Sei $L \supseteq K$ ein Zerfällungskörper des Polynoms $0 \neq f_1 \cdots f_r \in K[X]$. Es gibt also $\alpha_i \in L$ mit $f_i(\alpha_i) = 0$ für $1 \leqslant i \leqslant r$. Sei $\varphi\colon K \hookrightarrow L$ die Inklusionsabbildung. Wir definieren eine Familie $\underline{y} = \{y_f \mid f \in I\} \subseteq L$ durch $y_{f_i} := \alpha_i$ für $1 \leqslant i \leqslant r$ und $y_f := 0$ für $f \in I \setminus \{f_1, \ldots, f_r\}$. Wenden wir den Einsetzungs-Homomorphismus $\tilde{\varphi}_{\underline{y}}\colon K[I] \to L$ auf obige Gleichung an, so erhalten wir

$$\begin{aligned} 1 &= \tilde{\varphi}_{\underline{y}}(g_1)\tilde{\varphi}_{\underline{y}}(f_1(X_{f_1})) + \ldots + \tilde{\varphi}_{\underline{y}}(g_r)\tilde{\varphi}_{\underline{y}}(f_r(X_{f_r})) \\ &= \tilde{\varphi}_{\underline{y}}(g_1)\tilde{\varphi}_{\underline{y}}(\underbrace{f_1(\alpha_1)}_{=0}) + \ldots + \tilde{\varphi}_{\underline{y}}(g_r)\tilde{\varphi}_{\underline{y}}(\underbrace{f_r(\alpha_r)}_{=0}) = 0, \end{aligned}$$

Widerspruch. Also ist $J \subsetneqq K[I]$ und wir bilden den Faktorring $R := K[I]/J$; für $a \in K[I]$ schreiben wir $\bar{a} := a + J \in R$. Wegen $J \neq K[I]$ ist die Einschränkung $K \to R$, $x \mapsto \bar{x}$, injektiv; vermöge dieser Abbildung können wir K als Teilring von R auffassen (indem wir also einfach $x \in K$ mit $\bar{x} \in R$ identifizieren). Sei nun $f \in I = K[X]^{\#}$. Dann können wir $\alpha := \bar{X}_f \in R$ in f einsetzen und erhalten $f(\alpha) = f(\bar{X}_f) = \overline{f(X_f)} = \bar{0}$, weil $f(X(f)) \in J$. Also hat f tatsächlich eine Nullstelle in R — aber nach Konstruktion ist eben R nur ein Ring, und kein Körper. □

Eine geschickte Modifikation des obigen Argumentes liefert folgende stärkere Aussage[13].

Lemma 21.2 (Conrad)**.** *Es gibt einen kommutativen Ring* R *mit* 1*, so dass* $K \subseteq R$ *ein Teilring ist und jedes nicht-konstante Polynom in* $K[X]$ *in Linearfaktoren über* R *zerfällt.*

Beweis. Sei wieder $K[X]^{\#}$ die Menge der nicht-konstanten Polynome in $K[X]$. Für $f \in K[X]^{\#}$ sei $n_f := \mathrm{Grad}(f) \geqslant 1$ und $0 \neq a_{n_f} \in K$ der Leitkoeffizient von f. Wir bilden den Polynomring

$$A := K[I] \qquad \text{mit} \qquad I := \{(f,j) \mid f \in K[X]^{\#} \text{ und } 1 \leqslant j \leqslant n_f\}.$$

Diesmal gehen wir noch einen Schritt weiter und betrachten den Polynomring $A[X]$ in einer Unbestimmten X über A. Für ein beliebiges $f \in K[X]^{\#} \subseteq A[X]$ schreibe dann

$$f - a_{n_f} \prod_{j=1}^{n_f} (X - X_{f,j}) = \sum_{d=0}^{n_f-1} c_d(f) X^d \quad \in A[X] \qquad \text{mit} \qquad c_d(f) \in A.$$

Sei $J \trianglelefteq A$ das Ideal, das von allen $c_d(f)$ mit $f \in K[X]^{\#}$ und $0 \leqslant d \leqslant n_f - 1$ erzeugt wird. Behauptung: $J \neq A$. Dazu: Wäre $J = A$, so $1 \in J$, also gäbe es eine Gleichung

$$(*) \qquad 1 = g_1 c_{d_1}(f_1) + \ldots + g_r c_{d_r}(f_r) \qquad \text{mit } r \geqslant 1,$$

wobei $g_i \in A$, $f_i \in K[X]^{\#}$ und $d_i \in \mathbb{N}$ mit $0 \leqslant d_i \leqslant n_{f_i} - 1$. Sei $L \supseteq K$ ein Zerfällungskörper des Polynoms $0 \neq f_1 \cdots f_r \in K[X]$. Für jedes $i \in \{1, \ldots, r\}$ gilt dann also

$$f_i = a_{n_{f_i}} \prod_{j=1}^{n_{f_i}} (X - \alpha_{ij}) \qquad \text{mit } \alpha_{ij} \in L \text{ für } 1 \leqslant j \leqslant n_{f_i}.$$

Sei $I_0 := \{f_i, j) \mid 1 \leqslant i \leqslant r \text{ und } 1 \leqslant j \leqslant n_{f_i}\} \subseteq I$. Sei $\varphi\colon K \hookrightarrow L$ die Inklusionsabbildung. Wir definieren eine Familie $\underline{y} = \{y_{(f,j)} \mid (f,j) \in I\} \subseteq L$ durch $y_{(f_i,j)} := \alpha_{ij}$ für $(i,j) \in I_0$ und $y_{(f,j)} := 0$ für $(f,j) \in I \setminus I_0$. Sei $\tilde{\varphi}_{\underline{y}}\colon A \to L$ der zugehörige Einsetzungs-Homomorphismus. Mit Hilfe dieses Homomorphismus erhalten wir dann auch einen Einsetzungs-Homomorphismus

$$\Phi\colon A[X] \to L[X],$$

welcher $\tilde{\varphi}_{\underline{y}}$ auf die Koeffizienten eines Polynoms in $A[X]$ anwendet; insbesondere ist $\Phi(g) = g$ für alle $g \in K[X]$. Damit erhalten wir

$$\Phi\Big(a_{n_{f_i}} \prod_{j=1}^{n_{f_i}} (X - X_{f_i,j})\Big) = a_{n_{f_i}} \prod_{j=1}^{n_{f_i}} (X - \alpha_{ij}) = f_i = \Phi(f_i) \qquad \text{für } 1 \leqslant i \leqslant r.$$

Daraus folgt dann aber

$$\sum_{d=0}^{n_{f_i}-1} \tilde{\varphi}_{\underline{y}}\big(c_{d_i}(f_i)\big) X^d = \Phi\Big(\sum_{d=0}^{n_{f_i}-1} c_{d_i}(f_i) X^d\Big) = \Phi\Big(f_i - a_{n_{f_i}} \prod_{j=1}^{n_{f_i}} (X - X_{f_i,j})\Big) = 0,$$

[13] Siehe das "expository paper" zu "Constructing algebraic closures, I" auf Keith Conrad's Webseite **https://kconrad.math.uconn.edu/blurbs**; siehe auch Bourbaki [Bo, Chap. V, §4].

also $\tilde{\varphi}_{\underline{y}}\big(c_{d_i}(f_i)\big) = 0$ für $1 \leqslant i \leqslant r$. Wenden wir nun $\tilde{\varphi}_{\underline{y}}$ auf $(*)$ an, so erhalten wir

$$1 = \tilde{\varphi}_{\underline{y}}(g_1)\underbrace{\tilde{\varphi}_{\underline{y}}(c_{d_1}(f_1))}_{=0} + \ldots + \tilde{\varphi}_{\underline{y}}(g_r)\underbrace{\tilde{\varphi}_{\underline{y}}(c_{d_r}(f_r))}_{=0} = 0,$$

Widerspruch. Also ist $J \subsetneqq A$ und wir bilden wieder den Faktorring $R := A/J$; für $a \in A$ schreiben wir $\bar{a} := a + J \in R$. Wegen $J \neq A$ ist die Einschränkung $K \to R$, $x \mapsto \bar{x}$, injektiv; vermöge dieser Abbildung können wir K als Teilring von R auffassen (indem wir also einfach $x \in K$ mit $\bar{x} \in R$ identifizieren). Wir erhalten dann auch einen (Einsetzungs-) Homomorphismus $A[X] \to R[X]$, bei dem $a \mapsto \bar{a}$ auf die Koeffizienten eines Polynoms in $A[X]$ angewendet wird. Sei nun $f \in K[X]^{\#} \subseteq R[X]$. In $R[X]$ erhalten wir dann

$$f - a_{n_f}\prod_{j=1}^{n_f}(X - \overline{X}_{f,j}) = \sum_{d=0}^{n_f-1}\overline{c_d(f)}X^d = \bar{0};$$

also hat f nicht nur eine Nullstelle in R, sondern zerfällt vollständig in Linearfaktoren über R, mit Nullstellen $\bar{X}_{f,j} \in R$ für $1 \leqslant j \leqslant n_f$. □

Bemerkung 21.3. Ist K endlich oder abzählbar, so sind die in Lemma 21.1 und 21.2 konstruierten Ringe $R \supseteq K$ ebenfalls abzählbar.

Dazu: Sei S ein endlicher oder abzählbarer kommutativer Ring mit 1 und $S[X]^{\#}$ die Menge der nicht-konstanten Polynome in $S[X]$. Mit einem Zählargument völlig analog zu Beispiel 15.9 folgt, dass $S[X]^{\#}$ abzählbar ist. Wegen $S[X] = S \cup S[X]^{\#}$ ist dann auch $S[X]$ abzählbar. Mit Induktion nach n folgt, dass $S[X_1, \ldots, X_n]$ abzählbar ist für alle $n \in \mathbb{N}$.
Nun sei K endlich oder abzählbar. Dann ist $I := K[X]^{\#}$ wie im Beweis von Lemma 21.1 also ebenfalls abzählbar. Die Menge I im Beweis von Lemma 21.2 ist enthalten in $K[X]^{\#} \times \mathbb{N}$, also auch abzählbar. In beiden Fällen gibt es also eine Aufzählung $I = \{i_n \mid n \in \mathbb{N}\}$. Dann ist $K[I] = K[X_n \mid n \in \mathbb{N}]$, wobei $X_n := X_{i_n}$ für alle $n \in \mathbb{N}$. Für $n \in \mathbb{N}$ sei $P_n := K[X_1, \ldots, X_n] \subseteq K[I]$ der Polynomring in den Unbestimmten $X_1, \ldots, X_n$. Nun ist jedes $0 \neq a \in K[I]$ eine endliche Summe wie in (2); schreiben wir die Terme X_ξ wie in (3), so sehen wir, dass jeder dieser Terme in P_n liegt für ein $n \in \mathbb{N}$. Wählen wir n groß genug, so liegen alle Terme in P_n, also auch $a \in P_n$. Dies zeigt, dass $K[I] = \bigcup_{n\in\mathbb{N}} P_n$ gilt. Als abzählbare Vereinigung von abzählbaren Mengen ist auch $K[I]$ abzählbar. Schließlich ist auch $R = K[I]/J$ abzählbar, denn es gibt eine surjektive Abbildung $K[I] \to R$.

Satz 21.4 (Satz von Steinitz für abzählbare Körper)**.** *Ist* K *ein endlicher oder abzählbarer Körper, so gibt es einen algebraischen Abschluss* $L \supseteq K$*.*

Beweis. Sei $R \supseteq K$ wie in Lemma 21.2. Da R abzählbar ist (siehe Bemerkung 21.3), gibt es eine Aufzählung $R = \{a_n \mid n \in \mathbb{N}\}$. Wir definieren rekursiv eine Folge von Idealen $\{J_n \mid n \in \mathbb{N}_0\}$

in R wie folgt[14]: Sei $J_0 := \{0\}$; für $n \geqslant 1$ sei dann

$$J_n := \begin{cases} J_{n-1} + (a_n) & \text{falls } R \neq J_{n-1} + (a_n), \\ J_{n-1} & \text{sonst.} \end{cases}$$

Im ersten Fall ist J_n eine Summe von zwei Idealen, also selbst ein Ideal. Außerdem gilt nach Konstruktion $\{0\} = J_0 \subseteq J_1 \subseteq J_2 \subseteq \ldots$; daraus folgt sofort, dass die Vereinigung $J := \bigcup_{n\in\mathbb{N}_0} J_n \subseteq R$ ein Ideal ist. Behauptung: $M := R/J$ ist ein Körper.

Dazu: Wäre $J = R$, so $1 \in J$, also $1 \in J_n$ für ein $n \in \mathbb{N}$; sei n minimal mit dieser Eigenschaft. Dann ist $J_n \neq J_{n-1}$, also $J_n = J_{n-1} + (a_n) \neq R$. Wegen $1 \in J_n$ ist aber $R = J_n$, Widerspruch. Also gilt $J \subsetneqq R$. Wir müssen nun zeigen, dass $M^\times = M \setminus \{0\}$ gilt. Für $a \in R$ schreibe $\bar{a} := a + J \in M$. Sei nun $\bar{a} \neq \bar{0}$, also $a \in R \setminus J$. Dann ist $a = a_n$ für ein $n \in \mathbb{N}$. Wäre $R \neq J_{n-1} + (a_n)$, so $a = a_n \in J_n = J_{n-1} + (a_n) \subseteq J$, Widerspruch zur Annahme. Also ist $R = J_{n-1} + (a_n)$, d.h., es gibt ein $b \in J_{n-1} \subseteq J$ und ein $c \in R$ mit $1 = b + ca_n = b + ca$. Dann ist $\bar{1} = \bar{c}\bar{a}$, also $\bar{a} \in M^\times$.

Wegen $J \neq R$ ist die Einschränkung $K \to M$, $x \mapsto \bar{x}$, wieder injektiv, also können wir K als Teilkörper von M auffassen. Sei $f \in K[X]$ nicht-konstant mit $n = \mathrm{Grad}(f) \geqslant 1$. In $R[X]$ gilt dann $f = c\prod_{i=1}^{n}(X - a_i)$ mit $0 \neq c \in K$ und $a_i \in R$. Wiederum haben wir auch den (Einsetzungs-)Homomorphismus $R[X] \to M[X]$, bei dem $a \mapsto \bar{a}$ auf die Koeffizienten eines Polynoms in $R[X]$ angewendet wird. Damit erhalten wir $f = c\prod_{i=1}^{n}(X - \bar{a}_i) \in M[X]$, d.h.. f zerfällt in Linearfaktoren über M. Sei nun $L := \{\alpha \in M \mid \alpha \text{ algebraisch über } K\}$. Nach Folgerung 16.7 ist $L \subseteq M$ ein Teilkörper, also $L \supseteq K$ eine algebraische Erweiterung.

Dann folgt aber auch, dass L algebraisch abgeschlossen ist. Denn sei $f \in L[X]$ nicht-konstant und $f_1 \in L[X]$ ein normierter irreduzibler Faktor von f; wir müssen zeigen, dass $\mathrm{Grad}(f_1) = 1$ gilt. Dazu sei $L_1 \supseteq L$ ein Zerfällungskörper von f_1 und $\alpha \in L_1$ ein Element mit $f_1(\alpha) = 0$. Da $L_1 \supseteq L$ und $L \supseteq K$ algebraisch sind, ist auch $L_1 \supseteq K$ algebraisch. Sei $0 \neq g \in K[X]$ das Minimalpolynom von α. Fassen wir g als Polynom in $L[X]$ auf, so folgt $f_1 \mid g$ in $L[X]$, denn f_1 ist das Minimalpolynom von α über L. Wie wir oben gesehen haben, zerfällt jedes nicht-konstante Polynom in $K[X]$ in Linearfaktoren über $M \supseteq L$. Dies gilt insbesondere für g. Jede Nullstelle von g in M ist aber algebraisch über K, also in L enthalten (nach Definition von L). Also zerfällt g in Linearfaktoren über L, d.h., jeder irreduzible Faktor von g in $L[X]$ hat Grad 1; also ist auch $\mathrm{Grad}(f_1) = 1$. Also ist $L \supseteq K$ ein algebraischer Abschluss von K. □

[14] Für die Fans der Mengentheorie: Eine derartige rekursive Definition über alle $n \in \mathbb{N}$ mag zwar intuitiv klar erscheinen, benötigt aber streng genommen auch eine formale Rechtfertigung; diese wird durch den Rekursionssatz von Dedekind geleistet; siehe §12 im Buch von Halmos [Ha].

Beispiel 21.5. (a) Sei p eine Primzahl und $K = \mathbb{F}_p$. Da K endlich ist, liefert der obige Satz einen algebraischen Abschluß $L \supseteq \mathbb{F}_p$. Sei $F\colon L \to L$, $z \mapsto z^p$, der Frobenius-Homomorphismus (siehe Lemma 15.2). Für jedes $n \in \mathbb{N}$ ist dann auch $F^n\colon L \to L$, $z \mapsto z^{p^n}$, ein Homomorphismus. Wie im Beweis des Hauptsatzes über endliche Körper sieht man, dass $\mathbb{F}_{p^n} := \{z \in L \mid z^{p^n} = z\} \subseteq L$ ein Teilkörper mit p^n Elementen ist. Es gilt sogar

$$L = \bigcup_{n\in\mathbb{N}} \mathbb{F}_{p^n}.$$

Denn sei $z \in L$ beliebig. Dann ist z algebraisch über $\mathbb{F}_p$, also ist $\mathbb{F}_p(z) \supseteq \mathbb{F}_p$ eine endliche Erweiterung (Satz 15.5), und damit $\mathbb{F}_p(z)$ ein endlicher Körper; sei $|\mathbb{F}_p(z)| = p^n$ mit $n \geqslant 1$. Nach Lagrange ist jedes Element von $\mathbb{F}_p(z)$ eine Nullstelle von $X^{p^n} - X \in \mathbb{F}_p[X]$. Dieses Polynom hat aber nur höchstens p^n Nullstellen in L. Da bereits alle Elemente von $\mathbb{F}_{p^n}$ Nullstellen dieses Polynoms sind, folgt also $z \in \mathbb{F}_p(z) = \mathbb{F}_{p^n}$.

(b) Sei K_0 ein endlicher oder abzählbarer Körper, also etwa ein beliebiger endlicher Körper oder ein Erweiterungskörper von $\mathbb{Q}$ mit $[K_0 : \mathbb{Q}] < \infty$. Sei $n \geqslant 1$ und $R := K_0[X_1, \ldots, X_n]$ der Polynomring in den Unbestimmten $X_1, \ldots, X_n$. Wie in Bemerkung 21.3 gezeigt, ist auch R abzählbar. Sei

$$K := K_0(X_1, \ldots, X_n) \supseteq R \qquad \text{Quotientenkörper von } R.$$

Gemäß der Konstruktion in Bemerkung 13.9 erhält man K als Menge der Äquivalenzklassen bezüglich einer Äquivalenzrelation auf $R \times (R \setminus \{0\})$; es gibt also eine surjektive Abbildung $R \times (R \setminus \{0\}) \to K$ mit $(a, b) \mapsto a/b$. Da R abzählbar ist, ist auch $R \times (R \setminus \{0\})$ abzählbar und damit K. Also liefert der obige Satz auch einen algebraischen Abschluss $L \supseteq K$.

(c) Sei $K = \mathbb{R}$. Dann können wir zwar den obigen Satz nicht anwenden, aber wir wissen nach dem Fundamentalsatz der Algebra, dass $\mathbb{C}$ algebraisch abgeschlossen ist. Weil $\mathbb{C} \supseteq \mathbb{R}$ algebraisch ist, ist also $\mathbb{C}$ ein algebraischer Abschluss von $\mathbb{R}$.

Bemerkung 21.6. Mit einer Ausnahme funktioniert die ganze Strategie des Beweises von Satz 21.4 praktisch wörtlich auch für einen beliebigen Körper K. Die Ausnahme ist die Existenz eines Ideals $J \trianglelefteq R$, so dass $M := R/J$ ein Körper ist. (Dies heißt nichts Anderes, als dass J ein maximales Ideal in R ist; siehe Ü8A4.) Um die Existenz von J im Allgemeinen sicherzustellen, benötigt man aber gerade das ***Lemma von Zorn***; siehe zum Beispiel §15.8.3 im Buch von Karpfinger–Meyberg [KM][15]. Mit dem Lemma von Zorn kann man also völlig analog wie im obigen Beweis fortfahren, und erhält einen algebraischen Abschluss als $L = \{\alpha \in M \mid \alpha \text{ algebraisch über } K\} \supseteq K$, für beliebiges K.

[15] Das Lemma von Zorn ist sogar äquivalent dazu, dass es in jedem kommutativen Ring mit $1 \neq 0$ ein maximales Ideal gibt! Siehe W. HODGES, Krull implies Zorn, J. London Math. Soc. **19** (1979), 285–287.