Diskrete Strukturen

Helmut Schwichtenberg

Mathematisches Institut der Universität München
Sommersemester 2008

Inhaltsverzeichnis

KAPITEL 1

Grundbegriffe

In der Mathematik und der Informatik ist es wichtig, die Ausdrucksfähigkeit der Umgangssprache einzuschränken und stattdessen eine zwar abgemagerte, aber dafür mit präzisen Begriffen arbeitende und im Prinzip formalisierbare Sprache zu verwenden. Eine solche Sprache beschreiben wir im zweiten Abschnitt, nachdem wir im ersten Abschnitt die zugrunde liegenden rekursiven Definitionen untersucht haben.

Um einen sprachlichen Ausdruck als „wahr“ ansehen zu können, benötigt man eine Interpretation oder genauer ein „Modell“ der Sprache. Bei uns werden dies meist die natürlichen Zahlen oder ähnlich unproblematische Strukturen sein. Eine wichtige Rolle spielt dann der Begriff der Gleichheit. In Anwendungen möchte man oft die Gleichheit nur bis auf im aktuellen Kontext „unwesentliche“ Eigenschaften betrachten. Dies führt auf den Begriff der Äquivalenzrelation, den wir im dritten Abschnitt untersuchen.

1.1. Rekursive Definitionen

1.1.1. Datentypen. Wir betrachten (induktiv erzeugte) *Typen* (oft auch *Datentypen* genannt). Das für uns wichtigste Beispiel ist der Typ $\mathbf{N}$ der natürlichen Zahlen. Sie werden aus der Null 0 durch die einstellige Nachfolgeroperation S erzeugt. Daneben betrachten wir noch den Typ $\mathbf{P}$ der binär dargestellten positiven Zahlen, erzeugt aus der Eins 1 durch zwei einstellige Nachfolgeroperationen, S_0 und S_1, sowie den Typ $\mathbf{B}$ der Fregeschen Wahrheitswerte $\mathtt{tt}$ und $\mathtt{ff}$ (oft auch „boolesche Objekte“ genannt). Ferner erlauben wir Typen, die aus anderen, vorher erzeugten Typen aufgebaut sind:

- Produkte $\rho \times \sigma$, erzeugt durch Paarbildung $\langle x^\rho, y^\sigma \rangle$;
- Summen $\rho + \sigma$, erzeugt durch die Einbettungsoperationen $\mathrm{inl}(x^\rho)$ und $\mathrm{inr}(y^\sigma)$;
- Listen $\mathbf{L}(\rho)$, erzeugt aus der leeren Liste nil durch die Operation $x ::_\rho l$, die ein Objekt x vom Typ ρ vorne an die Liste l anhängt.

Unsere Schreibweise für die Mitteilung dieser Definitionen ist

$$\begin{aligned} \mathbf{B} &:= \mu_\alpha(\alpha, \alpha) \qquad \text{(Wahrheitswerte)}, \\ \mathbf{N} &:= \mu_\alpha(\alpha, \alpha \to \alpha) \qquad \text{(natürliche Zahlen, unär)}, \end{aligned}$$

$$\begin{aligned}
\mathbf{P} &:= \mu_\alpha(\alpha, \alpha \to \alpha, \alpha \to \alpha) \qquad \text{(positive Zahlen, binär)},\\
\rho \times \sigma &:= \mu_\alpha(\rho \to \sigma \to \alpha) \qquad \text{(Produkt)},\\
\rho + \sigma &:= \mu_\alpha(\rho \to \alpha, \sigma \to \alpha) \qquad \text{(Summe)},\\
\mathbf{L}(\rho) &:= \mu_\alpha(\alpha, \rho \to \alpha \to \alpha) \qquad \text{(Listen)}.
\end{aligned}$$

Hierbei bedeutet etwa $\mathbf{N} := \mu_\alpha(\alpha, \alpha \to \alpha)$, daß der Typ $\mathbf{N}$ der natürlichen Zahlen aus zwei *Konstruktoren* aufgebaut ist, die wir Null (0) und Nachfolgerfunktion (S) nennen.

Eine zentrale weitere Begriffsbildung ist die des *Funktionstyps* $\rho \to \sigma$. Unter einem Objekt dieses Typs stellen wir uns eine beliebige Funktione f vor, die einem Argument x des Typs ρ einen Wert $f(x)$ des Typs σ zuordnet. Zum Beispiel hat die Nachfolgerfunktion S den Typ $\mathbf{N} \to \mathbf{N}$. Oft werden wir auch mehrstellige Funktionen betrachten wollen; wir verwenden dann die folgende

SCHREIBWEISE. $\rho \to \sigma \to \tau$ steht für $\rho \to (\sigma \to \tau)$ und allgemein

$$\rho_1 \to \rho_2 \to \dots \rho_{n-1} \to \rho_n \quad \text{für} \quad \rho_1 \to (\rho_2 \to \dots (\rho_{n-1} \to \rho_n) \dots),$$

wir verwenden also Rechtsklammerung.

Beipiele sind die Addition und die Multiplikation vom Typ $\mathbf{N} \to \mathbf{N} \to \mathbf{N}$, oder auch die append-Funktion für Listen vom Typ $\mathbf{L}(\rho) \to \mathbf{L}(\rho) \to \mathbf{L}(\rho)$.

1.1.2. Explizite Definition von Funktionen. Wir wollen uns jetzt mit der Frage befassen, wie man Funktionen definieren kann. Unmittelbar haben wir die Konstruktoren der verwendeten Datentypen, wie etwa 0 und S im Fall von $\mathbf{N}$. Nur aus den Konstruktoren 0 und S für $\mathbf{N}$ aufgebaute Terme kürzen wir wie üblich ab, also $1 := \mathrm{S}(0)$, $2 := \mathrm{S}(\mathrm{S}(0))$ und so weiter. Wir nennen solche Terme *Numerale*. Für den Typ $\mathbf{P}$ der binär dargestellten positiven Zahlen, erzeugt aus der Eins 1 durch zwei einstellige Nachfolgeroperationen, S_0 und S_1, haben wir etwa folgende Numerale:

	Numeral des Typs $\mathbf{N}$	Binärdarstellung	Numeral des Typs $\mathbf{P}$
1	S(0)	1	1
2	S(S(0))	10	$S_0 1$
3	S(S(S(0)))	11	$S_1 1$
4	S(S(S(S(0))))	100	$S_0(S_0 1)$
5	S(S(S(S(S(0)))))	101	$S_1(S_0 1)$
6		110	$S_0(S_1 1)$
7		111	$S_1(S_1 1)$
8		1000	$S_0(S_0(S_0 1))$
9		1001	$S_1(S_0(S_0 1))$

Es ist unproblematisch, Funktionen durch *explizite Definitionen* einzuführen. Eine Funktion f heißt *explizit definiert*, wenn sie in der Form

$$f(x_1, \ldots, x_n) := r(x_1, \ldots, x_n)$$

gegeben ist, wobei $r(x_1, \ldots, x_n)$ ein Term ist, der aus den (verschiedenen) Variablen $x_1, \ldots, x_n$ und Konstanten für bereits eingeführte Funktionen aufgebaut ist (z.B. Konstruktoren). Beispiele für explizite Definitionen sind

(1) $f\colon \rho \to \mathbf{N}$, definiert durch $f(x) := 0$. Hier handelt es sich um die konstante Funktion auf ρ mit dem Wert 0.
(2) $f\colon \rho_1 \to \ldots \to \rho_n \to \rho_i$, definiert durch $f(x_1, \ldots, x_n) := x_i$. Man spricht hier von der *Projektion* auf die i-te Komponente.
(3) Sind $f\colon \rho \to \sigma$ und $g\colon \sigma \to \tau$ Funktionen, so definiert man die *Komposition* oder *Hintereinanderschaltung* $g \circ f\colon \rho \to \tau$ (gelesen g *nach* f) durch $(g \circ f)(x) := g(f(x))$.

Natürlich wollen wir auch allgemeinere Formen der Definition von Funktionen zulassen, zum Beispiel rekursive Definitionen. Mit ihnen befassen wir uns jetzt.

1.1.3. Primitive Rekursion. Die Addition für natürliche Zahlen ist eine zweistellige Funktion, also vom Typ $\mathbf{N} \to \mathbf{N} \to \mathbf{N}$. Man kann sie durch *Berechnungsregeln* (oder Rekursionsgleichungen) definieren:

$$x + 0 := x,$$
$$x + \mathrm{S}(y) := \mathrm{S}(x + y).$$

Diese Berechnungsregeln verstehen wir „operational“, also als Umformungsregeln in der Richtung von links nach rechts. Ein Beispiel einer solchen Umformung ist $3 + \mathrm{S}(\mathrm{S}(0)) = \mathrm{S}(3 + \mathrm{S}(0)) = \mathrm{S}(\mathrm{S}(3 + 0)) = \mathrm{S}(\mathrm{S}(3))$.

Ähnlich kann man die Multiplikation und die Exponentiation exp als Funktionen vom Typ $\mathbf{N} \to \mathbf{N} \to \mathbf{N}$ definieren:

$$x \cdot 0 := 0, \qquad x^0 := 1,$$
$$x \cdot \mathrm{S}(y) := (x \cdot y) + x \qquad x^{\mathrm{S}(y)} := (x^y) \cdot x,$$

wobei wir x^y für $\exp(x, y)$ geschrieben haben. Weitere Beispiele sind die Verdoppelungsfunktion und die daraus definierte Zweierpotenz, sowie die Vorgängerfunktion P vom Typ $\mathbf{N} \to \mathbf{N}$ und die daraus definierte „abgeschnittene“ Substraktion $\dot{-}$:

$$\mathrm{P}(0) := 0 \qquad x \mathbin{\dot{-}} 0 := x,$$
$$\mathrm{P}(\mathrm{S}(x)) := x \qquad x \mathbin{\dot{-}} \mathrm{S}(y) := \mathrm{P}(x \mathbin{\dot{-}} y),$$

Ein besonders einfaches, aber wichtiges Beispiel ist die *Fallunterscheidung* $\mathcal{C}$ vom Typ $\mathbf{B} \to \rho \to \rho \to \rho$, die wir definieren werden durch

$$\mathcal{C}(\mathtt{tt}, x, y) := x, \quad \mathcal{C}(\mathtt{ff}, x, y) := y.$$

Wir schreiben [**if** b **then** x **else** y] für $\mathcal{C}(b,x,y)$.

Allgemein haben wir das Schema der *primitiven Rekursion* (auch lineare Rekursion genannt) zur Definition einer Funktion $f\colon \rho_1 \to \ldots \to \rho_n \to \mathbf{N} \to \sigma$ aus gegebenen Funktionen $g\colon \rho_1 \to \ldots \to \rho_n \to \sigma$ und $h\colon \rho_1 \to \ldots \to \rho_n \to \mathbf{N} \to \sigma \to \sigma$:

$$\begin{aligned} f(x_1,\ldots,x_n,0) &:= g(x_1,\ldots,x_n),\\ f(x_1,\ldots,x_n,\mathrm{S}(y)) &:= h(x_1,\ldots,x_n,y,f(x_1,\ldots,x_n,y)). \end{aligned}$$

Als weiteres Beispiel betrachten wir die Definition der Fakultätsfunktion:

$$\begin{aligned} 0! &:= 1,\\ \mathrm{S}(x)! &:= \mathrm{S}(x)\cdot x!. \end{aligned}$$

Eine (für Berechnungen günstigere) „iterative“ Form dieser Definition ist

$$\begin{aligned} f(y,0) &:= y,\\ f(y,\mathrm{S}(x)) &:= f(y\cdot \mathrm{S}(x),x) \end{aligned} \qquad x! := f(1,x).$$

Ein anderes Beispiel für eine primitive Rekursion ist die Verkettungsfunktion für Listen, vom Typ $\mathbf{L}(\rho)\to\mathbf{L}(\rho)\to\mathbf{L}(\rho)$.

1.1.4. Ungeschachtelte Rekursion. Die folgenden rekursiven Definitionen lassen sich mit etwas Mühe auf (i.a. mehrere) Definitionen durch primitive Rekursion zurückführen. Da sie jedoch alle unter das in 1.1.6 einzuführende Schema der allgemeinen Rekursion fallen, wollen wir uns hier nicht näher mit dieser Reduktion befassen.

Addition und Multiplikation für Binärzahlen können wir als Funktionen vom Typ $\mathbf{P}\to\mathbf{P}\to\mathbf{P}$ durch eine Variante der primitiven Rekursion definieren, in der die linken Seiten aus Konstruktoren und verschiedenen Variablen aufgebaut sind, und in den rechten Seiten höchstens ein Aufruf der zu definierenden Funktion mit „kleinerem“ Argument auftritt. Vorbereitend dazu benötigen wir eine Nachfolgerfunktion S vom Typ $\mathbf{P}\to\mathbf{P}$. Die Berechnungsregeln für S sind

$$S1 := S_0 1, \quad S(S_0 p) := S_1 p, \quad S(S_1 p) := S_0(Sp).$$

Für die Addition sind die Berechnungsregeln

$$\begin{aligned} p+1 &:= Sp,\\ 1+S_0 q &:= S_1 q,\\ S_0 p+S_0 q &:= S_0(p+q),\\ S_1 p+S_0 q &:= S_1(p+q), \end{aligned} \qquad \begin{aligned} 1+S_1 q &:= S_0(Sq),\\ S_0 p+S_1 q &:= S_1(p+q),\\ S_1 p+S_1 q &:= S_0(S(p+q)). \end{aligned}$$

Die Multiplikation ist definiert durch die Berechnungsregeln

$$p\cdot 1 := p, \quad p\cdot S_0 q := S_0(p\cdot q), \quad p\cdot S_1 q := S_0(p\cdot q)+p.$$

Man beachte, daß damit Algorithmen für Addition und Multiplikation von Binärzahlen vollständig angegeben sind.

Entscheidbare Relationen können als booleschwertige Funktionen aufgefaßt werden. Beispiele sind die Funktionen $<$ und $=$, beide vom Typ $\mathbf{N} \to \mathbf{N} \to \mathbf{B}$. Sie sind definiert durch

$$\begin{aligned} (0 = 0) &:= \mathtt{tt}, & (n < 0) &:= \mathtt{ff}, \\ (\mathrm{S}(n) = 0) &:= \mathtt{ff}, & (0 < \mathrm{S}(m)) &:= \mathtt{tt}, \\ (0 = \mathrm{S}(m)) &:= \mathtt{ff}, & (\mathrm{S}(n) < \mathrm{S}(m)) &:= (n < m). \\ (\mathrm{S}(n) = \mathrm{S}(m)) &:= (n = m), \end{aligned}$$

Eine oft auftretende Form der Rekursion ist die sogenannte *Baumrekursion*. Hier dürfen mehrere ungeschachtelte Aufrufe der zu definierenden Funktion vorkommen. Ein typisches Beispiel dafür ist die Folge der *Fibonacci-Zahlen* vom Typ $\mathbf{N} \to \mathbf{N}$:

$$\begin{aligned} f(0) &:= 0, \\ f(1) &:= 1, \\ f(\mathrm{S}(\mathrm{S}(x))) &:= f(x) + f(\mathrm{S}(x)). \end{aligned}$$

Bei der Anwendung dieser Regeln muß man aus Effizienzgründen dafür sorgen, daß Mehrfachberechnungen vermieden werden. Ein weiteres Beispiel sind die *Binomialkoeffizienten*, die für natürliche Zahlen n und k definiert sind durch

$$\begin{aligned} \binom{n}{0} &:= 1, \\ \binom{n}{k} &:= 0 \quad \text{für } n < k, \\ \binom{n}{k} &:= \binom{n-1}{k-1} + \binom{n-1}{k} \quad \text{für } 1 \le k \le n. \end{aligned}$$

Es ist meist günstiger, eine (leicht als äquivalent nachzuweisende) andere Darstellung der Binomialkoeffizienten zu verwenden:

$$\binom{n}{k} = \prod_{j=1}^{k} \frac{n-j+1}{j} = \frac{n(n-1)\cdot\ldots\cdot(n-k+1)}{1\cdot 2\cdot\ldots\cdot k}.$$

Hieraus folgt unmittelbar

$$\begin{aligned} \binom{n}{k} &= 0 \quad \text{für } n < k, \\ \binom{n}{k} &= \frac{n!}{k!(n-k)!} = \binom{n}{n-k} \quad \text{für } 0 \le k \le n, \end{aligned}$$

was für Berechnungen vorteilhaft sein kann.

1.1.5. Geschachtelte Rekursion. Häufig hat man es auch mit der sogenannten *geschachtelten Rekursion* zu tun, in der mehrere geschachtelte Aufrufe der zu definierenden Funktion vorkommen. Ein typisches Beispiel ist die *Ackermann-Funktion*, die wie folgt definiert ist.

$$\begin{aligned} f(x,0) &:= 0,\\ f(0,y+1) &:= 2(y+1),\\ f(x+1,1) &:= 2,\\ f(x+1,y+2) &:= f(x,f(x+1,y+1)). \end{aligned}$$

1.1.6. Allgemeine Rekursion. Wir wollen das Schema der primitiven Rekursion noch weiter verallgemeinern und von den Rekursionsgleichungen nur noch verlangen, daß die linken Seiten aus Konstruktoren (wie zum Beispiel 0 und S) und verschiedenen Variablen aufgebaut sind. Für die rechten Seiten gibt es keine Einschränkungen. Der Einfachheit halber wollen wir (jedenfalls zunächst) fordern, daß zwei linke Seiten sich nicht „überlappen“ dürfen. Alle bisherigen Beispiele waren von dieser Form.

Bemerkung. Man kann auch zugelassen, daß zwei linke Seiten sich überlappen. Dann müssen jedoch die rechten Seiten bei jeder Substitution, die die linken Seiten gleich macht, auch gleich werden.

Ein Beipiel für eine solche Situation ist die Definition der booleschen Verknüpfungen andb, impb und orb, die alle vom Typ $\mathbf{B} \to \mathbf{B} \to \mathbf{B}$ sind:

$$\begin{aligned} \mathtt{tt}\ \text{andb}\ c &:= c,\\ b\ \text{andb}\ \mathtt{tt} &:= b,\\ \mathtt{ff}\ \text{andb}\ c &:= \mathtt{ff},\\ b\ \text{andb}\ \mathtt{ff} &:= \mathtt{ff}, \end{aligned} \qquad \begin{aligned} \mathtt{ff}\ \text{impb}\ c &:= \mathtt{tt},\\ \mathtt{tt}\ \text{impb}\ c &:= c,\\ b\ \text{impb}\ \mathtt{tt} &:= \mathtt{tt}, \end{aligned} \qquad \begin{aligned} \mathtt{tt}\ \text{orb}\ c &:= \mathtt{tt},\\ b\ \text{orb}\ \mathtt{tt} &:= \mathtt{tt},\\ \mathtt{ff}\ \text{orb}\ c &:= c,\\ b\ \text{orb}\ \mathtt{ff} &:= b. \end{aligned}$$

Wie bereits gesagt, verwenden wir die Rekursionsgleichungen – in der Richtung von links nach rechts – zum Berechnen von Funktionswerten oder allgemeiner zum Umformen von Termen; sie heißen deshalb Berechnungsregeln. Es wird *nicht* verlangt, daß diese Umformungen abbrechen. Man kann deshalb auch nicht terminierende Berechnungen haben, wie zum Beispiel bei

$$\begin{aligned} f(0) &:= \mathrm{S}(0),\\ f(\mathrm{S}(x)) &:= \mathrm{S}(f(\mathrm{S}(\mathrm{S}(x)))). \end{aligned}$$

Um dies auszuschließen, müssen wir nachweisen, daß der betreffende Term r ein Numeral als „Bedeutung“ oder „Wert“ hat. Wir werden dashalb ein entsprechendes Prädikatensymbol T in unsere logische Sprache aufnehmen; $T(r)$ ist dann zu lesen als „r ist total“.

$$\frac{\begin{array}{c}[u\colon A]\\ \mid M\\ B\end{array}}{A \to B}\,{\to}^{+}u \qquad \frac{\begin{array}{cc}\mid M & \mid N\\ A \to B & A\end{array}}{B}\,{\to}^{-}$$

ABBILDUNG 1. Einführungs- und Beseitigungsregeln für $\to$

1.2. Logische Formeln

1.2.1. Implikation und Allquantor. Wir wollen jetzt von Termen – die Objekte bezeichnen – zu Aussagen übergehen. Eine *Aussage* ist nach Aristoteles „ein sprachliches Gebilde, von dem es sinnvoll ist zu sagen, es sei wahr oder falsch“. In der Mathematik verwendet man anstelle der Umgangssprache künstliche, formale Sprachen, um Eindeutigkeit und Einfachheit zu gewährleisten. Aus gegebenen Aussagen $A, B, C \dots$ (etwa Gleichungen oder Totalitätsaussagen $T(r)$) bilden wir neue durch

- *Implikation* $A \to B$ (gelesen „wenn A, so B“), und der
- *All-Quantifizierung* $\forall_x A$ (gelesen „für alle x gilt A“).

SCHREIBWEISE. Wir schreiben $A \to B \to C$ für $A \to (B \to C)$ und allgemein

$$A_1 \to A_2 \to \dots A_{n-1} \to A_n \quad \text{für} \quad A_1 \to (A_2 \to \dots (A_{n-1} \to A_n) \dots),$$

verwenden also wieder Rechtsklammerung für die Implikation $\to$. In Formeln können wir Klammern sparen, wenn wir vereinbaren, daß $\forall$ stärker bindet als $\to$. Zum Beispiel ist $\forall_x A \to B$ zu lesen als $(\forall_x A) \to B$. Führende Allquantoren lassen wir oft weg.

Wir wollen hier Implikationen und Allquantoren intuitiv verstehen und keinen Logikformalismus einführen. Dies ist möglich, wenn wir Gerhard Gentzens (1934) Konzept des „natürlichen Schließens“ folgen und den Umgang mit $\to$ und $\forall$ durch je eine Einführungs- und Beseitigungsregel beschreiben. Die Einführungsregeln legen fest, unter welchen Umständen man eine Implikation bzw. Allaussage machen kann, und die Beseitigungsregeln sagen, wie man sie verwenden kann.

Für die Implikation $\to$ gibt es also eine Einführungsregel ${\to}^{+}u$ und eine Beseitigungsregel ${\to}^{-}$, die auch *modus ponens* genannt wird. Die linke Prämisse $A \to B$ in ${\to}^{-}$ heißt *Hauptprämisse*, und die rechte Prämisse A *Nebenprämisse*. Man beachte, daß bei Anwendung der ${\to}^{+}u$-Regel *alle* darüberstehenden mit $u\colon A$ markierten Annahmen gestrichen werden. Für den Allquantor $\forall$ gibt es eine Einführungsregel $\forall^{+}x$ und eine Beseitigungsregeln $\forall^{-}$, deren rechte Prämisse der zu substituierende Term r ist. Die Regel $\forall^{+}x$ unterliegt der folgenden *Variablenbedingung*: der Beweis M der

$$\frac{\begin{array}{c}|\,M\\ A\end{array}}{\forall_x A}\,\forall^+x \qquad \frac{\begin{array}{cc}|\,M & \\ \forall_x A(x) & r\end{array}}{A(r)}\,\forall^-$$

ABBILDUNG 2. Einführungs- und Beseitigungsregeln für $\forall$

Prämisse A darf keine offenen Annahmen enthalten, in denen x frei vorkommt.

1.2.2. Primformeln, Axiome. Elementare, nicht weiter zerlegbare Aussagen in unserer Sprache bilden wir mit den Prädikaten der *Totalität* T und der *Leibniz Gleichheit* Eq. Von beiden Prädikaten gibt es unendlich viele, für jeden Typ ρ eines. Wir sagen, daß T_ρ die „Stelligkeit" (ρ) und Eq_ρ die Stelligkeit (ρ,ρ) hat. Die Bedeutung dieser Prädikate legen wir durch Einführungs- und Beseitigungsaxiome fest. Die Leibniz Gleichheit Eq ist bestimmt durch die beiden Axiome

$$\mathrm{Eq}^+\colon \forall_x \mathrm{Eq}(x,x), \quad \mathrm{Eq}^-\colon \forall_{x,y}(\mathrm{Eq}(x,y) \to \forall_z C(z,z) \to C(x,y)).$$

Die wesentliche (und für Leibniz definierende) Eigenschaft von Eq ist, daß man im Fall von $\mathrm{Eq}(x,y)$ in einer beliebigen Aussage Vorkommen von x durch y ersetzen kann.

LEMMA (Verträglichkeit). $\forall_{x,y}\big(\mathrm{Eq}(x,y) \to A(x) \to A(y)\big)$.

BEWEIS. Verwende Eq^- mit $C(x,y) := A(x) \to A(y)$. □

Hieraus erhält man leicht:

LEMMA (Symmetrie und Transitivität der Leibniz Gleichheit).

$\forall_{x,y}(\mathrm{Eq}(x,y) \to \mathrm{Eq}(y,x)), \quad \forall_{x,y,z}(\mathrm{Eq}(x,y) \to \mathrm{Eq}(y,z) \to \mathrm{Eq}(x,z))$

BEWEIS. Die Beweise verwenden das Verträglichkeitslemma; sie seien dem Leser als Übung überlassen. □

Für die Totalität hat man die beiden Einführungsaxiome

$$T_0^+\colon T(0),$$
$$T_1^+\colon \forall_n(T(n) \to T(\mathrm{S}(n)))$$

und das Beseitigungsaxiom

$$T^-\colon \forall_m(T(m) \to A(0) \to \forall_n(T(n) \to A(n) \to A(\mathrm{S}(n))) \to A(m)).$$

Man beachte, daß die Allquantoren sämtlich auf T „relativiert" sind, also von der Form $\forall_n(T(n) \to \dots)$. Dies wird sehr oft der Fall sein; wir kürzen

derartige Aussagen deshalb ab mit $\forall_{n\in T}(\dots)$. Das Beseitigungsaxiom für T schreibt sich dann als

$$T^-\colon \forall_{m\in T}(A(0) \to \forall_{n\in T}(A(n) \to A(\mathrm{S}(n))) \to A(m)).$$

Man erhält also das bekannte Induktionsaxiom für natürliche Zahlen.

1.2.3. Atomare Formeln und Falschheit. Eine wichtige Verwendung der Leibniz Gleichheit Eq besteht darin, daß aus einem booleschen Term $r^{\mathbf{B}}$ eine Formel gemacht wird. Wir schreiben

$$\mathrm{atom}(r^{\mathbf{B}}) := \mathrm{Eq}(r^{\mathbf{B}}, \mathtt{tt}).$$

Damit ergibt sich ein bequemer Weg, mit der Gleichheit für Grundtypen umzugehen. In 1.1.4 hatten wir die (entscheidbare) Gleichheit für einen Datentyp ι als booleschwertige Funktion $=_\iota\colon \iota \to \iota \to \mathbf{B}$ eingeführt. Die Definitionsgleichungen stellen sicher, daß etwa der boolesche Term $\mathrm{S}(r) =_{\mathbf{N}} \mathrm{S}(s)$ identifiziert wird mit $r =_{\mathbf{N}} s$. Wir können jetzt diesen booleschen Term zu einer Formel $\mathrm{Eq}(\mathrm{S}(r) =_{\mathbf{N}} \mathrm{S}(s), \mathtt{tt})$ machen, die wir wieder durch $\mathrm{S}(r) =_{\mathbf{N}} \mathrm{S}(s)$ abkürzen, dieses Mal jedoch mit dem Verständnis, daß es eine Formel ist. Die beiden Formeln $\mathrm{S}(r) =_{\mathbf{N}} \mathrm{S}(s)$ und $r =_{\mathbf{N}} s$ sind also identifiziert, und infolgedessen müssen wir derartige einfache Aussagen nicht separat beweisen.

Eine zweite wichtige Verwendung der Leibniz Gleichheit ist die Definition der *Falschheit* $\mathbf{F}$ als

$$\mathbf{F} := \mathrm{Eq}(\mathtt{ff}, \mathtt{tt}).$$

Bei dieser Definition kann man das Schema $\mathbf{F} \to A$ des „ex-falso-quodlibet“ leicht beweisen.

Satz (Ex Falso Quodlibet). $\mathbf{F} \to A$.

Beweis. Wir zeigen zunächst, daß $\mathbf{F} \to \mathrm{Eq}(x^\rho, y^\rho)$. Um dies zu sehen beachte man, daß aus $\mathrm{Eq}(\mathtt{ff}, \mathtt{tt})$ mit der Verträglichkeit folgt

$$\mathrm{Eq}[\mathbf{if}\ \mathtt{tt}\ \mathbf{then}\ x\ \mathbf{else}\ y][\mathbf{if}\ \mathtt{ff}\ \mathbf{then}\ x\ \mathbf{else}\ y].$$

Also gilt $\mathrm{Eq}(x^\rho, y^\rho)$. Jetzt folgt $\mathbf{F} \to T(x)$ aus dem Axiom $T(0)$ wegen $\mathbf{F} \to \mathrm{Eq}(x, 0)$. Die Fälle $A \to B$ und $\forall_x A$ sind offensichtlich. □

1.2.4. Konjunktion, Disjunktion und Existenz. Wir führen die Konjunktion $\wedge$ durch Axiome ein. Das Einführungsaxiom für $\wedge$ ist

$$\wedge^+\colon A \to B \to A \wedge B$$

(mit Parametern A und B), und das Beseitigungsaxiom ist

$$\wedge^-\colon A \wedge B \to (A \to B \to C) \to C.$$

Die *Äquivalenz* $A \leftrightarrow B$ (gelesen „A äquivalent B“) führen wir als Abkürzung ein:

$$(A \leftrightarrow B) := (A \to B) \wedge (B \to A).$$

Bei der Disjunktion $\vee$ und dem Existenzquantor $\exists$ wollen wir zwischen der „starken“ und „schwachen“ Form unterscheiden, je nachdem, ob die gültige Alternative der Disjunktion bzw. das als existent nachzuweisende Objekt durch einen Beweis tatsächlich geliefert werden oder nicht. Als Beispiel betrachten wir die folgende Aussage.

Es gibt irrationale Zahlen a, b mit a^b rational.

Einen Beweis erhält man wie folgt durch Fallunterscheidung.

BEWEIS. *F*all $\sqrt{2}^{\sqrt{2}}$ ist rational. Man wähle $a = \sqrt{2}$ und $b = \sqrt{2}$. Dann sind a, b irrational, und nach Annahme ist a^b rational.

*F*all $\sqrt{2}^{\sqrt{2}}$ ist irrational. Man wähle $a = \sqrt{2}^{\sqrt{2}}$ und $b = \sqrt{2}$. Dann sind nach Annahme a, b irrational, und

$$a^b = \left(\sqrt{2}^{\sqrt{2}}\right)^{\sqrt{2}} = \left(\sqrt{2}\right)^2 = 2$$

ist rational. □

Solange wir nicht entschieden haben, ob $\sqrt{2}^{\sqrt{2}}$ nun rational ist oder nicht, wissen wir nicht, welche Zahlen a, b wir nehmen müssen. Damit haben wir ein Beispiel eines Existenzbeweises, der es nicht erlaubt, das als existent nachgewiesene Objekt tatsächlich anzugeben.

Die Disjunktion hat die beiden Einführungsaxiome:

$$\vee_0^+ \colon A \to A \vee B, \quad \vee_1^+ \colon B \to A \vee B,$$

und das Beseitigungsaxiom

$$\vee^- \colon A \vee B \to (A \to C) \to (B \to C) \to C.$$

Für $\exists$ hat man ein Einführungsaxiom (mit Parametern x und A)

$$\exists^+ \colon A \to \exists_x A$$

und das Beseitigungsaxiom

$$\exists^- \colon \exists_x A \to \forall_x (A \to B) \to B \quad (x \text{ nicht frei in } B).$$

1.2.5. Negation, schwache Disjunktion und Existenz. Negation, schwache (oder „klassische“) *Disjunktion*, und den schwachen („klassischen“) *Existenzquantor* definiert man durch

$$\begin{aligned} \neg A &:= A \to \mathbf{F}, \\ A \mathbin{\tilde{\vee}} B &:= \neg A \to \neg B \to \mathbf{F}, \\ \tilde{\exists}_x A &:= \neg \forall_x \neg A. \end{aligned}$$

Wir werden wann immer möglich die starken Formen der Disjunktion und Existenz verwenden.

1.2.6. Indirekte Beweise. Eine Formel A heißt *stabil*, wenn für sie das „Prinzip des indirekten Beweisens“ hergeleitet werden kann, also

$$\neg\neg A \to A.$$

Wir wollen uns überlegen, daß „entscheidbare“ Primformeln, also solche der Gestalt $\mathrm{atom}(r^{\mathbf{B}})$, stets stabil sind. Wir hatten gesehen, wie man mit Hilfe der Leibniz Gleichheit aus einem booleschen Term $r^{\mathbf{B}}$ eine Formel herstellen kann, nämlich als

$$\mathrm{atom}(r^{\mathbf{B}}) := \mathrm{Eq}(r^{\mathbf{B}}, \mathtt{tt}).$$

Mit Hilfe einer Fallunterscheidung zeigt man dann leicht, daß jede Formel der Gestalt $\mathrm{atom}(r^{\mathbf{B}})$ stabil ist, insbesondere also auch jede Gleichung $r =_{\mathbf{N}} s$ und die Falschheit $\mathbf{F}$.

Zur Formulierung des folgenden Satzes brauchen wir den Begriff der Menge der *Endkonklusionen* $\mathrm{End}(A)$ einer Formel A:

$$\begin{aligned}
\mathrm{End}(A \to B) &:= \mathrm{End}(B),\\
\mathrm{End}(A \wedge B) &:= \mathrm{End}(A) \cup \mathrm{End}(B),\\
\mathrm{End}(\forall_x A) &:= \mathrm{End}(A),\\
\mathrm{End}(A) &:= \{A\} \qquad \text{sonst.}
\end{aligned}$$

Insbesondere enthält eine Formel A nur stabile Endkonklusionen, wenn sie anstelle von $\vee$ und $\exists$ nur die schwachen Verknüpfungen $\tilde{\vee}$ und $\tilde{\exists}$ verwendet.

Satz. *Enthält A nur stabile Endkonklusionen, so ist A stabil.*

Beweis. Induktion über A. Für Primformeln gilt die Behauptung nach Annahme. Im Fall einer Implikation $A \to B$ verwendet man $(\neg\neg B \to B) \to \neg\neg(A \to B) \to A \to B$; ein Beweis ist

$$\dfrac{u\colon \neg\neg B \to B \qquad \dfrac{v\colon \neg\neg(A \to B) \qquad \dfrac{\dfrac{u_1\colon \neg B \qquad \dfrac{u_2\colon A \to B \qquad w\colon A}{B}}{\mathbf{F}}\to^+ u_2}{\neg(A \to B)}}{\mathbf{F}}}{\neg\neg B}\to^+ u_1}{B}$$

Im Fall $\forall_x A$ genügt $(\neg\neg A \to A) \to \neg\neg\forall_x A \to A$; ein Beweis ist

$$\dfrac{u\colon \neg\neg A \to A \qquad \dfrac{v\colon \neg\neg\forall_x A \qquad \dfrac{\dfrac{u_1\colon \neg A \qquad \dfrac{u_2\colon \forall_x A \qquad x}{A}}{\mathbf{F}}\to^+ u_2}{\neg\forall_x A}}{\mathbf{F}}}{\neg\neg A}\to^+ u_1}{A}$$

Für die Konjunktion ist der Beweis ähnlich. □

Man zeigt leicht, daß für die schwachen Verknüpfungen $\tilde{\vee}$ und $\tilde{\exists}$ fast dieselben Axiome hergeleitet werden können wir für $\vee$ und $\exists$: in den Beseitigungsaxiomen brauchen wir die Einschränkumg, daß die Konklusion C bzw. B stabil ist.

Daraus ergibt sich insbesondere, daß $\tilde{\vee}$ und $\tilde{\exists}$ tatsächlich schwächer sind als $\vee$ und $\exists$:

$$A \vee B \to A \mathbin{\tilde{\vee}} B, \quad \exists_x A \to \tilde{\exists}_x A.$$

Die Beweise dafür seien dem Leser als Übung überlassen.

1.2.7. Rechenregeln für die Negation. Für den Umgang mit der Negation gelten einige nützliche Rechenregeln, die man oft formal anwenden kann.

Lemma. (a) *Kontraposition*

$$(A \to B) \to (\neg B \to \neg A),$$
$$(\neg B \to \neg A) \to (A \to B) \qquad \textit{falls } B \textit{ stabil.}$$

(b) *Doppelte Negation*

$$A \to \neg\neg A,$$
$$\neg\neg A \to A \qquad \textit{falls } A \textit{ stabil.}$$

(c) *De Morgansche Regeln*

$$\neg(A \wedge B) \leftrightarrow (\neg A \mathbin{\tilde{\vee}} \neg B),$$
$$\neg(A \mathbin{\tilde{\vee}} B) \leftrightarrow (\neg A \wedge \neg B).$$

(d) *Negation von* $\forall$

$$\neg\forall_x A \to \tilde{\exists}_x \neg A \qquad \textit{falls } A \textit{ stabil,}$$
$$\tilde{\exists}_x \neg A \to \neg\forall_x A.$$

(e) *Negation von* $\tilde{\exists}$

$$\neg\tilde{\exists}_x A \leftrightarrow \forall_x \neg A.$$

Die Beweise dafür seien wieder dem Leser als Übung überlassen.

1.3. Mengen, Äquivalenzrelationen

Mengen fassen wir auf als gegeben durch eine Eigenschaft, genauer durch eine Formel mit einer ausgezeichneten Variablen. Wir bilden hier Mengen nur

durch Aussonderung aus Typen, also in der Form $\{ x^\rho \mid A \}$, und schreiben $r \in \{ x \mid A(x) \}$ für $A(r)$. Beispiele sind

$$\{ n^{\mathbf{N}} \mid n \text{ ist Primzahl} \},$$
$$\{ f^{\mathbf{N}\to\mathbf{N}} \mid \forall_n f(n) \le 1 \},$$
$$\mathbf{N}.$$

1.3.1. Teilmengen, Durchschnitt, Vereinigung. Sind

$$M := \{ x \mid A \}, \qquad N := \{ x \mid B \}$$

Mengen, so heißt M *Teilmenge* von N (geschrieben $M \subseteq N$), wenn jedes Element von M auch Element von N ist, d.h., wenn $A \to B$ gilt. Für $M \subseteq N$ schreibt man auch $N \supseteq M$. Sind $M := \{ x \mid A \}$ und $N := \{ x \mid B \}$ Mengen, so definieren wir

$$M \cap N := \{ x \mid A \wedge B \} \quad \textit{Durchschnitt} \text{ von } M \text{ und } N,$$
$$M \cup N := \{ x \mid A \vee B \} \quad \textit{Vereinigung} \text{ von } M \text{ und } N,$$
$$M \setminus N := \{ x \mid A \wedge \neg B \} \quad \textit{Differenz} \text{ von } M \text{ und } N.$$

Wir sprechen von einer *schwachen* Vereinigung und schreiben $M \mathbin{\tilde\cup} N$, wenn anstelle von $\vee$ die schwache Disjunktion $\mathbin{\tilde\vee}$ verwendet wurde.

Zwei Mengen M und N heißen *disjunkt*, wenn sie keine gemeinsamen Elemente haben, also wenn $M \cap N = \emptyset := \{ x \mid \mathbf{F} \}$.

Sind M und N Mengen, so nennt man die Menge

$$M \times N := \{ (x, y) \mid x \in M \wedge y \in N \}$$

das *kartesische Produkt* der Mengen M und N.

Unter einer *Relation* R zwischen (Elementen von) M und (Elementen von) N versteht man eine Teilmenge $R \subseteq M \times N$. Statt $(x, y) \in R$ schreibt man oft xRy. Ist speziell $M = N$, so spricht man von einer *Relation auf* M.

Beispiel. Die Teilbarkeitsrelation auf $\mathbf{N}$ ist die Menge

$$\{ (n, m) \mid \exists_k n \cdot k = m \} \subseteq \mathbf{N} \times \mathbf{N}.$$

1.3.2. Ein naiver Mengenbegriff; die Russellsche Antinomie. Cantor gab 1895 die folgende „allgemeinere Definition“:

> Unter einer *Menge* verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m unserer Anschauung oder unseres Denkens (welche die Elemente von M genannt werden) zu einem Ganzen.

Insbesondere ist eine Menge durch ihre Elemente vollständig bestimmt (Extensionalitätsprinzip).

Man kann versuchen, diese Definition wie folgt zu verstehen. Sei V die Gesamtheit aller Objekte (unserer Anschauung oder unseres Denkens). Dann kann man

$$\{\, x \mid A \,\}$$

bilden, die Menge aller Objekte x aus V mit der Eigenschaft A. Man beachte, daß $\{\, x \mid A \,\}$ wieder ein Objekt aus V ist. Da man hier alle Objekte mit einer gewissen Eigenschaft zusammenfaßt, spricht man von einem (naiven) *Komprehensionsprinzip*.

Cantors Definition – oder genauer unsere naive Auffassung davon – ist jedoch so nicht haltbar, da sie zu Widersprüchen führt. Am bekanntesten ist die *Russellsche Antinomie*: Sei

$$x_0 := \{\, x \mid \mathrm{Mg}(x) \wedge x \notin x \,\},$$

wobei $\mathrm{Mg}(x)$ die Eigenschaft „x ist Menge“ ausdrücken soll. Dann erhält man $x \in x_0 \leftrightarrow \mathrm{Mg}(x) \wedge x \notin x$ für alle Objekte x, also insbesondere

$$x_0 \in x_0 \leftrightarrow \mathrm{Mg}(x_0) \wedge x_0 \notin x_0 \leftrightarrow x_0 \notin x_0,$$

denn x_0 ist Menge. Einen Grund für diesen Widerspruch kann man darin sehen, daß wir hier – unter Verwendung des naiven Komprehensionsprinzips – von der Vorstellung einer fertigen Gesamtheit aller Mengen ausgegangen sind. Dies ist aber weder notwendig noch entspricht es dem Vorgehen in der Mathematik. Es reicht vollkommen aus, wenn man eine Menge nur dann bildet, wenn ihre Elemente bereits „zur Verfügung stehen“, etwa dadurch, daß sie Objekte eines festen Typs sind.

Für eine genauere Diskussion der historischen Entwicklung der Mengenlehre und insbesondere eine präzise axiomatische Entwicklung der Mengenlehre müssen wir auf die Literatur – etwa das Buch von Deiser (2004) – oder Vorlesungen über mathematische Logik verweisen.

1.3.3. Äquivalenzrelationen. In dieser Vorlesung und auch sonst in der Mathematik spielen Relationen eine besondere Rolle, die man als eine Art „Gleichheit“ ansehen kann. Die Verwendung solcher Relationen erlaubt es, von „unwesentlichen“ Eigenschaften abzusehen.

Definition. Sei $M = \{\, x^\rho \mid A \,\}$ eine Menge und $R(x, y)$ eine zweistellige Relation auf ρ, gegeben durch eine Formel mit zwei ausgezeichneten Variablen x, y vom Typ ρ. Wir schreiben $x \sim y$ für $R(x, y)$. $R(x, y)$ heißt *Äquivalenzrelation* auf M, wenn für alle $x, y, z \in M$ gilt

(a) $x \sim x$ (Reflexivität),
(b) $x \sim y \to y \sim x$ (Symmetrie),
(c) $x \sim y \to y \sim z \to x \sim z$ (Transitivität).

BEISPIEL. Sei n eine positive ganze Zahl. Auf den ganzen Zahlen $\mathbf{Z}$ definieren wir eine Relation $a \sim b$ durch die Eigenschaft, daß $a-b$ ein Vielfaches von n ist. Wir wollen uns überlegen, daß $\sim$ eine Äquivalenzrelation auf $\mathbf{Z}$ ist. $\sim$ ist reflexiv, da 0 ein Vielfaches von jeder Zahl ist. $\sim$ ist symmetrisch, denn ist $a-b$ ein Vielfaches von n, so auch $b-a$. $\sim$ ist auch transitiv, denn sind $a-b$ und $b-c$ beides Vielfache von n, so auch $(a-b)+(b-c)=a-c$.

BEISPIEL. Sei M die Menge aller Programme P, Q in irgendeiner festen Programmiersprache. $P \sim Q$ bedeute, daß für alle Eingaben I das Programm P auf I terminiert genau dann, wenn Q auf I terminiert, und daß in diesem Fall die Ausgaben gleich sind. Dann ist $\sim$ eine Äquivalenzrelation auf M.

Es ist möglich und oft üblich, das explizite Auftreten einer Äquivalenzrelation zu vermeiden und stattdessen mit der vertrauten Gleichheit zu arbeiten. Dazu faßt man äquivalente Objekte zusammen zu einer sogenannten Äquivalenzklasse (richtiger wäre „Äquivalenzmenge“, aber „Äquivalenzklasse“ hat sich eingebürgert).

Sei M eine Menge und $\sim$ eine Äquivalenzrelation auf M. Eine nicht leere Teilmenge N von M heißt eine *Äquivalenzklasse* der Relation $\sim$, wenn gilt:

$$\forall_{x,y\in M}\big(x \in N \to y \sim x \to y \in N\big) \quad (N \text{ ist } \sim\text{-}\textit{abgeschlossen}), \tag{1.1}$$

$$\forall_{x,y\in M}\big(x, y \in N \to x \sim y\big). \tag{1.2}$$

SATZ. *Sei M eine Menge und $\sim$ eine Äquivalenzrelation auf M. Jedes $x \in M$ liegt in genau einer Äquivalenzklasse; sie wird mit $[x]$ bezeichnet.*

BEWEIS. Existenz. Setze $[x] := \{\, y \in M \mid y \sim x \,\}$. Dann ist $x \in [x]$ wegen der Reflexivität von $\sim$. Zu zeigen bleibt, daß $[x]$ eine Äquivalenzklasse ist. Zu (1.1). Sei $y \in [x]$ und $z \sim y$. Dann gilt $y \sim x$, also auch $z \sim x$ wegen der Transitivität von $\sim$, also $z \in [x]$. Zu (1.2). Seien $y, z \in [x]$. Dann gilt $y \sim x$ und $z \sim x$. Aus $z \sim x$ folgt $x \sim z$ wegen der Symmetrie von $\sim$, also auch $y \sim z$ wegen der Transitivität von $\sim$.

Eindeutigkeit. Seien N_1, N_2 Äquivalenzklassen mit $x \in N_1$ und $x \in N_2$. Zu zeigen ist $N_1 = N_2$; aus Symmetriegründen genügt $N_1 \subseteq N_2$. Sei also $z \in N_1$. Zu zeigen ist $z \in N_2$. Wegen $z, x \in N_1$ folgt $z \sim x$ nach (1.2) für N_1, also auch $z \in N_2$ nach (1.1) für N_2. □

Aus dem Eindeutigkeitsteil dieser Aussage folgt, daß je zwei verschiedene Äquivalenzklassen disjunkt sind, d.h., daß aus $N_1 \neq N_2$ folgt $N_1 \cap N_2 = \emptyset$.

Eine Äquivalenzrelation auf M zerlegt also die „Trägermenge“ M vollständig in paarweise disjunkte Äquivalenzklassen.

1.3.4. Darstellung von Äquivalenzrelationen durch Graphen. Es ist manchmal nützlich, die eben eingeführten Begriffe graphisch zu veranschaulichen. Dazu verwenden wir den in 4.1.1 genauer untersuchten Begriff eines gerichteten Graphen. Den drei definierenden Eigenschaften einer Äquivalenzrelation entsprechen dann die folgenden graphischen Begriffe.

- Reflexivität: jeder Punkt in G hat eine Schleife;
- Symmetrie: wenn zwei Punkte in einer Richtung durch eine Kante verbunden sind, so gibt es auch eine Kante in umgekehrter Richtung;
- Transitivität: wenn x mit y und y mit z durch eine Kante verbunden sind, so gibt es auch eine Kante von x zu z.

Eine Äquivalenzrelation zerlegt also die zugrunde liegende Menge in Teile (genannt „Zusammenhangskomponenten“), so daß es keine Kanten zwischen Punkten in verschiedenen Teilen gibt, und alle Punkte in einem Teil miteinander verbunden sind.

Übung. Man zeichne den repräsentierenden Graphen für die durch die Äquivalenzklassen $\{a, b, c\}$, $\{d, e\}$, $\{f\}$ und $\{g\}$ auf $\{a, b, c, d, e, f, g\}$ gegebene Äquivalenzrelation.

1.3.5. Quotientenstrukturen. Ein wichtiger Grund für die Einführung von Äquivalenzrelationen ist die Möglichkeit, Quotientenstrukturen zu betrachten. Beispiele davon werden wir später sehen. Hier begnügen wir uns damit, auf das oben behandelte Beispiel der ganzen Zahlen modulo n zurückzukommen.

Sei $\sim$ auf $\mathbf{Z}$ definiert durch $a \sim b$ falls n ein Teiler von $a - b$ ist. Als Beispiel nehmen wir $n = 5$. Dann stehen a und b in der Relation $\sim$ wenn sie denselben Rest bei der Division durch 5 haben. Da es nur 5 mögliche solche Reste gibt, nämlich $0, 1, 2, 3$ und 4, hat man 5 Äquivalenzklassen:

$$\begin{aligned}
[0] &= \{\dots, -10, -5, 0, 5, 10, \dots\} \\
[1] &= \{\dots, -9, -4, 1, 6, 11, \dots\} \\
[2] &= \{\dots, -8, -3, 2, 7, 12, \dots\} \\
[3] &= \{\dots, -7, -2, 3, 8, 13, \dots\} \\
[4] &= \{\dots, -6, -1, 4, 9, 14, \dots\}
\end{aligned}$$

Diese Auffassung der ganzen Zahlen modulo n hat gegenüber einer Darstellung durch „Repräsentanten“ $0, 1, 2, 3$ und 4 den Vorteil, daß man etwa die Addition und Multiplikation in „natürlicher“ Weise definieren kann, indem man sie von den entsprechenden Operationen auf $\mathbf{Z}$ überträgt. Die Addition modulo 5 von 3 und 4 bildet man, indem man 3 und 4 in $\mathbf{Z}$ bildet und von dem Ergebnis 7 den Rest bei der Division durch 5 nimmt, also 2.

KAPITEL 2

Natürliche Zahlen

Das mit Abstand wichtigste Beweisprinzip in dieser Vorlesung ist das der mathematischen Induktion. Wir geben im ersten Abschnitt eine genaue Formulierung und Beispiele für verschiedene Varianten des Induktionsprinzips, einschließlich der Wertverlaufsinduktion. Für Zwecke der Datenverarbeitung sind nicht nur abstrakte Strukturen wichtig (wie etwa natürliche, ganze, rationale, reelle und komplexe Zahlen), sondern es kommt aus offensichtlichen Gründen besonders darauf an, wie solche Daten repräsentiert werden können. Wir behandeln diese Frage hier nur für die natürlichen Zahlen und entwickeln ihre sogenannte b-adische Darstellung für eine beliebige natürliche Zahl $b \geq 2$ als Basis. Im letzten Abschnitt behandeln wir den Euklidischen Algorithmus, den wir im nächsten Kapitel bei der Kongruenzenrechnung wesentlich verwenden werden.

2.1. Mathematische Induktion

2.1.1. Induktion. Beweise über natürliche Zahlen führt man meist durch Induktion. Darunter versteht man das Scheme

$$\mathrm{Ind}_{m,A}\colon \forall_m\big(A(0) \to \forall_n(A(n) \to A(\mathrm{S}n)) \to A(m^{\mathbf{N}})\big).$$

Die Induktionsschemata für den Typ $\mathbf{B}$ der booleschen Objekte, den Typ $\mathbf{L}(\rho)$ der Listen von Objekten des Typs ρ und den Typ $\mathbf{P}$ der binär dargestellten positiven Zahlen sind

$$\mathrm{Ind}_{b,A}\colon \forall_b\big(A(\mathtt{tt}) \to A(\mathtt{ff}) \to A(b^{\mathbf{B}})\big),$$
$$\mathrm{Ind}_{l,A}\colon \forall_l\big(A(\mathrm{nil}) \to \forall_{x,l'}(A(l') \to A(x :: l')) \to A(l^{\mathbf{L}(\rho)})\big),$$
$$\mathrm{Ind}_{q,A}\colon \forall_q\big(A(1) \to \forall_p(A(p) \to A(S_0 p)) \to \forall_p(A(p) \to A(S_1 p)) \to A(q^{\mathbf{P}})\big),$$

wobei $x :: l$ steht für $\mathrm{cons}\, x\, l$.

Man beachte, daß hier die Allquantoren sämtlich als auf T „relativiert“ zu lesen sind, also von der Form $\forall_n(T(n) \to \dots)$ oder kürzer $\forall_{n\in T}(\dots)$. Das Induktionsschema $\mathrm{Ind}_{m,A}$ für $\mathbf{N}$ ist also genau genommen

$$\mathrm{Ind}_{b,A}\colon \forall_{m\in T}(A(0) \to \forall_{n\in T}(A(n) \to A(\mathrm{S}(n))) \to A(m)).$$

Aus Gründen der Lesbarkeit lassen wir diese Relativierungen im hier und Folgenden weg.

Wir beginnen mit einfachen Beispielen für Induktionsbeweise.

BEISPIEL.

$$3\sum_{i=0}^{n} i(i+1) = n(n+1)(n+2).$$

BEWEIS. Induktion über n. *Basis.* Für $n = 0$ steht links und rechts 0. *Schritt* $n \mapsto n+1$.

$$\begin{aligned} 3\sum_{i=0}^{n+1} i(i+1) &= 3\sum_{i=0}^{n} i(i+1) + 3(n+1)(n+2) \\ &= n(n+1)(n+2) + 3(n+1)(n+2) \qquad \text{nach IH} \\ &= (n+1)(n+2)(n+3). \end{aligned}$$

Das war zu zeigen. □

BEISPIEL.

$$2\sum_{i=0}^{n} 3^i = 3^{n+1} - 1.$$

BEWEIS. Induktion über n. *Basis.* Für $n = 0$ steht links und rechts 2. *Schritt* $n \mapsto n+1$.

$$\begin{aligned} 2\sum_{i=0}^{n+1} 3^i &= 2\sum_{i=0}^{n} 3^i + 2 \cdot 3^{n+1} \\ &= 3^{n+1} - 1 + 2 \cdot 3^{n+1} \qquad \text{nach IH} \\ &= 3 \cdot 3^{n+1} - 1 \\ &= 3^{n+2} - 1. \end{aligned}$$

Das war zu zeigen. □

Man kann die Induktion auch mit anderen Anfangswerten als 0 beginnen:

BEISPIEL.

$$n^3 < 3^n \qquad \text{für } n \geq 4.$$

BEWEIS. Induktion über n. *Basis.* Für $n = 4$ steht links $4^3 = 64$ und rechts $3^4 = 81$. *Schritt* $n \mapsto n+1$.

$$\begin{aligned} (n+1)^3 &= n^3\Big(\frac{n+1}{n}\Big)^3 \\ &\leq n^3\Big(\frac{5}{4}\Big)^3 \qquad \text{da } n \geq 4 \end{aligned}$$

$$\begin{aligned} &= \frac{125}{64} n^3 \\ &< 3n^3 \\ &< 3 \cdot 3^n \qquad \text{nach IH} \\ &= 3^{n+1}. \end{aligned}$$

Das war zu zeigen. □

2.1.2. Induktion mit mehreren Rückgriffen. Ähnlich wie bei der Rekursion kann man auch bei der Induktion mehrere Rückgriffe erlauben. Sei a_n die n-te Fibonacci-Zahl, also $a_0 := 0$, $a_1 := 1$ und $a_{n+2} := a_n + a_{n+1}$.

Beispiel.

$$a_{n+2} \geq \left(\frac{3}{2}\right)^n \qquad \text{für } n \geq 4.$$

Beweis. Induktion über n. *Basis.* Für $n = 0, 1$ ist die Behauptung richtig. *Schritt* $n, n+1 \mapsto n+2$. Als IH haben wir

$$a_{n+2} \geq \left(\frac{3}{2}\right)^n \quad \text{und} \quad a_{n+3} \geq \left(\frac{3}{2}\right)^{n+1}$$

Man erhält

$$\begin{aligned} a_{n+4} &= a_{n+2} + a_{n+3} \\ &\geq \left(\frac{3}{2}\right)^n + \left(\frac{3}{2}\right)^{n+1} \qquad \text{nach IH} \\ &= \left(\frac{3}{2}\right)^n (1 + \frac{3}{2}) \\ &> \left(\frac{3}{2}\right)^n \cdot \frac{9}{4} \\ &= \left(\frac{3}{2}\right)^{n+2}. \end{aligned}$$

Das war zu zeigen. □

Die Anzahl und Lage der Rückgriffe kann beliebig sein; man spricht dann von einer *Wertverlaufsinduktion*. Als Beispiel beweisen wir, daß jede natürliche Zahl $n \geq 2$ als Produkt von Primfaktoren geschrieben werden kann.

Definition. Wir nennen eine natürliche Zahl n *zusammengesetzt*, wenn sie Produkt zweier kleinerer Faktoren ist, also

$$Z(n) := \exists_{m,k<n}\, (n = mk),$$

wobei $\exists_{m<n} A := \exists_m (m < n \wedge A)$. Eine natürliche Zahl n heißt *Primzahl*, wenn sie ≥ 2 und nicht zusammengesetzt ist, also

$$P(n) := 2 \leq n \wedge \neg Z(n).$$

SATZ. *Jede natürliche Zahl $n \geq 2$ kann als Produkt von Primfaktoren geschrieben werden.*

BEWEIS. Sei $n \geq 2$. Wir benutzen eine Wertverlaufsinduktion nach n, und verwenden eine Fallunterscheidung nach der (entscheidbaren) Eigenschaft, ob n eine Primzahl ist. *Fall* $P(n)$, also n ist Primzahl. Dann sind wir fertig. *Fall* $\neg P(n)$, also n ist keine Primzahl. Dann ist n zusammengesetzt, es gibt also $m, k < n$ mit $n = mk$. Wegen $2 \leq n$ folgt $2 \leq m, k$. Nach IH für m und k haben wir Darstellungen $m = p_1 \dots p_r$ und $k = q_1 \dots q_s$, also $n = mk = p_1 \dots p_r q_1 \dots q_s$. □

2.1.3. Wertverlaufsinduktion und das Prinzip vom kleinsten Element. Oft verwendet man das sogenannte „Prinzip vom kleinsten Element“, das wir gleich aus der Wertverlaufsinduktion beweisen werden. Es sagt aus, daß jede nicht leere Menge natürlicher Zahlen ein kleinstes Element haben muß. Man beachte, daß der hierbei verwendete Existenzquantor im schwachen Sinn zu verstehen ist: $\tilde{\exists}_n A(n)$ ist *definiert* durch $\neg \forall_n \neg A(n)$.

Als Anwendungsbeipiel geben wir einen zweiten Beweis des vorangehenden Satzes, jetzt aber mit dem schwachen Existenzquantor formuliert.

SATZ. *Jede natürliche Zahl $n \geq 2$ muß sich als Produkt von Primfaktoren schreiben lassen.*

BEWEIS. Angenommen, es gäbe eine natürlich Zahl $n \geq 2$, die sich nicht als Produkt von Primfaktoren schreiben läßt. Nach dem Prinzip vom kleinsten Element muß es dann auch eine kleinste derartige Zahl geben, sagen wir n_0. Offenbar kann n_0 keine Primzahl sein. Deshalb ist n_0 zusammengesetzt, also von der Form mk mit $m, k < n_0$. Da n_0 die kleinste Zahl ist, die sich nicht als Produkt von Primzahlen schreiben läßt, haben wir Darstellungen $m = p_1 \dots p_r$ und $k = q_1 \dots q_s$ mit Primzahlen p_i, q_j. Also ist $n_0 = mk = p_1 \dots p_r q_1 \dots q_s$, im Widerspruch zu unserer Annahme. □

Um einzusehen, daß es sich beim Prinzip vom kleinsten Element um eine Folgerung aus der Wertverlaufsinduktion handelt, brauchen wir eine genaue Formulierung der Wertverlaufsinduktion, als eine allgemeinere Form der Induktion über natürliche Zahlen.

DEFINITION. Unter dem Schema der allgemeinen Induktion oder der Wertverlaufsinduktion verstehen wir

$$\mathrm{GInd}_{n,A} \colon \forall_n \big(\mathrm{Prog}_n A(n) \to A(n) \big), \tag{2.1}$$

wobei $\mathrm{Prog}_n A(n)$ die „Progressivität“ bezüglich der Ordnung $<$ ausdrückt:

$$\mathrm{Prog}_n A(n) := \forall_n \big(\forall_{m<n} A(m) \to A(n) \big).$$

Wir wollen (2.1) aus dem gewöhnlichen Induktionsschema beweisen.

SATZ (Wertverlaufsinduktion). $\forall_n\big(\mathrm{Prog}_n A(n) \to A(n)\big)$.

BEWEIS. Wir fixieren n und nehmen $\mathrm{Prog}_n A(n)$ an. Zu zeigen ist $A(n)$. Dazu betrachten wir die Formel

$$B(n) := \forall_{m<n} A(m) := \forall_m(m < n \to A(m))$$

und beweisen $\forall_n B(n)$ aus $\mathrm{Prog}_n A(n)$ durch gewöhnliche (Null-Nachfolger-) Induktion. Dies genügt, denn aus $B(\mathrm{S}n)$ folgt $A(n)$ wegen $n < \mathrm{S}n$.

Basis. Zu zeigen ist $\forall_m(m < 0 \to A(m))$. Da $m < 0$ dasselbe ist wie $\mathbf{F}$, folgt dies aus Ex-Falso-Quodlibet.

Schritt $n \mapsto \mathrm{S}n$. Nach IH haben wir $\forall_m(m < n \to A(m))$. Zu zeigen ist $\forall_m(m < \mathrm{S}n \to A(m))$. Sei also m mit $m < \mathrm{S}n$ gegeben. Wir beweisen jetzt $A(m)$ durch Fallunterscheidung. *Fall* $m < n$. Dann gilt $A(m)$ nach der IH. *Fall* $m = n$. Aus der IH folgt mit der vorausgesetzten Progressivität $A(n)$, also wegen $m = n$ auch $A(m)$. □

SATZ (Prinzip vom kleinsten Element).

$$\tilde{\exists}_n A(n) \to \tilde{\exists}_n\big(A(n) \land \forall_{m<n} \neg A(m)\big).$$

BEWEIS. Ausgeschrieben lautet die Behauptung

$$\neg\forall_n \neg A(n) \to \neg\forall_n \neg\big(A(n) \land \forall_{m<n} \neg A(m)\big).$$

Wegen $(D \to \neg C) \leftrightarrow \neg(C \land D)$ genügt

$$\neg\forall_n \neg A(n) \to \neg\forall_n\big(\forall_{m<n} \neg A(m) \to \neg A(n)\big).$$

Wegen $(C \to D) \to \neg D \to \neg C$ genügt auch

$$\forall_n\big(\forall_{m<n} \neg A(m) \to \neg A(n)\big) \to \forall_n \neg A(n).$$

Setzen wir $B(n) := \neg A(n)$, so lautet unsere Behauptung

$$\forall_n\big(\forall_{m<n} B(m) \to B(n)\big) \to \forall_n B(n), \quad \text{also}$$
$$\mathrm{Prog}_n B(n) \to \forall_n B(n).$$

Dies ist aber gerade das Prinzip der Wertverlaufsinduktion für $B(n)$. □

2.2. Darstellung natürlicher Zahlen

Wenn man mit konkreten Zahlen rechnen will, ist es aus Effizienzgründen notwendig, eine andere Darstellung als die unäre zu verwenden. Meistens benutzt man die Dezimaldarstellung; wir wollen hier im allgemeinen die Binärdarstellung verwenden (also $b = 2$).

2.2.1. Die b-adische Darstellung natürlicher Zahlen.

SATZ. *Seien a, b natürliche Zahlen mit $a > 0$ und $b > 1$. Dann gibt es ein eindeutig bestimmtes n und eine eindeutig bestimmte Liste $c_0, \ldots, c_n$ natürlicher Zahlen $< b$ mit $c_0 > 0$ so daß*

$$a = \sum_{k=0}^{n} c_k b^{n-k}.$$

BEWEIS. Wir verwenden eine Wertverlaufsinduktion nach a. Sei also $a > 0$, und jedes a' mit $0 < a' < a$ habe eine eindeutige Darstellung der gewünschten Form. Ist $a < b$, so kann man $n = 0$ und $c_0 = a$ wählen. Zum Beweis der Eindeutigkeit nehmen wir an, wir hätten eine weitere Darstellung

$$a = \sum_{k=0}^{n'} c'_k b^{n'-k}.$$

Ist $n' > 0$, so folgt

$$a = c'_0 b^{n'} + \sum_{k=1}^{n'} c'_k b^{n'-k} \geq c'_0 b^{n'} \geq b^{n'} \geq b$$

im Widerspruch zur Annahme $a < b$. Also ist $n' = 0$ und wir erhalten

$$a = c'_0 b^0 = c'_0 = c_0.$$

Sei jetzt also $b \leq a$.

Existenz. Durch Division mit Rest erhalten wir q, r mit

$$a = bq + r \quad \text{und} \quad r < b.$$

Wegen $r < a$ ist $0 < q$, und wegen $1 < b$ ist $q < bq \leq a$. Nach IH für q findet man m und $d_0, \ldots, d_m$ mit $d_k < b$ und $d_0 > 0$ so daß

$$q = \sum_{k=0}^{m} d_k b^{m-k}, \quad \text{also} \quad a = \sum_{k=0}^{m} d_k b^{(m+1)-k} + r.$$

Wir können also setzen $n := m + 1$, $c_n := r$ und $c_k := d_k$ für $k < n$.

Eindeutigkeit. Nehmen wir an, wir hätten eine weitere Darstellung

$$a = \sum_{k=0}^{n'} c'_k b^{n'-k}.$$

Ist $n' = 0$, so erhalten wir

$$a = c'_0 b^0 = c'_0 < b$$

und damit einen Widerspruch. Also ist $n' > 0$ und deshalb

$$a = \underbrace{\Big(\sum_{k=0}^{n'-1} c'_k b^{(n'-1)-k}\Big)}_{q'} \cdot b + c'_{n'}.$$

Wegen $c'_{n'} < b$ folgt aus der Eindeutigkeit der Division mit Rest, daß $q' = q$ und $c'_{n'} = r$ sein muß. Aus der Eindeutigkeit der Darstellung für q folgt $n' - 1 = m$ und $c'_k = d_k$ für $k \leq m$. Deshalb ist

$$\begin{aligned} n' &= m + 1 = n, \\ c'_k &= d_k = c_k \quad \text{für } k \leq m. \end{aligned}$$

Das war zu zeigen. □

2.2.2. Änderung der Basis. Wir wollen uns in diesem Abschnitt mit der Frage befassen, wie man Zahldarstellungen zu verschiedenen Basen ineinander umrechnet.

Der Übergang von einer Basis b zur Basis 10 ist besonders einfach. Dazu gehen wir aus von der oben betrachteten Darstellung einer Zahl zur Basis b

$$(c_0 \dots c_n)_b := \sum_{k=0}^{n} c_k b^{n-k}.$$

Diesen Wert kann man direkt ausrechnen. Zum Beispiel ergibt sich

$$(54321)_7 = 5 \cdot 7^4 + 4 \cdot 7^3 + 3 \cdot 7^2 + 2 \cdot 7 + 1 = 13539.$$

Man kann hier zum Berechnen das sogenannte *Horner-Schema* verwenden:

$$\begin{aligned} &c_0 b^n + c_1 b^{n-1} + \dots + c_{n-2} b^2 + c_{n-1} b + c_n \\ &= ((\dots (c_0 b + c_1) b + \dots + c_{n-2}) b + c_{n-1}) b + c_n \end{aligned}$$

Im Beispiel erhält man

$$\begin{aligned} (54321)_7 &= (((5 \cdot 7 + 4)7 + 3)7 + 2)7 + 1 \\ &= ((39 \cdot 7 + 3)7 + 2)7 + 1 \\ &= (276 \cdot 7 + 2)7 + 1 = 13539. \end{aligned}$$

Vorteile bei der Verwendung des Horner-Schemas sind die geringere Anzahl der Multiplikationen und der kleinere Speicherbedarf.

Für die umgekehrte Richtung, also den Übergang von der Basis 10 zu einer Basis b, benötigt man fortgesetzte Divisionen (mit Rest) durch b. Es handelt sich um eine Art Umkehrung des Horner-Schemas. Aus

$$a = ((\dots (c_0 b + c_1) b + \dots + c_{n-2}) b + c_{n-1}) b + c_n$$

bestimmt man c_n als $a \mod b$. Dann subtrahiert man c_n und dividiert das Ergebnis durch b. Das Resultat ist

$$(\dots(c_0 b + c_1)b + \dots + c_{n-2})b + c_{n-1}$$

Dieses Verfahren kann man fortsetzen zur Bestimmung von $c_{n-1}, \dots, c_0$. Hierbei handelt es sich offensichtlich um den rechnerischen Gehalt des eben geführten Existenzbeweises für die b-adische Darstellung.

Als Beispiel berechnen wir die Darstellungen von 893 im Binär-, Oktal- und Hexadezimal-System. Bei der Hexadezimal-Darstellung muß man für die Zahlen 10 bis 15 neue Ziffern verwenden; wir nehmen dafür A, B, C, D, E, F.

$$\begin{aligned}
893 &= 446 \cdot 2 + 1 \\
446 &= 223 \cdot 2 + 0 \\
223 &= 111 \cdot 2 + 1 \\
111 &= 55 \cdot 2 + 1 \\
55 &= 27 \cdot 2 + 1 \\
27 &= 13 \cdot 2 + 1 \\
13 &= 6 \cdot 2 + 1 \\
6 &= 3 \cdot 2 + 0 \\
3 &= 1 \cdot 2 + 1 \\
1 &= 0 \cdot 2 + 1
\end{aligned}
\qquad
\begin{aligned}
893 &= 111 \cdot 8 + 5 \\
111 &= 13 \cdot 8 + 7 \\
13 &= 1 \cdot 8 + 5 \\
1 &= 0 \cdot 8 + 1
\end{aligned}
\qquad
\begin{aligned}
893 &= 55 \cdot 16 + 13 \\
55 &= 3 \cdot 16 + 7 \\
3 &= 0 \cdot 16 + 3
\end{aligned}$$

Binär:	1101111101
Oktal:	1575
Hexadezimal:	$37D$

Die Binärdarstellung hat für die Verwendung in Rechnern den offensichtlichen Vorteil, daß es nur zwei Ziffern gibt. Andererseits hat sie den Nachteil, daß die Darstellungen von natürlichen Zahlen schnell sehr lang werden, und man sich leicht verrechnet. Neben der vertrauten Dezimaldarstellung verwendet man deshalb auch gerne die Darstellungen zur Basis 8 (oktal) und 16 (hexadezimal). Da 8 und 16 Zweierpotenzen sind, haben sie den Vorteil, daß die Umrechnungen von und in Binärzahlen besonders einfach werden: man muß nur jeweils drei bzw. vier Binärziffern zu Gruppen zusammenfassen, wobei man von rechts anzufangen hat. Zum Beispiel erhält man die Binärdarstellung von $37D_{16}$ wie folgt:

$$\begin{aligned}
3 &= 0011_2 \\
7 &= 0111_2 \\
D_{16} &= 1101_2
\end{aligned}
\qquad\qquad
37D_{16} = 11 \mid 0111 \mid 1101_2$$

Die Oktaldarstellung der Binärzahl 1101111101_2 ergibt sich durch Bildung von Dreiergruppen $1 \mid 101 \mid 111 \mid 101$ als 1575_8.

2.3. Euklidischer Algorithmus

2.3.1. Größter gemeinsamer Teiler. Wir definieren die Teilbarkeit $m \mid n$ durch die Formel $\exists_q(n = mq)$. Aus dieser Definition erhält man

LEMMA. (a) $n \mid 0$.
(b) $0 \mid n \to n = 0$.
(c) $1 \mid n$.
(d) $n \mid 1 \to n = 1$.
(e) $n \mid n$.
(f) $n \mid m \to m \mid k \to n \mid k$.
(g) $n \mid m \to m \mid n \to n = m$.
(h) $n \mid m \to n \mid mk$.
(i) $n \mid m \to n \mid k \to n \mid m + k$.
(j) $n \mid m \to n \mid m + k \to n \mid k$.

DEFINITION. Eine natürliche Zahl d heißt *größter gemeinsamer Teiler* von n und m, wenn gilt
(a) $d \mid n$ und $d \mid m$;
(b) $\forall_q(q \mid n \to q \mid m \to q \mid d)$.

Die Eindeutigkeit des größten gemeinsamen Teilers ist eine einfache Folgerung aus Teil (g) des Lemmas. Zum Beweis der Existenz benötigen wir eine Wertverlaufsinduktion.

SATZ (Euklid). *Zu beliebigen natürlichen Zahlen n, m mit $n > m$ gibt es genau einen größten gemeinsamen Teiler d.*

BEWEIS. Die Eindeutigkeit hatten wir bereits bewiesen. Den Beweis der Existenz führen wir durch Wertverlaufsinduktion über n. Seien also n, m mit $n > m$ gegeben. Ist $m = 0$, so ist n größter gemeinsamer Teiler von n und m. Sei also $m > 0$. Division mit Rest ergibt

$$n = mq + r \quad \text{und} \quad r < m.$$

Wir zeigen zunächst $d \mid n \to d \mid m \to d \mid r$. Gelte also $d \mid n$ und $d \mid m$. Dann hat man k, l mit $n = dk$ und $m = dl$, also $dk = dlq + r$ und damit $d \mid r$ nach Teil (j) des Lemmas. Es gilt deshalb für beliebige d

$$d \mid n \wedge d \mid m \leftrightarrow d \mid m \wedge d \mid r.$$

Wir zeigen jetzt

$$d \text{ ist ggT von } n, m \leftrightarrow d \text{ ist ggT von } m, r. \tag{2.2}$$

Sei also d der größte gemeinsame Teiler von n, m. Dann gilt

(a) $d \mid n$ und $d \mid m$;
(b) $\forall_s(s \mid n \to s \mid m \to s \mid d)$.

Wir zeigen, daß d auch der größte gemeinsame Teiler von m, r ist. $d \mid r$ folgt aus der Vorüberlegung. Zu zeigen bleibt $\forall_t(t \mid m \to t \mid r \to t \mid d)$. Gelte also $t \mid m$ und $t \mid r$. Zu zeigen ist $t \mid d$. Wegen $n = mq + r$ folgt $t \mid n$ aus Teil (i) des Lemmas. Also gilt $t \mid d$ nach Voraussetzung. Die umgekehrte Richtung von (2.2) zeigt man ähnlich.

Nach IH haben wir einen größten gemeinsamen Teiler d von m, r. Aufgrund von (2.2) ist d auch größter gemeinsamer Teiler von n, m. □

DEFINITION. Den eindeutig bestimmten größten gemeinsamen Teiler von n und m bezeichnen wir mit $\mathrm{ggT}(n, m)$. Zwei natürliche Zahlen n, m heißen *teilerfremd*, wenn $\mathrm{ggT}(n, m) = 1$.

Das in dem Beweis verwendete Verfahren nennt man den *Euklidischen Algorithmus* zur Berechnung des größten gemeinsamen Teilers zweier Zahlen. Zum Beispiel erhält man $\mathrm{ggT}(66, 27)$ wie folgt.

$$66 = 27 \cdot 2 + 12$$
$$27 = 12 \cdot 2 + 3$$
$$12 = 3 \cdot 4.$$

Der größte gemeinsame Teiler von 66 und 27 ist also 3.

Wir zeigen jetzt, daß $\mathrm{ggT}(n, m)$ sich linear aus n und m kombinieren läßt.

SATZ. *Zu beliebigen natürlichen Zahlen n, m mit $n > m$ gibt es k, l mit* $\mathrm{ggT}(n, m) = |nk - ml|$.

BEWEIS. Wir führen den Beweis durch Wertverlaufsinduktion über n. Seien also n, m mit $n > m$ gegeben. Sei $d := \mathrm{ggT}(n, m)$.

Fall $m = 0$. Setze $k := 1$ und $l := 0$.

Fall $m > 0$. Division mit Rest ergibt

$$n = mq + r \quad \text{und} \quad r < m.$$

Nach der IH für m haben wir s, t mit $d = \mathrm{ggT}(m, r) = |ms - rt|$. Wir setzen $k := t$ und $l := qt + s$, und verwenden eine Fallunterscheidung. Ist $ms \leq rt$, so erhält man

$$ml = m(qt + s) = mqt + ms \leq mqt + rt = (mq + r)t = nt = nk$$

und

$$nk - ml = rt - ms = d.$$

Ist umgekehrt $ms > rt$, so erhält man

$$ml = m(qt + s) = mqt + ms > mqt + rt = (mq + r)t = nt = nk$$

und

$$ml - nk = ms - rt = d. \qquad \square$$

KOROLLAR. *Sind n und q teilerfremd und gilt $q \mid nm$, so folgt $q \mid m$.*

BEWEIS. Nach Annahme ist $\mathrm{ggT}(n, q) = 1$. Nach dem vorigen Satz gibt es also k, l mit $1 = |nk - ql|$. Sei etwa $nk \leq ql$. Dann hat man $1 = ql - nk$, also $1 + nk = ql$, also auch $m + mnk = mql$. Wegen $q \mid mn$ folgt $q \mid m$ aus Teil (j) des Lemmas. $\square$

2.3.2. Eindeutigkeit der Primfaktorzerlegung. Wir haben bereits gezeigt, daß jede natürliche Zahl $n \geq 2$ als Produkt von Primfaktoren geschrieben werden kann. Jetzt wollen wir beweisen, daß diese Darstellung bis auf die Reihenfolge eindeutig ist.

SATZ. *Jede natürliche Zahl $n \geq 2$ kann auf höchstens eine Weise (bis auf die Reihenfolge der Faktoren) als Produkt von Primfaktoren geschrieben werden.*

BEWEIS. Sei $n \geq 2$. Wir beginnen mit einer *Vorbemerkung*: Gilt $p \mid q_1 \dots q_s$ mit Primzahlen $p, q_1, \dots, q_s$, so ist $p = q_j$ für ein j. Den Beweis führen wir durch Induktion über s. *Basis.* Aus $p \mid 1$ folgt $\mathbf{F}$ und wir können Ex-Falso-Quodlibet verwenden. *Schritt* $s \mapsto s + 1$. *Fall* $p = q_{s+1}$. Dann ist die Behauptung offenbar richtig. *Fall* $p \neq q_{s+1}$. Dann sind p und q_{s+1} teilerfremd, also $p \mid q_1 \dots q_s$. Nach der IH folgt $p = q_j$ für ein j mit $1 \leq j \leq s$.

Sei jetzt $m = p_1 \dots p_r = q_1 \dots q_s$ mit Primzahlen $p_1, \dots, p_r, q_1, \dots, q_s$. Wir beweisen durch Induktion über r, daß dann $r = s$ gilt und und beide Darstellungen bis auf die Reihenfolge gleich sind. *Basis.* Dann ist $m = 1$ und deshalb $s = 0$. *Schritt* $r \mapsto r + 1$. Wir haben offenbar $p_{r+1} \mid q_1 \dots q_s$ und $s \geq 1$, also $p_{r+1} = q_j$ für ein j nach der Vorbemerkung. OBdA sei $p_{r+1} = q_s$. Man erhält $p_1 \dots p_r = q_1 \dots q_{s-1}$. Mit der IH folgt die Behauptung. $\square$

2.3.3. Existenz unendlich vieler Primzahlen.

SATZ (Euklid). *Es gibt unendlich viele Primzahlen, das heißt, zu jeder endlichen Liste $p_1, \dots, p_n$ von Primzahlen findet man eine von ihnen allen verschiedene Primzahl q.*

BEWEIS. Wir betrachten $p_1 \dots p_n + 1$ und führen den Beweis durch Fallunterscheidung. *Fall* $p_1 \dots p_n + 1$ ist Primzahl. Dann haben wir eine von $p_1, \dots, p_n$ verschiedene Primzahl gefunden. *Fall* $p_1 \dots p_n + 1$ ist keine Primzahl. Nach dem Satz über die Primfaktorzerlegung gibt es eine Primzahl q mit $q \mid (p_1 \dots p_n + 1)$. Wäre $q = p_i$ für ein i mit $1 \leq i \leq n$, so folgte $q \mid 1$ und damit ein Widerspruch. $\square$

DEFINITION. Mit p_n bezeichnen wir die (unendliche) Folge der Primzahlen, also $p_0 = 2$, $p_1 = 3$, $p_3 = 5$ und so weiter.

KOROLLAR. *Jede natürliche Zahl $n \geq 1$ läßt sich darstellen in der Form*

$$n = p_0^{i_0} p_1^{i_1} \dots p_r^{i_r}.$$

Diese Darstellung ist eindeutig, wenn man $0 < i_r$ verlangt.

KOROLLAR. *Es seien n, m natürliche Zahlen mit*

$$n = p_0^{i_0} p_1^{i_1} \dots p_r^{i_r}, \qquad m = p_0^{j_0} p_1^{j_1} \dots p_r^{j_r}.$$

Dann ist

$$\mathrm{ggT}(n, m) = p_0^{\min(i_0, j_0)} p_1^{\min(i_1, j_1)} \dots p_r^{\min(i_r, j_r)}.$$

BEWEIS. Übung. □

KAPITEL 3

Algebraische Grundbegriffe, Kongruenzen

Wir entwickeln die Anfänge der Gruppen- und der Ringtheorie in dem später benötigten Umfang. Ringe oder genauer Halbringe und die Matrizenmultiplikation über Halbringen werden wir zu Komplexitätsbetrachtungen in der Graphentheorie verwenden. In der Gruppentheorie beweisen wir den sogenannten kleinen Fermatschen Satz, den wir in der im letzten Abschnitt behandelten Arithmetik modulo n benötigen. Sie spielt an vielen Stellen in der Informatik eine wichtige, grundlegende Rolle spielt, etwa in der Kryptographie.

3.1. Gruppen

Wir erinnern an den aus der linearen Algebra bekannten Gruppenbegriff. Mit Hilfe von Nebenklassen führen wir den Begriff des Index einer Untergruppe ein und beweisen den Satz von Lagrange, der eine nützliche Anzahlaussage über endliche Gruppen und ihre Untergruppen macht. Mit seiner Hilfe beweisen wir den kleinen Fermatschen Satz.

3.1.1. Definition und einfache Eigenschaften von Gruppen.

DEFINITION. Seien G eine Menge und $\circ\colon G \to G \to G$ eine Abbildung. $(G, \circ)$ (oder oft nur kurz G) heißt *Gruppe*, wenn gilt

(a) $(x \circ y) \circ z = x \circ (y \circ z)$ für alle $x, y, z \in G$ (*Assoziativgesetz*).
(b) Es gibt ein $e \in G$ (genannt *neutrales Element* von G) mit
  (i) $\forall_{x \in G}(e \circ x = x)$,
  (ii) $\forall_{x \in G} \exists_{x' \in G}(x' \circ x = e)$ (x' heißt *inverses Element* zu x).

G heißt *abelsch*, wenn außerdem noch gilt $\forall_{x,y \in G}(x \circ y = y \circ x)$ (*Kommutativgesetz*).

BEISPIEL. $(\mathbf{Z}, +)$ ist eine abelsche Gruppe.

LEMMA. *Sei G eine Gruppe.*

(a) *(Linksinverse Elemente sind auch rechtsinvers). Sei $e \in G$ ein neutrales Element und $x, x' \in G$ derart, daß $x' \circ x = e$. Dann gilt auch $x \circ x' = e$.*
(b) *(Linksneutrale Elemente sind auch rechtsneutral). Sei $e \in G$ ein neutrales Element. Dann gilt $x \circ e = x$ für alle $x \in G$.*

(c) *Es gibt genau ein neutrales Element $e \in G$.*
(d) *Zu jedem $x \in G$ gibt es genau ein inverses Element $x' \in G$.*

BEWEIS. (a) Wähle x'' mit $x'' \circ x' = e$. Dann gilt

$$x \circ x' = e \circ x \circ x' = x'' \circ \underbrace{x' \circ x}_{e} \circ x' = x'' \circ x' = e.$$

(b) $x \circ e = x \circ x' \circ x = e \circ x = x$ nach (a).
(c) Seien e, e^* neutrale Elemente von G. Nach (b) gilt $e^* = e \circ e^* = e$.
(d) Seien x', x^* inverse Elemente zu $x \in G$. Dann gilt

$$x^* = x^* \circ e = x^* \circ x \circ x' = e \circ x' = x'.$$

Hierbei haben wir (a) - (c) benutzt. □

Das zu x eindeutig bestimmte inverse Element wird mit x^{-1} bezeichnet. Wir schreiben x^n für $x \circ \cdots \circ x$ und x^{-n} für $x^{-1} \circ \cdots \circ x^{-1}$, jeweils mit n Vorkommen von x bzw. x^{-1}. Ferner sei $x^0 := e$.

LEMMA. *Seien G eine nicht leere Menge und $\circ\colon G \to G \to G$ eine Abbildung. G ist eine Gruppe genau dann, wenn gilt:*
(a) $(x \circ y) \circ z = x \circ (y \circ z)$ *für alle $x, y, z \in G$ (*Assoziativgesetz*).*
(b) (i) $\forall_{x,y \in G} \exists_{z \in G} (x \circ z = y)$.
(ii) $\forall_{x,y \in G} \exists_{z \in G} (z \circ x = y)$.

BEWEIS. Übung.
$\rightarrow$. Sei G eine Gruppe. Zum Beweis von $\forall_{x,y \in G} \exists_{z \in G} (x \circ z = y)$ genügt es, $z = x^{-1} \circ y$ zu setzen, und zum Beweis von $\forall_{x,y \in G} \exists_{z \in G} (z \circ x = y)$ genügt es, $z = y \circ x^{-1}$ zu setzen.
$\leftarrow$. G erfülle die im Lemma aufgelisteten Eigenschaften. Zu zeigen ist, daß G eine Gruppe ist. Wir müssen also ein neutrales Element e finden derart, daß gilt

$$e \circ x = x \text{ für alle } x \in G, \text{ und}$$
$$\forall_{x \in G} \exists_{x' \in G} (x' \circ x = e).$$

Wir konstruieren zunächst ein solches e. Wähle $x_0 \in G$ fest (hier wird benötigt, daß G nicht leer ist). Wähle e so, daß $e \circ x_0 = x_0$ ist.

Wir zeigen jetzt die beiden obigen Eigenschaften. Sei $x \in G$ beliebig. Wähle $z \in G$ mit $x_0 \circ z = x$. Dann gilt

$$e \circ x = e \circ x_0 \circ z = x_0 \circ z = x.$$

Die zweite Eigenschaft folgt aus Teil (ii) von (b). □

LEMMA. *Seien G eine Gruppe und $x, y \in G$. Dann gilt*
(a) $(x^{-1})^{-1} = x$.
(b) $(x \circ y)^{-1} = y^{-1} \circ x^{-1}$.

BEWEIS. (a). Es gilt $(x^{-1})^{-1} \circ x^{-1} = e$ und auch $x \circ x^{-1} = e$. Die Behauptung folgt aus der Eindeutigkeit des Inversen.
(b). $y^{-1} \circ x^{-1} \circ x \circ y = y^{-1} \circ e \circ y = e$. □

BEISPIELE. Seien M eine nicht leere Menge und $S(M)$ die Menge der bijektiven Abbildungen von M auf sich selbst. Mit $\circ$ bezeichnen wir die Komposition (Hintereinanderausführung) von Abbildungen. Dann ist $(S(M), \circ)$ eine Gruppe. $(S(M), \circ)$ heißt die *symmetrische Gruppe* der Menge M. Neutrales Element ist die Identität $\mathrm{id}\colon M \to M$, und das inverse Element zu $f \in S(M)$ ist die Umkehrabbildung f^{-1}.

Man beachte, daß $S(M)$ i.a. nicht abelsch ist. Sei zum Beispiel

$$f\colon \{0,1,2\} \to \{0,1,2\} \qquad g\colon \{0,1,2\} \to \{0,1,2\}$$

$$f(a) = \begin{cases} 0 & \text{falls } a = 0 \\ 2 & \text{falls } a = 1 \\ 1 & \text{falls } a = 2 \end{cases} \quad \text{und} \quad g(a) = \begin{cases} 1 & \text{falls } a = 0 \\ 0 & \text{falls } a = 1 \\ 2 & \text{falls } a = 2. \end{cases}$$

Dann gilt $(g \circ f)(0) = 1$, aber $(f \circ g)(0) = 2$, also $g \circ f \neq f \circ g$.

Im Spezialfall $M = \{1, \ldots, n\}$ schreibt man S_n statt $S(M)$. Jede Abbildung $\sigma \in S_n$ heißt eine *Permutation* der Zahlen $1, \ldots, n$.

Ein wichtiges Beispiel einer endlichen abelschen Gruppe ist die Gruppe $\mathbf{Z}_n := \{\, x \in \mathbf{N} \mid x < n \,\}$ der *ganzen Zahlen modulo* n. Die Gruppenverknüpfung ist die *Addition modulo* n, die für $x, y \in \mathbf{Z}_n$ definiert ist als der (eindeutig bestimmte) Rest bei der Division von $x+y$ durch n. Man schreibt $x + y \equiv r \mod n$, falls $x + y = nq + r$ mit $q, r \in \mathbf{N}$ und $r < n$. Der Beweis der Gruppeneigenschaften ist einfach: für die Assoziativität verwendet man Fallunterscheidungen nach den Summen der jeweiligen Reste. Neutrales Element ist die 0, und das inverse Element zu x ist $n - x$.

3.1.2. Untergruppen. Eine Teilmenge $U \subseteq G$ wollen wir eine Untergruppe nennen, wenn die Gruppenstruktur auf G eine Gruppenstruktur auf U induziert. Genauer heißt das:

DEFINITION. Es sei G eine Gruppe und $U \subseteq G$ eine Teilmenge von G. U heißt *Untergruppe* von G, wenn gilt
(a) $xy \in U$ für alle $x, y \in U$,
(b) $e \in U$,
(c) $x^{-1} \in U$ für alle $x \in U$.

Da aus dem Zusammenhang klar ist, welche Gruppenverknüpfung gemeint ist, haben wir hier kurz xy anstelle von $x \circ y$ geschrieben. Dies werden wir auch im folgenden tun.

SATZ (Untergruppenkriterium). *Es sei G eine Gruppe und $U \subseteq G$ eine Teilmenge von G.*

(a) *U ist Untergruppe von G genau dann, wenn*
   (i) $U \neq \emptyset$,
   (ii) $xy^{-1} \in U$ *für alle* $x, y \in U$.
(b) *Sei U eine endliche Menge. U ist Untergruppe von G genau dann, wenn*
   (i) $U \neq \emptyset$,
   (ii) $xy \in U$ *für alle* $x, y \in U$.

BEWEIS. (a). $\rightarrow$. Dies folgt sofort aus der Definition von Untergruppen. $\leftarrow$. Da $U \neq \emptyset$, existiert ein $x_0 \in U$. Damit haben wir eine Einheit $e := x_0 x_0^{-1} \in U$. Sei jetzt $x \in U$. Dann ist auch $x^{-1} = ex^{-1} \in U$. Seien schließlich $x, y \in U$. Dann ist auch $y^{-1} \in U$ und damit $xy = x(y^{-1})^{-1} \in U$.

(b). $\rightarrow$. Dies folgt wieder aus der Definition von Untergruppen. $\leftarrow$. Sei $x \in U$. Betrachte die Abbildung

$$\hat{x}\colon U \to U \quad \text{mit } \hat{x}(y) := xy.$$

Diese Abbildung ist wohldefiniert nach Bedingung (ii). $\hat{x}$ ist injektiv, denn aus $\hat{x}(y) = \hat{x}(z)$ folgt $xy = xz$, also $x^{-1}xy = x^{-1}xz$ und damit auch $y = z$. Da U endlich ist, ist $\hat{x}$ damit schon bijektiv. Sei nun $x \in U$ beliebig, aber fest (die Existenz ist klar, da $U \neq \emptyset$). Da $\hat{x}$ bijektiv ist, existiert $y \in U$ mit $\hat{x}(y) = x$. Dann gilt $xy = x$ und somit $y = e$, also $e \in U$. Außerdem existiert $z \in U$ mit $\hat{x}(z) = e$; dann gilt $xz = e$ und damit $z = x^{-1}$, also $x^{-1} \in U$. Da $x \in U$ beliebig gewählt war, folgt die Behauptung. □

3.1.3. Homomorphismen für Gruppen.

DEFINITION. Seien G, H Gruppen und $f\colon G \to H$ eine Abbildung.

(a) f heißt *Homomorphismus* für Gruppen, wenn für alle $x, y \in G$ gilt
$$f(xy) = f(x)f(y).$$
(b) f heißt *Mono-, Epi- bzw. Isomorphismus*, wenn f Homomorphismus und injektiv, surjektiv bzw. bijektiv ist.
(c) f heißt *Endo- bzw. Automorphismus*, wenn $G = H$ und f Homo- bzw. Isomorphismus ist.
(d) G und H heißen *isomorph* ($G \cong H$), wenn es einen Isomorphismus $g\colon G \to H$ gibt.

LEMMA. *Seien G, H Gruppen und $f\colon G \to H$ ein Homomorphismus. e bzw. e' seien die neutralen Elemente von G bzw. H. Dann gilt*

(a) $f(e) = e'$, *und* $f(x^{-1}) = f(x)^{-1}$ *für alle* $x \in G$.
(b) *Ist U Untergruppe von G, so ist $f(U)$ Untergruppe von H. Ist V Untergruppe von H, so ist $f^{-1}(V)$ Untergruppe von G. Insbesondere ist*
$$\mathrm{Kern}(f) := f^{-1}(\{e'\})$$
eine Untergruppe von G.

(c) $\mathrm{Kern}(f) = \{e\}$ *genau dann, wenn* f *injektiv ist.*
(d) f *ist Isomorphismus genau dann, wenn ein Homomorphismus* $g\colon H \to G$ *mit* $g \circ f = \mathrm{id}_G$ *und* $f \circ g = \mathrm{id}_H$ *existiert.*

BEWEIS. (a). $f(e) = f(ee) = f(e)f(e)$, also $f(e) = e'$. Außerdem ist $f(x)f(x^{-1}) = f(xx^{-1}) = f(e) = e'$ also $f(x^{-1}) = f(x)^{-1}$.

(b) Sei U Untergruppe von G. Da $U \neq \emptyset$, folgt $f(U) \neq \emptyset$. Ferner gilt für alle $x, y \in U$

$$f(x)f(y)^{-1} = f(x)f(y^{-1}) = f(xy^{-1}) \in f(U).$$

Also ist nach dem Untergruppenkriterium $f(U)$ eine Untergruppe von H.

Sei V Untergruppe von H. Dann gilt $f^{-1}(V) \neq \emptyset$, da $f(e) = e' \in V$, also $e \in f^{-1}(V)$. Seien nun $x, y \in f^{-1}(V)$ und damit $f(x), f(y) \in V$. Da V Untergruppe ist, folgt $f(x)f(y)^{-1} = f(xy^{-1}) \in V$, also $xy^{-1} \in f^{-1}(V)$. Also ist $f^{-1}(V)$ Untergruppe von G nach dem Untergruppenkriterium.

(c) $\rightarrow$. Seien $x, y \in G$ mit $f(x) = f(y)$. Dann gilt $e' = f(x)f(y)^{-1} = f(xy^{-1})$ und somit $xy^{-1} = e$, also $x = y$.

$\leftarrow$. $f(e) = e'$, also $e \in \mathrm{Kern}(f)$. Sei $x \in \mathrm{Kern}(f)$. Dann gilt $f(x) = e'$ und damit $x = e$, da f injektiv ist. Also ist $\mathrm{Kern}(f) = \{e\}$.

(d). $\rightarrow$. Setze $g := f^{-1}$. Zu zeigen bleibt, daß f^{-1} ein Homomorphismus ist. Seien $x, y \in H$ mit $x = f(u)$, $y = f(v)$ für $u, v \in G$. Dann gilt $f^{-1}(x)f^{-1}(y) = uv = f^{-1}(f(uv)) = f^{-1}(f(u)f(v)) = f^{-1}(xy)$. Also ist f^{-1} Homomorphismus.

$\leftarrow$. Klar, da aus der Existenz von g die Bijektivität von f folgt. □

3.1.4. Nebenklassen.

DEFINITION. Sei G eine Gruppe und $U \subseteq G$ eine Untergruppe.

(a) Sei $x \in G$. Dann heißt

$$xU := \{\, xu \mid u \in U \,\}$$

die von x erzeugte *Linksnebenklasse* bzgl. U. Analog heißt

$$Ux := \{\, ux \mid u \in U \,\}$$

die von x erzeugte *Rechtsnebenklasse* bzgl. U.
(b) $G/U := \{\, xU \mid x \in G \,\}$. (Sprechweise: „$G$ modulo U“).

LEMMA. *Sei* G *eine Gruppe und* $U \subseteq G$ *eine Untergruppe. Dann gilt*

(a) $xU = yU$ *genau dann, wenn* $x^{-1}y \in U$. *Definiert man für* $x, y \in G$

$$(x \sim_U y) := (x^{-1}y \in U)$$

(Schreibweise: $x \equiv y \mod U$, *„*x *kongruent* y *modulo* U *“), so ist also* $\sim_U$ *eine Äquivalenzrelation auf* G.
(b) $xU = \{\, y \mid y \sim_U x \,\}$ *(=:* $[x]$, *Äquivalenzklasse von* x*).*

(c) *Wähle in jeder Linksnebenklasse ein Element (Repräsentant) x_i. $(x_i)_{i\in I}$ sei die Familie der Repräsentanten. Dann gilt*

$$G = \bigcup_{i\in I} x_i U \quad \textit{(disjunkte Vereinigung).}$$

BEWEIS. (a). Es genügt, die erste Aussage zu beweisen. $\rightarrow$. Da $y = ye \in yU = xU$, gilt $y = xu$ für ein $u \in U$. Damit folgt $x^{-1}y = u \in U$.

$\leftarrow$. $x^{-1}y = u \in U$, also $y = xu$. Dann folgt für alle $v \in U$, daß $yv = xuv \in xU$, also $yU \subseteq xU$. Wegen $(x^{-1}y)^{-1} = y^{-1}x$ folgt analog $xU \subseteq yU$.

(b). $\subseteq$. Sei $u \in U$. Dann gilt $(xu)^{-1}x = u^{-1}x^{-1}x = u^{-1} \in U$, also $xu \sim_U x$.

$\supseteq$. Sei $y \sim_U x$, also $y^{-1}x \in U$. Dann gibt es ein $u \in U$ mit $y^{-1}x = u$. Folglich gilt $y = xu^{-1}$, also $y \in xU$.

(c). Dies gilt bekanntlich für jede Äquivalenzrelation (siehe 1.3.3). □

DEFINITION (Index einer Untergruppe). Für eine Menge M bezeichnen wir mit $|M|$ die Anzahl der Elemente von M ($|M| = \infty$ ist zugelassen). Sei nun G eine Gruppe und $U \subseteq G$ eine Untergruppe.

$$[G\colon U] := |G/U| \quad \text{heißt } \textit{Index} \text{ von } U \text{ in } G.$$

$[G\colon U]$ gibt die Anzahl der Linksnebenklassen an.

SATZ (Lagrange). *Sei G eine Gruppe und $U \subseteq G$ eine Untergruppe. Dann gilt*

$$|G| = |U| \cdot [G\colon U].$$

BEWEIS. Falls $|U| = \infty$, so gilt die Behauptung. Falls $|U| < \infty$, genügt es zu zeigen $|xU| = |U|$ für alle $x \in G$ (mit Teil (c) des obigen Lemmas folgt dann die Behauptung). Betrachte die Abbildungen

$$\hat{x}\colon U \to xU \qquad\qquad \bar{x}\colon xU \to U,$$
$$\hat{x}(u) = xu \qquad\qquad \bar{x}(v) = x^{-1}v.$$

($\bar{x}$ ist wohldefiniert, denn aus $v \in xU$ folgt $x^{-1}v \in U$). Dann ist $\bar{x} \circ \hat{x} = \mathrm{id}_U$ und $\hat{x} \circ \bar{x} = \mathrm{id}_{xU}$. Also ist $\hat{x}$ bijektiv, und damit $|xU| = |U|$. □

KOROLLAR (Kleiner Fermatscher Satz). *Sei G eine endliche Gruppe. Dann gilt für alle $x \in G$*

$$x^{|G|} = e.$$

BEWEIS. Sei $x \in G$ beliebig, aber fest gewählt. $U := \{\, x^n \mid n \in \mathbf{N} \,\}$ ist abelsche Untergruppe von G, da $U \neq \emptyset$ und $x^n x^m = x^{n+m} \in U$ (Untergruppenkriterium für endliche Gruppen). Es gilt $x^{|U|} = e$, denn

$$\prod_{y\in U}(xy) = x^{|U|} \prod_{y\in U} y \quad \text{da } U \text{ abelsch,}$$

$$\prod_{y\in U}(xy)=\prod_{y\in U}y,\qquad \text{da } \hat{x}\colon U\to U,\ \hat{x}(y)=xy \text{ bijektiv ist.}$$

Also folgt $x^{|G|}=x^{|U|\cdot[G\colon U]}=(x^{|U|})^{[G\colon U]}=e$. □

3.1.5. Zyklische Gruppen.

DEFINITION. Sei $(G,\circ)$ eine Gruppe. G heißt *zyklisch*, wenn es ein $x\in G$ gibt mit $G=\{\,x^i\mid i\in\mathbf{Z}\,\}=:\langle x\rangle$.

BEISPIELE. (a) $\mathbf{Z}$ ist zyklisch, denn $\mathbf{Z}=\langle 1\rangle$.
(b) $n\mathbf{Z}:=\{\,nx\mid x\in\mathbf{Z}\,\}$ $(n\in\mathbf{N})$ ist ebenfalls zyklisch, denn $n\mathbf{Z}=\langle n\rangle$.
(c) $\mathbf{Z}_n:=\{0,1,\dots,n-1\}$ $(n\in\mathbf{N},\ n>0)$ mit der Addition modulo n als Gruppenverknüpfung ist zyklisch, denn $\mathbf{Z}_n=\langle 1\rangle$.

SATZ (Bestimmung der Untergruppen von $\mathbf{Z}$). *Die Untergruppen von* $\mathbf{Z}$ *sind genau alle* $n\mathbf{Z}$, $n\in\mathbf{N}$.

BEWEIS. Alle $n\mathbf{Z}$ sind offenbar Untergruppen von $\mathbf{Z}$. Sei nun $U\subseteq\mathbf{Z}$ eine Untergruppe von $\mathbf{Z}$. Zu zeigen ist

$$\tilde{\exists}_{n\in\mathbf{N}}(U=n\mathbf{Z}).$$

Sei o.B.d.A. $U\neq\{0\}$ (sonst setze $n:=0$). Sei n die kleinste positive Zahl in U (da U Untergruppe ist, existieren positive Zahlen in U). Dann gilt $U=n\mathbf{Z}$. Begründung: Da $n\in U$ und U Untergruppe, gilt $n\mathbf{Z}\subseteq U$. Sei nun $x\in U$. Es existieren $s\in\mathbf{Z}$ und $r\in\{0,\dots,n-1\}$ mit $x=s\cdot n+r$. Daraus folgt $r=x-s\cdot n\in U$, also $r=0$ nach Wahl von n. Damit ist auch $x=s\cdot n\in n\mathbf{Z}$ und folglich $U\subseteq n\mathbf{Z}$, insgesamt also $U=n\mathbf{Z}$. □

SATZ (Bestimmung der zyklischen Gruppen). *Sei G zyklische Gruppe. Dann gilt*

$$G\cong\mathbf{Z}\ \textit{oder}\ G\cong\mathbf{Z}_n\ \textit{für ein}\ n\in\mathbf{N},\ n>0.$$

BEWEIS. Sei G eine zyklische Gruppe und $G=\langle x\rangle=\{\,x^i\mid i\in\mathbf{Z}\,\}$. Betrachte die Abbildung

$$f\colon\mathbf{Z}\to\langle x\rangle,\quad f(i)=x^i.$$

f ist Homomorphismus, da $x^{i+j}=x^i\cdot x^j$. f ist auch surjektiv, denn es gilt $f(\mathbf{Z})=\{\,x^i\mid i\in\mathbf{Z}\,\}=\langle x\rangle$.

Fall $\mathrm{Kern}(f)=\{0\}$. Dann ist f injektiv, also bijektiv, und damit gilt $\mathbf{Z}\cong\langle x\rangle$.

Fall $\mathrm{Kern}(f)\neq\{0\}$. Da $\mathrm{Kern}(f)$ Untergruppe von $\mathbf{Z}$ ist, existiert nach dem vorigen Satz ein $n>0$ mit $\mathrm{Kern}(f)=n\mathbf{Z}$. Damit folgt $x^n=e$ und $x^m\neq e$ für alle $m\in\{1,\dots,n-1\}$. Nun sind alle x^i für $i\in\{0,\dots,n-1\}$ verschieden (denn für $0\le r<s<n$ ist $0<s-r<n$, also $x^{s-r}\neq e$ und damit $x^r\neq x^s$). Also ist $\langle x\rangle=\{x^0,x^1,\dots,x^{n-1}\}$. □

DEFINITION. Sei G eine Gruppe und $x \in G$. Dann heißt

$$\operatorname{ord}(x) := \begin{cases} \text{kleinstes } n > 0 \text{ mit } x^n = e & \text{falls eines existiert,} \\ \infty & \text{sonst} \end{cases}$$

die *Ordnung* von x.

SATZ. *Sei G eine endliche Gruppe und $x \in G$. Dann ist* $\operatorname{ord}(x)$ *Teiler der Gruppenordnung $|G|$. Ist $|G|$ eine Primzahl, so ist G zyklisch.*

BEWEIS. $\langle x \rangle$ ist eine Untergruppe von G. Mit dem Satz von Lagrange folgt, daß $|\langle x \rangle|$ ein Teiler von $|G|$ ist, und es ist $|\langle x \rangle| = \operatorname{ord}(x)$. Falls $|G|$ Primzahl ist, so ist demnach $\operatorname{ord}(x) = |G|$ für alle $x \neq e$. Damit ist G zyklisch. □

3.1.6. Normalteiler. Ist G eine Gruppe und $U \subseteq G$ eine Untergruppe, so trägt G/U im allgemeinen noch keine Gruppenstruktur. Um dies zu erreichen, benötigen wir eine weitere Eigenschaft von U, daß nämlich U ein Normalteiler ist.

DEFINITION. Eine Untergruppe N einer Gruppe G heißt *Normalteiler*, wenn für alle $x \in G$ gilt

$$xN = Nx,$$

wobei $xN := \{ xy \mid y \in N \}$ und $Nx := \{ yx \mid y \in N \}$.

LEMMA. *Sei G eine Gruppe und $U \subseteq G$ eine Teilmenge. U ist genau dann Normalteiler von G, wenn U eine Untergruppe von G ist und für alle $x \in G$ und $y \in U$ auch $xyx^{-1} \in U$ ist.*

BEWEIS. $\rightarrow$. Aus $xU = Ux$ folgt $xyx^{-1} \in U$. $\leftarrow$. Für alle $x \in G$ gilt $xU \subseteq Ux$. Damit hat man auch $x^{-1}U \subseteq Ux^{-1}$, woraus $Ux \subseteq xU$ folgt. Also ist $xU = Ux$. □

LEMMA. *Seien G, H Gruppen, $f\colon G \to H$ ein Gruppenhomomorphismus. Ist V Normalteiler von H ist, so ist $f^{-1}(V)$ Normalteiler von G.*

BEWEIS. Sei $x \in G$ und $y \in f^{-1}(V)$. Dann gilt

$$f(xyx^{-1}) = f(x)f(y)f(x)^{-1} \in V,$$

da $f(y) \in V$ und V Normalteiler ist. Also ist $xyx^{-1} \in f^{-1}(V)$. □

BEMERKUNG. Sei U Normalteiler in G und $f\colon G \to H$ ein Gruppenhomomorphismus. Dann muß $f(U)$ nicht Normalteiler in H sein. Ist nämlich $G \subseteq H$ Untergruppe, aber kein Normalteiler, und $f\colon G \to H$ die Einbettung von G nach H, so ist $f(G) = G$ kein Normalteiler in H, obwohl G Normalteiler in G ist.

3.1.7. Faktorgruppen. Sei G eine Gruppe und $N \subseteq G$ ein Normalteiler. Wir konstruieren die Faktorgruppe G/N, zusammen mit dem natürlichen (oder „kanonischen“) Gruppenhomomorphismus $\mathrm{id}\colon G \to G/N$. Es zeigt sich, daß hierfür die Eigenschaft von N, Normalteiler zu sein, notwendig ist. Als einfache Folgerung erhalten wir eine Charakterisierung von Normalteilern als Kerne von Gruppenhomomorphismen.

DEFINITION. Sei $(G, \cdot)$ eine Gruppe und $N \subseteq G$ ein Normalteiler. Die *Faktorgruppe* G/N ist $(G, \cdot, \sim_N)$, wobei $(x \sim_N y) := (x^{-1}y \in N)$.

SATZ (Konstruktion der Faktorgruppe). *Sei G eine Gruppe und $N \subseteq G$ Normalteiler. Wir betrachten die* kanonische Abbildung

$$\mathrm{id}\colon G \to G/N, \quad x \mapsto x.$$

Auf G/N gibt es genau eine Gruppenstruktur, so daß $\mathrm{id}\colon G \to G/N$ *ein Gruppenhomomorphismus wird. In diesem Fall gilt* $\mathrm{Kern}(\mathrm{id}) = N$.

BEWEIS. *Existenz.* Wir müssen zeigen, daß die Gruppenverknüpfung von G mit $\sim_N$ verträglich ist. Seien also $x, \hat{x}, y, \hat{y} \in G$ mit $x \sim_N \hat{x}$ und $y \sim_N \hat{y}$. Zu zeigen ist $xy \sim_N \hat{x}\hat{y}$. Dies folgt aus

$$\begin{aligned}
(xy)^{-1}(\hat{x}\hat{y}) &= y^{-1}\underbrace{x^{-1}\hat{x}}_{\in N}\hat{y} \\
&= y^{-1}\hat{y}n && \text{für ein } n \in N \text{ wegen } N\hat{y} = \hat{y}N, \\
&\in Nn && \text{wegen } y \sim_N \hat{y} \\
&\subseteq N.
\end{aligned}$$

id ist offenbar Gruppenhomomorphismus. Ferner gilt $\mathrm{Kern}(\mathrm{id}) = N$, denn $x \sim_N e \leftrightarrow x^{-1}e \in N \leftrightarrow x \in N$.

Eindeutigkeit. Eine Verknüpfung $\circ$ auf G/N, bezüglich derer id Gruppenhomomorphismus ist, muß die Bedingung $\mathrm{id}(xy) \sim_N \mathrm{id}(x) \circ \mathrm{id}(y)$ erfüllen, also $xy \sim_N x \circ y$. □

KOROLLAR. *Sei G eine Gruppe und $N \subseteq G$ Untergruppe. Dann ist N Normalteiler von G genau dann, wenn es eine Gruppe H und einen Gruppenhomomorphismus $f\colon G \to H$ gibt mit $N = \mathrm{Kern}(f)$.*

BEWEIS. $\to$ folgt aus dem eben bewiesenen Satz.
$\leftarrow$. Es ist $N = \mathrm{Kern}(f) = f^{-1}(\{e\})$, und $\{e\}$ ist Normalteiler. □

3.2. Ringe

3.2.1. Definition und einfache Eigenschaften von Ringen.

DEFINITION. Sei A eine Menge und $+\colon A \to A \to A$, $\cdot\colon A \to A \to A$ Abbildungen. A (oder genauer $(A, +, \cdot)$) heißt *Ring*, wenn gilt

(a) $(A, +)$ ist eine abelsche Gruppe.
(b) $x(yz) = (xy)z$ für alle $x, y, z \in A$.
(c) $x(y + z) = (xy) + (xz)$ und $(x + y)z = (xz) + (yz)$ für alle $x, y, z \in A$.

A heißt *kommutativ*, wenn $xy = yx$ für alle $x, y \in A$. Ein Element $1 \in A$ heißt *Einselement* von A, wenn für alle $x \in A$ gilt $1x = x1 = x$.

Zur Vereinfachung beschränken wir uns meistens auf kommutative Ringe mit Eins.

SCHREIBWEISE. (a) „$\cdot$“ bindet stärker als „$+$“; beispielsweise steht $xy + xz$ für $(xy) + (xz)$.
(b) $x - y$ steht für $x + (-y)$.
(c) nx steht für $\underbrace{x + x + \cdots + x}_{n\text{-mal}}$, und x^n steht für $\underbrace{x \cdot x \cdot \cdots \cdot x}_{n\text{-mal}}$.

LEMMA. *Sei A ein kommutativer Ring mit 1. Dann besitzt A genau ein Einselement. (Beweis: Seien $1, 1^*$ Einselemente. Dann gilt $1 = 1 \cdot 1^* = 1^*$). Ferner gilt für alle $x, y \in A$:*

(a) $0 \cdot x = 0$
(b) $-x = (-1) \cdot x$
(c) $(-x) \cdot (-y) = xy$

BEWEIS. (a). $0x = (0 + 0)x = 0x + 0x$ und damit $0x = 0$.

(b). Es gilt $x + (-1)x = 1x + (-1)x = (1 + (-1))x = 0x = 0$, also $-x = (-1)x$.

(c). Es ist $1+(-1) = 0$ und damit $1(-1)+(-1)(-1) = 0$, also $(-1)(-1) = 1$. Daraus folgt $(-x)(-y) = (-1)x(-1)y = (-1)(-1)xy = xy$. □

BEISPIELE. (a). $(\mathbf{Z}, +, \cdot)$ ist ein kommutativer Ring mit 1.

(b). Sei A ein kommutativer Ring mit 1 und X eine nicht-leere Menge. Auf

$$A^X := \{\, f \colon X \to A \mid f \text{ Abbildung} \,\}$$

erklärt man $+, \cdot$ komponentenweise, also durch

$$(f + g)(x) := f(x) + g(x),$$
$$(f \cdot g)(x) := f(x) \cdot g(x)$$

für alle $x \in A$. Damit wird A^X zu einem kommutativen Ring mit 1. Einselement ist die Abbildung $X \to A$, $x \mapsto 1$.

(c). Die Menge $\mathbf{Z}^{n\times n}$ der $n \times n$-Matrizen über den ganzen Zahlen mit der üblichen Addition und (Matrizen)-multiplikation ist ein Ring mit der Einheitsmatrix E als Einselement. Er ist für $n \geq 2$ *nicht* kommutativ.

(d). Seien $A_1, \ldots, A_n$ kommutative Ringe mit 1. Auf $A_1 \times \cdots \times A_n$ erkläre man $+, \cdot$ komponentenweise. Damit wird $A_1 \times \cdots \times A_n$ zu einem kommutativen Ring mit 1, dessen Einselement die Spalte aus lauter Einsen ist (*direktes Produkt* von $A_1, \ldots, A_n$).

(e). $\{0\}$ ist ein kommutativer Ring mit Einselement 0.

DEFINITION. Sei A ein kommutativer Ring mit 1. Man nennt A einen *Integritätsbereich* (oder *Integritätsring*), wenn gilt

(a) $1 \neq 0$.
(b) A ist nullteilerfrei, d.h. für alle $x, y \in A$ folgt aus $xy = 0$ stets $x = 0$ oder $y = 0$.

BEISPIELE. (a). $(\mathbf{Z}, +, \cdot)$ ist ein Integritätsbereich.

(b). Sei A ein Integritätsbereich, X eine Menge mit mindestens zwei Elementen a und b. Dann ist A^X kein Integritätsbereich, denn betrachtet man die Abbildungen

$$f\colon X \to A, \qquad f(x) = \begin{cases} 0 & \text{falls } x = a \\ 1 & \text{falls } x = b \\ 0 & \text{sonst} \end{cases}$$

$$g\colon X \to A, \qquad g(x) = \begin{cases} 1 & \text{falls } x = a \\ 0 & \text{falls } x = b \\ 0 & \text{sonst,} \end{cases}$$

so gilt $f, g \neq 0$, aber $f \cdot g = 0$.

3.2.2. Ideale.

DEFINITION. Seien A ein kommutativer Ring mit 1 und $U \subseteq A$ eine Teilmenge. U heißt *Unterring* mit 1 von A, wenn gilt

(a) $U \neq \emptyset$;
(b) $x + y, xy \in U$ für alle $x, y \in U$;
(c) U bildet zusammen mit den Abbildungen

$$U \to U \to U \qquad (x, y) \mapsto x + y$$

$$U \to U \to U \qquad (x, y) \mapsto xy$$

einen kommutativen Ring mit 1.

DEFINITION. Sei A ein kommutativer Ring mit 1 und $\mathfrak{a} \subseteq A$ eine Teilmenge. $\mathfrak{a}$ heißt *Ideal* von A, wenn gilt

(a) $\mathfrak{a}$ ist Untergruppe der additiven Gruppe von A.
(b) $A\mathfrak{a} \subseteq \mathfrak{a}$ (wobei $A\mathfrak{a} := \{\, yx \mid y \in A, x \in \mathfrak{a} \,\}$).

BEISPIELE. Sei A ein kommutativer Ring mit 1.

(a). $\{0\}$ und A sind Ideale von A (triviale Ideale).

(b). Ist $x \in A$, so ist $(x) := Ax := \{\, yx \mid y \in A \,\}$ ein Ideal von A (Hauptideal).

(c). $\mathfrak{a}$ ist Ideal von $\mathbf{Z}$ genau dann, wenn ein $n \in \mathbf{N}$ existiert mit $\mathfrak{a} = n\mathbf{Z}$. Dies beweist man wie folgt. $\leftarrow$. Gilt nach (b). $\rightarrow$. $\mathfrak{a}$ ist eine Untergruppe von $(\mathbf{Z}, +)$, also $\mathfrak{a} = n\mathbf{Z}$ nach 3.1.5, wo wir die Untergruppen von $\mathbf{Z}$ bestimmt haben.

DEFINITION. Seien A, B kommutative Ringe mit 1 und $f\colon A \to B$ eine Abbildung. f heißt *Homomorphismus* für Ringe, wenn für alle $x, y \in A$ gilt

(a) $f(x + y) = f(x) + f(y)$,
(b) $f(xy) = f(x)f(y)$,
(c) $f(1) = 1$.

Die Begriffe *Mono-, Iso-, Endo-* und *Automorphismus* erklärt man wie für Gruppen.

LEMMA. *Seien A, B kommutative Ringe mit* 1 *und $f\colon A \to B$ ein Ringhomomorphismus. Dann gilt*

(a) *Ist U Unterring von A, so ist $f(U)$ Unterring von B, und ist V Unterring von B, so ist $f^{-1}(V)$ Unterring von A.*
(b) *Sei wieder* $\mathrm{Kern}(f) := f^{-1}(0)$. *Dann ist* $\mathrm{Kern}(f) = \{0\}$ *genau dann, wenn f injektiv ist.*

BEWEIS. (a). Nach einem Lemma über Gruppenhomomorphismen (in 3.1.3) sind $f(U)$ bzw. $f^{-1}(V)$ additive Untergruppen. Ferner folgt aus $f(1) = 1$ stets $1 \in f(U)$ bzw. $1 \in f^{-1}(V)$. Seien nun $x, y \in f(U)$. Dann existieren $u, v \in U$ so daß $x = f(u)$ und $y = f(v)$, und es gilt $xy = f(u)f(v) = f(uv) \in f(U)$. Seien $w, z \in f^{-1}(V)$. Dann gilt $f(wz) = f(w)f(z) \in V$, also $wz \in f^{-1}(V)$. Daher sind $f(U)$ und $f^{-1}(V)$ Unterringe von B bzw. A.

(b). Dies folgt direkt aus dem entsprechenden Lemma in 3.1.3 über Gruppenhomomorphismen. □

3.2.3. Restklassenringe. Sei A ein kommutativer Ring mit 1 und $\mathfrak{a} \subseteq A$ ein Ideal. Wir konstruieren den Restklassenring $A/\mathfrak{a}$, zusammen mit dem natürlichen (oder „kanonischen“) Ringhomomorphismus $\mathrm{id}\colon A \to A/\mathfrak{a}$. Es zeigt sich, daß hierfür die Eigenschaft von $\mathfrak{a}$, Ideal zu sein, notwendig ist. Als einfache Folgerung erhalten wir eine Charakterisierung von Idealen als Kerne von Ringhomomorphismen.

DEFINITION. Sei $(A, +, \cdot)$ ein kommutativer Ring mit 1 und $\mathfrak{a} \subseteq A$ Ideal. Der *Restklassenring* $A/\mathfrak{a}$ ist $(A, +, \cdot, \sim_{\mathfrak{a}})$, wobei $(x \sim_{\mathfrak{a}} y) := (y - x \in \mathfrak{a})$.

SATZ (Konstruktion des Restklassenrings). *Sei A ein kommutativer Ring mit* 1 *und $\mathfrak{a} \subseteq A$ Ideal. Wir betrachten die* kanonische Abbildung

$$\mathrm{id}\colon A \to A/\mathfrak{a}, \quad x \mapsto x.$$

Auf $A/\mathfrak{a}$ gibt es genau eine Ringstruktur, so daß $\mathrm{id}\colon A \to A/\mathfrak{a}$ *ein Ringhomomorphismus wird. In diesem Fall gilt* $\mathrm{Kern}(\mathrm{id}) = \mathfrak{a}$.

BEWEIS. *Existenz.* $(A/\mathfrak{a}, +, \sim_\mathfrak{a})$ ist nach dem Satz über Faktorgruppen in 3.1.7 eine abelsche Gruppe. Wir müssen nur noch zeigen, daß die Ringmultiplikation von A mit $\sim_\mathfrak{a}$ verträglich ist. Seien also $x, \hat{x}, y, \hat{y} \in A$ mit $x \sim_\mathfrak{a} \hat{x}$ und $y \sim_\mathfrak{a} \hat{y}$. Zu zeigen ist $xy \sim_\mathfrak{a} \hat{x}\hat{y}$. Dies folgt aus

$$\hat{x}\hat{y} - xy = \hat{x}\hat{y} - x\hat{y} + x\hat{y} - xy = \underbrace{(\hat{x} - x)}_{\in\mathfrak{a}}\hat{y} + x\underbrace{(\hat{y} - y)}_{\in\mathfrak{a}} \in \mathfrak{a}.$$

id ist offenbar Ringhomomorphismus. Ferner ist $\mathrm{Kern}(\mathrm{id}) = \mathfrak{a}$, denn es gilt $x \sim_\mathfrak{a} 0 \leftrightarrow 0 - x \in \mathfrak{a} \leftrightarrow x \in \mathfrak{a}$.

Eindeutigkeit. Für $+$ ist dies klar nach dem Satz über Faktorgruppen. Eine Multiplikation $\circ$ auf $A/\mathfrak{a}$, bezüglich derer id Ringhomomorphismus ist, muß die Bedingung $\mathrm{id}(xy) \sim_\mathfrak{a} \mathrm{id}(x) \circ \mathrm{id}(y)$ erfüllen, also $xy \sim_\mathfrak{a} x \circ y$. □

KOROLLAR. *Sei A ein kommutativer Ring mit 1 und $\mathfrak{a} \subseteq A$ Teilmenge. $\mathfrak{a}$ ist Ideal genau dann, wenn $\mathfrak{a}$ Kern eines Ringhomomorphismus $f\colon A \to B$ ist.*

BEWEIS. $\rightarrow$ folgt aus dem eben bewiesenen Satz.
$\leftarrow$. Es ist $\mathfrak{a} = \mathrm{Kern}(f) = f^{-1}(\{0\})$, und $\{0\}$ ist Ideal. □

3.2.4. Summe und Durchschnitt von Idealen. Wir zeigen, daß Summe und Durchschnitt von Idealen wieder Ideale ergeben. Auf diesem Wege ergibt sich ein nützlicher Zusammenhang zwischen den Begriffen des größten gemeinsamen Teilers und des kleinsten gemeinsamen Vielfachen einerseits und dem Idealbegriff andererseits.

Zunächst stellen wir fest, daß man den in 2.3.1 eingeführten Begriff des größten gemeinsamen Teilers genauso wie früher auch für mehr als zwei Argumente definieren kann. Den Beweis der Existenz (Satz von Euklid) kann man dann leicht durch Induktion über die Anzahl der Argumente führen. Ebenso beweist man, daß $\mathrm{ggT}(x_1, \ldots, x_k)$ als eine Linearkombination von $x_1, \ldots, x_k$ geschrieben werden kann.

Einen anderen Zugang zum Begriff des größten gemeinsamen Teilers erhält man durch die Theorie der Ideale in einem Ring, und zwar in unserem Fall im Ring $\mathbf{Z}$ der ganzen Zahlen. Dieser Ring hat die besondere Eigenschaft, daß jedes Ideal von einem Ringelement erzeugt ist, sich also in der Form (x) schreiben läßt. Solche Ringe nennt man *Hauptidealringe*; viele unserer Betrachtungen lassen sich auf beliebige Hauptidealringe übertragen.

DEFINITION. Sei A ein kommutativer Ring mit 1.

(a) Für Ideale $\mathfrak{a}, \mathfrak{b} \subseteq A$ definiert man

$$\mathfrak{a} + \mathfrak{b} := \{\, x + y \mid x \in \mathfrak{a}, y \in \mathfrak{b} \,\}, \quad \mathfrak{a} \cap \mathfrak{b} := \{\, x \mid x \in \mathfrak{a}, x \in \mathfrak{b} \,\}.$$

(b) Für eine Teilmenge $M \subseteq A$ sei

$$(M) := \{\, x_1a_1 + \cdots + x_na_n \mid n \geq 0, a_1, \ldots, a_n \in M, x_1, \ldots, x_n \in A \,\}.$$

Wir schreiben $(a_1, \ldots, a_n)$ für$(\{a_1, \ldots, a_n\})$.

LEMMA. *Unter den Voraussetzungen der Definition gilt*

(a) $\mathfrak{a} + \mathfrak{b}$ *und* $\mathfrak{a} \cap \mathfrak{b}$ *sind Ideale von* A.
(b) (M) *ist das kleinste* M *umfassende Ideal von* A, *d.h. es gilt*
  (i) (M) *ist ein Ideal.*
  (ii) $(M) \supseteq M$.
  (iii) *Ist* $\mathfrak{a} \supseteq M$ *Ideal von* A, *so ist* $\mathfrak{a} \supseteq (M)$.
(c) *Für* $a, b \in A$ *ist* $(a) + (b) = (a, b)$.

BEWEIS. (a). Unter Benützung des Untergruppenkriteriums läßt sich leicht zeigen, daß $\mathfrak{a} + \mathfrak{b}$ und $\mathfrak{a} \cap \mathfrak{b}$ Untergruppen von $(A, +)$ sind. Seien nun $a \in A$, $w \in \mathfrak{a} + \mathfrak{b}$, $z \in \mathfrak{a} \cap \mathfrak{b}$. Dann existieren $x \in \mathfrak{a}$ und $y \in \mathfrak{b}$ so, daß $w = x + y$ und es gilt $aw = a(x + y) = ax + ay \in \mathfrak{a} + \mathfrak{b}$, da $ax \in \mathfrak{a}$ und $ay \in \mathfrak{b}$ ($\mathfrak{a}$ und $\mathfrak{b}$ sind Ideale), sowie $az \in \mathfrak{a} \cap \mathfrak{b}$, da $az \in \mathfrak{a}$ und $az \in \mathfrak{b}$. Damit folgt, daß $\mathfrak{a} + \mathfrak{b}$ und $\mathfrak{a} \cap \mathfrak{b}$ Ideale von A sind.

(b). Wieder folgt nach dem Untergruppenkriterium sofort, daß (M) Untergruppe von $(A, +)$ ist. Seien $a \in A$ und $x = x_1 a_1 + \cdots + x_n a_n \in (M)$. Dann folgt $ax = (ax_1)a_1 + \cdots + (ax_n)a_n \in (M)$, da $ax_i \in A$ für alle i, also ist (M) ein Ideal. $(M) \supseteq M$ ist klar. Sei $\mathfrak{a} \supseteq M$ Ideal von A. Dann folgt $\mathfrak{a} \supseteq AM (:= \{\, yx \mid y \in A, x \in M \,\})$, da $\mathfrak{a}$ Ideal ist, und damit $\mathfrak{a} \supseteq (M)$, da $\mathfrak{a}$ Untergruppe von $(A, +)$ ist.

(c). Seien $a, b \in A$. Es gilt $(a) \subseteq (a, b)$ und $(b) \subseteq (a, b)$, also $(a) + (b) \subseteq (a, b)$, da (a, b) Untergruppe von $(A, +)$ ist. Nach (a) ist $(a) + (b)$ ein Ideal und damit $(a) + (b) \supseteq (a, b)$ nach (b). Also gilt $(a) + (b) = (a, b)$. □

BEISPIEL. Im Ring $\mathbf{Z}$ der ganzen Zahlen betrachten wir die Ideale (n), (m) mit $n, m \neq 0$. OBdA können wir $n, m > 0$ annehmen. Es gilt

(a) $(m) \supseteq (n)$ genau dann, wenn $m \mid n$ $(:= \exists_{q \in \mathbf{N}}\, mq = n)$.
(b) $(m) + (n) = (m, n) = (d)$ für ein eindeutig bestimmtes $d > 0$ (nach dem Satz über die Untergruppen von $\mathbf{Z}$). (d) ist charakterisiert durch

$$(d) \supseteq (m), (n),$$
$$\forall_{q \in \mathbf{N}}\big((q) \supseteq (m), (n) \to (q) \supseteq (d)\big).$$

Also ist d charakterisiert durch

$$d \mid m, n,$$
$$\forall_{q \in \mathbf{N}}(q \mid m, n \to q \mid d).$$

Dieses d hatten wir den *größten gemeinsamen Teiler* von m und n genannt und ihn mit $\mathrm{ggT}(m, n)$ bezeichnet.

(c) $(m) \cap (n) = (v)$ für ein eindeutig bestimmtes $v > 0$ (nach der Charakterisierung der Untergruppen von $\mathbf{Z}$). (v) ist charakterisiert durch

$$(m), (n) \supseteq (v),$$

$$\forall_{q\in\mathbf{N}}\big((m),(n)\supseteq(q)\to(v)\supseteq(q)\big).$$

Also ist v charakterisiert durch

$$m,n \mid v,$$
$$\forall_{q\in\mathbf{N}}(m,n \mid q \to v \mid q).$$

Dieses v hatten wir das *kleinste gemeinsame Vielfache* von m und n genannt und es mit $\mathrm{kgV}(m,n)$ bezeichnet.

3.3. Kongruenzen

In 3.1.1 hatten wir $\mathbf{Z}_n := \{\, a \in \mathbf{N} \mid a < n \,\}$ als Beispiel einer endlichen abelschen Gruppe eingeführt; sie hieß die Gruppe der *ganzen Zahlen modulo* n. Die Gruppenverknüpfung ist die *Addition modulo* n, die für $a, b \in \mathbf{Z}_n$ definiert war als der (eindeutig bestimmte) Rest bei der Division von $a+b$ durch n. Genauso kann man die *Multiplikation modulo* n definieren, als den Rest bei der Division von ab durch n. Es ist leicht zu sehen, daß $\mathbf{Z}_n$ mit dieser Addition und Multiplikation einen kommutativen Ring mit Einselement bildet, und daß er isomorph ist zu dem Restklassenring $\mathbf{Z}/n\mathbf{Z}$. Die hierbei verwendete Äquivalenzrelation auf der Trägermenge $\mathbf{Z}$ nennt man Kongruenz modulo n. Wir wollen diesen Kongruenzbegriff jetzt genauer untersuchen; seine Bedeutung wurde zuerst von Gauß erkannt (Disquisitiones arithmeticae, 1801).

3.3.1. Charakterisierung der Kongruenz modulo n.

Lemma. *Für $a, b \in \mathbf{Z}$ und $n \in \mathbf{N}$, $n > 0$ sind äquivalent*

(a) *a und b haben denselben Rest bei der Division durch n.*
(b) *$n \mid a - b$.*

Beweis. Seien $a = np + r$ und $b = nq + s$ mit $p, q \in \mathbf{Z}$ und $r, s \in \mathbf{N}$, $r, s < n$. $\to$. Gelte $r = s$. Dann ist $a - b = n(p - q)$, also $n \mid a - b$. $\leftarrow$. Es ist $a - b = n(p-q) + (r - s)$. Aus $n \mid a - b$ folgt also $n \mid r - s$. Wegen $0 \le r, s < n$ folgt $r = s$. □

Falls eine (und damit beide) des Bedingungen des Lemmas erfüllt sind, sagen wir, daß „a kongruent zu b modulo n“ ist. Dafür verwenden wir die Bezeichnung $a \equiv b \mod n$ (oder auch $a \equiv b(n)$ oder $a \equiv_n b$).

Lemma. *Für $a, b, c, d \in \mathbf{Z}$ und $n \in \mathbf{N}$, $n > 0$ gelte $a \equiv b \mod n$ und $c \equiv d \mod n$. Dann gilt auch*

(a) *$a + c \equiv b + d \mod n$,*
(b) *$-a \equiv -b \mod n$,*
(c) *$ac \equiv bd \mod n$.*

Beweis. Übung. □

BEMERKUNG. Eine einfache Folgerung aus Teil (c) des Lemmas ist, daß für $a, b \in \mathbf{Z}$ und $n, k \in \mathbf{N}$ mit $n > 0$ aus $a \equiv b \mod n$ stets $a^{k+1} \equiv b^{k+1} \mod n$ folgt.

Mit Hilfe dieser Eigenschaften der Kongruenz lassen sich anscheinend schwierige Fragen nach dem Rest einer großen Zahl modulo einer kleinen relativ leicht beantworten. Wichtig ist nur, daß man die gegebene große Zahle faktorisieren, also als Produkt hinreichend kleiner Zahlen darstellen kann. Eine Faktorisierung zufälliger großer Zahlen läßt sich allerdings auch mit modernen und leistungsfähigen Rechnern nicht in vernünftiger Zeit durchführen; auf der praktischen Unmöglichkeit der Faktorisierung beruht das mit einem öffentlichen und einem privaten Schlüssel arbeitende sogenannte RSA-Verfahren, das wir in 3.3.5 besprechen.

Als Beispiel betrachten wir das Problem, den Rest von 2^{16} bei der Division durch 11 zu bestimmen. Dazu stellen wir 2^{16} als Produkt $2^4 \cdot 2^4 \cdot 2^4 \cdot 2^4$ dar. Wir verwenden jetzt, daß $2^4 \equiv 5 \mod 11$ ist, also $2^{16} \equiv 5^4 \mod 11$. Ferner ist $5^2 \equiv 3 \mod 11$, also $5^4 \equiv 3^2 \mod 11$. Insgesamt ergibt sich $2^{16} \equiv 9 \mod 11$.

3.3.2. Simultane Kongruenzen. Wir wollen uns überlegen, daß und wie sich ein System simultaner Kongruenzen mit teilerfremden Moduln immer lösen läßt. Ferner werden wir zeigen können, daß das Lösungsverfahren nützliche strukturelle Eigenschaften hat: es ist ein Ringisomorphismus. Davon werden wir im nächsten Abschnitt Gebrauch machen.

SATZ (Chinesischer Restsatz). *Seien $n_1, \dots n_r$ teilerfremd, $n := \prod_i n_i$. Das System der Kongruenzen $x \equiv a_i \mod n_i$ hat eine „simultane“ Lösung; sie ist modulo n eindeutig bestimmt. Genauer gilt: Sei $q_i := \prod_{j \neq i} n_j$. Wähle y_i mit $1 \equiv q_i y_i \mod n_i$ und setze $e_i := q_i y_i$.*

(a) *Die Abbildung*

$$f\colon \mathbf{Z}_{n_1} \times \cdots \times \mathbf{Z}_{n_r} \to \mathbf{Z}_n, \quad (a_1, \dots, a_r) \mapsto a_1 e_1 + \cdots + a_r e_r$$

hat die Eigenschaft $f(a_1, \dots, a_r) \equiv a_i \mod n_i$.

(b) *f ist ein Ringisomorphismus.*

BEWEIS. Eindeutigkeit. Aus $x \equiv a_i \mod n_i$ und $x' \equiv a_i \mod n_i$ folgt $x - x' \equiv 0 \mod n_i$, also $x - x' \equiv 0 \mod n$.

Existenz. Zur Konstruktion von y_i mit $1 \equiv q_i y_i \mod n_i$ benutzt man, daß q_i und n_i teilerfremd sind, sich also 1 linear aus q_i und n_i kombinieren läßt. Für $e_i := q_i y_i$ erhält man

$$e_i \equiv 1 \mod n_i, \qquad e_i \equiv 0 \mod n_j \quad \text{für } i \neq j. \tag{3.1}$$

Also ist $a_1 e_1 + \cdots + a_r e_r \equiv a_i \mod n_i$. Damit ist auch (a) bewiesen.

(b). Wir zeigen, daß f ein injektiver Ringhomomorphismus ist. Seien $(a_1, \ldots, a_r), (b_1, \ldots, b_r) \in \mathbf{Z}_{n_1} \times \cdots \times \mathbf{Z}_{n_r}$. Dann gilt

$$\begin{aligned} f(a_1, \ldots, a_r) + f(b_1, \ldots, b_r) &= a_1 e_1 + \cdots + a_r e_r + b_1 e_1 + \cdots + b_r e_r \\ &= (a_1 + b_1) e_1 + \cdots + (a_r + b_r) e_r \\ &= f(a_1 + b_1, \ldots, a_r + b_r). \end{aligned}$$

Das heißt, daß f ein Gruppenhomomorphismus von der additiven Gruppe $(\mathbf{Z}_{n_1} \times \cdots \times \mathbf{Z}_{n_r}, +)$ in die additive Gruppe $(\mathbf{Z}_n, +)$ ist. Zum Beweis, daß f auch ein Ringhomomorphismus ist, beachte man $e_i^2 - e_i = (e_i - 1)e_i \equiv 0 \mod n$ wegen (3.1). Man erhält

$$\begin{aligned} &f(a_1, \ldots, a_r) \cdot f(b_1, \ldots, b_r) \\ &= (a_1 e_1 + \cdots + a_r e_r) \cdot (b_1 e_1 + \cdots + b_r e_r) \\ &= \sum_{i,j} a_i b_j e_i e_j \\ &\equiv \sum_i a_i b_i e_i^2 \mod n \quad \text{denn } e_i e_j \equiv 0 \mod n \text{ für } i \neq j \text{ nach (3.1)} \\ &\equiv \sum_i a_i b_i e_i \mod n \quad \text{denn } e_i^2 \equiv e_i \mod n \text{ nach Vorbemerkung} \\ &= f(a_1 b_1, \ldots, a_r b_r). \end{aligned}$$

Schließlich ist $\mathrm{Kern}(f) = 0$, denn aus $f(a_1, \ldots, a_r) \equiv 0 \mod n$ folgt $a_i \equiv 0 \mod n_i$ nach (a).

Damit ist gezeigt, daß f ein injektiver Ringhomomorphismus ist. f ist bijektiv, da $\mathbf{Z}_{n_1} \times \cdots \times \mathbf{Z}_{n_r}$ und $\mathbf{Z}_n$ dieselbe Anzahl von Elementen enthalten, nämlich n. Also ist f ein Ringisomorphismus. □

Beispiel. Sei $n_1 := 3$, $n_2 := 7$, $n_3 := 11$. Dann ist $n := n_1 n_2 n_3 = 231$. Ferner ist $q_1 := n_2 n_3 = 77$, $q_2 := n_1 n_3 = 33$ und $q_3 := n_1 n_2 = 21$. Wir müssen y_i bestimmen mit $1 \equiv q_i y_i \mod n_i$.

y_1 bestimmen wir aus $1 \equiv 77 y_1 \equiv 2 y_1 \mod 3$, also $y_1 := 2$.

y_2 bestimmen wir aus $1 \equiv 33 y_2 \equiv 5 y_2 \mod 7$, also $y_2 := 3$.

y_3 bestimmen wir aus $1 \equiv 21 y_3 \equiv (-1) y_3 \mod 11$, also $y_3 := -1$.

Daraus ergibt sich $e_1 := q_1 y_1 = 154$, $e_2 := q_2 y_2 = 99$, $e_3 := q_3 y_3 = -21$. Eine Lösung etwa der simultanen Kongruenzen $x \equiv 1 \mod 3$, $x \equiv 2 \mod 7$, $x \equiv 6 \mod 11$ erhält man nach dem Satz wie folgt. Es ist $a_1 = 1$, $a_2 = 2$ und $a_3 = 6$. Eine Lösung ist also $a_1 e_1 + a_2 e_2 + a_3 e_3 = 1 \cdot 154 + 2 \cdot 99 - 6 \cdot 21 = 154 + 198 - 126 = 154 + 72 = 226$. Zur Kontrolle verifiziert man leicht $226 \equiv 1 \mod 3$, $226 \equiv 2 \mod 7$ und $226 \equiv 6 \mod 11$.

Man beachte, daß das angegebene Lösungsverfahren besonders vorteilhaft ist, wenn mehrere Systeme von Kongruenzen nach denselben Moduln zu

lösen sind. Die e_i hängen nämlich nicht von den a_i ab, müssen also nur einmal berechnet werden. Will man etwa noch die simultanen Kongruenzen $x \equiv 2 \mod 3$, $x \equiv 4 \mod 7$, $x \equiv -3 \mod 11$ lösen, so ist jetzt $a_1 = 2$, $a_2 = 4$ und $a_3 = -3$. Eine Lösung ist also $a_1e_1 + a_2e_2 + a_3e_3 = 2 \cdot 154 + 4 \cdot 99 + 3 \cdot 21 = 308 + 396 + 63 = 767 \equiv 74 \mod 231$. Zur Kontrolle verifiziert man leicht $74 \equiv 2 \mod 3$, $74 \equiv 4 \mod 7$ und $74 \equiv -3 \mod 11$.

3.3.3. Einheitengruppe; prime Reste modulo n. Wir betrachten jetzt die multiplikative Struktur des Rings $\mathbf{Z}_n$, genauer seine Einheitengruppe. Sie ist auch unter dem Namen Gruppe der primen Reste modulo n bekannt. Die Anzahl ihrer Elemente bezeichnet man mit $\varphi(n)$; diese Bezeichnung geht auf Euler zurück. Von besonderem Interesse ist der Fall, daß n eine Primzahl p ist. Dann sind alle Zahlen $1, 2, \ldots, p-1$ prime Reste modulo p, also $\varphi(p) = p - 1$. Als Folgerung ergibt sich, daß für jede ganze Zahl $x \neq 0$ gilt $x^p \equiv x \mod p$.

Insbesondere besteht die Einheitengruppe des Rings $\mathbf{Z}_p$ für eine Primzahl p aus allen von 0 verschiedenen Ringelementen. Solche Strukturen nennt man *Körper*.

Definition. Sei A ein kommutativer Ring mit 1. Dann heißt $u \in A$ *Einheit*, wenn es ein $v \in A$ gibt mit $vu = 1$. Die Menge der Einheiten von A wird mit A^* bezeichnet.

Zum Beispiel ist $\mathbf{Z}^* = \{1, -1\}$.

Lemma. *Sei A ein kommutativer Ring mit* 1. *Dann ist $(A^*, \cdot)$ eine Gruppe, die* Einheitengruppe *von A.*

Beweis. Dies folgt leicht aus der Gruppendefinition:

A^* ist abgeschlossen unter der Ringmultiplikation: Seien $u_1, u_2 \in A^*$. Dann haben wir $v_1, v_2 \in A$ mit $v_iu_i = 1$ für $i = 1, 2$. Es folgt $v_2v_1u_1u_2 = v_2u_2 = 1$, also $u_1u_2 \in A^*$.

1 ist offenbar neutrales Element von A^*.

A^* ist abgeschlossen unter Inversenbildung: Sei $u \in A^*$. Dann haben wir ein $v \in A$ mit $vu = 1$. Jetzt folgt $v \in A^*$ aus $vu = uv = 1$. □

Satz. *Sei A ein Integritätsbereich und $a, b \in A$. Dann gilt $(a) = (b)$ genau dann, wenn es ein $u \in A^*$ gibt mit $a = ub$.*

Beweis. $\rightarrow$. Im Fall $a = b = 0$ ist die Behauptung trivial. Sei also oBdA $a \neq 0$. Wegen $a \in (b)$ gibt es ein $u \in A$ mit $a = ub$. Wegen $b \in (a)$ gibt es ein $v \in A$ mit $b = va$. Also ist $a = uva$ und deshalb $a(1 - uv) = 0$. Da A ein Integritätsbereich ist und $a \neq 0$, folgt $1 - uv = 0$ und somit $1 = uv$, also $u \in A^*$.

$\leftarrow$. Sei $a = ub$ mit $u \in A^*$. Es gilt $(a) = Aa = Aub \subseteq Ab = (b)$ und damit $(a) \subseteq (b)$. Analog hat man $(b) \subseteq (a)$, da $b = u^{-1}a$. □

DEFINITION. Die Einheitengruppe $\mathbf{Z}_n^*$ von $\mathbf{Z}_n$ nennt man die *Gruppe der primen Reste modulo n*.

DEFINITION. $\varphi(m) :=$ Anzahl der zu m teilerfremden $n \in \{1, \ldots, m-1\}$ (d.h. für die gilt $\mathrm{ggT}(m,n) = 1$) heißt *Eulersche φ-Funktion*.

Die Behandlung der Eulerschen φ-Funktion ordnet sich hier ein, da nach dem folgenden Satz $\varphi(n) = |\mathbf{Z}_n^*|$ gilt.

SATZ (Euler-Funktion). *Sei $m \in \mathbf{N}$, $m > 0$. Dann gilt*

(a) $\mathbf{Z}_m^* = \{\, n \in \mathbf{Z}_m \mid m, n \text{ teilerfremd} \,\}$.
(b) $|\mathbf{Z}_m^*| = \varphi(m)$.
(c) *Sei $m \in \mathbf{N}$, $n > 0$ und m, n teilerfremd. Dann ist $\varphi(mn) = \varphi(m)\varphi(n)$.*
(d) $\varphi(p^r) = p^{r-1}(p-1)$ *für p Primzahl und $r > 0$.*

BEWEIS. (a). Sei $n \in \{\, 1, \ldots, m-1 \,\}$. Dann sind folgende Aussagen äquivalent:

$$\begin{aligned} & n \in \mathbf{Z}_m^* \\ & nq \equiv 1 \mod m \quad \text{für ein } q \in \mathbf{Z}_m \\ & \mathrm{ggT}(m,n) = 1. \end{aligned}$$

(b). Folgt aus (a) nach Definition der Eulerschen φ-Funktion.

(c). Seien $n, m > 0$ und teilerfremd. Nach (b) genügt es zu zeigen, daß $\mathbf{Z}_{mn}^* \cong \mathbf{Z}_m^* \times \mathbf{Z}_n^*$. (Hierbei ist das direkte Produkt von Gruppen gemeint: die Verknüpfung geschieht komponentenweise). Nach Teil (b) des chinesischen Restsatzes wissen wir bereits $\mathbf{Z}_{mn} \cong \mathbf{Z}_m \times \mathbf{Z}_n$. Zu zeigen bleibt $(\mathbf{Z}_m \times \mathbf{Z}_n)^* = \mathbf{Z}_m^* \times \mathbf{Z}_n^*$. Dies folgt aus der Äquivalenz der folgenden Aussagen.

$$\begin{aligned} & (x,y) \in (\mathbf{Z}_m \times \mathbf{Z}_n)^* \\ & (x,y) \cdot (x',y') = (1,1) \quad \text{für geeignete } (x',y') \in \mathbf{Z}_m \times \mathbf{Z}_n \\ & xx' = 1 \text{ und } yy' = 1 \text{ für geeignete } x' \in \mathbf{Z}_m \text{ und } y' \in \mathbf{Z}_n \\ & (x,y) \in \mathbf{Z}_m^* \times \mathbf{Z}_n^*. \end{aligned}$$

(d). Betrachte

$$0 \ldots p \ldots 2p \ldots lp \ldots (l+1)p \ldots p^{r-1}p.$$

Es genügt zu zeigen, daß jedes m mit $lp < m < (l+1)p$ zu p teilerfremd ist. (Denn da es genau p^{r-1} Intervalle $\{\, m \mid lp < m < (l+1)p \,\}$ gibt und in jedem $p-1$ Elemente liegen, folgt daraus $\varphi(p^r) = p^{r-1}(p-1)$). Nehmen wir $p \mid m$ an. Dann folgt $m = kp$ für ein $k \in \mathbf{Z}$ und damit $lp < kp < (l+1)p$; dies ist aber nicht möglich. □

In 3.1.4 hatten wir den kleinen Fermatschen Satz bewiesen. Er sagte für eine beliebige endliche Gruppe G aus, daß für alle $x \in G$ gilt

$$x^{|G|} = e.$$

Wir wollen diesen allgemeinen gruppentheoretischen Satz jetzt auf die multiplikative Gruppe $\mathbf{Z}_m^*$ anwenden. Dadurch ergeben sich interessante zahlentheoretische Aussagen.

SATZ (Fermat). *Sei p eine Primzahl. Dann gilt $x^{p-1} \equiv 1 \mod p$ für jedes $x \in \{1, \ldots, p-1\}$.*

BEWEIS. Die multiplikative Gruppe $\mathbf{Z}_p^*$ der primen Reste modulo p hat genau $p-1$ Elemente, nämlich $\{1, \ldots, p-1\}$. □

Eine unmittelbare Folgerung ist, daß für Primzahlen p gilt $x^p \equiv x \mod p$ für jedes $x \in \{1, \ldots, p-1\}$. Da wir auch für beliebiges $n > 0$ die Anzahl der Elemente der multiplikativen Gruppe $\mathbf{Z}_n^*$ bestimmt hatten, erhalten wir den allgemeineren Satz von Euler-Fermat:

SATZ (Euler-Fermat). *Sei $n > 0$. Dann gilt $x^{\varphi(n)} \equiv 1 \mod n$ für jeden primen Rest x modulo n.*

3.3.4. Der Satz von Wilson. Als Anwendung der Gruppeneigenschaften von $\mathbf{Z}_n^*$, also der Gruppe der primen Reste modulo n, wollen wir ein nützliches Kriterium für die Primzahleigenschaft beweisen.

SATZ (Wilson). *Sei $p \in \mathbf{N}$, $p \geq 1$. Dann ist p Primzahl genau dann, wenn $(p-1)! \equiv -1 \mod p$ ist.*

BEWEIS. $\leftarrow$. Gelte $(p-1)! \equiv -1 \mod p$. Dann ist p Teiler von $(p-1)!+1$, und kein n mit $1 < n < p$ ist Teiler von $(p-1)!+1$. Also ist p Primzahl.

$\rightarrow$. Für $p = 2, 3$ ist die Behauptung offenbar richtig. Sei also $p \geq 5$. Betrachte $\mathbf{Z}_p^* = \{1, 2, \ldots, p-1\}$, und ein beliebiges a darin. Es gibt dann ein eindeutig bestimmtes $a' \in \mathbf{Z}_p^*$ mit $aa' \equiv 1 \mod p$. Im Fall $a \equiv a' \mod p$ muß $a \equiv \pm 1 \mod p$ sein, denn aus $a \equiv a' \mod p$ folgt $p \mid a^2 - 1$, also $p \mid (a-1)(a+1)$. Aus den Gruppeneigenschaften von $\mathbf{Z}_p^*$ (insbesondere der Eindeutigkeit des Inversen) folgt, daß sich alle Elemente aus $\mathbf{Z}_p^* \setminus \{1, p-1\}$ zu $\frac{p-1}{2}$ Paaren a, a' mit $aa' \equiv 1 \mod p$ zusammenfassen lassen. Also ist

$$(p-1)! \equiv (+1)(-1) \prod aa' \equiv -1 \mod p. \qquad \square$$

3.3.5. Anwendung: RSA-Verfahren. Eine Faktorisierung zufälliger großer Zahlen läßt sich auch mit leistungsfähigen Großrechnern nicht in vernünftiger Zeit durchführen. Auf dieser praktischen Unmöglichkeit der Faktorisierung beruht eine gängige Verschlüsselungstechnik, nämlich das nach seinen Erfindern Ronald Rivest, Adi Shamir und Leonard Adleman

benannte RSA-Verfahren. Dessen Grundidee besteht darin, daß die Verschlüsselung mit Hilfe eines sogenannten „öffentlichen Schlüssels“ einfach durchzuführen ist, jedoch das Entschlüsseln nur bei Verwendung eines geheimen „privaten“ Schlüssels in vernünftiger Zeit möglich ist.

Seien p, q verschiedene Primzahlen und $N := pq$. Man wählt sich ein k mit $1 < k < \varphi(N)$ so daß $\mathrm{ggT}(\varphi(N), k) = 1$; φ ist die Eulerfunktion, also $\varphi(N) = (p-1)(q-1)$. Aus k und $\varphi(N)$ kann man mit Hilfe des Euklidischen Algorithmus eine Zahl a mit $ka \equiv 1 \mod \varphi(N)$ berechnen. Das Paar (a, N) nennt man den *öffentlichen Schlüssel*. Man verwendet ihn zur Verschlüsselung der Nachricht m mit $1 \leq m < pq$ mittels der Funktion

$$f(m) := m^a \mod N.$$

Die Entschlüsselung erfolgt dann mit Hilfe des *privaten Schlüssels* k, durch Berechnung von $f(m)^k \mod N$. Dies ist möglich, weil folgendes gilt.

Lemma. *Seien p, q verschiedene Primzahlen und $1 \leq m < pq$. Dann ist*

$$(m^a)^k \equiv m \mod pq.$$

Beweis. Da p und q verschiedene Primzahlen sind, genügt es, die Kongruenz modulo p und modulo q zu beweisen. Aus Symmetriegründen können wir uns auf den Beweis für p beschränken. Im Fall $p \mid m$ ist die Behauptung offenbar richtig (man braucht $0 < ak$, was aus $ak \equiv 1 \mod \varphi(N)$ folgt). Sei also p kein Teiler von m. Nach dem kleinen Fermatschen Satz gilt dann $m^{p-1} \equiv 1 \mod p$. Wegen $ka \equiv 1 \mod \varphi(N)$ haben wir b mit

$$ak + b(p-1)(q-1) = 1,$$

also durch Potenzieren mit Basis m

$$m^{ak}(m^{p-1})^{b(q-1)} = m.$$

Wegen $m^{p-1} \equiv 1 \mod p$ folgt $m^{ak} \equiv m \mod p$. □

Beispiel. Seien $p := 53$ und $q := 61$ die gegebenen Primzahlen, also $N = 3233$ und $\varphi(N) = 3120$. Wir wählen eine Zahl $k := 1013$ mit $\mathrm{ggT}(k, \varphi(N)) = 1$ (k ist sogar eine Primzahl). Aus k und $\varphi(N) = 3120$ erhält man mit Hilfe des Euklidischen Algorithmus $77k = 78001 = 25 \cdot 3120 + 1$, also $a = 77$. Der öffentliche Schlüssel ist also $(a, N) = (77, 3233)$. Ein Zahl $m < 3233$ kann man jetzt verschlüsseln als $m^{77} \mod 3233$. Etwa für $m := 10$ erhält man $10^{77} \mod 3233 = 2560$. Zur Entschlüsselung benötigt man den privaten Schlüssel $k = 1013$: es ist $2560^{1013} \mod 3233 = 10$.

KAPITEL 4

Relationen

Eine Relation R über einer endlichen Menge kann man durch eine Matrix über dem Halbring $(\{0,1\}, \max, \cdot)$ darstellen. Der Verkettung $R \circ S$ zweier Relationen entspricht dann die Matrizenmultiplikation. Da der bekannte Algorithmus zur Multiplikation zweier $n \times n$-Matrizen die Komplexität $O(n^3)$ besitzt, ergibt sich daraus ein Algorithmus zur Berechnung der transitiven Hülle einer Relation von der Komplexität $O(n^4)$. Wir überlegen uns in diesem Kapitel, daß sich dieses Problem auch mit der Komplexität $O(n^3)$ lösen läßt. Dies leistet der Warshall-Algorithmus; er beruht auf der einfachen Idee, Wiederholungen in Pfaden zu vermeiden.

4.1. Verknüpfung von Relationen und Matrizenmultiplikation

4.1.1. Relationen und ihre Potenzen. Sei X eine Menge. Eine *Relation* auf X ist eine Teilmenge $R \subseteq X \times X$. R heißt

(a) *reflexiv*, wenn $\forall_x Rxx$;
(b) *antireflexiv*, wenn $\forall_x \neg Rxx$;
(c) *symmetrisch*, wenn $\forall_{x,y}(Rxy \to Ryx)$;
(d) *transitiv*, wenn $\forall_{x,y,z}(Rxy \to Ryz \to Rxz)$.

Unter der *transitiven Hülle* R^+ einer Relation R versteht man die kleinste R umfassende transitive Relation. Entsprechend definiert man den Begriff der *reflexiv-transitiven Hülle* R^* einer Relation R.

Definition. Seien R und S zwei Relationen auf der Menge X. Die *Verknüpfung* $R \circ S$ ist wie folgt definiert.

$$\big((x,y) \in R \circ S\big) := \exists_{z \in X}((x,z) \in R \wedge (z,y) \in S).$$

Insbesondere kann man die *Potenzen* R^n einer Relation R auf X bilden:

$$\begin{aligned} R^0 &:= \Delta_X := \{\, (x,x) \mid x \in X \,\}, \\ R^1 &:= R, \\ R^2 &:= R \circ R, \\ &\vdots \end{aligned}$$

$$R^{k+1} := R^k \circ R \quad \text{für } k \geq 1.$$

LEMMA. *Sei R eine Relation auf X. Dann gilt für die transitive Hülle R^+ von R*

$$R^+ = \bigcup_{n=1}^{\infty} R^n.$$

Für die reflexiv-transitive Hülle R^* von R gilt

$$R^* = \bigcup_{n=0}^{\infty} R^n.$$

BEWEIS. Übung. □

Ist R reflexiv (d.h., gilt $\Delta_X \subseteq R$), so hat man $R \subseteq R^2 \subseteq R^3 \dots$.

DEFINITION (Matrix einer Relation). Sei R eine Relation auf der Menge $X = \{x_1, x_2, \dots, x_n\}$, x_i paarweise verschieden. Dann kann R durch folgende $n \times n$-Matrix $M_R = (r_{ij})$ beschrieben werden:

$$r_{ij} := \begin{cases} 1 & \text{falls } (x_i, x_j) \in R \\ 0 & \text{sonst.} \end{cases}$$

BEISPIEL. Sei

$$X := \{x_1, x_2, x_3, x_4\},$$
$$R := \{(x_1, x_2), (x_2, x_3), (x_2, x_4), (x_3, x_4), (x_4, x_1), (x_4, x_4)\}.$$

Dann ist

$$M_R = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

Eine bequeme graphische Darstellung einer Relation $R \subseteq M \times M$ erhält man mittels *gerichteter Graphen*. Die Elemente der zugrunde liegenden Menge M werden durch Punkte wiedergegeben, und zwei Punkte werden durch eine *gerichtete Kante* (einen geraden oder gekrümmten Pfeil) verbunden wenn sie in der Relation R stehen. Als Beispiel kann man sich etwa die Teilbarkeitsrelation auf der Menge $\{1, 2, 3, 6\}$ auf diese Weise veranschaulichen.

4.1.2. Halbringe. Wir wollen jetzt die Verknüpfung zweier Relationen mit der Matrizenmultiplikation über Halbringen in Verbindung bringen.

DEFINITION. Sei A eine Menge und $\vee\colon A \to A \to A$, $\wedge\colon A \to A \to A$ Abbildungen. A (oder genauer $(A, \vee, \wedge)$) heißt kommutativer *Halbring*, wenn gilt

(a) $(A, \vee)$ ist eine abelsche Halbgruppe mit einem neutralen Element $0 \in A$, d.h. es gilt
  (i) $(x \vee y) \vee z = x \vee (y \vee z)$ für alle $x, y, z \in A$,
  (ii) $x \vee y = y \vee x$ für alle $x, y \in A$,
  (iii) $0 \vee x = x$ für alle $x \in A$.

(b) $(A, \wedge)$ ist eine abelsche Halbgruppe, d.h. es gilt
  (i) $(x \wedge y) \wedge z = x \wedge (y \wedge z)$ für alle $x, y, z \in A$.
  (ii) $x \wedge y = y \wedge x$ für alle $x, y \in A$,

(c) Es gilt das Distributivgesetz $(x \vee y) \wedge z = (x \wedge z) \vee (y \wedge z)$ für alle $x, y, z \in A$. Ferner gilt $0 \wedge x = 0$ für alle $x \in A$.

Ein Element $1 \in A$ heißt *Einselement* von A, wenn gilt $1 \wedge x = x$ für alle $x \in A$.

Beispiele. Sei X eine beliebige Menge und $A = \mathcal{P}(X)$ die Menge aller Teilmengen von X (also die *Potenzmenge* von X). Dann ist $(A, \cup, \cap)$ ein kommutativer Halbring. Neutrales Element ist die leere Menge $\emptyset$, und Einselement ist die Gesamtmenge X. Ein weiteres Beispiel erhält man aus $A = \{0, 1\}$ und

$$x \vee y := \max(x, y),$$
$$x \wedge y := \min(x, y) \quad (= xy \text{ mit der gewöhnlichen Multiplikation in } \mathbf{N}).$$

Wir können 0 und 1 als falsch und wahr verstehen, und dann $\wedge$ und $\vee$ als logisches „und“ und „oder“. Ebenso ist eine mengentheoretische Interpretation möglich: Für $X := \{*\}$ betrachte man $\mathcal{P}(X) = \{\emptyset, \{*\}\}$. Dann entspricht die leere Menge $\emptyset$ der 0 und die Einermenge $\{*\}$ der 1.

4.1.3. Matrizenmultiplikation.

Definition (Matrizenmultiplikation über Halbringen). Sei $(A, \vee, \cdot)$ ein kommutativer Halbring mit Einselement 1. Seien $M_1 = (a_{ij})_{1 \le i,j \le n}$ und $M_2 = (b_{ij})_{1 \le i,j \le n}$ zwei $n \times n$-Matrizen mit Elementen $a_{ij}, b_{ij} \in A$. Dann wird das *Produkt*

$$M := M_1 M_2$$

definiert als

$$M = (c_{ij})_{1 \le i,j \le n} \quad \text{mit} \quad c_{ij} = \bigvee_{k=1}^{n} a_{ik} b_{kj}.$$

Dabei ist $\bigvee_{k=1}^{n} x_k$ eine Abkürzung für $x_1 \vee \cdots \vee x_k$.

Sei $(A, \vee, \cdot)$ ein kommutativer Halbring mit Einselement 1. Dann ist das Matrizenprodukt in $M(n \times n, A)$ offenbar assoziativ; neutrales Element ist

$$E_n := \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}.$$

SATZ. *Seien R und S zwei Relationen auf einer endlichen Menge $X = \{x_1, \dots, x_n\}$ und M_R, M_S die zugehörigen Matrizen aus $M(n \times n, A)$ mit $A = \{0, 1\}$. Dann gilt für die Matrix $M_{R \circ S}$ der Verknüpfung von R mit S*

$$M_{R \circ S} = M_R M_S,$$

wobei die Matrizenmultiplikation über dem Halbring $(\{0, 1\}, \vee, \cdot)$ mit $x \vee y := \max(x, y)$ zu verstehen ist.

BEWEIS. Seien $M_R = (a_{ij})$ und $M_S = (b_{ij})$ mit

$$a_{ij} = \begin{cases} 1 & \text{falls } (x_i, x_j) \in R, \\ 0 & \text{sonst,} \end{cases}$$

und analog für b_{ij}. Ferner sei $M_{R \circ S} = (c_{ij})$. Dann gilt

$$\begin{aligned} c_{ij} = 1 &\leftrightarrow (x_i, x_j) \in R \circ S \\ &\leftrightarrow \exists_k((x_i, x_k) \in R \wedge (x_k, x_j) \in S) \\ &\leftrightarrow \exists_k(a_{ik} = 1 \wedge b_{kj} = 1). \end{aligned}$$

Sei andererseits $M_R M_S = (c'_{ij})$. Dann gilt $c'_{ij} = \bigvee_k a_{ik} b_{kj} = \max_k(a_{ik} b_{kj})$ und

$$\begin{aligned} (\max_k(a_{ik} b_{kj}) = 1) &\leftrightarrow \exists_k(a_{ik} b_{kj} = 1) \\ &\leftrightarrow \exists_k(a_{ik} = 1 \wedge b_{kj} = 1). \end{aligned}$$

Also ist $c_{ij} = c'_{ij}$. □

4.1.4. Ein naiver Algorithmus zur Matrizenmultiplikation. Wir beschreiben jetzt einen Algorithmus zur Multiplikation von Matrizen über dem Halbring $(\{0, 1\}, \vee, \cdot)$. Sei $M = (a_{ij})_{1 \le i,j \le n}$, also

$$M = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Die Zeilen $a_i = (a_{i1}, \dots, a_{in})$ werden als Bitvektoren aufgefaßt. Für zwei Bitvektoren $x = (x_1, \dots, x_n) \in \{0, 1\}^n$ und $y = (y_1, \dots, y_n) \in \{0, 1\}^n$ bezeichne

$x \vee y \in \{0,1\}^n$ den Bitvektor

$$x \vee y := (x_1 \vee y_1, \ldots, x_n \vee y_n).$$

Gegeben seien also zwei Matrizen $M_1, M_2 \in M(n \times n, \{0,1\})$ mit den Zeilen

$$M_1 = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \quad \text{und} \quad M_2 = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Gesucht ist

$$M_1 \cdot M_2 = M_3 = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}.$$

Die c_i können durch den folgenden Algorithmus berechnet werden.

```
for i := 1 to n do
   c_i := 0
   for j := 1 to n do
      if a_ij = 1 then
         c_i := c_i ∨ b_j
      end;
   end;
end;
```

Wir wollen uns jetzt die Korrektheit dieses Algorithmus überlegen. In den n Durchläufen der äußeren Schleife werden die Zeilen c_i unabhängig voneinander berechnet. Es genügt also, nur einen Durchlauf für ein festes i zu betrachten. Sei $c_i^{(j)}$ der Inhalt des Speicherplatzes für c_i nach dem j-ten Durchlauf der inneren Schleife. Vor dem ersten Durchlauf der inneren Schleife (mit Laufvariable j) hat man $c_i^{(0)} = 0$. Für $j > 0$ gilt

$$c_{ik}^{(j)} = c_{ik}^{(j-1)} \vee a_{ij} b_{jk} \quad \text{für } k = 1 \ldots n.$$

Daraus folgt durch Induktion $c_{ik}^{(n)} = \bigvee_{j=1}^{n} a_{ij} b_{jk}$. Dies ist aber die (i,k)-te Komponente des gesuchten Matrizenprodukts.

Die Komplexität des Algorithmus ist $O(n^3)$. (Dies gilt, falls jede Bitoperation einzeln gezählt wird. Hat man die Möglichkeit, die Bitoperationen auf einem Bitvektor parallel durchzuführen, so ist die Komplexität $O(n^2)$). Man beachte, daß die Komplexität der gewöhnlichen Matrizenmultiplikation für $n \times n$-Matrizen ebenfalls $O(n^3)$ ist.

4.2. Berechnung der transitiven Hülle

4.2.1. Ein naiver Algorithmus. Wir wollen nun die Komplexität eines naiven Algorithmus zur Berechnung der transitiven Hülle abschätzen.

LEMMA. *Sei $X = \{x_1, \ldots, x_n\}$ eine endliche Menge und R eine Relation auf X. Dann gilt*

$$R^+ = \bigcup_{k=1}^{n} R^k.$$

BEWEIS. Es gilt

$$(x, y) \in R^+ \leftrightarrow \exists_{z_1,\ldots,z_m}(xRz_1Rz_2R \ldots z_{m-1}Rz_mRy).$$

OBdA sind die z_i untereinander und von x, y verschieden. Also folgt aus $(x, y) \in R^+$, daß es ein k mit $1 \leq k \leq n$ gibt mit $(x, y) \in R^k$. □

Sei M_R die zu R gehörige Matrix. Dann ist offenbar

$$M_{R^+} = \bigvee_{k=1}^{n} M_R^k.$$

Dabei wird $\vee$ (= max) komponentenweise auf die einzelnen Matrizen angewandt. Aufgrund der obigen Abschätzung des Matrizenmultiplikations-Algorithmus kommt man also auf die Komplexität $O(n^4)$.

4.2.2. Der Warshall-Algorithmus. Der Warshall-Algorithmus berechnet die transitive Hülle R^+ einer Relation R. Während der naive Algorithmus eine Komplexität $O(n^4)$ hat, kommt der Warshall-Algorithmus mit der Komplexität $O(n^3)$ aus.

Wir beschreiben zunächst die Idee des Warshall-Algorithmus. Gegeben ist eine Relation R auf einer n-elementigen Menge $X = \{x_1, \ldots, x_n\}$. Zu vorgelegtem x_j, x_k wollen wir entscheiden, ob es einen *R-Pfad* von x_j nach x_k gibt, d.h., ob es $z_1, \ldots, z_m \in X$ gibt mit $x_jRz_1Rz_2R \ldots z_{m-1}Rz_mRx_k$. Die Grundidee des Warshall-Algorithmus ist nun folgende:

Vermeide Wiederholungen bei den Zwischenpunkten $z_1, \ldots, z_m$.

Wir definieren zunächst Relationen $R^{(i)}$ so daß $(x_j, x_k) \in R^{(i)}$ genau dann, wenn es einen R-Pfad von x_j nach x_k gibt, der nur Zwischenpunkte aus $\{x_1, \ldots, x_i\}$ verwendet. Sei also $R^{(0)} := R$ und

$$(x_j, x_k) \in R^{(i)} := (x_j, x_k) \in R^{(i-1)} \vee x_jR^{(i-1)}x_iR^{(i-1)}x_k.$$

Offenbar gilt dann wie gewünscht

$$(x, y) \in R^{(i)} \leftrightarrow (x, y) \in R \vee \exists_{z_1,\ldots,z_m \in \{x_1,\ldots,x_i\}}(xRz_1Rz_2 \ldots z_{m-1}Rz_mRy),$$

also auch $R^{(n)} = R^+$.

Der Warshall-Algorithmus konstruiert Matrizen $A^{(i)}$, die die Relationen $R^{(i)}$ darstellen. Nach n Schritten hat man dann also die gesuchte Matrix $A^{(n)}$ der Relation $R^{(n)} = R^+$. Sei $A^{(0)}$ die darstellende Matrix A der gegebenen Relation R, d.h. $A \in M(n \times n, \{0,1\})$ mit den Zeilen $a_1, \ldots, a_n$:

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Die Zeilen sind Bitvektoren; sie werden durch den Algorithmus verändert.

```
for i := 1 to n do
  for j := 1 to n do
    if a_ji = 1 then
      a_j := a_j ∨ a_i
    end;
  end;
end;
```

Hier sind in $a_j := a_j \vee a_i$ mit a_j, a_i auf der rechten Seite die Bitvektoren des *vorigen* Durchlaufs (also für $i-1$) der äußeren Schleife gemeint.

Sei $A^{(i)}$ der Inhalt des Speicherplatzes für A nach dem i-ten Durchlauf der äußeren Schleife. Wir zeigen $A^{(i)} = M_{R^{(i)}}$ durch Induktion über i. Nach Annahme ist $A = A^{(0)} = M_{R^{(0)}}$. Aufgrund des Algorithmus haben wir

$$a_{jk}^{(i)} = a_{jk}^{(i-1)} \vee a_{ji}^{(i-1)} a_{ik}^{(i-1)}$$

für alle j, i, k. Mit Hilfe der IH $A^{(i-1)} = M_{R^{(i-1)}}$ erhält man

$$\begin{aligned} a_{jk}^{(i)} = 1 \quad &\leftrightarrow \quad a_{jk}^{(i-1)} = 1 \quad \text{oder} \quad \left(a_{ji}^{(i-1)} = 1 \quad \text{und} \quad a_{ik}^{(i-1)} = 1\right) \\ &\leftrightarrow \quad (x_j, x_k) \in R^{(i-1)} \quad \text{oder} \quad x_j R^{(i-1)} x_i R^{(i-1)} x_k \quad \text{nach IH} \\ &\leftrightarrow \quad (x_j, x_k) \in R^{(i)} \quad \text{nach Definition der } R^{(i)}, \end{aligned}$$

also $A^{(i)} = M_{R^{(i)}}$. Wegen $R^{(n)} = R^+$ ist $A^{(n)} = M_{R^{(n)}}$ die darstellende Matrix der transitiven Hülle von R. Damit haben wir die Korrektheit des Warshall-Algorithmus gezeigt; seine Komplexität ist offenbar $O(n^3)$ (bzw. $O(n^2)$, wenn man die Möglichkeit hat, die Bitoperationen auf einem Bitvektor parallel durchzuführen).

KAPITEL 5

Graphen

Einen ungerichteten Graphen G über einer endlichen Menge kann man durch seine Adjazenzmatrix über dem Halbring $(\{0,1\}, \max, \cdot)$ darstellen. Aus dem Warshall-Algorithmus ergibt sich demnach ein Algorithmus der Komplexität $O(n^3)$ zur Feststellung des Zusammenhangs eines Graphen.

Mit Hilfe des Begriffs des Grades eines Knotens geben wir ein hinreichendes und notwendiges Kriterium dafür an, daß Eulersche Wege und Zyklen in einem Graphen existieren. Aus der Notwendigkeit des Kriteriums folgt, daß das Königsberger Brückenproblem unlösbar ist.

Weiter behandeln wir Abstände in bewerteten Graphen. Der einfachste Abstandsbegriff ist die Länge des kürzesten Pfades zwischen zwei Knoten. Mit dem linearen Algorithmus von Moore bestimmen wir zu einem festen Knoten v_0 und einem gegebenen Abstand k alle Knoten, die in diesem Sinn von v_0 den Abstand k haben, also die von v_0 aus durch einen Pfad der Länge höchstens k erreicht werden können.

Ein feinerer Abstandsbegriff stützt sich auf eine vorgegebene Bewertungsfunktion, die jeder Kante eines Graphen eine feste rationale Zahl (oder ∞) als Kantenlänge zuordnet. Mit einer einfachen Verallgemeinerung des Warshall-Algorithmus kann man dann zu je zwei Knoten ihren Abstand berechnen, d.h., die Länge der kürzesten (im Sinne der aufsummierten Kantenlängen) Verbindung zwischen ihnen; die Komplexität ist wieder $O(n^3)$. Abschließend behandeln wir den Dijkstra-Algorithmus zur Berechnung der Abstände von einem festen Knoten v_0; seine Komplexität ist $O(n^2)$.

5.1. Allgemeine Begriffe

5.1.1. Adjazenz- und Inzidenzmatrizen.

Definition. Ein (ungerichteter) *Graph* ist ein Paar $\Gamma = (V, E)$ bestehend aus einer Menge V von *Knoten* oder *Ecken* (englisch *vertex*) und einer Menge $E \subseteq \mathcal{P}(V)$ von zweielementigen Teilmengen von V. Die Elemente von E heißen (ungerichtete) *Kanten* (englisch *edge*). Ist $e = \{A_0, A_1\}$, so sagt man, daß die Kante e die Knoten A_0 und A_1 verbindet.

Ein Beispiel eines Graphen ist in Abbildung 1 skizziert.

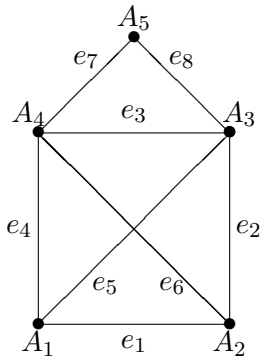


ABBILDUNG 1. Beispiel eines Graphen

Man kann endliche Graphen auf verschiedene Weise durch Matrizen darstellen.

DEFINITION (*Adjazenzmatrix* eines endlichen Graphen). Sei $\Gamma = (V, E)$ ein endlicher Graph mit $V = \{A_1, \ldots, A_n\}$. Dann kann Γ durch folgende Matrix $M_\Gamma = (a_{ij})_{1 \le i,j \le n} \in M(n \times n, \{0, 1\})$ beschrieben werden:

$$a_{ij} := \begin{cases} 1 & \text{falls } \{A_i, A_j\} \in E, \\ 0 & \text{sonst.} \end{cases}$$

Die Adjazenzmatrix eines (ungerichteten) Graphen ist offenbar symmetrisch. Ferner sind alle Diagonalelemente $a_{ii} = 0$. Im Beispiel des Graphen in Abbildung 1 ist die Adjazenzmatrix

$$M_\Gamma = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}.$$

Man beachte, daß die Adjazenzmatrix eines Graphen $\Gamma = (V, E)$ auch die Matrix einer gewissen symmetrischen, antireflexiven Relation auf V ist, nämlich der Relation R_Γ definiert durch

$$R_\Gamma := \{\, (A, B) \mid \{A, B\} \in E \,\}.$$

DEFINITION (*Inzidenzmatrix* eines endlichen Graphen). Sei $\Gamma = (V, E)$ ein endlicher Graph mit $V = \{A_1, \ldots, A_n\}$, $E = \{e_1, \ldots, e_m\}$. Dann kann Γ durch folgende Matrix $\mathrm{IM}_\Gamma = (c_{ij}) \in M(n \times m, \{0, 1\})$ beschrieben werden:

$$c_{ij} := \begin{cases} 1 & \text{falls } A_i \in e_j, \\ 0 & \text{sonst.} \end{cases}$$

Im Beispiel des Graphen in Abbildung 1 ist die Inzidenzmatrix

$$\mathrm{IM}_\Gamma = \begin{array}{cccccccc|c} e_1 & e_2 & e_3 & e_4 & e_5 & e_6 & e_7 & e_8 & \\ \hline 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & A_1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & A_2 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & A_3 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & A_4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & A_5 \end{array}$$

5.1.2. Zusammenhang.

DEFINITION. Sei $\Gamma = (V, E)$ ein Graph. Unter einem *Kantenzug* versteht man eine Liste $(A_0, \ldots, A_n)$ von (nicht notwendig verschiedenen) Knoten so daß $\{A_i, A_{i+1}\} \in E$ für $i < n$. Ein Kantenzug heißt *geschlossen*, wenn $A_0 = A_n$. Ein Kantenzug $(A_0, \ldots, A_n)$ heißt ein *Weg*, wenn keine Kante mehrfach auftritt, also wenn $\{A_i, A_{i+1}\} \neq \{A_j, A_{j+1}\}$ für $i \neq j$. Ein geschlossener Weg heißt ein *Zykel* (oder *Zyklus*). Ein Weg $(A_0, \ldots, A_n)$ in einem Graph heißt *einfach*, wenn $A_i \neq A_j$ für alle $i \neq j$, mit eventueller Ausnahme der Indexpaare $(0, n)$ und $(n, 0)$. Ein Graph $\Gamma = (V, E)$ heißt *zusammenhängend*, wenn es zu je zwei Knoten $A, B \in V$ einen Kantenzug mit Anfangspunkt A und Endpunkt B gibt.

Man beachte, daß ein Graph $\Gamma = (V, E)$ schon dann zusammenhängend ist, wenn es einen Knoten $A_0 \in V$ gibt, so daß sich jeder andere Knoten $B \in V$ mit A_0 durch einen Kantenzug verbinden läßt.

Aus dem Warshall-Algorithmus erhält man den folgenden *Algorithmus zur Feststellung des Zusammenhangs eines Graphen* Γ: Man stelle die Adjazenzmatrix M_Γ des Graphen auf und bestimme mit Hilfe des Warshall-Algorithmus die reflexiv-transitive Hülle $M_\Gamma^* = \bigvee_{k \geq 0} M_\Gamma^k$ der durch M_Γ gegebenen Relation. Der Graph ist genau dann zusammenhängend, wenn die Matrix M_Γ^* nur aus Einsen besteht. Man beachte, daß dies schon dann der Fall ist, wenn eine Zeile von M_Γ^* nur aus Einsen besteht.

DEFINITION. Zwei Knoten eines Graphen $\Gamma = (V, E)$ heißen *äquivalent*, wenn sie sich durch einen Kantenzug verbinden lassen. Die Äquivalenzklassen dieser Äquivalenzrelation heißen *Zusammenhangskomponenten* von Γ.

Besteht eine Zusammenhangskomponente eines Graphen nur aus einem einzigen Knoten, so nennt man diesen Knoten einen *isolierten Punkt* des Graphen.

5.1.3. Isomorphie.

DEFINITION. Zwei Graphen $\Gamma = (V, E)$ und $\Delta = (V', E')$ heißen *isomorph*, wenn es eine bijektive Abbildung $\varphi\colon V \to V'$ gibt, so daß für die induzierte Abbildung $\bar{\varphi}\colon \mathcal{P}(V) \to \mathcal{P}(V')$ gilt $\bar{\varphi}(E) = E'$.

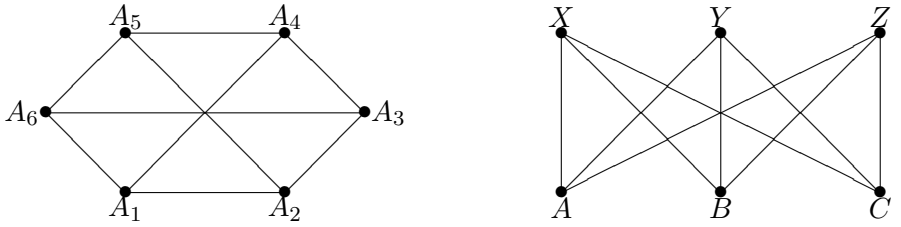


ABBILDUNG 2. Zwei isomorphe Darstellungen des $K33$

BEISPIEL. Man betrachte die beiden Graphen aus Abbildung 2. Beide Graphen sind isomorph, und zwar mittels der Abbildung φ:

$$\begin{array}{rcl} A & \mapsto & A_1 \\ B & \mapsto & A_3 \\ C & \mapsto & A_5 \\ X & \mapsto & A_2 \\ Y & \mapsto & A_4 \\ Z & \mapsto & A_6 \end{array}$$

Hieraus ergibt sich für $\bar{\varphi}$:

$$\begin{array}{cccccccccc} AX & AY & AZ & BX & BY & BZ & CX & CY & CZ \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ A_1A_2 & A_1A_4 & A_1A_6 & A_3A_2 & A_3A_4 & A_3A_6 & A_5A_2 & A_5A_4 & A_5A_6 \end{array}$$

5.1.4. Bipartite Graphen.

DEFINITION. Ein Graph $\Gamma = (V, E)$ heißt *bipartit*, wenn es eine Zerlegung $V = V_1 \cup V_2$ von V in disjunkte Teilmengen V_1 und V_2 gibt so, daß für jede Kante $e \in E$ gilt $e = \{A_1, A_2\}$ mit $A_1 \in V_1$, $A_2 \in V_2$.

BEISPIEL. Zur graphentheoretischen Behandlung des bekannten Problems vom Fährmann mit Wolf, Ziege und Kohl verwendet man am geeignetsten einen bipartiten Graphen mit den folgenden Eckpunkten:

	V_1	V_2	
A_1	$FWZK \mid$	$\mid FWZK$	B_1
A_2	$FZK \mid W$	$W \mid FZK$	B_2
A_3	$FWK \mid Z$	$Z \mid FWK$	B_3
A_4	$FWZ \mid K$	$K \mid FWZ$	B_4
A_5	$FZ \mid WK$	$WK \mid FZ$	B_5

Der Graph in Abbildung 3 beschreibt die Lösung des Problems:

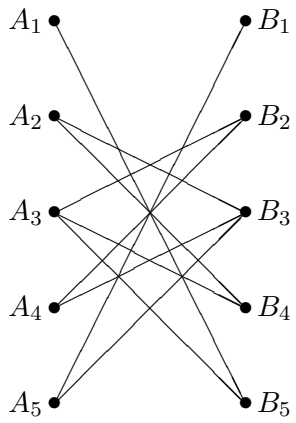
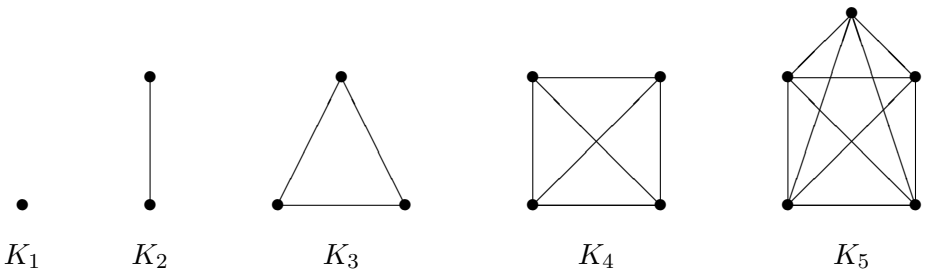


Abbildung 3. Graph zum Fährmann-Wolf-Ziege-Kohl Problem

Definition. Ein Graph $\Gamma = (V, E)$ heißt *vollständig*, wenn E aus *allen* zweielementigen Teilmengen von V besteht. Ein Graph $\Gamma = (V, E)$ heißt *vollständig bipartit*, wenn es eine Zerlegung von V in disjunkte Teilmengen V_1 und V_2 gibt, so daß $E = \{\, \{A_1, A_2\} \mid A_1 \in V_1 \text{ und } A_2 \in V_2 \,\}$.

Der in Abbildung 2 angegebene Garph ist vollständig-bipartit.

Man sieht leicht, daß je zwei vollständige Graphen mit n Eckpunkten isomorph sind. Es gibt also bis auf Isomorphie nur einen vollständigen Graphen mit n Eckpunkten; er wird mit K_n bezeichnet. Die ersten vollständigen Graphen sind



Ebenfalls sind je zwei vollständig-bipartite Graphen mit $\#(V_1) = n$ und $\#(V_2) = m$ isomorph. Der bis auf Isomorphie eindeutig bestimmte vollständig-bipartite Graph wird mit $K_{n,m}$ bezeichnet. In Abbildung 2 ist der $K_{3,3}$ angegeben.

5.2. Eulersche Wege und Zyklen

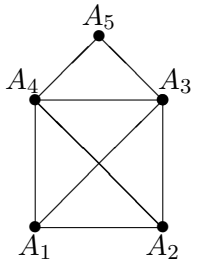
In diesem Abschnitt behandeln wir Eulersche und Hamiltonsche Wege und Zyklen. Wir definieren den Begriff des Grads eines Knotens und geben ein hinreichendes und notwendiges Kriterium dafür an, daß Eulersche Wege und Zyklen in einem Graphen existieren. Mit Hilfe der Notwendigkeit des Kriteriums zeigen wir, daß das Königsberger Brückenproblem unlösbar ist.

5.2.1. Die Euler-Formel.

DEFINITION. Ein *Eulerscher Weg* in einem Graphen $\Gamma = (V, E)$ ist ein Weg, in dem jede Kante $e \in E$ genau einmal vorkommt. Ein *Eulerscher Zyklus* ist ein Eulerscher Weg, der gleichzeitig ein Zyklus ist.

DEFINITION. Sei $\Gamma = (V, E)$ ein Graph. Unter dem *Grad* eines Knotens $A \in V$, geschrieben $\deg(A)$, versteht man die Anzahl derjenigen Kanten $e \in E$, die mit A inzidieren, d.h. die A als ein Element besitzen.

BEISPIEL.



Dann ist

$$\deg(A_1) = \deg(A_2) = 3,\ \deg(A_3) = \deg(A_4) = 4,\ \deg(A_5) = 2.$$

Sei $\Gamma = (V, E)$ ein Graph mit $V = \{A_1, \ldots, A_n\}$ und $M_\Gamma = (a_{ij}) \in M(n \times n, \{0, 1\})$ die Adjazenzmatrix von Γ. Es ist also $a_{ij} = 1$ genau dann, wenn A_i und A_j durch eine Kante verbunden werden. Damit ergibt sich

$$\begin{aligned} \deg(A_i) &= \text{Anzahl der } j \text{ mit } a_{ij} = 1 \\ &= \text{Zeilensumme der } i\text{-ten Zeile von } M_\Gamma. \end{aligned}$$

Ferner gibt es einen entsprechenden Zusammenhang der Gradbegriffs mit der Inzidenzmatrix IM_Γ von Γ. Sei also noch $E = \{e_1, \ldots, e_m\}$ und $\mathrm{IM}_\Gamma = (b_{ij}) \in M(n \times m, \{0, 1\})$ die Inzidenzmatrix von Γ. Dann ist nach Definition der Inzidenzmatrix $b_{ij} = 1$ genau dann, wenn der Knoten A_i mit der Kante e_j inzidiert, also wieder

$$\begin{aligned} \deg(A_i) &= \text{Anzahl der } j \text{ mit } b_{ij} = 1 \\ &= \text{Zeilensumme der } i\text{-ten Zeile von } \mathrm{IM}_\Gamma. \end{aligned}$$

LEMMA (Euler-Formel). *Sei $\Gamma = (V, E)$ ein endlicher Graph. Dann gilt*

$$\sum_{A \in V} \deg(A) = 2 \cdot \#(E).$$

Insbesondere ist also die Anzahl der Knoten $A \in V$ mit $\deg(A)$ ungerade eine gerade Zahl.

BEWEIS. Induktion nach der Anzahl $m = \#(E)$ der Kanten von Γ. *Basis* $m = 0$. Dann besteht der Graph nur aus isolierten Punkten, also $\deg(A) = 0$ für alle $A \in V$. *Schritt* $m - 1 \mapsto m$. Sie Γ' der Graph, der aus Γ durch Wegnahme einer (beliebigen) Kante e entsteht, also $\Gamma' = (V, E \setminus \{e\})$. Für Γ' gilt nach IH

$$\sum \deg_{\Gamma'}(A) = 2 \cdot \#(E \setminus \{e\}) = 2 \cdot \#(E) - 2.$$

Sei $e = \{B, C\}$. Dann ist

$$\begin{aligned} \deg_{\Gamma'}(B) &= \deg_{\Gamma}(B) - 1, \\ \deg_{\Gamma'}(C) &= \deg_{\Gamma}(C) - 1, \\ \deg_{\Gamma'}(P) &= \deg_{\Gamma}(P) \qquad \text{für } P \notin \{B, C\}. \end{aligned}$$

Es folgt

$$\sum \deg_{\Gamma}(A) = \Big(\sum \deg_{\Gamma'}(A)\Big) + 2$$

und damit die Behauptung. □

5.2.2. Eulersche Wege und Zyklen. Wir geben ein notwendiges Kriterium dafür an, daß α ein Eulerscher Weg bzw. Zyklus in einem Graphen ist.

LEMMA (Notwendiges Kriterium für Eulersche Wege und Zyklen).

(a) *Sei α ein Eulerscher Weg in einem endlichen Graphen $\Gamma = (V, E)$ mit verschiedenem Anfangs- und Endpunkt A und B. Dann gilt*

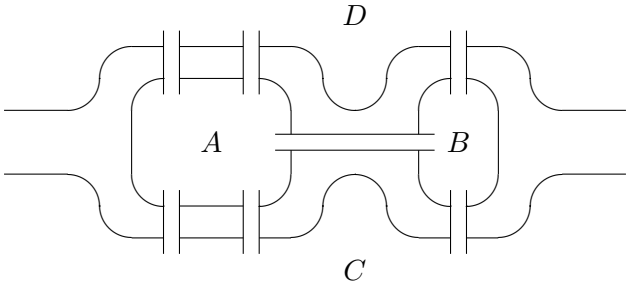
$$\begin{aligned} \deg(A) &\equiv 1 \mod 2, \\ \deg(B) &\equiv 1 \mod 2, \\ \deg(P) &\equiv 0 \mod 2 \quad \textit{für alle } P \in V \setminus \{A, B\}. \end{aligned}$$

(b) *Ist α ein Eulerscher Zyklus in einem endlichen Graphen $\Gamma = (V, E)$, so ist*

$$\deg(P) \equiv 0 \mod 2 \quad \textit{für alle } P \in V.$$

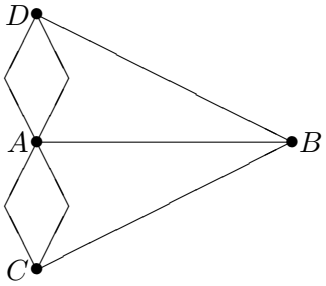
BEWEIS. (a). Beim Durchlaufen eines Knotens gibt ein innerer Punkt einen Beitrag 2 zum Grad des Knotens, und der Anfangs- und Endpunkt je einen Beitrag 1. Teil (b) ergibt sich mit demselben Argument. □

Beispiel (Königsberger Brückenproblem). Zu finden ist ein Spaziergang, der jede der 7 Brücken genau einmal benützt und zum Ausgangspunkt zurückkehrt.



Euler löste dieses Problem im negativen Sinne.

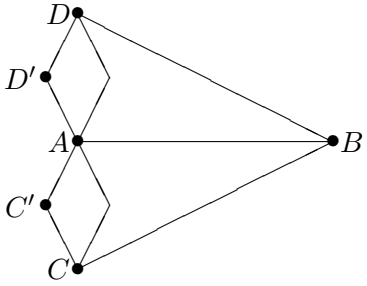
Wir haben folgenden Multi-Graphen Γ (d.h. Graphen mit evtl. mehr als einer Kante zwischen je zwei Knoten) zu betrachten:



Es ist

$$\deg(A) = 5, \quad \deg(B) = \deg(C) = \deg(D) = 3.$$

Durch Einführung zusätzlicher Knoten läßt sich das Problem auf eines für (einfache) Graphen zurückführen. Sei also Γ' der folgende Graph:



Hier hat man dieselben Grade für A, B, C, D und zusätzlich $\deg(C') = \deg(D') = 2$. Offenbar gibt es genau dann einen Eulerschen Zyklus in Γ',

wenn es einen solchen in Γ gibt. Aus dem vorangehenden Lemma folgt aber, daß es in Γ' keinen Eulerschen Zyklus geben kann.

Wir zeigen jetzt, daß das angegebene notwendige Kriterium für die Existenz Eulerscher Wege bzw. Zykeln auch hinreichend ist.

LEMMA. *Sei $\Gamma = (V, E)$ ein zusammenhängender Graph und $e \in E$. Dann besteht der Graph $\Gamma' = (V, E \setminus \{e\})$ aus höchstens zwei Zusammenhangskomponenten.*

BEWEIS. Sei $e = \{A, B\}$ mit $A \neq B$. Es genügt zu zeigen, daß in $\Gamma' = (V, E \setminus \{e\})$ jeder Knoten $C \in V$ mit mindestens einem der Knoten A oder B durch einen Kantenzug verbunden werden kann. Sei also $C \in V$. Da Γ zusammenhängend ist, kann in Γ der Knoten C mit A durch einen Kantenzug α verbunden werden. Kommt die Kante e in α nicht vor, so sind wir fertig. Andernfalls gibt es eine erstes Auftreten von e in α, d.h. α ist von der Form

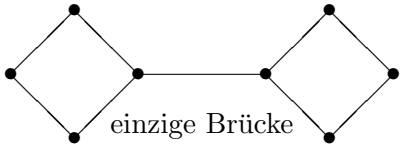
$$\alpha = (C, D_1, \ldots, D_n, \underbrace{A, B}_{e}, \ldots, A) \quad \text{oder} \quad \alpha = (C, D_1, \ldots, D_n, \underbrace{B, A}_{e}, \ldots, A).$$

Dann ist aber $\alpha' = (C, D_1, \ldots, D_n, A)$ bzw. $\alpha' = (C, D_1, \ldots, D_n, B)$ ein Kantenzug von C nach A bzw. nach B in Γ. □

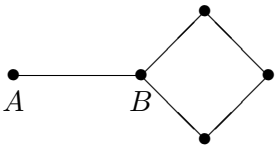
DEFINITION. Sei $\Gamma = (V, E)$ ein zusammenhängender Graph. Eine Kante $e \in E$ heißt *Brücke*, wenn der Graph $\Gamma' = (V, E \setminus \{e\})$, der aus Γ durch Wegnahme der Kante e entsteht, unzusammenhängend ist.

Sei $e = \{A, B\}$ eine Brücke im Graphen $\Gamma = (V, E)$, so daß der Graph $\Gamma' = (V, E \setminus \{e\})$ den Punkt A als isolierten Punkt besitzt. Dann heißt e *hängende Kante* von Γ und A *hängender Knoten* von Γ.

BEISPIEL.



Ein Beispiel für eine hängende Kante ist

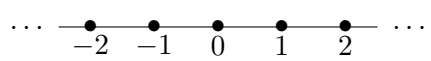


LEMMA. *Sei $\Gamma = (V, E)$ ein endlicher zusammenhängender Graph mit*

$$\deg(P) \equiv 0 \mod 2 \quad \textit{für alle } P \in V.$$

Dann besitzt Γ keine Brücke.

Man beachte, daß dies falsch wird, wenn man die Voraussetzung der Endlichkeit fallen läßt. Ein Gegenbeispiel ist der Graph $\Gamma = (\mathbb{Z}, E)$ mit $E = \{\, \{n, n+1\} \mid n \in \mathbb{Z} \,\}$, also



BEWEIS. Sei $e = \{A, B\}$ eine Brücke in Γ. Dann zerfällt $\Gamma' = (V, E \setminus \{e\})$ in zwei Zusammenhangskomponenten. Sei $\Gamma_1 = (V_1, E_1)$ die Zusammenhangskomponente von A. Es gilt dann

$$\begin{aligned} \deg_{\Gamma_1}(A) &= \deg_\Gamma(A) - 1, \\ \deg_{\Gamma_1}(P) &= \deg_\Gamma(P) \quad \text{für alle } P \in V_1 \setminus \{A\}. \end{aligned}$$

Es gibt also in Γ_1 genau einen Knoten (nämlich A) mit ungeradem Grad; alle anderen haben geraden Grad. Dies widerspricht aber der Euler-Formel. □

LEMMA. *Sei $\Gamma = (V, E)$ ein endlicher zusammenhängender Graph, der genau zwei Knoten mit ungeradem Grad besitzt. Sei $A \in V$ ein Knoten mit $\deg_\Gamma(A)$ ungerade und > 1. Dann gibt es mindestens eine Kante $e = \{A, B\} \in E$, die keine Brücke ist.*

BEWEIS. Sei $e = \{A, B\}$ eine beliebige von A ausgehende Kante. Wenn e keine Brücke ist, so sind wir fertig. Sei also e eine Brücke. Dann zerfällt Γ in zwei Zusammenhangskomponenten. Sei $\Gamma_1 = (V_1, E_1)$ die Zusammenhangskomponente von A. Alle Knoten in Γ_1 haben dann (nach Voraussetzung und der Euler-Formel) einen geraden Grad. Aus dem vorigen Lemma folgt, daß Γ_1 keine Brücke besitzt. □

Jetzt können wir zeigen, daß das oben angegebene Kriterium für die Existenz Eulerscher Wege und Zykeln auch hinreichend ist.

SATZ (Hinreichendes Kriterium für Eulersche Wege und Zyklen). *Sei $\Gamma = (V, E)$ ein endlicher zusammenhängender Graph.*

(a) *Sind $A, B \in V$ Knoten mit*

$$\begin{aligned} \deg(A) &\equiv 1 \mod 2, \\ \deg(B) &\equiv 1 \mod 2, \\ \deg(P) &\equiv 0 \mod 2 \quad \textit{für alle } P \in V \setminus \{A, B\}, \end{aligned}$$

so gibt es in Γ einen Eulerschen Weg von A nach B.

(b) *Gilt*

$$\deg(P) \equiv 0 \mod 2 \quad \textit{für alle } P \in V,$$

so besitzt Γ einen Eulerschen Zyklus.

BEWEIS. Induktion über die Anzahl der Kanten von Γ.

(a). *Fall* $\deg(A) = 1$. Dann ist A ein hängender Knoten und die einzige von A ausgehende Kante e ist eine hängende Kante. Wir betrachten den Graphen $\Gamma' = (V \setminus \{A\}, E \setminus \{e\})$; sei $e = \{A, C\}$.

UFall $C = B$. Dann haben alle Knoten in Γ' einen geraden Grad. Nach IH(b) gibt es deshalb in Γ' einen Eulerschen Zyklus mit Anfangs- und Endpunkt B. Zusammengesetzt mit der Kante e liefert dies einen Eulerschen Weg in Γ von A nach B.

UFall $C \neq B$. Dann gilt

$$\begin{aligned} &\deg_{\Gamma'}(C) \equiv 1 \mod 2, \\ &\deg_{\Gamma'}(B) \equiv 1 \mod 2, \\ &\deg_{\Gamma'}(P) \equiv 0 \mod 2 \quad \text{für alle } P \in V \setminus \{A, B, C\}. \end{aligned}$$

Nach IH(a) gibt es deshalb in Γ' einen Eulerschen Weg von C nach B. Zusammengesetzt mit der Kante e liefert dies einen Eulerschen Weg in Γ von A nach B.

Fall $\deg(A) \geq 3$. Dann gibt es nach dem vorigen Lemma eine von A ausgehende Kante $e = \{A, C\}$, die keine Brücke ist. Also ist $\Gamma' = (V, E \setminus \{e\})$ zusammenhängend.

UFall $C = B$. Dann haben alle Knoten in Γ' einen geraden Grad. Nach IH(b) gibt es deshalb in Γ' einen Eulerschen Zyklus mit Anfangs- und Endpunkt B. Zusammengesetzt mit der Kante e liefert dies einen Eulerschen Weg in Γ von A nach B.

UFall $C \neq B$. Dann haben in Γ' die Knoten C und B einen ungeraden und alle anderen Knoten (einschließlich A) einen geraden Grad. Nach IH(a) gibt es einen Eulerschen Weg in Γ' von C nach B. Zusammengesetzt mit e liefert dies wieder einen Eulerschen Weg in Γ von A nach B.

(b). Wir können annehmen, daß Γ mindestens eine Kante $e = \{A, B\}$ hat. Dann ist nach dem vorigen Lemma diese Kante keine Brücke, also der Graph $\Gamma' = (V, E \setminus \{e\})$ zusammenhängend. Es gilt

$$\begin{aligned} &\deg_{\Gamma'}(A) = \deg_{\Gamma}(A) - 1 \equiv 1 \mod 2, \\ &\deg_{\Gamma'}(B) = \deg_{\Gamma}(B) - 1 \equiv 1 \mod 2, \\ &\deg_{\Gamma'}(P) = \deg_{\Gamma}(P) \quad \equiv 0 \mod 2 \quad \text{für alle } P \in V \setminus \{A, B\}. \end{aligned}$$

Nach IH(a) gibt es in Γ' einen Eulerschen Weg von A nach B. Zusammengesetzt mit der Kante e liefert dies einen Eulerschen Zyklus in Γ. □

5.2.3. Hamiltonsche Wege.

DEFINITION. Ein *Hamiltonscher Weg* in einem Graphen $\Gamma = (V, E)$ ist ein Weg, in dem jeder Knoten $v \in V$ genau einmal vorkommt. Ein *Hamiltonscher Zyklus* ist ein Hamiltonscher Weg, der gleichzeitig ein Zyklus ist.

BEMERKUNG. Es ist kein effizienter Algorithmus zur Auffindung der Hamiltonschen Wege in einem endlichen Graphen bekannt. Effizient soll hier bedeuten. daß die Komplexität höchstens polynomial in der Anzahl der Kanten ist.

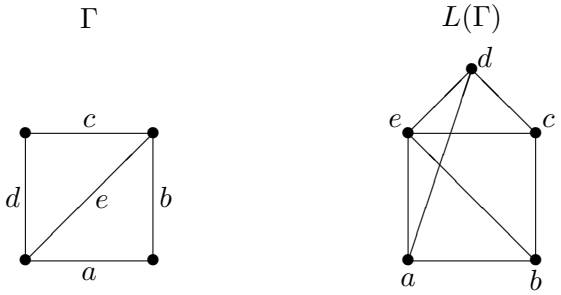
DEFINITION. Sei $\Gamma = (V, E)$ ein Graph. Unter dem *Kantengraph* (englisch line graph) $L(\Gamma)$ zu Γ versteht man den Graphen

$$L(\Gamma) = (E, E')$$

mit

$$E' := \{\, \{e, f\} \subseteq E \mid e \neq f \text{ und es gibt einen gemeinsamen Eckpunkt } A \text{ von } e \text{ und } f \text{ in } \Gamma \,\}$$

BEISPIEL.



LEMMA. *Sei Γ ein Graph, der einen Eulerschen Zyklus besitzt. Dann besitzt der zugehörige Kantengraph $L(\Gamma)$ einen Hamiltonschen Zyklus.*

BEWEIS. Sei

ein Eulerscher Zyklus in Γ. Dann sind nach Definition die Kanten e_i paarweise verschieden. Der zugehörige Hamiltonsche Zyklus in $L(\Gamma)$ verbindet die Knoten $e_1, \ldots, e_n, e_1$ von $L(\Gamma)$. Dies ist ein geschlossener Weg in $L(\Gamma)$, da $\{e_i, e_{i+1}\} \in E'$ und ebenso $\{e_n, e_1\} \in E'$. □

5.3. Abstände in bewerteten Graphen

Behandelt werden

- der Algorithmus von Moore zur Berechnung der Abstände der Knoten in einem Graphen von einem gegebenen Knoten v_0,
- eine Verallgemeinerung des Warshall-Algorithmus zur simultanen Bestimmung aller kürzesten Abstände in einem bewerteten Graphen $\Gamma = (V, E, d)$ (Komplexität $O(n^3)$),
- der Algorithmus von Dijkstra zur Bestimmung der kürzesten Abstände von einem festen Knoten v_0 in einem bewerteten Graphen $\Gamma = (V, E, d)$ (Komplexität $O(n^2)$).

5.3.1. Der Algorithmus von Moore. Wir beginnen mit einem einfachen Abstandsbegriff. Sei $\Gamma = (V, E)$ ein Graph und $A, B \in V$. Man setzt $d(A, B) := \infty$, wenn A und B in verschiedenen Zusammenhangskomponenten von Γ liegen, und $d(A, B) := k$, wenn A und B durch einen Weg mit k Kanten, aber durch keinen Weg mit weniger Kanten verbunden werden können. Insbesondere hat man also $d(A, A) = 0$ für alle $A \in V$.

Offenbar gilt dann

$$\begin{aligned} &d(A, B) = 0 \quad \text{genau dann, wenn} \quad A = B, \\ &d(A, B) = d(B, A) \quad \text{(Symmetrie)}, \\ &d(A, C) \leq d(A, B) + d(B, C) \quad \text{(Dreiecksungleichung)}. \end{aligned}$$

Wir verwenden sogenannte *Adjazenzmengen* oder *Adjazenzlisten* eines Graphen $\Gamma = (V, E)$. Darunter versteht man die Menge $A_\Gamma(v)$ aller Knoten $w \in V$ mit $\{v, w\} \in E$. Zur Adjazenzmatrix besteht ein einfacher Zusammenhang: Sei $M_\Gamma = (a_{ij}) \in M(n \times n, \{0, 1\})$ die Adjazenzmatrix von Γ (bzgl. einer vorgegebenen Numerierung der Knoten). Dann ist

$$\begin{aligned} A_\Gamma(v_i) &= \{\, v_j \mid a_{ij} = 1 \,\}, \\ \#(A_\Gamma(v_i)) &= \deg_\Gamma(v_i) = \text{Zeilensumme der } i\text{-ten Zeile von } M_\Gamma. \end{aligned}$$

Ferner ist $\sum_{v \in V} \#(A_\Gamma(v)) = \sum \deg_\Gamma(v) = 2 \cdot \#(E)$ nach der Euler-Formel.

Im *Algorithmus von Moore* benötigen wir einen Array $d(v)$, $v \in V$. Am Ende steht in $d(v)$ die Länge des kürzesten Weges von v_0 nach v. Weiter werden Mengen $W(k)$ konstruiert, so daß am Ende gilt $W(k) = \{\, v \in V \mid d(v) = k \,\}$. Nach dem k-ten Durchlauf der Schleife sind alle Knoten richtig bestimmt, die von v_0 einen Abstand $\leq k$ haben; dies zeigt man leicht durch Induktion über k.

5.3.2. Eine Verallgemeinerung des Warshall-Algorithmus. Den restlichen in diesem Abschnitt behandelten Algorithmen liegt ein feinerer Abstandsbegriff zugrunde, bei dem jeder Kante $e \in E$ des betrachteten

```
d(v_0) := 0; d(v) := ∞ für v ∈ V \ {v_0};
W(0) := {v_0};
for k := 1 to n do (n := #(V))
   if W(k − 1) = ∅ then return end;
   W(k) := ∅;
   for v ∈ ⋃{ A(w) | w ∈ W(k − 1) } do
      if d(v) = ∞ then
         d(v) := k;
         W(k) := W(k) ∪ {v};
         end;
      end;
   end;
end;
```

Abbildung 4. Der Algorithmus von Moore

Graphen $\Gamma = (V, E)$ eine nicht negative rationale Zahl oder ∞ als Abstand $d(e)$ zugeordnet ist.

Definition. Ein *bewerteter Graph* ist ein Tripel (V, E, d), wobei (V, E) ein (ungerichteter) Graph ist und d eine Funktion $d\colon E \to \mathbb{Q}^+ := \{\, r \in \mathbb{Q} \mid r \geq 0 \,\}$; d heißt *Bewertungsfunktion*. Wir erweitern d auf alle $\{u, v\} \notin E$ mit $u \neq v$, indem wir setzen

$$d(v_i, v_j) := \infty \quad \text{für } \{v_i, v_j\} \notin E \text{ mit } i \neq j.$$

Die *Länge* eines Weges $\alpha = (v_0, \ldots, v_n)$ ist

$$|\alpha| := \sum_{i<n} d(v_i, v_{i+1}).$$

Um den Warshall-Algorithmus und seine Analyse auf diese Situation verallgemeinern zu können, führen wir zunächst eine Halbringstruktur auf

$$X := \mathbb{Q}^+ \cup \{\infty\} = \{\, r \in \mathbb{Q} \mid r \geq 0 \,\} \cup \{\infty\}$$

ein. $(X, \vee, \wedge)$ wird zu einem Halbring mit Einselement 0 und Nullelement ∞, wenn man $\vee, \wedge\colon X \times X \to X$ definiert durch

$$r \vee s := \min(r, s) := \begin{cases} \text{übl. Min. in } \mathbb{Q} & \text{falls } r, s \in \mathbb{Q} \\ r & \text{falls } r \in \mathbb{Q}, s = \infty \\ s & \text{falls } s \in \mathbb{Q}, r = \infty \\ \infty & \text{falls } r = s = \infty, \end{cases}$$

$$r \wedge s := r + s := \begin{cases} \text{übl. Summe in } \mathbb{Q} & \text{falls } r, s \in \mathbb{Q} \\ \infty & \text{falls } r = \infty \text{ oder } s = \infty. \end{cases}$$

Dies läßt sich leicht beweisen.

Einem bewerteten Graphen $\Gamma = (V, E, d)$ mit $d\colon E \to \mathbb{Q}^+ \cup \{\infty\}$ und $V = \{v_1, \ldots, v_n\}$ ordnen wir eine Matrix $D = (d_{ij}) \in M(n \times n, \mathbb{Q}^+ \cup \{\infty\})$ zu durch

$$d_{ij} := \begin{cases} 0 & \text{für } i = j, \\ d(v_i, v_j) & \text{für } \{v_i, v_j\} \in E, \\ \infty & \text{sonst.} \end{cases}$$

Wir geben zunächst eine graphentheoretische Interpretation der Potenzen D^k der Matrix D über dem Halbring $(\mathbb{Q}^+ \cup \{\infty\}, \min, +)$.

LEMMA. *Sei D die dem bewerteten Graphen $\Gamma = (V, E, d)$ zugeordnete Matrix und $D^k = (a_{ij})$ die k-te Potenz von D bzgl. der Matrizenmultiplikation über dem Halbring $(\mathbb{Q}^+ \cup \{\infty\}, \min, +)$. Dann ist a_{ij} das Minimum der Längen aller Wege von v_i nach v_j mit höchstens k Kanten ($k \geq 1$).*

BEWEIS. Induktion über k. Der Induktionsanfang $k = 1$ ist trivial. Im Schritt $k-1 \mapsto k$ haben wir $D^{k-1} = (b_{ij})$ und $D^k = D^{k-1} \cdot D = (a_{ij})$. Dann gilt

$$a_{ij} = \bigvee_{\nu=1,\ldots,n} (b_{i\nu} \wedge d_{\nu j}) = \min_{\nu=1,\ldots,n} (b_{i\nu} + d_{\nu j}).$$

Offensichtlich ist also a_{ij} die Länge des kürzesten Weges α von v_i nach v_j mit höchstens k Kanten. □

LEMMA. *Der folgende* verallgemeinerte Warshall-Algorithmus *verändert die eben angegebene Matrix D so, daß am Ende der Matrix-Eintrag d_{jk} die kürzeste Länge eines Weges von v_j nach v_k ist:*

```
for i := 1 to n do
  for j := 1 to n do
    for k := 1 to n do
      d_jk := min(d_jk, d_ji + d_ik)
    end;
  end;
end;
```

Die Komplexität ist $O(n^3)$.

BEWEIS. Sei $D^{(i)}$ der Inhalt des Speicherplatzes für D nach dem i-ten Durchlauf der äußeren Schleife. Wir zeigen durch Induktion über i, daß $d_{jk}^{(i)}$ die Länge des kürzesten i-Weges (d.h. Weg mit Zwischenpunkten aus $\{v_1, \ldots, v_i\}$) von v_j nach v_k ist.

Nach Annahme ist die Behauptung für $i = 0$ richtig. *Schritt* $i - 1 \mapsto i$: Aufgrund des Algorithmus haben wir

$$d_{jk}^{(i)} = \min(d_{jk}^{(i-1)}, d_{ji}^{(i-1)} + d_{ik}^{(i-1)}).$$

für alle j, i, k. Man betrachte einen i-Weg $\alpha = v_j A_1 \dots A_n v_k$ von v_j nach v_k, der unter allen solchen i-Wegen kürzeste Länge hat. *Fall* 1. v_i kommt in $A_1 \dots A_n$ nicht vor. Dann ist

$$\begin{aligned} d_{jk}^{(i)} &= \min(d_{jk}^{(i-1)}, d_{ji}^{(i-1)} + d_{ik}^{(i-1)}) \\ &= d_{jk}^{(i-1)} && \text{nach Annahme} \\ &= |\alpha| && \text{nach IH.} \end{aligned}$$

Fall 2. v_i kommt in $A_1 \dots A_n$ vor. Wir können annehmen, daß v_i genau einmal in $A_1 \dots A_n$ vorkommt. Dann hat α die Form

$$\alpha = v_j A_1 \dots A_{m-1} v_i A_{m+1} \dots A_n v_j$$

mit $A_1, \dots, A_{m-1}, A_{m+1}, \dots, A_n \in \{v_1, \dots, v_{i-1}\}$, und nach der IH ist dann $d_{ji}^{(i-1)} = |(v_j, A_1, \dots, A_{m-1}, v_i)|$ und $d_{ik}^{(i-1)} = |(v_i, A_{m+1}, \dots, A_n, v_k)|$. Es folgt

$$\begin{aligned} d_{jk}^{(i)} &= \min(d_{jk}^{(i-1)}, d_{ji}^{(i-1)} + d_{ik}^{(i-1)}) \\ &= d_{ji}^{(i-1)} + d_{ik}^{(i-1)} && \text{nach Annahme} \\ &= |\alpha| && \text{nach IH.} \end{aligned}$$

Damit haben wir die Korrektheit des verallgemeinerten Warshall-Algorithmus gezeigt; seine Komplexität ist offenbar wieder $O(n^3)$. □

5.3.3. Dijkstras Algorithmus. Wir behandeln jetzt den Dijkstra-Algorithmus zur Bestimmung der Abstände von einem festen Knoten. Gegeben ist ein bewerteter Graph $\Gamma = (V, E, w)$ mit $w\colon E \to \mathbb{Q}^+ \cup \{\infty\}$; wir setzen wieder $w(u, v) := \infty$ falls $u \neq v$ und $\{u, v\} \notin E$.

Gesucht ist eine Abbildung $d\colon V \to \mathbb{Q}^+ \cup \{\infty\}$ so daß $d(v)$ die Länge des kürzesten Weges von v_0 nach v ist.

$d(v_0) := 0$; $d(v) := \infty$ für $v \in V \setminus \{v_0\}$;
$U := V$ (= set of undone nodes);
while $U \neq \emptyset$ do
  $\langle$pick $u \in U$ such that $d(u) = \min_{v \in U} d(v)\rangle$;
  for $v \in U$ do
    $d(v) := \min(d(v), d(u) + w(u, v))$;
  end;
  $U := U \setminus \{u\}$;
end;

Die Komplexität ist offenbar $O(n^2)$.

Wir zeigen jetzt die Korrektheit dieses Dijkstra-Algorithmus. Es werden der Reihe nach Knoten $v_0 = u_1, u_2, \ldots, u_n$ aus U entfernt. Für $v \in V$ sei

$\bar{d}(v) :=$ tatsächlicher kürzester Abstand von v_0 nach v,

$d(v) :=$ Ergebnis des Dijkstra-Algorithmus,

$d^{(i)}(v) :=$ Inhalt von $d(v)$ nach dem i-tem Durchlauf der while-Schleife.

Man beachte, daß offenbar gilt

(5.1)
$$d^{(0)}(u_i) \geq \cdots \geq d^{(i)}(u_i) = d^{(i+1)}(u_i) = \cdots = d^{(n)}(u_i) = d(u_i) \geq \bar{d}(u_i),$$

(5.2)
$$d^{(k)}(u_k) + w(u_k, u_{i+1}) \geq d^{(k)}(u_{i+1}) \quad \text{für } k \leq i.$$

(5.1) ist klar, und (5.2) sieht man wie folgt. Im k-ten Durchlauf wird erst u_k ausgewählt und dann für u_k und alle anderen unerledigten Knoten v – insbesondere u_{i+1} – die $d^{(k)}$-Abstände so festgelegt, daß

$$d^{(k)}(v) = \min\{d^{(k-1)}(v),\, d^{(k)}(u_k) + w(u_k, v)\} \leq d^{(k)}(u_k) + w(u_k, v).$$

Wir zeigen

$$\bar{d}(u_i) \geq d(u_i)$$

durch Induktion nach i; wegen $\bar{d}(u_i) \leq d(u_i)$ folgt daraus die Behauptung.

Beweis. *Basis* $i = 1$. Es ist $u_1 = v_0$ und $d(v_0) = \bar{d}(v_0) = 0$. *Schritt* $\{\, k \mid k \leq i \,\} \mapsto i + 1$. Sei $u_1 = v_0, v_1, \ldots, v_m, u_{i+1}$ ein kürzester Weg von v_0 nach u_{i+1}.

Fall 1. Alle inneren Knoten $v_1, \ldots, v_m$ liegen in $\{u_1, \ldots, u_i\}$. Sei etwa $v_m = u_k$ mit $k \leq i$. Dann gilt

$$\begin{aligned}
\bar{d}(u_{i+1}) &= \bar{d}(u_k) + w(u_k, u_{i+1}) \\
&= d(u_k) + w(u_k, u_{i+1}) && \text{nach IH für } k \\
&= d^{(k)}(u_k) + w(u_k, u_{i+1}) && \text{nach (5.1)} \\
&\geq d^{(k)}(u_{i+1}) && \text{wegen } k \leq i \text{ nach (5.2)} \\
&\geq d(u_{i+1}) && \text{nach (5.1).}
\end{aligned}$$

Fall 2. Seien $v_1, \ldots, v_j \in \{u_1, \ldots, u_i\}$, aber $v_{j+1} \notin \{u_1, \ldots, u_i\}$. Sei etwa $v_j = u_k$ mit $k \leq i$. Annahme: $\bar{d}(u_{i+1}) < d(u_{i+1})$. Dann erhält man

$$\begin{aligned}
d(u_{i+1}) &> \bar{d}(u_{i+1}) \\
&= \bar{d}(u_k) + w(u_k, v_{j+1}) + \ldots \\
&= d(u_k) + w(u_k, v_{j+1}) + \ldots && \text{nach IH für } k
\end{aligned}$$

$$
\begin{aligned}
&\geq d(u_k) + w(u_k, v_{j+1}) && \\
&= d^{(k)}(u_k) + w(u_k, v_{j+1}) && \text{nach (5.1)} \\
&\geq d^{(k)}(v_{j+1}) && \text{nach (5.2) wegen } k \leq i \\
& && \text{und } v_{j+1} \notin \{u_1, \ldots, u_i\} \\
&\geq d^{(i+1)}(v_{j+1}) && \text{wegen } k \leq i \text{ nach (5.1).}
\end{aligned}
$$

Also ist $d(u_{i+1}) = d^{(i+1)}(u_{i+1}) > d^{(i+1)}(v_{j+1})$. Mit $v_{j+1} \notin \{u_1, \ldots, u_i\}$ folgt, daß v_{j+1} vor u_{i+1} aus U hätte entnommen werden müssen, und wir haben einen Widerspruch. □

Tabelle 1 enthält ein Beispiel für den Ablauf des Dijkstra-Algorithmus.

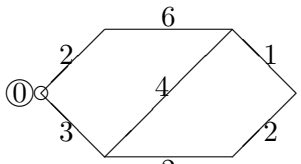
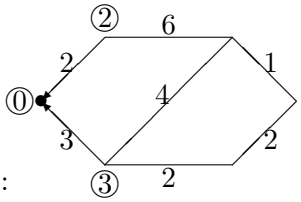
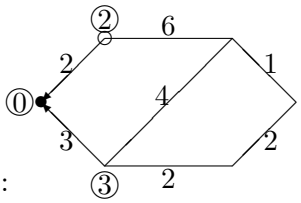
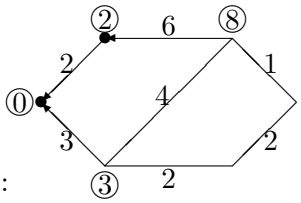
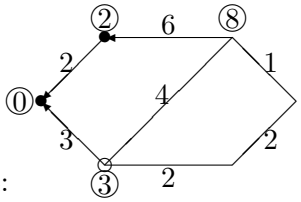
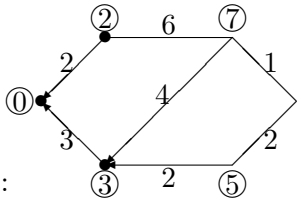
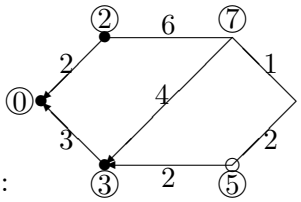
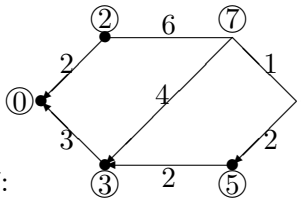
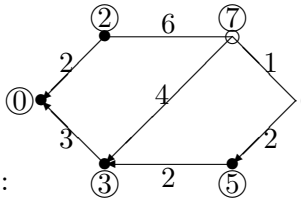
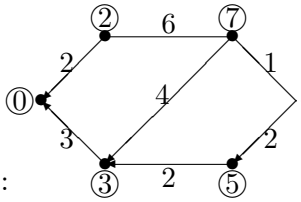
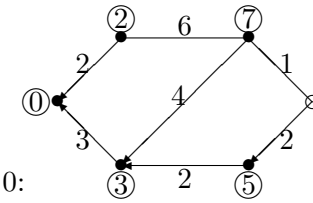
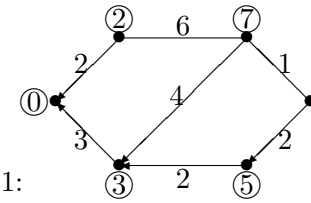
Auswahl	Aktualisierung
 $i = 0$:	 $i = 1$:
 $i = 2$:	 $i = 3$:
 $i = 4$:	 $i = 5$:
 $i = 6$:	 $i = 7$:
 $i = 8$:	 $i = 9$:
 $i = 10$:	 $i = 11$:

TABELLE 1. Ein Beispiel für den Dijkstra-Algorithmus

KAPITEL 6

Bäume

Ein (ungerichteter) Graph G heißt Baum, wenn er azyklisch und zusammenhängend ist. Dies ist äquivalent damit, daß je zwei Knoten durch einen eindeutig bestimmten Weg verbunden werden können, oder auch damit, daß die Anzahl der Kanten um eins größer als die Anzahl der Knoten ist.

Wir betrachten wieder bewertete Graphen, bei denen jeder Kante eine feste rationale Zahl (oder ∞) als Kantenlänge zuordnet ist. Der Kruskal-Algorithmus bestimmt einen minimalen aufspannenden Baum zu einem gegebenen zusammenhängenden bewerteten Graphen. Im Beweis verwenden wir das Austauschlemma der Graphentheorie, das eine einfache Folgerung aus dem Austauschlemma der linearen Algebra ist.

6.1. Allgemeine Begriffe

Ein Graph heißt *azyklisch*, wenn er keine Zykeln besitzt. Ein *Baum* ist ein zusammenhängender azyklischer Graph. Ein azyklischer Graph heißt ein *Wald*. Ein Graph ist also genau dann ein Wald, wenn alle seine Zusammenhangskomponenten Bäume sind.

Lemma (Kantenabschätzung). *Für jeden endlichen zusammenhängenden Graphen* $\Gamma = (V, E)$ *gilt* $\#(E) \geq \#(V) - 1$.

Beweis. Induktion über $\#(E) =: n$. *Basis* $n = 0$. Trivial. *Schritt* $n \mapsto n + 1$. Sei e eine Kante, $E' := E \setminus \{e\}$ und $\Gamma' := (V, E')$.

Fall Γ' zusammenhängend. Nach IH ist dann $\#(E') \geq \#(V) - 1$, also auch $\#(E) = \#(E') + 1 \geq \#(V) - 1$.

Fall Γ' nicht zusammenhängend. Dann besteht Γ' aus höchstens zwei Zusammenhangskomponenten, etwa $\Gamma_1 = (V_1, E_1)$ und $\Gamma_2 = (V_2, E_2)$, und es ist $V = V_1 \cup V_2$ und $E \setminus \{e\} = E_1 \cup E_2$. Auf die beiden Teilgraphen kann man die IH anwenden und erhält $\#(E_i) \geq \#(V_i) - 1$, also $\#(E) = \#(E_1) + \#(E_2) + 1 \geq \#(V_1) + \#(V_2) - 1 = \#(V) - 1$. □

Satz (Charakterisierung von Bäumen). *Für einen endlichen zusammenhängenden Graphen* $\Gamma = (V, E)$ *sind die folgenden Aussagen äquivalent.*

(a) Γ *ist ein Baum (d.h.,* Γ *ist azyklisch).*

(b) *Je zwei Knoten in* Γ *sind durch einen eindeutig bestimmten Weg verbindbar.*
(c) $\#(E) = \#(V) - 1$.

BEWEIS. (a) $\Rightarrow$ (b). Gelte (a); wir nehmen an, daß (b) falsch ist. Dann gibt es zwei verschiedene Wege

$$v_1, w_1, w_2, \ldots, w_n, v_2 \quad \text{und} \quad v_1, u_1, u_2, \ldots, u_m, v_2.$$

Sei i_0 maximaler Index mit $w_1 = u_1, w_2 = u_2, \ldots, w_{i_0} = u_{i_0}$. Dann ist

$$v_2, w_n, w_{n-1}, \ldots, w_{i_0} = u_{i_0}, \ldots, u_{m-1}, u_m, v_2$$

ein Zyklus.

(b) $\Rightarrow$ (a). Gelte (b); wir nehmen an, daß (a) falsch ist. Es gibt also einen Zyklus $w_1, w_2, \ldots, w_n = w_1$ mit $n \geq 4$. Dann gibt es aber zwei verschiedene Wege von w_1 nach w_2, nämlich w_1, w_2 und $w_1 = w_n, w_{n-1}, \ldots, w_3, w_2$.

(a) $\Rightarrow$ (c). Sei also $\Gamma = (V, E)$ ein Baum mit $\#(V) = n$. Wir zeigen $\#(E) = n - 1$ durch Induktion über n. Für $n = 1, 2$ ist die Behauptung trivial. *Schritt* $n \mapsto n+1$. Durch Entfernen einer Kante aus $\Gamma = (V, E)$ bilde man $\Gamma' = (V, E \setminus \{e\})$. Dann ist Γ' nicht zusammenhängend, denn andernfalls gäbe es einen Zyklus in Γ. Da Γ zusammenhängend ist, zerfällt Γ' in zwei Zusammenhangskomponenten $\Gamma_1 = (V_1, E_1)$ und $\Gamma_2 = (V_2, E_2)$, und es ist $V = V_1 \cup V_2$ und $E \setminus \{e\} = E_1 \cup E_2$. Auf die beiden Teilgraphen, die wieder Bäume sind, kann man die IH anwenden und erhält $\#(E_i) = \#(V_i) - 1$, also $\#(E) = \#(E_1) + \#(E_2) + 1 = \#(V_1) + \#(V_2) - 1 = \#(V) - 1$.

(c) $\Rightarrow$ (a). Gelte $\#(E) = \#(V) - 1$. Annahme: Γ ist kein Baum. Dann gibt es einen Zyklus in Γ. Wir entfernen eine Kante e aus diesem Zyklus; sei $\Gamma' := (V, E')$ mit $E' := E \setminus \{e\}$). Dann ist Γ' zusammenhängend. Nach dem Kantenabschätzungslemma gilt $\#(E') \geq \#(V) - 1$, also $\#(E) = \#(E') + 1 \geq \#(V) - 1 + 1 = \#(V)$, also $\#(E) \neq \#(V) - 1$. □

DEFINITION. Ein *Wurzelbaum* $\Gamma = (V, E, r)$ besteht aus einem Baum (V, E) mit einem ausgezeichneten Knoten r, der *Wurzel* genannt wird.

Man beachte, daß es zu einem gegebenen Baum im allgemeinen mehrere nicht isomorphe Wurzelbäume gibt. Dabei heißen zwei Wurzelbäume $\Gamma = (V, E, r)$ und $\Gamma' = (V', E', r')$ isomorph, wenn es einen Graphisomorphismus $\varphi\colon V \to V'$ gibt (d.h., φ induziert eine Bijektion von E auf E'), so daß $\varphi(r) = r'$.

DEFINITION. Sei $\Gamma = (V, E, r)$ ein Wurzelbaum.

(a) Für jeden Knoten $v \in V$ wird die *Tiefe* definiert als die Länge des eindeutig bestimmten Weges zur Wurzel minus 1. Unter der *Höhe* oder *Tiefe* $\mathrm{dp}(\Gamma)$ des Wurzelbaums versteht man die maximale Tiefe eines Knotens.

(b) Die *partielle Ordnung* auf Γ ist erklärt durch $v_1 \leq v_2$ genau dann, wenn v_1 auf dem eindeutig bestimmten Weg von der Wurzel nach v_2 liegt.
(c) Ein Knoten $v \in V$ heißt *Blatt*, wenn v maximales Element der in (b) definierten partiellen Ordnung ist. Die übrigen Knoten heißen *innere Knoten*. Die Blätterzahl (engl. leaf size) $\mathsf{ls}(\Gamma)$ von Γ ist die Anzahl der Blätter des Wurzelbaums.
(d) Sei $v \in V$ ein Knoten. Ein Knoten $v_1 \in V$ heißt *Kind* von v, wenn $v \leq v_1$ und $\{v, v_1\}$ eine Kante des Baums ist. v heißt dann *Elternteil* von v_1. Zwei Knoten heißen *Geschwister* (engl. siblings), wenn sie denselben Elternteil haben. Γ heißt *höchstens k-fach verzweigt* (*k-fach verzweigt*), wenn jeder innere Knoten höchstens k (genau k) Nachfolger hat.

Die *Länge* oder *Größe* eines endlichen Graphen $\Gamma = (V, E)$ ist die Anzahl $\#(V)$ der Knoten des Graphen. Wir schreiben $\|\Gamma\|$ für die Größe von Γ.

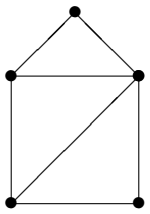
Sei Γ ein höchstens k-fach verzweigter endlicher Wurzelbaum, d.h., jeder Knoten hat höchstens k unmittelbare Nachfolger. Dann gilt offenbar

(a) $\|\Gamma\| \leq k^{\mathrm{dp}(\Gamma)+1} - 1$ für $k \geq 2$;
(b) $\|\Gamma\| = 2\mathsf{ls}(\Gamma) - 1$ für genau 2-fach verzweigte Wurzelbäume;
(c) $\mathsf{ls}(\Gamma) \leq \|\Gamma\|$.

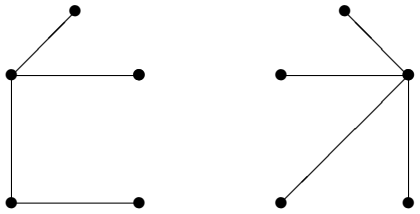
6.2. Aufspannende Bäume in Graphen: der Kruskal-Algorithmus

Definition. Sei $\Gamma = (V, E)$ ein zusammenhängender Graph. Ein *aufspannender Baum* von Γ ist ein Baum $T = (V, B)$ mit $B \subseteq E$, der dieselbe Knotenmenge wie Γ hat. Ist $\Gamma = (V, E)$ ein nicht notwendig zusammenhängender Graph, so versteht man unter einem *aufspannenden Wald* von Γ einen Wald $T = (V, B)$ mit $B \subseteq E$, der dieselbe Knotenmenge wie Γ hat.

Beispiel. Man betrachte den Graphen



Aufspannende Bäume sind etwa



DEFINITION. Sei $\Gamma = (V, E, w)$ ein endlicher zusammenhängender Graph mit Bewertungsfunktion $w\colon E \to \mathbb{Q}^+$. Ein *minimal aufspannender Baum* ist ein aufspannender Baum $T = (V, B)$ mit $w^*(T) := \sum_{b\in B} w(b) \leq w^*(T')$ für alle aufspannenden Bäume $T' = (V, B')$.

Aus dem nächsten Lemma wird folgen, daß der gleich anzugebende Kruskal-Algorithmus einen aufspannenden Baum liefert.

LEMMA. *Sei $\Gamma = (V, E)$ ein endlicher zusammenhängender Graph und $\mathcal{F} \subseteq \mathcal{P}(E)$ die Menge aller azyklischen Teilmengen von E. Offenbar ist $\mathcal{F}$ durch $\subseteq$ partiell geordnet. Sei $F_0 \in \mathcal{F}$ maximales Element bzgl. $\subseteq$. Dann ist (V, F_0) ein aufspannender Baum.*

BEWEIS. Sei F_0 ein maximales Element und nehmen wir an, daß (V, F_0) kein aufspannender Baum ist. Dann gibt es einen Knoten $v_0 \in V$, der auf keiner Kante von F_0 liegt. Da (V, E) zusammenhängend ist, gibt es eine Kante $e = \{v_0, v\} \in E$ (also $e \notin F_0$). Setze $F_1 := F_0 \cup \{e\}$. F_1 ist azyklisch, denn da F_0 nach Annahme azyklisch ist, müßte ein Zyklus e enthalten und deshalb läge v_0 doch auf einer Kante von F_0. Damit wäre aber F_0 nicht maximal. □

Wir zeigen jetzt das Austauschlemma der Graphentheorie; es ist eine Folgerung aus dem bekannten Austauschlemma der linearen Algebra.

LEMMA (Austauschlemma der linearen Algebra). *Sei V ein Vektorraum über einem Körper K. Sei $v_1, \ldots, v_k$ ein k-Tupel linear unabhängiger Vektoren aus V und $w_1, \ldots, w_l$ ein l-Tupel linear unabhängiger Vektoren aus V mit $l > k$. Dann gibt es einen Index $j \in \{1, \ldots, l\}$ so daß $v_1, \ldots, v_k, w_j$ linear unabhängig ist.*

LEMMA (Austauschlemma der Graphentheorie). *Sei $\Gamma = (V, E)$ ein endlicher zusammenhängender Graph. Ferner seien $F_0 \subseteq E$ und $F_1 \subseteq E$ azyklische Teilmengen mit $\#(F_0) < \#(F_1)$. Dann gibt es eine Kante $f \in F_1 \setminus F_0$ so daß $F_0 \cup \{f\}$ azyklisch ist.*

BEWEIS. Sei $V = \{v_1, \ldots, v_n\}$ und $E = \{e_1, \ldots, e_m\}$. Die Inzidenzmatrix $\mathrm{IM}_\Gamma = (a_{ij}) \in M(n \times m, \{0,1\})$ von Γ ist gegeben durch

$$a_{ij} := \begin{cases} 1 & \text{falls } v_i \in e_j, \\ 0 & \text{sonst.} \end{cases}$$

Die Koeffizienten der Matrix seien als Elemente des Körpers $\mathbb{F}_2 = \{0,1\}$ betrachtet. Jede Spalte der Matrix $\mathrm{IM}(\Gamma)$ hat genau zwei Komponenten mit dem Wert 1, alle andere Komponenten sind $= 0$.

Wir zeigen, daß ein k-Tupel $e_{j_1}, \ldots, e_{j_k} \in E$ genau dann einen Zyklus enthält, wenn die zugehörigen Spaltenvektoren $s_{j_1}, \ldots, s_{j_k} \in \mathbb{F}_2^n$ linear abhängig sind. Die Behauptung folgt dann aus dem Austauschlemma der linearen Algebra.

$\Longrightarrow$. Nach einer eventuellen Umnumerierung können wir annehmen, daß $e_{j_1}, \ldots, e_{j_p}$ einen Zyklus bilden. Also gibt es Knoten $w_0, w_1, \ldots, w_p = w_0$ mit $e_{j_i} = \{w_{i-1}, w_i\}$. Dann ist aber $s_{j_1} + \cdots + s_{j_p} = 0$, da wir in $\mathbb{F}_2$ arbeiten: In der w_i entsprechenden Zeile von IM_Γ steht im Fall $1 \leq i < p$ genau in den Spalten s_{j_i} und $s_{j_{i+1}}$ eine 1, und im Fall $i = p$ genau in den Spalten s_{j_p} und s_{j_1} eine 1.

$\Longleftarrow$. Sind $s_{j_1}, \ldots, s_{j_k} \in \mathbb{F}_2^n$ linear abhängig, so gibt es eine nichttriviale Linearkombination, die den Nullvektor darstellt. Man wähle eine minimale nichttriviale Darstellung des Nullvektors. Nach eventueller Umnumerierung ist dies $s_{j_1} + \cdots + s_{j_p} = 0$. Dann enthält die Kantenmenge $e_{j_1}, \ldots, e_{j_p}$ einen Zyklus (wegen der Minimalität ist sie sogar ein Zyklus). □

Insbesondere ergibt sich: Sind $F_0, F_1 \subseteq E$ azyklische Teilmengen mit $\#(F_0) = \#(F_1)$, so gibt es zu jeder Kante $e \in F_0$ eine Kante $f \in F_1$ derart daß $(F_0 \setminus \{e\}) \cup \{f\}$ azyklisch ist.

SATZ (Kruskal-Algorithmus). *Der folgende Algorithmus bestimmt einen minimalen aufspannenden Baum in einem endlichen zusammenhängenden Graphen $\Gamma = (V, E, w)$ mit Bewertungsfunktion w.*

1. *Ordne die Kanten nach aufsteigendem Gewicht: $w(e_1) \leq w(e_2) \leq \cdots \leq w(e_n)$.*
2. *Aufbau des Baumes:*

```
F:=0;
for i := 1 to n do
   if (F ∪ {e_i} is acyclic) then
      F := F ∪ {e_i}
   end;
end;
```

BEWEIS. Offenbar erzeugt der Kruskal-Algorithmus eine maximale azyklische Teilmenge von E, und nach dem obigen Lemma ist eine maximale azyklische Teilmenge eines zusammenhängenden Graphen ein aufspannender Baum. Zu zeigen bleibt, daß dieser Baum minimales Gewicht hat, d.h., daß es keinen aufspannenden Baum von kleinerem Gewicht gibt.

Der vom Kruskal-Algorithmus gelieferte Baum habe etwa die Kanten $f_1, \ldots, f_k$ mit $w(f_1) \leq w(f_2) \leq \cdots \leq w(f_k)$. Wir betrachten jetzt einen beliebigen aufspannenden Baum; er muß ebenfalls die Kantenzahl $k = \#(V)-1$ haben. Seine Kanten seien etwa $g_1, \ldots, g_k$ mit $w(g_1) \leq w(g_2) \leq \cdots \leq w(g_k)$. Es reicht zu zeigen $\sum_{\nu<i} w(f_\nu) \leq \sum_{\nu<i} w(g_\nu)$ für alle $i \leq k$; für $i = k$ folgt daraus die Behauptung. Beweis durch Induktion über i. Basis $i = 1$. Es ist $f_1 = e_1$, also $w(f_1) \leq w(e_j)$ für $1 \leq j \leq k$. Schritt $i > 1$. Zu $\{f_1, \ldots, f_{i-1}\}$ und $\{g_1, \ldots, g_i\}$ (beide azyklisch) gibt es nach dem Austauschlemma ein $j \leq i$ mit $g_j \notin \{f_1, \ldots, f_{i-1}\}$ und $\{f_1, \ldots, f_{i-1}, g_j\}$ azyklisch. Nach Wahl von f_i im Algorithmus ist $w(f_i) \leq w(g_j)$, also

$$\sum_{\nu\leq i} w(f_\nu) = \sum_{\nu<i} w(f_\nu) + w(f_i) \leq \sum_{\nu<i} w(g_\nu) + w(g_j) \leq \sum_{\nu\leq i} w(g_\nu). \qquad \square$$

BEMERKUNG. Wie entscheidet man im Kruskal-Algorithmus effizient, ob durch Hinzufügen einer Kante ein Zyklus entsteht? Man beachte dazu, daß bei der Ausführung des Kruskal-Algorithmus durch fortgesetztes Hinzufügen von Kanten stets Wälder entstehen. Sei F ein solcher Zwischenwald, und e eine Kante, die nicht zu F gehört. F besteht aus endlich vielen Zusammenhangskomponenten $T_1, \ldots, T_n$, die Bäume sind. Es gilt dann: $F \cup \{e\}$ enthält einen Zyklus genau dann, wenn beide Eckpunkte von e in demselben T_i liegen.

Literaturverzeichnis

O. Deiser. *Einführung in die Mengenlehre*. Springer Verlag, Berlin, Heidelberg, New York, 2nd edition, 2004.

R. Diestel. *Graphentheorie*. Springer Verlag, Berlin, Heidelberg, New York, 2nd edition, 2000.

G. Gentzen. Untersuchungen über das logische Schließen. *Mathematische Zeitschrift*, 39:176–210, 405–431, 1934.

J. K. Truss. *Discrete Mathematics for Computer Scientists*. Addison Wesley, 2nd edition, 1999.

Index