Skript zur Vorlesung
Grundlagen und Einzelfragen der Mathematik
für Lehramtskandidaten

Wintersemester 2009/2010

Prof. Dr. Helmut Maier

19.10.2009

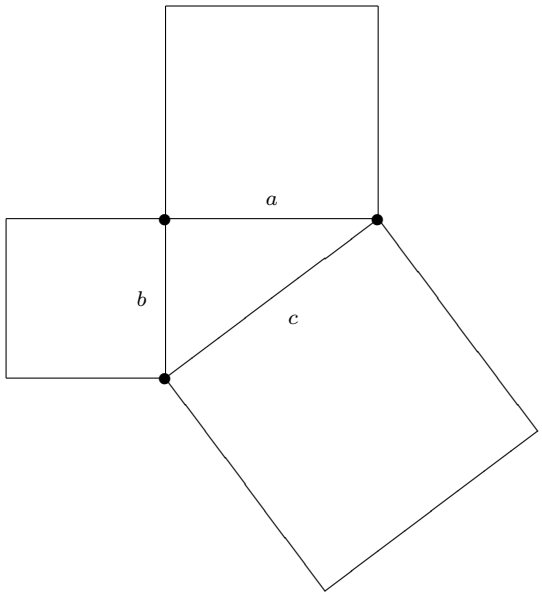
Inhaltsverzeichnis

Einleitung

Die Mathematik gilt als die exakteste aller Wissenschaften. Heute liegt jeder mathematischen Theorie ein System von (meist wenigen) Axiomen zugrunde. Dabei handelt es sich um nicht beweisbare Aussagen, aus denen dann die übrigen Aussagen - die Lehrsätze - durch eine Kette von logischen Schlüssen - den Beweisen - gefolgert werden.

Diese Auffassung der Mathematik wurde zuerst von den alten Griechen im Teilgebiet der Geometrie entwickelt. Zuvor waren geometrische Gegenstände, wie Geraden, Kreise, Ebenen usw. als Idealisierungen von in der Natur vorkommenden Objekten entstanden, und Aussagen darüber als Erfahrungstatsachen betrachtet worden.

Als Beispiel kann der Satz von Pythagoras dienen: im rechtwinkligen Dreieck ist das Quadrat über der Hypothenuse gleich der Summe der Quadrate über den Katheten.



Dieser Satz wurde ursprünglich wohl einfach durch Abmessung zahlreicher rechtwinkliger Dreiecke entdeckt. Die Axiomatisierung der Geometrie wurde systematisch zuerst wohl von Euklid in seinen Elementen (um 300 v.C.) vorgenommen. Auch in seiner Anschauung waren die Axiome (nicht mehr die Lehrsätze, die aus ihnen hergeleitet wurden) unmittelbar einleuchtende Aussagen über die Natur. Die moderne Auffassung von der Geometrie als einer von der Natur völlig unabhängigen Theorie kam erst mit Hilbert (nachdem im 19. Jahrhundert die Nichteuklidischen Geometrien entwickelt worden waren) im 20. Jahrhundert voll zum Durchbruch: man muss statt Geraden, Ebenen, Punkte stets auch Stühle, Tische, Bierseidel sagen können.

Die axiomatische Grundlegung anderer mathematischer Theorien erfolgte zum Teil erst viel später, beispielsweise in der Wahrscheinlichkeitstheorie erst ab 1933 durch Kolmogoroff.

Der strenge Aufbau der Analysis wurde von Cauchy um 1820 begonnen. Die grundlegenden Objekte der Analysis, die Menge der reellen Zahlen und die darauf definierten Rechenoperationen, erhielten jedoch erst gegen Ende des 19. Jahrhunderts eine axiomatische Theorie - hauptsächlich durch Dedekind. In unserem ersten Kapitel geben wir diesen Aufbau wieder.

Kapitel 1

Grundlagen der Analysis

1.1 Die natürlichen Zahlen, Peano-Axiome

Die einfachsten Zahlen, die auch Kinder als erstes lernen, sind die natürlichen Zahlen. Peano hat für sie ein Axiomensystem gegeben, das wir in diesem Abschnitt vorstellen.

Definition 1.1.1. $(\mathbb{N}, 0, S)$ heißt ein System natürlicher Zahlen, falls gilt:

(P1) $0 \in \mathbb{N}$,

(P2) $S : \mathbb{N} \to \mathbb{N}$, $n \mapsto S(n) = n'$ ist eine Abbildung (die Nachfolgerabbildung) mit

(P3) 0 ist kein Nachfolger: es gibt kein $n \in \mathbb{N}$, so dass $0 = n'$,

(P4) die Nachfolgerabbildung ist injektiv,

(P5) $\mathbb{N}$ besitzt keine echte induktive Teilmenge, d. h. $A \subseteq \mathbb{N}$ induktiv $\Rightarrow A = \mathbb{N}$.

Definition 1.1.2. Für $A \subseteq \mathbb{N}$ sei $A' = \{k' \mid k \in A\}$. Eine Teilmenge $A \subseteq \mathbb{N}$ heißt induktiv, falls die Induktionsverankerung $0 \in A$ und der Induktionsschritt $A' \subseteq A$ gelten.

Definition 1.1.3. Wir schreiben $1 := 0'$, $2 := 1'$, $3 := 2'$, etc.

Satz 1.1.4. *Es sei* $(\mathbb{N}, 0, S)$ *ein System natürlicher Zahlen, dann gilt:*

(i) $\mathbb{N}' + \{0\} = \mathbb{N}$ *(disjunkte Vereinigung),*

(ii) $\forall n \in \mathbb{N} : n' \neq n$, *speziell also* $1 \neq 0$.

Beweis.
Zu (i): Die Vereinigung ist disjunkt nach Axiom (P3). Es sei $K = \mathbb{N}' + \{0\}$, wir zeigen: K ist induktiv. $0 \in K$ ist klar, und aus $n \in K$ folgt $n' \in \mathbb{N}' \subseteq K$. Also ist $K = \mathbb{N}$ nach Axiom (P5). Zu (ii): Wir betrachten die Menge $M = \{n \in \mathbb{N} \mid n' \neq n\}$ und zeigen: M ist induktiv. Zunächst ist $0 \in M$ wegen (P3). Sei $n \in M$, also $n \neq n'$. Nach (P4) folgt daraus $n' \neq (n')'$, also auch $n' \in M$. Damit ist M induktiv, und nach (P5) ist $M = \mathbb{N}$. □

1.2 Anfangsabschnitte

Der Definition des Anfangs liegt die folgende Idee zugrunde: ein Anfang „bis n“ ist die Menge aller natürlichen Zahlen, die beim Abzählen bis n vorkommen.

Definition 1.2.1. Es sei $(\mathbb{N}, 0, S)$ ein System natürlicher Zahlen:

(i) Eine Teilmenge $A \subseteq \mathbb{N}$ heißt Anfang, wenn $0 \in A$ ist und für alle $k \in \mathbb{N}$ die Implikation $k' \in A \Rightarrow k \in A$ gilt.

(ii) Es sei $n \in \mathbb{N}$ beliebig. Ein Anfang A heißt Anfang bis n, falls gilt: $n \in A$ und $n' \notin A$.

(iii) Wir werden im nächsten Lemma zeigen, dass es für jedes $n \in \mathbb{N}$ genau einen Anfang bis n gibt. Dieser wird mit A_n bezeichnet.

Lemma 1.2.2. *Für jedes System natürlicher Zahlen und jedes Element $n \in \mathbb{N}$ gilt:*

(i) *A ist Anfang bis n $\Rightarrow$ $B := A + \{n'\}$ ist Anfang bis n',*

(ii) *A ist Anfang bis n' $\Rightarrow$ $B := A - \{n'\}$ ist Anfang bis n,*

(iii) *$\forall n \in \mathbb{N}$: es gibt einen Anfang bis n,*

(iv) *$\forall n \in \mathbb{N}$: es gibt genau einen Anfang bis n (bezeichnet mit A_n),*

(v) *$A_{n'} = A_n + \{n'\}$.*

Beweis.
Zu (i): B ist ein Anfang wegen $0 \in A \subseteq B$, und für $k' \in B = A + \{n'\}$ gilt entweder $k' = n'$ (dann ist $k = n \in A \subseteq B$ nach Voraussetzung), oder $k' \in A$ (dann ist $k \in A \subseteq B$, da A ein Anfang ist). Nach Voraussetzung ist $n' \in B$, aber $(n')' \notin B$, andernfalls wäre $(n')' \in A$ wegen $(n')' \neq n'$, also auch $n' \in A$, im Widerspruch zu $n' \notin A$. Zu (ii): B ist ein Anfang, denn es ist $0 \in B = A - \{n'\}$ ($0 \neq n'$ nach P3), und aus $k' \in B = A - \{n'\} \subset A$ folgt $k \in A$, da A Anfang ist. Wegen $k \neq n'$ ist dann auch $k = n \in B$ (denn aus $k = n'$ folgt $(n')' = k' \in B \subset A$ und damit $(n')' \in A$ im Widerspruch zu $(n')' \notin A$). Auch ist $n \in B$, denn $n' \in A \Rightarrow n \in A \Rightarrow n \in B$ wegen $n \neq n'$. Und natürlich ist $n' \notin B$ nach Definition. Zu (iii): Es sei $E = \{n \in \mathbb{N} \mid \exists \text{ Anfang bis } n\}$. Wir zeigen, dass E induktiv ist (damit $E = \mathbb{N}$ nach P5). Zunächst ist $\{0\}$ ein Anfang bis 0, also $0 \in E$. Ist A_n ein Anfang bis n, so ist nach (ii) die Menge $B := A_n + \{n'\}$ ein Anfang bis n'. Also $n \in E \Rightarrow n' \in E$ und damit $E = \mathbb{N}$. Zu (iv): Es sei $F = \{n \in N \mid \exists! \text{ Anfang bis } n\}$. Auch F ist induktiv: $\{0\}$ ist der einzige Anfang bis 0, denn für jeden weiteren Anfang A bis 0 ist $\{0\} + A^c$ induktiv, also $\{0\} + A^c = \mathbb{N}$. Nach Satz 1.1.4 ist dann $A^c = \mathbb{N}'$ und damit $A = \{0\}$ (die Induktivität von $N = \{0\} + A^c$ folgt, da für $k \neq 0$ in $\mathbb{N}$ stets $k \in A^c$ gilt, also $k \notin A$, damit $k' \notin A$ da A Anfang ist, also $k' \in A^c$). Also ist $0 \in F$. Sei nun $n \in F$ beliebig und $A^{(1)}, A^{(2)}$ beliebige Anfänge bis n' (es gibt mindestens einen nach der vorigen Rechnung). Nach (ii) sind dann $B^{(1)} = A^{(1)} - \{n'\}$ und $B^{(2)} = A^{(2)} - \{n'\}$ Anfänge bis n. Da $n \in F$ ist folgt $B^{(1)} = B^{(2)}$ und damit $A^{(1)} = A^{(2)}$, damit auch $n' \in F$. Also ist F induktiv und $F = \mathbb{N}$. Zu (v): $A_{n'} = A_n + \{n'\}$ folgt aus (iv) und (i). □

1.3 Anordnung

Wir erinnern zunächst an eine Definition aus Algebra I:

Definition 1.3.1. Eine Relation $\leq$ auf einer Menge M heißt Ordnungsrelation falls die folgenden Eigenschaften erfüllt:

(i) Reflexivität: $\forall a \in M : \ a \leq a$,

(ii) Transitivität: $\forall a, b, c \in M : \ a \leq b \leq c \ \Rightarrow \ a \leq c$,

(iii) Antisymmetrie: $\forall a, b \in M : \ a \leq b \leq a \ \Rightarrow \ a = b$.

Falls zusätzlich die Vergleichbarkeit ($\forall a, b \in M : \ a \leq b \,\text{oder}\, b \leq a$) gilt, heißt $\leq$ eine totale Ordnungsrelation.

Das Konzept der Anfänge ermöglicht es nun, eine Anordnung auf $\mathbb{N}$ einzuführen.

Lemma 1.3.2. *Es gilt:*

(i) *Der (beliebige) Durchschnitt von Anfängen ist ein Anfang,*

(ii) *die (beliebige nicht leere) Vereinigung von Anfängen ist ein Anfang,*

(iii) *für alle $m, n \in \mathbb{N}$ gilt: $n \in A_m \ \Rightarrow \ A_n = A_n \cap A_m$ (also $A_n \subseteq A_m$),*

(iv) *für alle $m, n \in \mathbb{N}$ gilt: $n \notin A_m \ \Rightarrow \ A_n = A_n \cup A_m$ (also $A_m \subseteq A_n$),*

(v) $A_n = A_m \ \Leftrightarrow \ n = m$.

Beweis.
Zu (i): Es sei

$$A \ := \bigcap_{\alpha} A(\alpha)$$

für Anfänge $A(\alpha)$ von $\mathbb{N}$. Zunächst ist $0 \in A$, da $0 \in A(\alpha)$ für alle α ist. Ist $k' \in A$, so liegt k' auch in jedem $A(\alpha)$, dann liegt aber auch k in jedem $A(\alpha)$, da alle $A(\alpha)$ Anfänge sind, also $k \in A$. Zu (ii): Sei nun

$$A \ := \bigcup_{\alpha} A(\alpha)$$

gesetzt. Da 0 in jedem $A(\alpha)$ liegt, ist auch $0 \in A$. Ist $k' \in A$, so gibt es mindestens ein α mit $k' \in A(\alpha)$. Dann ist auch $k \in A(\alpha)$ und damit $k \in A(\alpha) \subseteq A$. Zu (iii): Die Menge $A := A_n \cap A_m$ ist nach (i) ein Anfang. Es ist $n \in A$ wegen $n \in A_n$ und $n \in A_m$, aber $n' \notin A$ wegen $n' \notin A_n$. Also ist A ein Anfang bis n, nach Lemma 1.2.2(iv) also $A = A_n$. Zu (iv): $A := A_n \cup A_m$ ist nach (ii) ein Anfang, und es ist $n \in A$ wegen $A \in A_n$. Andererseits ist $n' \notin A$, da sonst (wegen $n' \notin A_n$) auch $n' \in A_m$ folgen würde (und damit der Widerspruch $n \in A_m$ da A_m Anfang ist). Also ist $A = A_n$ der eindeutige Anfang bis n. Zu (v): Es sei $M = \{n \in \mathbb{N} \mid A_n = A_m \Rightarrow n = m\}$, wir zeigen die Induktivität von M und damit $M = \mathbb{N}$. Angenommen $\{0\} = A_0 = A_m$, dann folgt aus $m \in A_m$ sofort $m = 0$, also $0 \in M$. Sei nun $n \in M$ beliebig, dann gilt

$$A_n + \{n'\} = A_{n'} = A_m \ \Rightarrow \ \begin{cases} m = n' & \text{oder} \\ m \in A_n \end{cases}$$

wobei die Annahme $m \in A_n$ aber zum Widerspruch $A_m \subseteq A_n \subset A_{n'} = A_m$ führt, also bleibt nur $m = n'$, damit $n' \in M$, und M ist induktiv. Die Richtung $\Leftarrow$ aus (v) ist klar. □

Bemerkung 1.3.3. Lemma 1.3.2 besagt, dass für je zwei natürliche Zahlen m und n die zugehörigen Anfangsabschnitte mengentheoretisch vergleichbar sind: Es gilt stets $A_n \subseteq A_m$ oder $A_m \subseteq A_n$.

Definition 1.3.4 (Anordnung der natürlichen Zahlen)**.** Seien $m, n \in \mathbb{N}$. Wir schreiben

(i) $n \leq m$ für $A_n \subseteq A_m$,

(ii) $n < m$ für $A_n \subset A_m$.

Wie üblich steht der Ausdruck $m \geq n$ für $n \leq m$ (bzw. $m > n$ für $n < m$).

Lemma 1.3.5. *Es gilt die Äquivalenz* $[n \leq m] \Leftrightarrow$ [*entweder* $n < m$ *oder* $n = m$].

Beweis.
Nach der vorigen Definition gilt

$$n \leq m \Leftrightarrow A_n \subseteq A_m \Leftrightarrow \begin{cases} A_n \subset A_m & \Leftrightarrow \quad n < m \\ A_n = A_m & \Leftrightarrow \quad n = m \quad \text{(Lemma 1.3.2)} \end{cases}.$$

□

Satz 1.3.6. *Die Anordnung* $\leq$ *von* $\mathbb{N}$ *ist eine totale Ordnungsrelation.*

Beweis.
Aus der Bemerkung zu Lemma 1.3.2 und Lemma 1.3.5 folgt, dass stets mindestens eine der Aussagen $a < b$, $a = b$ oder $a > b$ zutrifft. Dass jeweils höchstens eine der Aussagen zutrifft folgt aus mengentheoretischen Gründen, da die drei Aussagen nach Definition 1.3.4 und Lemma 1.3.2(iv) äquivalent sind zu $A_n \subset A_m$, $A_n = A_m$, $A_m \subset A_n$. Transitivität und Antisymmetrie erbt die Relation $\leq$ von der Relation $\subseteq$. □

Die Ordnungsrelation $\leq$ ist mit der Nachfolgerbildung in folgendem Sinne verträglich:

Lemma 1.3.7. *Es seien* $m, n \in \mathbb{N}$*, dann gilt:*

(i) $0 \leq m$,

(ii) $m < m'$,

(iii) $n < m \Rightarrow n' \leq m$,

(iv) $n < m \Rightarrow n' < m'$.

Beweis.
Zu (i): Es sei $L := \{m \in \mathbb{N} \mid A_0 \subseteq A_m\}$, wir zeigen die Induktivität von L. Für $m = 0$ ist $A_0 \subseteq A_0$ klar, also $0 \in L$. Sei $m \in L$ beliebig. Nach Lemma 1.2.2 ist dann $A_{m'} = A_m + \{m'\}$ und $A_0 \subseteq A_m \subseteq A_{m'}$, also $m' \in L$ und damit $L = \mathbb{N}$. Zu (ii): $A_{m'} = A_m + \{m'\} \Rightarrow A_m \subset A_{m'}$. Zu (iii): $n < m \Leftrightarrow A_n \subset A_m \Leftrightarrow A_{n'} \subseteq A_m \Leftrightarrow n' \leq m$. Zu (iv): $n < m \Rightarrow n' \leq m < m'$ nach (ii) und (iii). □

Definition 1.3.8. Sei A eine Teilmenge von $\mathbb{N}$:

(i) Ein Minimum/kleinstes Element ist ein $k \in A$, so dass $n \in A \Rightarrow k \leq n$ für alle $n \in A$ gilt.

(ii) Ein Maximum/größtes Element ist ein $k \in A$, so dass $n \in A \Rightarrow k \geq n$ für alle $n \in A$ gilt.

Lemma 1.3.9. *Jedes* $A \subseteq \mathbb{N}$ *besitzt höchstens ein Minimum und höchstens ein Maximum.*

Beweis.
Seien $k_1, k_2 \in A$ Minima von A, so gilt nach der vorigen Definition $k_1 \leq k_2$, aber auch $k_2 \leq k_1$, damit $k_1 = k_2$ wegen der Antisymmetrie von $\leq$. Das Gleiche gilt für das Maximum. □

Im Falle der Existenz schreiben wir wie gewohnt $\min(A)$ bzw. $\max(A)$ für das eindeutige Minimum bzw. Maximum von $A \subseteq \mathbb{N}$.

Lemma 1.3.10. *Für alle $n \in \mathbb{N}$ gilt: Jede nichtleere Teilmenge A von A_n besitzt ein Minimum.*

Beweis.
Es sei L die Menge der $n \in \mathbb{N}$, für welche die obige Aussage zutrifft. Wir zeigen die Induktivität von L: Im Fall $n = 0$ gilt $\emptyset \neq A \subseteq A_0 = \{0\} \Rightarrow A = \{0\} \Rightarrow 0$ ist ein Minimum. Sei nun $n \in L$ beliebig und A eine nichtleere Teilmenge von $A_{n'}$. Falls $A \cap A_n = \emptyset$ ist bleibt nur die Möglichkeit $A = \{n'\}$, und das gesuchte Minimum ist n'. Andernfalls sei $a^* := \min(A \cap A_n)$. Das Minimum existiert, weil $n \in L$ angenommen ist, und $A \cap A_n$ eine nichtleere Teilmenge von A_n ist. Dann ist natürlich $a^* \in A$, und es gilt $k \in A \Rightarrow a^* \leq k$, wegen

$$k \in A \Rightarrow \begin{cases} k \in A \cap A_n & \Rightarrow \quad a^* \leq k & \text{da minimal in } A \cap A_n \\ k = n' & \Rightarrow \quad a^* \leq n \leq n' = k & \text{(Lemmata 1.3.2/1.3.7)} \end{cases}.$$

□

Satz 1.3.11 (Wohlordnungssatz)**.** *Jede nichtleere Teilmenge A von $\mathbb{N}$ besitzt ein Minimum.*

Beweis.
Sei $k \in A$ beliebig. Nach dem vorigen Lemma existiert $m := \min(A \cap A_k)$, und m ist sogar ein Minimum von A wegen $m \in A$ und $n \in A \Rightarrow m \leq n$, denn es gilt

$$n \in A \Rightarrow \begin{cases} n \in A \cap A_k & \Rightarrow \quad m \leq n & \text{da minimal in } A \cap A_k \\ n \notin A_k & \Rightarrow \quad n \geq k \geq m & \text{(Lemma 1.3.2)} \end{cases}.$$

□

1.4 Das Rekursionstheorem

In diesem Abschnitt führen wir das zentrale Hilfsmittel zur induktiven Definition (beispielsweise des Summenzeichens) ein.

Satz 1.4.1 (Endliches Rekursionstheorem)**.** *Es sei M eine beliebige Menge, $a \in M$ und für alle $k \in \mathbb{N}$ sei $\varphi_k : M \to M$ eine Abbildung. Dann gibt es für alle $n \in \mathbb{N}$ genau eine Abbildung $f_n : A_n \to M$ mit*

(i) $f_n(0) = a$,

(ii) $\forall k \in \mathbb{N} : k' \in A_n \Rightarrow f_n(k') = \varphi_k(f_n(k))$.

Bemerkung 1.4.2. Ausformuliert bedeutet das gerade

$f_n(0) = a$,
$f_n(1) = f_n(0') = \varphi_0(f_n(0)) = \varphi_0(a)$,
$f_n(2) = f_n(1') = \varphi_1(f_n(1)) = \varphi_1(\varphi_0(a))$, usw.

Beweis.
Es sei L die Menge aller $n \in \mathbb{N}$, für die es genau eine Abbildung $f_n : A_n \to M$ gibt, so dass die Behauptungen (i) und (ii) gelten. Wir zeigen die Induktivität von L: Für $n = 0$ ist $f_0 : \{0\} \to M$ mit $f_0(0) = a$ offensichtlich die einzige solche Abbildung, also ist $0 \in L$. Nun sei $n \in L$ beliebig, wir zeigen, dass es genau ein $f_{n'}$ mit den geforderten Eigenschaften gibt:

Eindeutigkeit:
Sei $f_{n'}$ eine solche Abbildung auf $A_{n'}$, dann ist die Restriktion von $f_{n'}$ auf $A_n \subset A_{n'}$ auch eine Abbildung, die (i) und (ii) erfüllt. Wegen $n \in L$ folgt $\forall k \in A_n : \; f_{n'}(k) = f_n(k)$. Also gilt auch $f_{n'}(n') = \varphi_n(f_{n'}(n)) = \varphi_n(f_n(n))$, womit $f_{n'}$ für alle Elemente von $A_{n'}$ durch f_n und φ_n eindeutig festgelegt ist.

Existenz:
Für $n \in L$ erklären wir $f_{n'}$ auf $A_{n'} = A_n + \{n'\}$ durch

$$f_{n'}(k) := \begin{cases} f_n(k) & \text{falls} \quad k \in A_n \\ \varphi_n(f_n(n)) & \text{falls} \quad k = n' \end{cases} .$$

Dann folgt einerseits für $k' \in A_n$ die Gleichheit

$$f_{n'}(k') \; = \; f_n(k') \; = \; \varphi_k(f_n(k)) \; = \; \varphi_k(f_{n'}(k))$$

da $k \in A_n$ ist und die Abbildungen $f_n, f_{n'}$ wegen $n \in L$ und der Eindeutigkeitsbehauptung auf A_k übereinstimmen müssen. Für die verbleibende Möglichkeit $k' = n'$ folgt andererseits $k = n$, und damit

$$f_{n'}(k') \; = \; f_{n'}(n') \; = \; \varphi_n(f_n(k)) \; = \; \varphi_n(f_{n'}(n)) \; = \; \varphi_k(f_{n'}(k))$$

womit $f_{n'}$ in jedem Fall die Eigenschaften (i) und (ii) erfüllt und damit auch $n' \in L$ ist. □

Insbesondere stimmen für jedes n die durch den Satz festgelegten Abbildungen f_n und $f_{n'}$ auf dem Anfang A_n überein.

Satz 1.4.3 (Rekursionstheorem)**.** *Sei eine Menge M, ein Element $a \in M$ und für jedes $k \in \mathbb{N}$ eine Abbildung $\varphi_k : M \to M$ gegeben. Dann gibt es genau eine Abbildung $f : \mathbb{N} \to M$ mit*

(i) $f(0) = a$,

(ii) $\forall k \in \mathbb{N} : \; f(k') = \varphi_k(f(k))$,

(iii) *Zusatz: sind alle φ_k identisch zu $\varphi : M \to M$, so folgt $\forall k \in \mathbb{N} : \; f(k') = \varphi(f(k))$.*

Beweis.
Wir verwenden die Bezeichnungen aus Satz 1.4.1:

Existenz:
Man setzt (mit f_k aus Satz 1.4.1)

$$(*) : \;\; f : \mathbb{N} \to M, \; n \mapsto f_n(n)$$

und folgert induktiv

$$f(0) \underset{(*)}{=} f_0(0) \; = \; a$$

wie in Satz 1.4.1 bzw.

$$f(k') \underset{(*)}{=} f_{k'}(k') = \varphi_k(f_{k'}(k)) = \varphi_k(f_k(k)) \underset{(*)}{=} \varphi_k(f(k))$$

da f_k und $f_{k'}$ auf A_k übereinstimmen.

Eindeutigkeit:
Ist $\tilde{f}$ irgend eine Lösung, so ist ihre Restriktion auf A_n ein Lösung des endlichen Rekursionsproblems auf A_n, also gerade (das nach Satz 1.4.1 eindeutige) f_n. Für jedes $n \in \mathbb{N}$ gilt also $\tilde{f}(n) = f_n(n) = f(n)$, damit ist $\tilde{f} = f$ auf ganz $\mathbb{N}$. □

1.5 Die Addition

Satz 1.5.1. *Es existiert genau eine (zweistellige und mit + bezeichnete) Operation* $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ *mit*

(i) $\forall n \in \mathbb{N} : n + 0 = n$,

(ii) $\forall m, n \in \mathbb{N} : (n + m)' = n + (m')$.

Beweis.
Eindeutigkeit:
Es seien + und $\boxplus$ zwei Operationen mit den gegebenen Eigenschaften. Für jedes feste $n \in \mathbb{N}$ ist die Menge $M_n = \{m \in \mathbb{N} \mid n + m = n \boxplus m\}$ induktiv:

$$n + 0 \underset{\text{(i)}}{=} n \underset{\text{(i)}}{=} n \boxplus 0$$

und damit $0 \in M_n$. Sei nun $m \in M_n$ beliebig, dann gilt

$$n + (m') \underset{\text{(ii)}}{=} (n + m)' = (n \boxplus m)' \underset{\text{(ii)}}{=} n \boxplus (m')$$

und somit $m' \in M_n$. Also ist $M_n = \mathbb{N}$ für jedes n.

Existenz:
Für festes $n \in \mathbb{N}$ liefert das Rekursionstheorem 1.4.3 mit $M := \mathbb{N}$, $a := n$ und $\varphi_k := S$ (die Nachfolgerabbildung) für alle k eine Abbildung $g_n : \mathbb{N} \to \mathbb{N}$ mit $g_n(0) = n$ und $g_n(m') = (g_n(m))'$. Wir setzen $n + m := g_n(m)$. Es folgt $n + 0 = g_n(0) = n$ und $n + (m') = g_n(m') = (g_n(m))' = (n + m)'$. □

Lemma 1.5.2. *Es gilt:*

(i) $\forall n \in \mathbb{N} : 0 + n = n$,

(ii) $\forall m, n \in \mathbb{N} : (n') + m = (n + m)'$,

(iii) $\forall n \in \mathbb{N} : n' = n + 1$.

Beweis.
Zu (i): Sei $A := \{n \in \mathbb{N} \mid 0 + n = n\}$, wir zeigen wieder Induktivität von A: $0 + 0 = g_0(0) = 0$, also $0 \in A$. Für $n \in A$ gilt $0 + n' = (0 + n)' = n'$ nach Satz 1.5.1. Zu (ii): Für festes n sei $B_n := \{m \in \mathbb{N} \mid (n') + m = (n + m)'\}$. Auch B_n ist induktiv: $(n') + 0 = n' = (n + 0)'$, also $0 \in B_n$. Für $m \in B_n$ gilt $(n') + (m') = ((n') + m)' = ((n + m)')' = (n + (m'))'$, also auch $m' \in B_n$. Zu (iii): $n' = (n + 0)' = n + (0') = n + 1$ nach Satz 1.5.1 und Definition 1.1.3. □

Satz 1.5.3. *Die Addition + besitzt auf* $\mathbb{N}$ *die folgenden Eigenschaften:*

(i) *Kommutativität:* $\forall m, n \in \mathbb{N} : \; n + m = m + n$,

(ii) *Assoziativität:* $\forall k, m, n \in \mathbb{N} : \; (k + m) + n = k + (m + n)$,

(iii) *Eindeutige Subtraktion:* $\forall k, m, n \in \mathbb{N} : \; n + k = m + k \;\Rightarrow\; n = m$.

Beweis.
Zu (i): Für festes $n \in \mathbb{N}$ sei $K_n := \{m \in \mathbb{N} \mid n + m = m + n\}$. K_n ist induktiv: $0 \in K_n$ folgt aus Satz 1.5.1: $n + 0 = n = 0 + n$. Sei $m \in K_n$ beliebig, dann gilt

$$n + (m') \;=\; (n + m)' \;=\; (m + n)' \;=\; (m') + n$$

wegen Satz 1.5.1 und Lemma 1.5.2. Zu (ii): Es seien $m, n \in \mathbb{N}$ fest und $A_{m,n} = \{k \in \mathbb{N} \mid (k + m) + n = k + (m + n)\}$. Wieder per Induktion: $0 \in A_{m,n}$ wegen $(0 + m) + n = m + n = 0 + (m + n)$. Für $k \in A_{m,n}$ gilt

$$((k') + m) + n \;=\; ((k + m)') + n \;=\; ((k + m) + n)' \;=\; (k + (m + n))' \;=\; (k') + (m + n)$$

und damit $k' \in A_{m,n}$. Zu (iii): Es sei $K_{m,n} := \{k \in \mathbb{N} \mid n + k = m + k \;\Rightarrow\; n = m\}$. Es ist $0 \in K_{m,n}$ wegen $n + 0 = n$ und $m + 0 = m$. Für $k \in K_{m,n}$ gilt

$$n + (k') = m + (k') \;\Leftrightarrow\; (n + k)' = (m + k)' \;\Leftrightarrow_{(\mathrm{P4})}\; n + k = m + k \;\Rightarrow\; n = m$$

und damit $k' \in K_{m,n}$. □

Satz 1.5.4 (Existenz der Differenz)**.** *Es gilt die Äquivalenz* $[n \leq m] \;\Leftrightarrow\; [\exists k \in \mathbb{N} : \; n + k = m]$.

Beweis.
Hinrichtung:
Suche ein Gegenbeispiel mit minimalem n (das existiert nach dem Wohlordnungssatz): Es ist $n > 0$, da $n = 0$ kein Gegenbeispiel liefert. Also gibt es $p \in \mathbb{N}$ mit $p' = n$ nach Satz 1.1.4. Ist $0 < p' = n \leq m$ (etwa mit $q' = m$), so folgt aus Lemma 1.3.7 auch $p \leq q$. Nun gibt es ein $k \in \mathbb{N}$ mit $p + k = q$, da $p < n$ kein Gegenbeispiel ist wegen der minimalen Wahl von n. Nun gilt aber $p + k = q \;\Rightarrow\; n + k = (p') + k = (p + k)' = q' = m$, also war schon n kein Gegenbeispiel, und die Hinrichtung ist für alle $n \in \mathbb{N}$ richtig.

Rückrichtung:
Suche wieder ein Gegenbeispiel mit minimalem k: da $k = 0$ kein Gegenbeispiel ist folgt $k > 0$. Aus $0 < k = p'$ mit $p < p'$ folgt

$$n + p = q \;\Leftrightarrow\; n \leq q \leq q' = m$$

und n ist auch kein Gegenbeispiel. □

1.6 Die Multiplikation

Satz 1.6.1. *Es existiert genau eine (zweistellige und mit* $\cdot$ *bezeichnete) Operation* $\cdot : \; \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ *mit*

(i) $\forall n \in \mathbb{N} : \; n \cdot 0 = 0$,

(ii) $\forall m, n \in \mathbb{N} : \; n \cdot (m') = (n \cdot m) + n.$

Hierbei soll die Konvention „Punkt vor Strich“ gelten: $a \cdot b + c = (a \cdot b) + c$.

Beweis.
Eindeutigkeit:
Es seien $\cdot$ und $\boxdot$ zwei Operationen mit den gegebenen Eigenschaften. Für jedes feste $n \in \mathbb{N}$ ist die Menge $E_n := \{m \in \mathbb{N} \mid n \cdot m = n \boxdot m\}$ induktiv: $n \cdot 0 = 0 = n \boxdot 0$, also $0 \in E_n$, und für $m \in E_n$ gilt

$$n \cdot (m') \; = \; n \cdot m + n \; = \; n \boxdot m + n \; = \; n \boxdot (m')$$

und damit $m' \in E_n$.

Existenz:
Für jedes feste $n \in \mathbb{N}$ liefert das Rekursionstheorem mit $M := \mathbb{N}$, $a := 0$ und $\varphi_n(k) := k+n$ eine Abbildung $g_n : \mathbb{N} \to \mathbb{N}$ mit $g_n(0) = 0$ und $g_n(m') = \varphi_n(g_n(m)) = g_n(m) + n$. Man setzt $n \cdot m := g_n(m)$. □

Lemma 1.6.2. *Es gilt:*

(i) $\forall m \in \mathbb{N} : \; 0 \cdot m = 0,$

(ii) $\forall n \in \mathbb{N} : \; n \cdot 1 = n,$

(iii) $\forall n \in \mathbb{N} : \; 1 \cdot n = n,$

(iv) $\forall m, n \in \mathbb{N} : \; (n') \cdot m = (n \cdot m) + m.$

Beweis.
Zu (i): Die Menge $O := \{m \in \mathbb{N} \mid 0 \cdot m = 0\}$ ist induktiv, denn ist $0 \cdot 0 = 0$ nach dem vorigen Satz, also $0 \in O$. Sei $m \in O$, dann ist $0 \cdot (m') = 0 \cdot m + 0 = 0 + 0 = 0$ nach dem vorigen Satz und Satz 1.5.1, also $m' \in O$. Zu (ii): $n \cdot 1 = n \cdot (0') = n \cdot 0 + n = 0 + n = n$ (Lemma 1.5.2). Zu (iii): Die Menge $E := \{n \in \mathbb{N} \mid 1 \cdot n = n\}$ ist induktiv: $1 \cdot 0 = 0$ ist klar, und für $m \in E$ gilt

$$1 \cdot (n') \; = \; 1 \cdot n + 1 \; = \; n + 1 \; = \; n' \, .$$

Zu (iv): Für festes $n \in \mathbb{N}$ sei $D_n := \{m \in \mathbb{N} \mid (n') \cdot m = (n \cdot m) + m\}$. $(n') \cdot 0 = 0 = 0 + 0 = n \cdot 0 + 0$ folgt aus den Sätzen 1.6.1 und 1.5.1, also $0 \in D_n$. Sei $m \in D_n$ beliebig, dann gilt wegen der Kommutativität von $+$ auch

$$(n') \cdot (m') \; = \; (n') \cdot m + (n') \; = \; n \cdot m + m + (n') \; = \; n \cdot m + n + (m') \; = \; n \cdot (m') + (m')$$

und damit $m' \in D_n$. □

Satz 1.6.3. *Die Multiplikation $\cdot$ besitzt auf $\mathbb{N}$ die folgenden Eigenschaften:*

(i) *Kommutativität:* $\forall m, n \in \mathbb{N} : \; n \cdot m = m \cdot n,$

(ii) *Distributivität:* $\forall k, m, n \in \mathbb{N} : \; k \cdot (m + n) = (k \cdot m) + (k \cdot n),$

(iii) *Assoziativität:* $\forall k, m, n \in \mathbb{N} : \; (n \cdot m) \cdot k = n \cdot (m \cdot k).$

Beweis.
Zu (i): Die Menge $K := \{n \in \mathbb{N} \mid n \cdot m = m \cdot n\}$ ist induktiv. $0 \cdot m = 0 = m \cdot 0$ nach Lemma 1.6.2, und für $m \in K$ gilt

$$(n') \cdot m \;=\; n \cdot m + m \;=\; m \cdot n + m \;=\; m \cdot (n')$$

nach Satz 1.6.1 und Lemma 1.6.2, also folgt $K = \mathbb{N}$. Zu (ii): Es sei $D_{m,n} = \{k \in \mathbb{N} \mid k \cdot (m+n) = (k \cdot m) + (k \cdot n)\}$. Wegen $0 \cdot (m+n) = 0 = 0 + 0 = (0 \cdot m) + (0 \cdot n)$ ist $0 \in D_{m,n}$. Für $k \in D_{m,n}$ gilt

$$(k') \cdot (m+n) \;=\; k \cdot (m+n) + (m+n) \;=\; (k \cdot m) + (k \cdot n) + m + n$$

$$=\; (k \cdot m) + m + (k \cdot n) + n \;=\; (k') \cdot m + (k') \cdot n$$

und damit $k' \in D_{m,n}$. Zu (iii): Für feste $m, n \in \mathbb{N}$ sei $A_{m,n} = \{k \in \mathbb{N} \mid (n \cdot m) \cdot k = n \cdot (m \cdot k)\}$. Es ist $(n \cdot m) \cdot 0 = 0 = n \cdot 0 = n \cdot (m \cdot 0)$, also $0 \in A_{m,n}$. Für $k \in A_{m,n}$ gilt

$$n \cdot (m \cdot (k')) \;=\; n \cdot (m \cdot k + m) \underset{\text{(ii)}}{=} n \cdot (m \cdot k) + n \cdot m \;=\; (n \cdot m) \cdot k + n \cdot m \;=\; (n \cdot m) \cdot (k')$$

und damit $k' \in A_{m,n}$. □

Satz 1.6.4 (Nullteilerfreiheit)**.** *Aus $m, n \neq 0$ folgt $n \cdot m \neq 0$.*

Beweis.
Nach Satz 1.1.4 gibt es $p, q \in \mathbb{N}$ mit $n = p'$ und $m = q'$. Damit gilt

$$n \cdot m \;=\; (p') \cdot (q') \;=\; (p') \cdot q + p' \;=\; ((p') \cdot q + p)' \;\neq\; 0\,.$$

□

Satz 1.6.5 (Kürzungsregel)**.** *Ist $k \neq 0$ und $n \cdot k = m \cdot k$, so folgt $n = m$.*

Beweis.
Wir führen einen Widerspruchsbeweis, es sei daher $n \neq m$ angenommen. Nach Satz 1.3.6 ist $n < m$ oder $m < n$. O.B.d.A. sei $n < m$. Nach Satz 1.5.4 gibt es ein $l \in \mathbb{N}$ mit $n + l = m$, und es ist $l \neq 0$. Dann gilt:

$$n \cdot k = m \cdot k \;\Rightarrow\; n \cdot k = (n+l) \cdot k = n \cdot k + l \cdot k\,,$$

das ist aber nach Satz 1.5.3(iii) nur mit $l \cdot k = 0$ möglich, im Widerspruch zur Nullteilerfreiheit. □

1.7 Die ganzen Zahlen

Wir wiederholen zunächst einige Grundstrukturen aus der Algebra:

Definition 1.7.1. Es sei $X \neq \emptyset$. Auf X sei eine (innere) Verknüpfung $\circ$, d. h. eine Abbildung $\circ : X \times X \to X, (x, y) \mapsto x \circ y$ definiert.

(i) $(X, \circ)$ heißt Halbgruppe, falls die Verknüpfung $\circ$ dem Assoziativgesetz genügt:

$$\forall x, y, z \in X : (x \circ y) \circ z = x \circ (y \circ z)\,.$$

(ii) Ein $e \in X$ heißt neutrales Element der Halbgruppe, falls $e \circ x = x \circ e = x$ für alle $x \in X$ gilt.

(iii) Ein $x \in X$ heißt Einheit, falls es ein Inverses $x^{-1} \in X$ mit $x \circ x^{-1} = x^{-1} \circ x = e$ gibt.

(iv) $(X, \circ)$ heißt abelsche/kommutative Halbgruppe, falls $x \circ y = y \circ x$ für alle $x, y \in X$ gilt.

(v) Jede Halbgruppe mit neutralem Element, die nur Einheiten besitzt, heißt Gruppe.

$(\mathbb{N}, +)$ besitzt nach Satz 1.5.4 nur eine einzige Einheit (nämlich 0), ist also keine Gruppe. $(\mathbb{N}, +)$ kann nun (wie jede kommutative Halbgruppe mit Kürzungsregel) zu einer Gruppe erweitert werden, der Gruppe $(\mathbb{Z}, +)$ der ganzen Zahlen. Diese Konstruktion soll in diesem Abschnitt durchgeführt werden. Mit den Verknüpfungen $+$ und $\cdot$ bildet $\mathbb{Z}$ sogar einen Ring:

Definition 1.7.2. Es sei $X \neq \emptyset$ eine Menge mit zwei inneren Verknüpfungen, der Addition $+$ und der Multiplikation $\cdot$, dann definiert man:

(i) $(X, +, \cdot)$ heißt ein Ring, falls gilt:

- $(X, +)$ ist eine abelsche Gruppe,
- $(X, \cdot)$ ist eine Halbgruppe,
- es gelten die Distributivgesetze:

$$\begin{aligned} (a+b)\cdot c &= (a\cdot c)+(b\cdot c) \\ c\cdot(a+b) &= (c\cdot a)+(c\cdot b) \quad . \end{aligned}$$

(ii) Das neutrale Element von $(X, +)$ heißt Null (Schreibweise: 0).

(iii) Besitzt $(X, \cdot)$ ein neutrales Element, so heißt dieses Eins (Schreibweise: 1).

(iv) Ist $(X, \cdot)$ abelsch, so heißt $(X, +, \cdot)$ ein kommutativer Ring.

(v) Der Ring $(X, +, \cdot)$ heißt nullteilerfrei, falls für alle $x, y \in X$ gilt:

$$x \cdot y = 0 \;\Rightarrow\; x = 0 \text{ oder } y = 0\,.$$

(vi) Ein nullteilerfreier und kommutativer Ring, der mindestens zwei verschiedene Elemente enthält, heißt Integritätsring.

(vii) Ein Integritätsring mit 1, in dem jedes $a \neq 0$ eine (multiplikative) Einheit ist, heißt Körper.

Wir beschreiben jetzt die Konstruktion der ganzen Zahlen. Wir beginnen mit der Einführung einer Äquivalenzrelation auf $\mathbb{N} \times \mathbb{N}$:

Definition 1.7.3. Zwei Paare (n_1, m_1) und (n_2, m_2) heißen äquivalent , Schreibweise $(n_1, m_1) \sim (n_2, m_2)$, genau dann, wenn $n_1 + m_2 = m_1 + n_2$ gilt.

Lemma 1.7.4. *Die Relation $\sim$ ist eine Äquivalenzrelation auf $\mathbb{N} \times \mathbb{N}$:*

(i) *Reflexivität:* $(n, m) \sim (n, m)$ *für alle* $(n, m) \in \mathbb{N} \times \mathbb{N}$,

(ii) *Symmetrie:* $(n_1, m_1) \sim (n_2, m_2) \;\Leftrightarrow\; (n_2, m_2) \sim (n_1, m_1)$,

(iii) *Transitivität: Aus* $(n_1, m_1) \sim (n_2, m_2)$ *und* $(n_2, m_2) \sim (n_3, m_3)$ *folgt* $(n_1, m_1) \sim (n_3, m_3)$.

Beweis.
(i) und (ii) folgen aus der Kommutativität der Addition. Zu (iii): Es gilt mit der additiven Kürzungsregel nach Satz 1.5.3(iii)

$$\begin{aligned}(n_1,m_1)\sim(n_2,m_2)\,,\ (n_2,m_2)\sim(n_3,m_3) \quad &\Rightarrow \quad n_1+m_2=m_1+n_2\,,\ n_2+m_3=m_2+n_3\\ &\Rightarrow \quad n_1+m_2+n_2+m_3=m_1+n_2+m_2+n_3\\ &\Rightarrow \quad n_1+m_3=m_1+n_3\\ &\Rightarrow \quad (n_1,m_1)\sim(n_3,m_3)\,.\end{aligned}$$

□

Zu jeder Äquivalenzrelation $\sim$ auf einer Menge M gehört eine Klasseneinteilung, eine Partition, von M in Äquivalenzklassen. Diese sind von der Form

$$\bar{n} \;=\; \{m\in M \mid m\sim n\}\,.$$

Je zwei Äquivalenzklassen sind disjunkt oder identisch, und M ist die Vereinigung der Äquivalenzklassen.

Definition 1.7.5 (Die ganzen Zahlen)**.** Es ist $\mathbb{Z} := \{\overline{(n,m)} \mid n,m\in\mathbb{N}\}$ mit der Äquivalenzklasse $\overline{(n,m)}$ von (n,m) bzgl. der in Definition 1.7.3 gegebenen Äquivalenzrelation.

Im Folgenden seien $\alpha=\overline{(n,m)}$ und $\beta=\overline{(l,k)}$ Elemente von $\mathbb{Z}$.

Definition 1.7.6. Für $\alpha,\beta\in\mathbb{Z}$ ist

$$\alpha\oplus\beta \;:=\; \overline{(n+l,m+k)}\ ,\ \ \alpha\odot\beta \;:=\; \overline{(n\cdot l+m\cdot k, n\cdot k+m\cdot l)}\ ,\ \ \alpha<\beta \;\Leftrightarrow\; n+k<m+l$$

für irgend welche $(n,m)\in\alpha$ und $(l,k)\in\beta$.

Lemma 1.7.7. *Addition, Multiplikation und Anordnung ganzer Zahlen sind wohldefiniert, d. h. sie sind unabhängig von der Wahl der Repräsentanten.*

Beweis.
Addition:
Seien $\alpha=\overline{(n_1,m_1)}=\overline{(n_2,m_2)}$ und $\beta=\overline{(l_1,k_1)}=\overline{(l_2,k_2)}$ vorgelegt. Es ist $(n_1,m_1)\sim(n_2,m_2)$ und $(l_1,k_1)\sim(l_2,k_2)$, also $n_1+m_2=m_1+n_2$ und $l_1+k_2=k_1+l_2$. Daraus folgt $n_1+l_1+m_2+k_2=n_2+l_2+m_1+k_1$, und damit $\overline{(n_1+l_1,m_1+k_1)}=\overline{(n_2+l_2,m_2+k_2)}$.

Multiplikation:
Mit den gleichen Repräsentanten für α und β ergibt sich

$$n_1\cdot l_1+m_1\cdot k_1+n_2\cdot k_2+m_2\cdot l_2=n_1\cdot k_1+m_1\cdot l_1+n_2\cdot l_2+m_2\cdot k_2$$

und damit

$$\overline{(n_1\cdot l_1+m_1\cdot k_1, n_1\cdot k_1+m_1\cdot l_1)}=\overline{(n_2\cdot l_2+m_2\cdot k_2, n_2\cdot k_2+m_2\cdot l_2)}\,,$$

also ist auch die Multiplikation unabhängig von der Repräsentantenwahl.

Anordnung:
Sei $(n_1,m_1)<(l_1,k_1)$, also $n_1+k_1<l_1+m_1$. Dann folgt durch Einsetzen der Äquivalenz zu den mit 2 indizierten Repräsentanten

$$n_1+k_1+n_2+l_2+k_2+m_2 \;<\; l_1+m_1+n_2+l_2+k_2+m_2$$

und damit durch Subtraktion der obigen Gleichungen

$$n_2 + k_2 \ < \ l_2 + m_2 \, .$$

Damit ist auch die Anordnung unabhängig von der Repräsentantenwahl. □

Definition 1.7.8. Wir definieren die Null durch $0 := \overline{(0,0)}$ und die Eins durch $1 := \overline{(1,0)}$.

Künftig bezeichnen wir Addition und Multiplikation der ganzen Zahlen wie üblich mit $+$ und $\cdot$.

Satz 1.7.9. $(\mathbb{Z}, +, \cdot)$ *ist ein Integritätsring mit der Null* 0 *und der Eins* 1 *aus der vorigen Definition.*

Beweis.
Die Rechenregeln für Addition und Multiplikation (Kommutativität, Assoziativität, Distributivität) folgen aus den entsprechenden Regeln für die natürlichen Zahlen. Zudem gilt

$$\overline{(n,m)} + \overline{(0,0)} \ = \ \overline{(n+0, m+0)} \ = \ \overline{(n,m)} \quad \text{und}$$

$$\overline{(n,m)} \cdot \overline{(1,0)} \ = \ \overline{(n \cdot 1 + m \cdot 0, n \cdot 0 + m \cdot 1)} \ = \ \overline{(n,m)} \, .$$

Nun sei

$$(*): \ \overline{(n,m)} \cdot \overline{(l,k)} = \overline{(n \cdot l + m \cdot k, n \cdot k + m \cdot l)} = \overline{(0,0)}$$

und $\overline{(l,k)} \neq \overline{(0,0)}$ angenommen. Offenbar ist $\overline{(l,k)} = \overline{(0,0)}$ gleichbedeutend mit $l = k$, also ist entweder $l < k$ oder $l > k$, wir nehmen o.B.d.A $l < k$ an. Nach Satz 1.5.4 (Existenz der Differenz) gibt es ein $t \in \mathbb{N} - \{0\}$, so dass $l + t = k$ ist. Aus der Gleichheit $(*)$ folgt dann $n \cdot l + m \cdot k = n \cdot k + m \cdot l$, also $n \cdot l + m \cdot (l+t) = n \cdot (l+t) + m \cdot l$. Aus den Distributivgesetzen folgt $n \cdot l + m \cdot l + m \cdot t = n \cdot l + n \cdot t + m \cdot l$, und mit der additiven Kürzungsregel $m \cdot t = n \cdot t$, mit Satz 1.6.5 und $t \neq 0$ folgt schließlich $n = m$, also $\overline{(n,m)} = \overline{(0,0)}$. Damit ist $(\mathbb{Z}, +, \cdot)$ nullteilerfrei. □

Das so konstruierte $\mathbb{Z}$ ist bisher keine Erweiterung der natürlichen Zahlen $\mathbb{N}$. Wir können $\mathbb{N}$ in $\mathbb{Z}$ einbetten durch die injektive Abbildung

$$\Phi : \begin{cases} \mathbb{N} & \to & \mathbb{Z} \\ n & \mapsto & \overline{(n,0)} \end{cases}$$

Durch Nachrechnen zeigt man leicht

Satz 1.7.10. *Die Abbildung* Φ *ist relationstreu bzgl. der Operationen* $+$, $\cdot$ *und* $\leq$. *Die Anordnung auf* $\mathbb{Z}$ *ist eine totale Ordnungsrelation.*

1.8 Die rationalen Zahlen

Auch die rationalen Zahlen definieren wir durch einen Paarbildungsprozess. Wir erklären eine Äquivalenzrelation auf $\mathbb{Z} \times (\mathbb{Z} - \{0\})$ wie folgt:

Definition 1.8.1. Zwei Paare $(p,q), (r,s) \in \mathbb{Z} \times (\mathbb{Z} - \{0\})$ heißen äquivalent, Schreibweise $(p,q) \sim (r,s)$, falls $p \cdot s = q \cdot r$.

Dahinter steckt natürlich die Idee, rationale Zahlen als Brüche von ganzen Zahlen aufzufassen: Die Menge $\mathbb{Q}$ der rationalen Zahlen ist die Menge der Äquivalenzklassen bzgl. dieser Äquivalenzrelation. Für $q \neq 0$ bezeichnet man die Äquivalenzklasse $\overline{(p,q)}$ auch kurz mit $\frac{p}{q}$. Wir verzichten auf die Details, und erinnern an die Definition des Quotientenkörpers aus der Vorlesung Algebra I: Die Verknüpfungen von $\mathbb{Z}$ übertragen sich auf $\mathbb{Q}$ gemäß

$$\frac{p}{q} + \frac{r}{s} \;=\; \frac{sp+rq}{qs} \quad , \quad \frac{p}{q} \cdot \frac{r}{s} \;=\; \frac{pr}{qs} \, .$$

Die Anordnung auf $\mathbb{Q}$ ist gegeben durch $\frac{p}{q} < \frac{r}{s} \;\Leftrightarrow\; p \cdot s < r \cdot q$ für $(p,q),(r,s) \in \mathbb{Z} \times (\mathbb{Z} - \{0\})$.

Satz 1.8.2. *Addition, Multiplikation und Anordnung sind auf $\mathbb{Q}$ wohldefiniert, d. h. unabhängig von der Wahl der Repräsentanten. $\mathbb{Q}$ bildet mit diesen Operationen und den neutralen Elementen $\frac{0}{1}$ bzw. $\frac{1}{1}$ einen Körper.*

So wie die Menge $\mathbb{N}$ der natürlichen Zahlen in die Menge $\mathbb{Z}$ der ganzen Zahlen eingebettet werden konnte, kann nun auch die Menge $\mathbb{Z}$ in die Menge $\mathbb{Q}$ der rationalen Zahlen eingebettet werden. Wir definieren die injektive Abbildung Φ durch

$$\Phi : \mathbb{Z} \longrightarrow \mathbb{Q} \;\; , \;\; g \longmapsto \overline{(g,1)} \;=\; \frac{g}{1} \, .$$

Durch Nachrechnen zeigt man:

Satz 1.8.3. *Die Abbildung $\Phi : \mathbb{Z} \to \mathbb{Q}$ ist relationstreu bzgl. der Operationen $+, \cdot$ und $<$, d. h.*

$$\begin{array}{ccc} g_1 + g_2 = g_3 & \Leftrightarrow & \Phi(g_1) + \Phi(g_2) = \Phi(g_3) \\ g_1 \cdot g_2 = g_3 & \Leftrightarrow & \Phi(g_1) \cdot \Phi(g_2) = \Phi(g_3) \\ g_1 < g_2 & \Leftrightarrow & \Phi(g_1) < \Phi(g_2) \, . \end{array}$$

Die Relation $<$ ist auf $\mathbb{Q}$ eine totale Ordnungsrelation.

Satz 1.8.4. *Für $\alpha, \beta, \gamma \in \mathbb{Q}$ gilt:*

(i) $\alpha < \beta \;\Leftrightarrow\; \alpha + \gamma < \beta + \gamma$,

(ii) $\alpha < \beta$ *und* $\gamma > 0 \;\Rightarrow\; \alpha \cdot \gamma < \beta \cdot \gamma$,

(iii) *Gilt* $0 < \alpha < \beta$, *so ist* $\beta^{-1} < \alpha^{-1}$.

Beweis.
Der Beweis für (i) sei zunächst für $\alpha, \beta, \gamma \in \mathbb{Z}$ geführt, also $\alpha = \overline{(n,m)}$, $\beta = \overline{(k,l)}$ und $\gamma = \overline{(i,j)}$ mit der Äquivalenzklassenbildung über $\mathbb{N} \times \mathbb{N}$. Dann ist $\alpha + \gamma = \overline{(n+i, m+j)}$ und $\beta + \gamma = \overline{(k+i, l+j)}$ mit den Bezeichnungen wie im vorigen Abschnitt. Nach Definition ist

$$\alpha < \beta \;\Leftrightarrow\; n + l < k + m \, .$$

Nach Satz 1.5.4 (Existenz der Differenz) gibt es ein $r \in \mathbb{N}$, so dass $k + m = n + l + r$ gilt:

$$\Rightarrow\; k + m + i + j = n + l + r + i + j \;\Rightarrow\; n + l + i + j < k + m + i + j \;\Rightarrow\; \alpha + \gamma < \beta + \gamma \, .$$

Rückrichtung: Aus $\alpha+\gamma<\beta+\gamma$ folgt wegen $\gamma\in\mathbb{Z}$ auch $\alpha+\gamma+(-\gamma)<\beta+\gamma+(-\gamma)$, also $\alpha<\beta$. Seien nun $\alpha=\overline{(p,q)}$, $\beta=\overline{(r,s)}$ und $\gamma=\overline{(t,u)}$ mit $q,s,u>0$ in $\mathbb{Q}$ gegeben, d. h. $p,q,r,s,t,u\in\mathbb{Z}$. Dann gilt

$$\begin{aligned}
\alpha<\beta \quad &\Leftrightarrow \quad p\cdot s &&< \quad r\cdot q\\
&\Leftrightarrow \quad p\cdot s\cdot u &&< \quad r\cdot q\cdot u\\
&\Leftrightarrow \quad psu+tqs &&< \quad rqu+tqs\\
&\Leftrightarrow \quad \overline{(psu+tqs,qsu)} &&< \quad \overline{(rqu+tqs,qsu)}\\
&\Leftrightarrow \quad \overline{(pu+tq,qu)} &&< \quad \overline{(ru+ts,su)} \quad \Leftrightarrow \quad \alpha+\gamma \;<\; \beta+\gamma\,.
\end{aligned}$$

Die Rückrichtung folgt wie im ganzzahligen Fall. Teil (iii) folgt wegen $\alpha=\overline{(p,q)}\Rightarrow\alpha^{-1}=\overline{(q,p)}$. □

Definition 1.8.5 (Der Betrag)**.** Für $\alpha\in\mathbb{Q}$ setzen wir

$$|\alpha| \;:=\; \begin{cases} \alpha & \text{falls} \quad 0\le\alpha \\ -\alpha & \text{falls} \quad \alpha<0 \end{cases}.$$

Satz 1.8.6. *Es seien $\alpha,\beta\in\mathbb{Q}$, dann gilt:*

(i) *Es ist stets $|\alpha|\ge 0$ und $|\alpha|=0\Leftrightarrow\alpha=0$,*

(ii) *es ist $|-\alpha|=|\alpha|$,*

(iii) *es gilt $|\alpha|\le|\beta| \;\Leftrightarrow\; -|\beta|\le\alpha\le|\beta|$.*

Beweis.
(i) und (ii) zeigt man sofort durch Unterscheidung der Fälle $0\le\alpha$ und $\alpha<0$. (iii) zeigt man durch Unterscheidung der vier Fälle

$$\begin{array}{lll} \text{a)} & 0\le\alpha & , \quad 0\le\beta \\ \text{b)} & 0\le\alpha & , \quad \beta<0 \\ \text{c)} & \alpha<0 & , \quad 0\le\beta \\ \text{d)} & \alpha<0 & , \quad \beta<0 \quad . \end{array}$$

□

Satz 1.8.7 (Dreiecksungleichung)**.** *Für alle $\alpha,\beta\in\mathbb{Q}$ gilt $|\alpha+\beta|\le|\alpha|+|\beta|$.*

Beweis.
Fall a):
Es gilt $0\le\alpha+\beta$ und daher $|\alpha+\beta|=\alpha+\beta=|\alpha|+|\beta|$.

Fall b):
Es gilt $\alpha+\beta=\alpha-|\beta|$, also $-|\beta|\le\alpha+\beta\le\alpha$, mit Satz 1.8.6 daher $|\alpha+\beta|\le\max(|\alpha|,|\beta|)$ bzw. $|\alpha+\beta|\le|\alpha|+|\beta|$.

Die weiteren Fälle sind ähnlich zu zeigen. □

Satz 1.8.8 (Gesetz des Archimedes)**.** *Zu $\alpha,\beta\in\mathbb{Q}$ mit $\alpha,\beta>0$ existiert stets ein $n\in\mathbb{N}$ mit $n\cdot\alpha>\beta$.*

Beweis.
Es sei $\alpha = \overline{(p,q)}$ und $\beta = \overline{(r,s)}$ mit $p,q,r,s > 0$ in $\mathbb{Z}$. Da $\mathbb{N}$ kein Maximum besitzt gibt es ein $n \in \mathbb{N}$ mit $n > q \cdot r$. Wegen $p \cdot s \geq 1$ folgt $n \cdot (p \cdot s) > q \cdot r$, das ist aber gleichbedeutend mit $\overline{(np,q)} > \overline{(r,s)}$, also mit $n \cdot \alpha > \beta$. □

1.9 Definition der reellen Zahlen mittels Fundamentalfolgen

Die grundlegende Idee für die Konstruktion der reellen Zahlen durch Fundamentalfolgen (oder Cauchyfolgen) wird im Cauchykriterium der Analysis sichtbar.

Definition 1.9.1. Eine rationale Zahlenfolge (kurz: Folge) ist eine Abbildung $S : \mathbb{N} \to \mathbb{Q},\ n \mapsto s_n$.

Definition 1.9.2 (Konvergenz rationaler Zahlenfolgen)**.** Es sei $S : \mathbb{N} \to \mathbb{Q},\ n \mapsto s_n$ eine rationale Zahlenfolge:

(i) Die Folge S konvergiert gegen $s \in \mathbb{Q}$ genau dann, wenn es für alle $\varepsilon > 0$ ($\varepsilon \in \mathbb{Q}$) ein $N \in \mathbb{N}$ gibt, so dass für alle $n \in \mathbb{N}$ mit $n > N$ gilt: $|s_n - s| < \varepsilon$.

(ii) S heißt (rationale) Nullfolge genau dann, wenn S gegen 0 konvergiert.

Definition 1.9.3 (Addition von rationalen Folgen)**.** Sind $S : \mathbb{N} \to \mathbb{Q},\ n \mapsto s_n$ und $T : \mathbb{N} \to \mathbb{Q},\ n \mapsto t_n$ rationale Folgen, so verstehen wir unter ihrer Summe $S + T$ (bzw. Differenz $S - T$) die Folgen $S + T :\ n \mapsto s_n + t_n$ (bzw. $S - T :\ n \mapsto s_n - t_n$).

Definition 1.9.4. Eine rationale Folge heißt Fundamentalfolge (oder Cauchyfolge), falls es für alle $\varepsilon > 0$ ($\varepsilon \in \mathbb{Q}$) ein $N \in \mathbb{N}$ gibt, so dass für alle $m, n \in \mathbb{N}$ mit $m, n > N$ gilt: $|s_m - s_n| < \varepsilon$.

Wir definieren nun wiederum die reellen Zahlen als Äquivalenzklassen von Fundamentalfolgen.

Definition 1.9.5. Zwei Fundamentalfolgen S und T heißen äquivalent (geschrieben $S \sim T$) genau dann, wenn $S - T$ eine Nullfolge bildet.

Satz 1.9.6. *Die Relation $\sim$ ist eine Äquivalenzrelation.*

Beweis.
Reflexivität:
Es ist $S - S :\ n \mapsto 0$, und nach Definition 1.9.2 konvergiert $S - S$ gegen 0.

Symmetrie:
Es seien $S : n \mapsto s_n$ und $T :\ n \mapsto t_n$ Fundamentalfolgen. Sei $S - T :\ n \mapsto s_n - t_n$ eine Nullfolge. Damit gibt es für alle $\varepsilon > 0$ ($\varepsilon \in \mathbb{Q}$) ein $N \in \mathbb{N}$, so dass $|s_n - t_n| < \varepsilon$ für alle $n > N$. Dann ist aber auch $|t_n - s_n| = |s_n - t_n| < \varepsilon$ für alle $n > N$, also ist auch $T - S :\ n \mapsto t_n - s_n$ eine Nullfolge.

Transitivität:
Es seien $S : n \mapsto s_n$, $T :\ n \mapsto t_n$ und $U :\ n \mapsto u_n$ Fundamentalfolgen, sowie $S \sim T$ und $T \sim U$, also $T - S :\ n \mapsto t_n - s_n$ und $U - T :\ n \mapsto u_n - t_n$ Nullfolgen. Wir zeigen, dass auch $U - S$ eine Nullfolge ist, und damit $S \sim U$. Es sei $\varepsilon > 0$ ($\varepsilon \in \mathbb{Q}$) beliebig. Dann gibt es $N_1, N_2 \in \mathbb{N}$, so dass $|t_n - s_n| < \frac{1}{2}\varepsilon$ für alle $n > N_1$ und $|u_n - t_n| < \frac{1}{2}\varepsilon$ für alle $n > N_2$ ist. Es sei $N_3 := \max(N_1, N_2)$. Für ein $n > N_3$ ist dann nach der Dreiecksungleichung die Ungleichung $|u_n - s_n| \leq |u_n - t_n| + |t_n - s_n| < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon$, also ist $U - S$ eine Nullfolge. □

Definition 1.9.7. Die Menge $\mathbb{R}$ der reellen Zahlen ist die Menge aller Äquivalenzklassen bzgl. dieser Äquivalenzrelation. Die Äquivalenzklasse, der eine Fundamentalfolge S angehört, bezeichnen wir mit $\overline{S}$.

Die Menge der reellen Zahlen kann nun wieder durch Einführung der Rechenoperationen Addition und Multiplikation, sowie einer Anordnung, zu einem angeordneten Körper gemacht werden.

Lemma 1.9.8. *Es gilt:*

(i) *Die (elementweise) Summe zweier Nullfolgen ist eine Nullfolge.*

(ii) *Das (elementweise) Produkt einer Fundamentalfolge mit einer Nullfolge ist eine Nullfolge.*

(iii) *Eine Fundamentalfolge S ist beschränkt, d. h. es gibt $G \in \mathbb{Q}$, so dass $|s_n| < G$ für alle $n \in \mathbb{N}$ ist.*

(iv) *Ist S keine Nullfolge, so gibt es eine Fundamentalfolge T mit $S \cdot T \sim (1, 1, \ldots)$.*

Beweis.
Diese Aussagen zeigt man wie die entsprechenden Aussagen aus Analysis I für reelle Zahlenfolgen. Wir zeigen exemplarisch (iv): Es sei $S :\mapsto s_n$ keine Nullfolge, d. h. es gibt ein $\varepsilon_0 > 0$ ($\varepsilon_0 \in \mathbb{Q}$), so dass für alle $N \in \mathbb{N}$ ein $n \geq N$ existiert mit $|s_n| \geq \varepsilon_0$. Weiter gibt es für $\varepsilon_1 := \frac{1}{2}\varepsilon_0$ ein N_0, so dass $|s_m - s_n| < \varepsilon_1$ für alle $m, n \geq N_0$. Sei $d \geq N_0$ so gewählt, dass $|s_d| \geq \varepsilon_0$. Dann ist nach der Dreiecksungleichung $|s_n| \geq |s_d| - |s_d - s_n| > \varepsilon_0 - \varepsilon_1 = \varepsilon_1$ für alle $n \geq N_0$. Wir definieren die Folge $T = (t_n)$ durch:

$$t_n := \begin{cases} 1 & \text{falls} \quad n < N_0 \\ s_n^{-1} & \text{sonst} \end{cases}$$

Für alle $m, n \geq N_0$ gilt dann:

$$|t_m - t_n| = |s_m^{-1} - s_n^{-1}| = |s_m - s_n| \cdot |s_m s_n|^{-1} \leq \frac{1}{\varepsilon_1^2} \cdot |s_m - s_n| .$$

Also ist T eine Fundamentalfolge. Wegen $s_n t_n = 1$ für $n \geq N_0$ folgt $S \cdot T \sim (1, 1, \ldots)$. □

Wir definieren im Folgenden Addition, Multiplikation und Anordnung auf $\mathbb{R}$. Dazu seien $S :\ n \mapsto s_n$ und $T :\ n \mapsto t_n$ Fundamentalfolgen:

Definition 1.9.9. Wir setzen

$$\overline{S} + \overline{T} := \overline{S + T} \quad , \quad \overline{S} \cdot \overline{T} := \overline{S \cdot T} \quad , \quad \overline{S} < \overline{T} \Leftrightarrow \exists \varepsilon > 0 \, \exists n_0 \, \forall n \geq n_0 \in \mathbb{N} :\ t_n - s_n > \varepsilon .$$

Wir definieren zudem die Null $\overline{0} := \overline{(0, 0, \ldots)}$ und die Eins $\overline{1} := \overline{(1, 1, \ldots)}$.

Satz 1.9.10. $(\mathbb{R}, +, \cdot, \overline{0}, \overline{1})$ *ist ein Körper.*

Beweis.
Wir zeigen zunächst die Wohldefiniertheit von Addition, Multiplikation und Anordnung. Dazu seien im Folgenden

$$\begin{array}{ll} S_1 :\ n \mapsto s_n^{(1)} & S_2 :\ n \mapsto s_n^{(2)} \\ T_1 :\ n \mapsto t_n^{(1)} & T_2 :\ n \mapsto t_n^{(2)} \end{array}$$

Fundamentalfolgen, so dass $S_1 \sim S_2$ und $T_1 \sim T_2$ gilt, d. h. $S_1 - S_2$ und $T_1 - T_2$ sind Nullfolgen.

Addition:
Dann ist auch

$$(S_2 + T_2) - (S_1 + T_1) \ : \ n \mapsto (s_n^{(2)} + t_n^{(2)}) - (s_n^{(1)} + t_n^{(1)}) \ = \ (s_n^{(2)} - s_n^{(1)}) + (t_n^{(2)} - t_n^{(1)})$$

eine Nullfolge nach Lemma 1.9.8.

Multiplikation:
Es ist $s_n^{(2)} t_n^{(1)} - s_n^{(1)} t_n^{(1)} = (s_n^{(2)} - s_n^{(1)}) \cdot t_n^{(1)}$ nach Lemma 1.9.8 eine Nullfolge. Also ist $S_1 T_1 \sim S_2 T_1$. Ebenso zeigt man $S_2 T_1 \sim S_2 T_2$. Wegen der Transitivität von $\sim$ folgt $S_1 T_1 \sim S_2 T_2$.

Anordnung:
Wir betrachten die Bedingungen U_j $(j = 1, 2)$:

$$U_j : \ \exists \varepsilon > 0, N \in \mathbb{N} : \ t_n^{(j)} - s_n^{(j)} > \varepsilon \ (\forall n > N) \, .$$

Wir zeigen: Aus der Bedingung U_1 folgt die Bedingung U_2. Denn es gibt ein N, so dass

$$\begin{array}{rcll} |t_n^{(1)} - t_n^{(2)}| & < & \frac{1}{4}\varepsilon & \text{und} \\ |s_n^{(1)} - s_n^{(2)}| & < & \frac{1}{4}\varepsilon & \forall n > N \end{array}$$

gilt. Dann ist $t_n^{(1)} - s_n^{(1)} = (t_n^{(1)} - t_n^{(2)}) + (t_n^{(2)} - s_n^{(2)}) + (s_n^{(2)} - s_n^{(1)})$. Annahme: Es gibt $n > N$, so dass $|t_n^{(2)} - s_n^{(2)}| \leq \frac{1}{4}\varepsilon$ ist. Dann gilt nach der Dreiecksungleichung $|t_n^{(1)} - s_n^{(1)}| < \frac{1}{4}(\varepsilon + \varepsilon + \varepsilon) < \varepsilon$, ein Widerspruch. Also gilt $t_n^{(2)} - s_n^{(2)} > \frac{1}{4}\varepsilon$.

Die Rechenregeln von $\mathbb{Q}$ übertragen sich auf $\mathbb{R}$. Die Neutralitätseigenschaften von $\overline{0}$ und $\overline{1}$ bestätigt man durch Nachrechnen. Das Negative von $\alpha = \overline{S}$ ist $-\alpha := \overline{(0, 0, \ldots) - S}$. Die Existenz des multiplikativen Inversen folgt aus Lemma 1.9.8(iv). □

Als nächstes führen wir wiederum die Einbettung von $\mathbb{Q}$ in $\mathbb{R}$ ein. Den Nachweis der Ordnungseigenschaften verschieben wir auf die Übungsaufgaben. Wir definieren die injektive Abbildung Φ durch

$$\Phi : \ \mathbb{Q} \to \mathbb{R} \ \ r \mapsto \overline{(r, r, \ldots)} \, .$$

Satz 1.9.11. *Die Abbildung Φ ist relationstreu (im Sinn von Satz 1.8.3).*

Beweis.
Durch Nachrechnen. □

Damit ist der Körper der reellen Zahlen vollständig konstruiert. Man kann nun versuchen, den Prozess der Vergrößerung durch Fundamentalfolgen, der von $\mathbb{Q}$ zu $\mathbb{R}$ geführt hat, fortzusetzen, und durch Betrachtung von reellen Fundamentalfolgen zu einem noch größerem Bereich zu gelangen. Eine solche Vergrößerung ist wegen der Vollständigkeitseigenschaft der reellen Zahlen nicht mehr möglich. Die Vollständigkeitseigenschaft der reellen Zahlen, auch als Cauchykriterium bekannt, die von fundamentaler Wichtigkeit für den Aufbau der Analysis ist, besagt: Jede reelle Fundamentalfolge (Cauchyfolge) konvergiert.

Wir beginnen mit dem Begriff des Betrags:

Definition 1.9.12 (Betrag)**.** Für alle $\alpha \in \mathbb{R}$ setzen wir:

$$|\alpha| := \begin{cases} \alpha & \text{falls} \quad 0 \leq \alpha \\ -\alpha & \text{falls} \quad \alpha < 0 \end{cases}.$$

Satz 1.9.13. *Es seien $\alpha, \beta \in \mathbb{R}$, dann gilt:*

(i) *Es ist stets $|\alpha| \geq 0$ und $|\alpha| = 0 \Leftrightarrow \alpha = 0$,*

(ii) *es ist $|-\alpha| = |\alpha|$,*

(iii) $|\alpha| \leq |\beta| \Leftrightarrow -|\beta| \leq \alpha \leq |\beta|$.

Beweis.
Wie Satz 1.8.6 durch Fallunterscheidung. □

Satz 1.9.14 (Dreiecksungleichung)**.** *Für $\alpha, \beta \in \mathbb{R}$ gilt: $|\alpha + \beta| \leq |\alpha| + |\beta|$.*

Beweis.
Durch Fallunterscheidung. □

Satz 1.9.15. *Es sei $\varepsilon > 0$ ($\varepsilon \in \mathbb{Q}$) und $\alpha = \overline{S}$ mit einer rationalen Fundamentalfolge $S : n \mapsto s_n$. Dann gibt es ein $N \in \mathbb{N}$, so dass für alle $m \geq N$ gilt: $|\alpha - s_m| < \varepsilon$.*

Beweis.
Es sei $\varepsilon_1 > 0$ ($\varepsilon_1 \in \mathbb{Q}$) gegeben. Es ist $\alpha - s_m = \overline{(s_1 - s_m, s_2 - s_m, \ldots)}$. Wir haben zu zeigen, dass $\overline{(s_1 - s_m, s_2 - s_m, \ldots)} < \varepsilon_1$ und $\overline{(s_m - s_1, s_m - s_2, \ldots)} < \varepsilon_1$ gilt. Die beiden Ungleichungen sind nach Definition der Anordnung äquivalent dazu, dass es ein $\varepsilon_2 > 0$ gibt, so dass

$$(*)\ s_m + \varepsilon_1 - s_n > \varepsilon_2 \text{ und } s_n + \varepsilon_1 - s_m > \varepsilon_2$$

für genügend große n gilt. Wir setzen $\varepsilon_2 := \frac{1}{2}\varepsilon_1$. Da S eine Fundamentalfolge ist, gibt es $N \in \mathbb{N}$, so dass $|s_m - s_n| < \varepsilon_2$ ist für alle $m, n \geq N$. Das ist gleichbedeutend mit $-\varepsilon_2 < s_n - s_m < \varepsilon_2$. Damit gilt

$$\begin{array}{lllll} s_m + \varepsilon_1 - s_n & > & \varepsilon_1 - \varepsilon_2 & = & \varepsilon_2 \quad \text{und} \\ s_n + \varepsilon_1 - s_m & > & \varepsilon_1 - \varepsilon_2 & = & \varepsilon_2 \end{array}$$

und damit $(*)$. □

Definition 1.9.16. Eine reelle Zahlenfolge S ist eine Abbildung $S : \mathbb{N} \to \mathbb{R},\ n \mapsto \alpha_n$.

Definition 1.9.17 (Konvergenz reeller Zahlenfolgen)**.** Sei S eine reelle Zahlenfolge $S : \mathbb{N} \to \mathbb{R},\ n \mapsto \alpha_n$.

(i) S konvergiert gegen $\alpha \in \mathbb{R}$ genau dann, wenn es für alle $\varepsilon > 0$ ($\varepsilon \in \mathbb{R}$) ein $N \in \mathbb{N}$ gibt, so dass für alle $n > N$ gilt: $|\alpha_n - \alpha| < \varepsilon$.

(ii) S heißt Fundamentalfolge (oder Cauchyfolge) genau dann, wenn es zu jedem $\varepsilon > 0$ ($\varepsilon \in \mathbb{R}$) ein $N \in \mathbb{N}$ gibt, so dass für alle $m, n > N$ gilt: $|\alpha_m - \alpha_n| < \varepsilon$.

Satz 1.9.18. *Jede reelle Cauchyfolge konvergiert.*

Beweis.
Es sei $\Sigma : \mathbb{N} \to \mathbb{R},\ n \mapsto \sigma_n$ eine reelle Fundamentalfolge mit reellen Elementen $\sigma_n = \overline{S_n}$ und deren Vertretern S_n. Dann sind die S_n rationale Fundamentalfolgen $S_n : \mathbb{N} \to \mathbb{Q},\ p \mapsto s_{p,n}$. Zu jedem σ_n gibt es ein $p_n \in \mathbb{N}$ mit

$$(*) : |\sigma_n - s_{p_n,n}| < \frac{1}{3n} .$$

Wir betrachten die Folge $S : \mathbb{N} \to \mathbb{Q},\ n \mapsto s_{p_n,n}$ und zeigen, dass es sich um eine (rationale) Fundamentalfolge handelt. Dazu sei $\varepsilon > 0$ ($\varepsilon \in \mathbb{Q}$) gegeben. Dann gibt es $N_0 \in \mathbb{N}$, so dass für alle $m, n > N_0$ gilt: $|\sigma_m - \sigma_n| < \frac{1}{3}\varepsilon$, und es gibt nach dem Gesetz des Archimedes ein $N_1 \in \mathbb{N}$, so dass für alle $n > N_1$ gilt: $\frac{1}{n} < \frac{1}{3}\varepsilon$. Für $m, n > N_1$ gilt dann nach der Dreiecksungleichung und $(*)$ auch

$$|s_{p_m,m} - s_{p_n,n}| < |s_{p_m,m} - \sigma_m| + |\sigma_m - \sigma_n| + |\sigma_n - s_{p_n,n}| < \varepsilon .$$

Also konvergiert die Folge S gegen eine reelle Zahl σ. Wir zeigen, dass auch die Folge Σ gegen σ konvergiert. Es sei $\varepsilon > 0$ ($\varepsilon \in \mathbb{R}$) gegeben. Nach Satz 1.9.15 gibt es ein $N \in \mathbb{N}$, so dass für alle $m \geq N$ gilt: $|\sigma - s_{p_m,m}| < \frac{1}{2}\varepsilon$ und $\frac{1}{3m} < \frac{1}{2}\varepsilon$. Dann folgt nach $(*)$ für alle $m \geq N$:

$$|\sigma_m - \sigma| \leq |\sigma_m - s_{p_m,m}| + |s_{p_m,m} - \sigma| < \varepsilon .$$

Damit konvergiert Σ gegen σ. □

1.10 Definition der reellen Zahlen mittels Dedekindscher Schnitte

Die Methode von Dedekind wollen wir nur skizzieren. Die Heuristik ist die Folgende: Jede reelle Zahl zerlegt die rationale Zahlengerade in zwei Klassen:

die Unterklasse $U = \{r \in \mathbb{Q} \mid r < \alpha\}$ und
die Oberklasse $O = \{r \in \mathbb{Q} \mid \alpha \leq r\}$.

Wir arbeiten nur mit Unterklassen.

Definition 1.10.1. Eine reelle Zahl α ist eine Menge α mit:

(i) $\emptyset \subset \alpha \subset \mathbb{Q}$,

(ii) α besitzt kein Maximum, d. h. zu jedem $r \in \alpha$ gibt es $s \in \alpha$ mit $s > r$,

(iii) α ist „abgeschlossen nach unten“, d. h. $s \in \alpha$ und $r < s$ impliziert $r \in \alpha$.

$\mathbb{R}$ ist die Menge dieser reellen Zahlen. Wir definieren zunächst die Anordnung auf $\mathbb{R}$:

Definition 1.10.2. $\alpha \leq \beta$ bedeutet $\alpha \subseteq \beta$, und $\alpha < \beta$ bedeutet $\alpha \subset \beta$.

Lemma 1.10.3. *Für jedes reelle α gilt:*

(i) *$a \in \alpha$ und $a' \in \alpha^c$ impliziert $a < a'$,*

(ii) *$a' \in \alpha^c$ und $a' < x$ impliziert $x \in \alpha^c$.*

Beweis.
Das folgt durch Widerspruch, denn aus $a' \leq a \in \alpha$ folgt $a' \in \alpha$ nach Definition 1.10.1. Ebenso impliziert $a' < x \in \alpha$ auch $a' \in \alpha$. □

Die Anordnung bildet eine totale Ordnungsrelation:

Satz 1.10.4. *Seien $\alpha, \beta, \gamma \in \mathbb{R}$. Die Relation $\leq$ erfüllt die folgenden Eigenschaften:*

(i) *Vergleichbarkeit: Es gilt genau eine der drei Aussagen $\alpha < \beta$, $\alpha = \beta$, $\alpha > \beta$,*

(ii) *Transitivität: Aus $\alpha \leq \beta$ und $\beta \leq \gamma$ folgt $\alpha \leq \gamma$,*

(iii) *Antisymmetrie: Gilt $\alpha \leq \beta$ und $\beta \leq \alpha$, so ist $\alpha = \beta$.*

(iv) *Reflexivität: Es gilt $\alpha \leq \alpha$.*

Beweis.
Zu (i): Es ist klar, dass die drei Relationen einander ausschließen. Sei $\alpha \not\leq \beta$, dann gibt es ein $x \in \alpha \cap \beta^c$. Wegen Lemma 1.10.3 folgt $\forall b \in \beta : \ b < x \in \alpha$, also $\beta \subseteq \alpha$ bzw. $\beta \leq \alpha$. Die beiden anderen Behauptungen folgen aus den entsprechenden Eigenschaften der Relation $\subseteq$. □

Während die Vollständigkeit von $\mathbb{R}$ bei der Einführung mittels Fundamentalfolgen relativ schwierig zu zeigen ist, ergibt sie sich bei der Einführung durch Dedekindsche Schnitte ziemlich leicht, allerdings in Form des Supremumprinzips. Aus der Analysis weiß man, dass daraus das Cauchyprinzip abgeleitet werden kann.

Definition 1.10.5. Mit der Definition der reellen Zahlen nach 1.10.1 setzen wir:

(i) Ein $\gamma \in \mathbb{R}$ heißt obere Schranke von $M \subseteq \mathbb{R}$ genau dann wenn gilt: $\forall \alpha \in M : \ \alpha \leq \gamma$.

(ii) $\gamma \in \mathbb{R}$ heißt Supremum oder kleinste obere Schranke von M, falls gilt: $\gamma \leq \beta$ für alle oberen Schranken β von M.

(iii) Eine Menge $M \subseteq \mathbb{R}$ heißt nach oben beschränkt, wenn sie eine obere Schranke besitzt.

Satz 1.10.6 (Vollständigkeit von $\mathbb{R}$)**.** *Jede nichtleere und nach oben beschränkte Menge $M \subseteq \mathbb{R}$ besitzt eine (eindeutig bestimmte) kleinste obere Schranke.*

Beweis.
Es sei $M \subseteq \mathbb{R}$ und $\beta^* \in \mathbb{R}$ eine obere Schranke von M. Jedes $\alpha \in M$ ist nach Definition 1.10.2 eine Teilmenge von $\mathbb{Q}$. Wir definieren $\gamma \subseteq \mathbb{Q}$ durch

$$\gamma := \bigcup_{\alpha \in M} \alpha$$

und zeigen zunächst, dass γ eine reelle Zahl (nach Definition 1.10.1) ist, und dann, dass γ ein Supremum von M ist (die Eindeutigkeit des Supremums ist klar). Es ist $\gamma \in \mathbb{R}$, denn

(i) $\gamma \neq \emptyset$ wegen $\exists \alpha \in M \ \Rightarrow \ \emptyset \subset \alpha \subset \gamma$.

(ii) Es ist $\gamma \neq \mathbb{Q}$, denn aus $\forall \alpha \in M : \ \alpha \subseteq \beta^*$ folgt $\gamma \subseteq \beta^* \neq \mathbb{Q}$.

(iii) γ ist nach unten abgeschlossen:
$s < r \in \gamma \ \Rightarrow \ \exists \alpha \in M : \ s \in \alpha \ \Rightarrow \ \exists r \in \alpha : \ s < r \in \alpha \subseteq \gamma \ \Rightarrow \ s \in \gamma$.

(iv) γ ist eine obere Schranke von M, denn $\forall \alpha \in M : \ \alpha \subseteq \gamma$.

(v) γ ist die kleinste obere Schranke von M, denn für jede obere Schranke β von M gilt:

$$\gamma = \bigcup_{\alpha \in M} \alpha \subseteq \bigcup_{\alpha \in M} \beta = \beta$$

und damit $\gamma \leq \beta$.

□

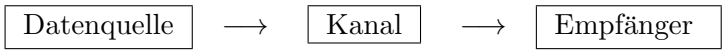
Im Gegensatz zur Vollständigkeit sind Addition und Multiplikation schwierig zu begründen. Wir verschieben dies auf die Übungen.

Kapitel 2

Fehlerkorrigierende Codes

2.1 Allgemeines

Codes werden bei der Übertragung und Umsetzung von Nachrichten aller Art verwendet. Das grundlegende Modell der Nachrichtenübertragung kann wie folgt dargestellt werden:



Viele Kanäle haben dabei folgende Eigenschaften:

(i) Übertragen werden Zeichenfolgen aus Zeichen eines endlichen Alphabets.

(ii) Nur gewisse Zeichenfolgen sind Codeworte, andere nicht.

Das Problem der Theorie der fehlerkorrigierenden Codes ist, Fehler in der Übertragung zu erkennen und zu korrigieren. Dies ist natürlich nur dann möglich, wenn nicht zu viele Fehler bei der Übertragung auftreten.

Beispiel 2.1.1. Ein Absender kommuniziert mit einem Partner per eMail. Der Empfänger erhält folgende Botschaft:

Ich möchte Sie morgen um zehn Uhr in `Stuttgalt` treffen.

Bei der Übertragung ist ein Fehler passiert. Das verwendete Alphabet sind die lateinischen Buchstaben, die Codeworte die Worte der deutschen Sprache. Anstelle des Codeworts `Stuttgart` wurde das Nicht-Codewort `Stuttgalt` empfangen. Dieser Fehler ist jedoch korrigierbar: Erstens kann der Fehler erkannt werden, denn es gibt kein Wort `Stuttgalt` in der deutschen Sprache. Zweitens kann der Fehler korrigiert werden, denn es gibt ein Codewort - und auch nur eines, nämlich `Stuttgart` - das sich von der gesendeten Zeichenfolge nur an einer Stelle unterscheidet. Wir sagen: die beiden Zeichenfolgen haben Hammingabstand 1. Wir betrachten ein anderes Beispiel: Der Absender möchte dem Empfänger folgende Botschaft schicken:

Ich möchte Sie Anfang `Juni` besuchen.

Bei der Übertragung entsteht ein einziger Fehler, und der Empfänger erhält:

Ich möchte Sie Anfang `Juli` besuchen.

Hier kann der Fehler weder erkannt noch korrigiert werden: Die Codeworte `Juni` und `Juli` besitzen den Hammingabstand 1, sie können durch einen einzigen Fehler ineinander übergehen.

Aus diesem Beispiel ergibt sich eine erste Forderung an einen guten Code: Um Fehler bei der Übertragung erkennen und möglichst korrigieren zu können, sollten zwei verschiedene Codeworte einen nicht zu kleinen Hammingabstand haben.

Beispiel 2.1.2. Die einfachsten Codes, die dies erreichen, sind die so genannten Repetition-Codes, deren Codeworte durch mehrfache Aneinanderhängung der ursprünglichen Buchstaben der Nachricht gebildet werden. Der Nachricht `Juni` wird im Falle eines Repetition-Codes der Ordnung 3 der Code `JJJuuunnniii` zugeordnet. Wird nun die Nachricht `JJJuuunlniii` empfangen, so kann der Fehler erkannt werden, denn `JJJuuunlniii` ist kein Codewort (Codeworte sind dann die Wörter der deutschen Sprache, in denen jeder Buchstabe dreimal geschrieben wird). Er kann auch korrigiert werden: Das Codewort zur Nachricht `Juli` ist `JJJuuullliii`, es besitzt Hammingabstand 2 zum empfangenen Wort. Das Codewort zur Nachricht `Juni` ist dagegen `JJJuuunnniii`, es besitzt den Hammingabstand 1. Der Empfänger wird also schließen, dass die ursprüngliche Nachricht `Juni` gewesen ist.

Die Repetition-Codes haben den Nachteil, dass sie sehr ineffizient sind. Im obigen Beispiel wird beim Repetition-Code der Ordnung 3 pro Zeichen nur ein Drittel der Information übertragen. Der Repetition-Code enthält im Verhältnis zu seiner Länge nur wenige Codeworte. Gute Codes sollten folgende Eigenschaften haben:

1. Um möglichst viel Information übertragen zu können, sollte der Code möglichst viele Codeworte (im Vergleich zur Länge) enthalten.
2. Um Fehler bei der Übertragung nachweisen und korrigieren zu können, sollten zwei verschiedene Codeworte einen nicht zu kleinen Hammingabstand besitzen.
3. Die Fehlerkorrektur sollte einfach durchführbar sein.

2.2 Grundlegende Sätze und Definitionen

Wir betrachten im Folgenden nur Codes, in denen alle Codeworte die gleiche Länge besitzen (so genannte Blockcodes).

Definition 2.2.1. Es X eine endliche Menge. Der Hammingabstand auf dem Kartesischen Produkt X^n zweier Tupel $\vec{x} = (x_1, \ldots, x_n)$ und $\vec{y} = (y_1, \ldots, y_n)$ ist definiert durch

$$d(\vec{x}, \vec{y}) \;=\; |\{i \mid x_i \neq y_i\}| \,.$$

Satz 2.2.2. *Der Hammingabstand ist eine Metrik auf X^n, d. h. es gilt*

$$\begin{aligned} d(\vec{x}, \vec{z}) &\leq\; d(\vec{x}, \vec{y}) + d(\vec{y}, \vec{z}) \;, \\ d(\vec{x}, \vec{y}) &=\; 0 \quad \Leftrightarrow \quad \vec{x} = \vec{y} \;. \end{aligned}$$

Beweis.
Es seien $\vec{x} = (x_1, \ldots, x_n)$, $\vec{y} = (y_1, \ldots, y_n)$ und $\vec{z} = (z_1, \ldots, z_n)$ Elemente von X^n. Gilt $z_i \neq x_i$, so ist notwendigerweise $y_i \neq x_i$ oder $z_i \neq y_i$, also gilt $\{i \mid z_i \neq x_i\} \subseteq \{i \mid y_i \neq x_i\} \cup \{i \mid z_i \neq y_i\}$. und damit die erste Aussage. Die zweite Aussage ist trivial. □

Definition 2.2.3. Es sei X eine endliche Menge der Mächtigkeit $q \in \mathbb{N}$. Ein $(n, M, d; q)$-Code (oder kürzer (n, M, d)-Code, wenn q klar ist) auf X ist eine M-elementige Teilmenge $C \subseteq X^n$, für die $d(\vec{x}, \vec{y}) \geq d$ für alle $\vec{x}, \vec{y} \in C$ mit $\vec{x} \neq \vec{y}$ gilt, und für die es auch $\vec{x}, \vec{y} \in C$ mit $d(\vec{x}, \vec{y}) = d$ gibt. Die Zahl d heißt Minimalabstand des Codes C.

Beispiel 2.2.4. Sei X eine nichtleere Menge mit m Elementen. Der Repetition-Code der Ordnung n über X wird mit RP_n bezeichnet. Er ist ein $(n, m, n; m)$-Code: dem Buchstaben $x \in X$ wird das Codewort $(x, x, \ldots, x) \in X^n$ zugeordnet, d. h. $\mathrm{RP}_n = \{(x, \ldots, x) \in X^n\}$. Sein Minimalabstand ist gerade n, denn verschiedene Codeworte unterscheiden sich in allen n Stellen voneinander (dies ist auch der einzig mögliche Hammingabstand von verschiedenen Wörtern in diesem Code).

Beispiel 2.2.5. Sei $X = \{0, 1\}$, d. h. $q = 2$, und $n \in \mathbb{N}$ beliebig mit $n \geq 2$. Der Parity-Check-Code der Ordnung n über X wird mit PC_n bezeichnet. Er ist definiert über $\mathrm{PC}_n = \{(x_1, \ldots, x_n) \mid \sum x_i \text{ gerade}\}$. PC_n ist ein $(n, 2^{n-1}, 2; 2)$-Code. Sein Minimalabstand ist 2, da er die Wörter $(0, \ldots, 0, 1, 1)$ und $(0, \ldots, 0, 0, 0)$ enthält, andererseits aber keine zwei Wörter den Hammingabstand 1 besitzen können, da sonst mindestens eines eine ungerade Quersumme (d. h. „ungerade Parität") besitzen würde. Zur Ordnung n besitzt er 2^{n-1} Codeworte, da X^n genau 2^n Tupel enthält, von denen genau die Hälfte eine gerade Quersumme besitzt.

In gewissem Sinne sind RP-Codes und PC-Codes Extremfälle:

1. RP-Codes besitzen den höchstmöglichsten Mindestabstand $d = n$, bieten aber nur m Codeworte bei m^n möglichen Wörtern in X^n.
2. PC-Codes haben den kleinstmöglichen noch sinnvollen Mindestabstand 2, schöpfen mit 2^{n-1} Codeworten den maximal möglichen Raum von 2^n Worten aber gut aus.
3. Beide Codes sind sehr einfach zu dekodieren.

Eine Möglichkeit der Abstufung zwischen diesen beiden Extremen bieten die Hamming-Codes:

Beispiel 2.2.6 (Der $(7, 4)$-Hamming-Code)**.** Es sei $X = \{0, 1\}$. Wir wählen $n = 7$ als Länge für die Codeworte und setzen $C = \{\vec{x} = (x_1, \ldots, x_7) \in X^n \mid (*)\}$ mit

$$(*): \begin{cases} (1) & x_2 + x_3 + x_4 + x_5 & & & \text{ist} & \text{gerade} \\ (2) & x_1 + x_3 + x_4 & + x_6 & & \text{ist} & \text{gerade} \\ (3) & x_1 + x_2 & + x_4 & + x_7 & \text{ist} & \text{gerade} \end{cases}$$

Wir bestimmen zunächst die Mächtigkeit M des Codes C. Sind $x_1, \ldots, x_4 \in \{0, 1\}$ vorgegeben, so sind dadurch x_5, x_6, x_7 eindeutig bestimmt (man sagt, dass der Code 4 Informationsbits und 3 Prüfbits pro Codewort besitzt). Also ist $M = 16$. Als nächstes bestimmen wir den Minimalabstand. Es seien $\vec{x} = (x_1, \ldots, x_7)$ und $\vec{y} = (y_1, \ldots, y_7)$ mit $\vec{x} \neq \vec{y}$ in C gegeben. Dann muss $x_i \neq y_i$ für mindestens ein $i \in \{1, 2, 3, 4\}$ gelten, d. h. $\vec{x}$ und $\vec{y}$ müssen sich in mindestens einer der ersten vier Koordinaten unterscheiden. Es sei $d^* = |\{1 \leq i \leq 4 \mid x_i \neq y_i\}|$.

Fall 1: $d^* = 1$
Da jede der vier Variablen $x_1, \ldots, x_4$ in mindestens zwei Gleichungen vorkommt, muss $x_i \neq y_i$ für mindestens zwei Werte $i \in \{5, 6, 7\}$ sein.

Fall 2: $d^* = 2$
Es sei also $x_i \neq y_i$ und $x_j \neq y_j$ für verschiedene $i, j \in \{1, 2, 3, 4\}$. Es gibt eine unter den Gleichungen aus $(*)$, in der nur eine der beiden Variablen x_i, y_i auftritt. In dieser Gleichung kommt auch x_k mit $k \in \{5, 6, 7\}$ vor. Dann muss auch $x_k \neq y_k$ gelten.

Man sieht: der Minimalabstand ist $d \geq 3$. Andererseits sind $\vec{x} = (0, \ldots, 0)$ und $\vec{y} = (1, 1, 1, 0, 0, 0, 0)$ beide in C mit Abstand $d(\vec{x}, \vec{y}) = 3$. Der Minimalabstand ist also $d = 3$, und der $(7, 4)$-Hamming-Code ist also ein $(7, 16, 3; 2)$-Code. Wir werden später diese Tatsache sehr viel einfacher erhalten, indem wir die Zeichen als Elemente eines Körpers $\mathbb{F}_2 = \{0, 1\}$ und damit X^7 als Vektorraum über $\mathbb{F}_2$ ansehen. Der $(7, 4)$-Hamming-Code wird sich als linearer Code, d. h. als Untervektorraum von X^7 erweisen.

Als nächstes diskutieren wir den Zusammenhang zwischen Minimalabstand und Fehlerkorrektur. Wird bei der Übertragung einer Nachricht eine Zeichenfolge $\vec{x}' \notin C$ empfangen, die kein Codewort ist, so wird sie bei der Korrektur durch ein Codewort $\vec{x}_1 \in C$ ersetzt, für das $d(\vec{x}_1, \vec{x}')$ minimal ist. Der Minimalabstand des Codes C sei $d = 2e + 1$ für ein $e \in \mathbb{N}$, und es seien bei der Übertragung von $\vec{x}$ eine Anzahl $\leq e$ Fehler passiert, d. h. $d(\vec{x}_1, \vec{x}') \leq e$. Ist $\vec{x}_2 \in C$ ein von $\vec{x}_1$ verschiedenes Codewort, so gilt:

$$2e + 1 \;\leq\; d(\vec{x}_1, \vec{x}_2) \;\leq\; d(\vec{x}_1, \vec{x}') + d(\vec{x}', \vec{x}_2)$$

und damit $d(\vec{x}', \vec{x}_2) \geq (2e+1) - d(\vec{x}_1, \vec{x}') \geq e+1$. Damit wird bei dieser Art von Decodierung $\vec{x}'$ durch das richtige Codewort $\vec{x}_1$ ersetzt. Diese Beobachtung lässt die folgende Definition sinnvoll erscheinen:

Definition 2.2.7. Ein (n, M, d)-Code mit $d = 2e + 1$ heißt ein e-fehlerkorrigierender Code

Beispiel 2.2.8. Der betrachtete Hamming-Code ist ein 1-fehlerkorrigierender Code, da der Minimalabstand $d = 3 = 2 \cdot 1 + 1$ ist. Es werde beispielsweise an Stelle des Codeworts $\vec{x}_1 = (1, 1, 1, 0, 0, 0, 0)$ die Zeichenfolge $\vec{x}' = (1, 1, 1, 0, 1, 0, 0)$ empfangen, d. h. an der 5. Stelle tritt ein Fehler auf. Es ist $d(\vec{x}_1, \vec{x}') = 1$. Nach der obigen Diskussion ist damit $\vec{x}_1$ das eindeutig bestimmte Codewort mit minimalem Hammingabstand zu $\vec{x}'$. Bei der Fehlerkorrektur wird damit $\vec{x}'$ durch das richtige Codewort $\vec{x}_1$ ersetzt. Wir werden später diskutieren, wie die Fehlerkorrektur mit wenig Rechenaufwand durchgeführt werden kann.

Im betrachteten Hamming-Code sind in einer Zeichenfolge die drei letzten Ziffern bestimmt, wenn die ersten vier Ziffern vorgegeben sind, d. h. von den 7 Zeichen wird nur durch 4 Information übertragen. Die Rate ist somit

$$\frac{4}{7} \;=\; \frac{1}{n} \cdot \log_2(2^4) \;=\; \frac{1}{n} \log_q M \,.$$

Dies lässt die folgende Definition sinnvoll erscheinen:

Definition 2.2.9. Die Rate eines $(n, M, d; q)$-Codes ist die Zahl $\frac{1}{n} \log_q M$.

Bei gleichem Minimalabstand werden wir von zwei Codes den mit der größeren Rate, also der größeren Mächtigkeit, als besser ansehen. Daher ist es von Interesse, nach der maximalen Mächtigkeit, für die ein $(n, M, d; q)$-Code existiert, zu fragen.

Definition 2.2.10. Es sei $A(n, d; q) = \max\{M \in \mathbb{N}_0 \mid \exists\ (n, M, d; q) - \text{Code}\}$.

Beispiel 2.2.11. Die Existenz des $(7, 4)$-Hamming-Codes zeigt, dass $A(7, 3; 2) \geq 16$ ist. Wir werden später sehen, dass $A(7, 3; 2) = 16$, der $(7, 4)$-Hamming-Code also in gewissem Sinn bestmöglich ist. Im Folgenden werden wir einige grundlegende Ungleichungen für $A(n, d; q)$ diskutieren.

Lemma 2.2.12. *Es gilt $A(n, d; q) \leq A(n - 1, d - 1; q)$ für $d \geq 2$.*

Beweis.
Gegeben sei ein $(n, M, d; q)$-Code C. Aus C erhält man einen ,,punktierten" Code C_i, indem man aus jedem Codewort in C die i-te Koordinate weglässt. Dabei werde $\vec{x} \in C$ zu $\vec{x}_i \in C_i$. Sind nun $\vec{x}, \vec{y} \in C$ mit $d(\vec{x}, \vec{y}) = d$, so ist $d(\vec{x}_i, \vec{y}_i) \geq d-1$. Für $d \geq 2$ ist C_i ein $(n-1, M, d-1; q)$-Code (durch Punktieren aus C konstruiert, falls man für i eine Koordinate wählt, in der sich zwei Codeworte aus C mit dem Minimalabstand d unterscheiden). □

Satz 2.2.13 (Singleton-Schranke)**.** *Es gilt die Abschätzung* $A(n, d; q) \leq q^{n-d+1}$.

Beweis.
Indem man Lemma 2.2.12 genau $(d-1)$-mal anwendet, ergibt sich ein $(d-1)$-fach punktierter Code, nämlich ein $(n-d+1, M, 1; q)$-Code, d. h. eine Teilmenge der Mächtigkeit M von X^{n-d+1}. Es ist $|X^{n-d+1}| \leq q^{n-d+1}$. □

Satz 2.2.14 (Kugelpackungsschranke)**.** *Es gilt*

$$A(n, 2e+1; q) \;\leq\; \frac{q^n}{\sum\limits_{i=0}^{e} \binom{n}{i} (q-1)^i} \,.$$

Beweis.
Es sei C ein $(n, M, 2e+1; q)$-Code. Für jedes $\vec{x} \in C$ konstruieren wir die „Kugel"

$$K_e(\vec{x}) \;=\; \{\vec{u} \in X^n \mid d(\vec{u}, \vec{x}) \leq e\} \,.$$

Wir zeigen, dass für verschiedene $\vec{x}, \vec{z} \in C$ gilt:

$$(1): \; K_e(\vec{x}) \cap K_e(\vec{z}) \;=\; \emptyset \,.$$

Dazu sei $\vec{y} \in K_e(\vec{x}) \cap K_e(\vec{z})$. Dann gilt nach Satz 2.2.2: $d(\vec{x}, \vec{z}) \leq d(\vec{x}, \vec{y}) + d(\vec{y}, \vec{z}) \leq 2e$, ein Widerspruch zu $d = 2e+1$. Es gibt offenbar genau

$$\binom{n}{i} \cdot (q-1)^i$$

Möglichkeiten, ein Codewort an i Stellen abzuändern. Damit ist

$$|K_e(\vec{x})| \;=\; 1 + \binom{n}{1}(q-1) + \cdots + \binom{n}{e} \cdot (q-1)^e \,.$$

Andererseits ist

$$|C| \cdot |K_e(\vec{x})| \;=\; |C| \cdot \left(1 + \binom{n}{1}(q-1) + \cdots + \binom{n}{e}(q-1)^e\right) \;\leq\; |X|^n \;=\; q^n$$

und damit

$$|C| \;\leq\; \frac{q^n}{\sum_{i=0}^{e} \binom{n}{i} (q-1)^i} \,.$$

□

Definition 2.2.15. Ein Code C, der die Kugelpackungsschranke aus Satz 2.2.14 mit Gleichheit erfüllt, heißt perfekt.

Beispiel 2.2.16. Im $(7,4)$-Hamming-Code ist $n=7$, $q=2$, $d=2e+1=3$ und somit

$$\sum_{i=0}^{e}\binom{n}{i}(q-1)^i \;=\; 1+\binom{7}{1} \;=\; 8 \;,\; \text{sowie } |C| = \frac{2^7}{8} \;=\; 16\,.$$

Der $(7,4)$-Hamming-Code ist also perfekt.

Satz 2.2.17 (Plotkin-Schranke)**.** *Es sei $\theta = \frac{q-1}{q}$ und $d > \theta \cdot n$. Dann gilt:*

$$A(n,d;q) \;\leq\; \frac{d}{d-\theta n}\,,$$

insbesondere gilt

$$A(n,d;2) \;\leq\; \frac{2d}{2d-n} \text{ falls } n < 2d\,.$$

Beweis.
Wir berechnen die Schranke des durchschnittlichen Abstands zweier Codeworte in einem $(n,M,d;q)$-Code. Wir schreiben die M Codeworte $\vec{x}_1 = (x_{11},\ldots,x_{1n}),\ldots,\vec{x}_M = (x_{M1},\ldots,x_{Mn})$ als die Zeilen einer (M,n)-Matrix $\mathbf{M}$:

$$\mathbf{M} \;=\; \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{M1} & \cdots & x_{Mn} \end{pmatrix}\,.$$

In der k-ten Spalte dieser Matrix komme das Symbol $s \in X$ genau $m_{k,s}$-mal vor. Es sei

$$S \;=\; \sum_{\substack{i,j=1\\ i\neq j}}^{M} d(\vec{x}_i,\vec{x}_j)$$

die Summe aller Abstände von je zwei verschiedenen Codeworten. Wir schätzen S auf zwei Arten ab: Wegen $d(\vec{x}_i,\vec{x}_j) \geq d$ ist

$$(1):\; S \;\geq\; M\cdot(M-1)\cdot d\,.$$

Zur anderen Art der Auswertung von S setzen wir

$$c_{i,j,k}(s) \;=\; \begin{cases} 0 & \text{falls} \quad x_{ik} = x_{jk} = s \\ 1 & \text{sonst} \end{cases}\,.$$

Es ist

$$d(\vec{x}_i,\vec{x}_j) \;\leq\; \sum_{s\in X}\sum_{k=1}^{n} c_{i,j,k}(s) \text{ und damit}$$

$$S \;\leq\; \sum_{\substack{i,j=1\\ i\neq j}}^{M}\sum_{s\in X}\sum_{k=1}^{n} c_{i,j,k}(s) \;=\; \sum_{k=1}^{n}\sum_{s\in X}\sum_{\substack{i,j=1\\ i\neq j}}^{M} c_{i,j,k}(s)\,,$$

und es ist

$$\sum_{\substack{i,j=1\\ i\neq j}}^{M} c_{i,j,k}(s) \;\leq\; m_{k,s}\cdot(M-m_{k,s}) \;,\; \text{also}$$

$$(2):\; S \leq \sum_{k=1}^{n}\sum_{s\in X} m_{k,s}(M-m_{k,s}) \;=\; \sum_{k=1}^{n}\left(M\cdot\sum_{s\in X} m_{k,s} - \sum_{s\in X} m_{k,s}^2\right)\,.$$

Wir setzen

$$\overline{m}_k \;=\; \frac{1}{q} \cdot \sum_{s \in X} m_{k,s} \,.$$

Aus der Ungleichung

$$\sum_{s \in X} (m_{k,s} - \overline{m}_k)^2 \;\geq\; 0 \text{ folgt}$$

$$\sum_{s \in X} m_{k,s}^2 \;-\; 2 \cdot \left(\sum_{s \in X} m_{k,s} \right) \overline{m}_k \;+\; \sum_{s \in X} \overline{m}_k^2 \;\geq\; 0 \text{ , also}$$

$$\sum_{s \in X} m_{k,s}^2 \;\geq\; \frac{1}{q} \cdot \left(\sum_{s \in X} m_{k,s} \right)^2 \,.$$

Daraus und aus (2) folgt

$$(3): \; S \;\leq\; \sum_{k=1}^{n} \left(M^2 - \frac{1}{q} (\sum_{s \in X} m_{k,s})^2 \right) \;\leq\; n \cdot \theta \cdot M^2 \,.$$

Aus (1) und (3) folgt nun $M \cdot (M-1) \cdot d \leq n \cdot \theta \cdot M^2$. Auflösen nach M ergibt die Behauptung des Satzes. □

2.3 Lineare Codes

Bei linearen Codes weist das verwendete Alphabet eine algebraische Struktur auf, es wird als ein endlicher Körper mit q Elementen betrachtet. Man kann zeigen, dass ein solcher genau dann existiert (und bis auf die Namen der Elemente eindeutig ist), falls q eine Primzahlpotenz ist. Er wird mit $\mathrm{GF}(q)$ bezeichnet[1]. In der Praxis ist fast nur der Fall $q = 2^m$ von Bedeutung, weil dann die Elemente des Alphabets als Folge von Nullen und Einsen der Länge m geschrieben werden können.

Ein linearer Code C lässt sich dann als linearer Unterraum des Vektorraums $\mathrm{GF}(q)^n$ beschreiben.

Zur Fehlerkorrektur können Methoden der Linearen Algebra verwendet werden. Auch grundlegende Parameter, wie beispielsweise der Minimalabstand, können einfach bestimmt werden.

Definition 2.3.1. Ein $(n, M, d; q)$-Code C über $\mathrm{GF}(q)$ heißt linear, wenn C ein Unterraum des Vektorraums $\mathrm{GF}(q)^n$ ist.

Satz 2.3.2. *Es ist stets $M = q^k$ mit $k = \dim(C)$.*

Beweis.
Aus der Linearen Algebra ist bekannt, dass C eine Basis $B = \{\vec{b}_1, \ldots, \vec{b}_k\}$ mit $\vec{b}_i \in C$ besitzt. Dann ist jedes $\vec{x} \in C$ eindeutig als Linearkombination

$$\vec{x} \;=\; \sum_{j=1}^{k} \lambda_j \vec{b}_j \;\;,\;\; \lambda_j \in \mathrm{GF}(q)$$

darstellbar. Für die Wahl eines jeden λ_j gibt es q Möglichkeiten. Insgesamt erhalten wir $|C| = q^k$. □

[1] Als Abkürzung für „Galois Field", benannt nach Évariste Galois. Auch die Bezeichnung $\mathbb{F}_q$ ist üblich.

Beispiel 2.3.3. Der $(7,4)$-Hamming-Code kann als linearer Code gedeutet werden, wenn das Alphabet $X = \{0,1\}$ mit dem Körper $\mathrm{GF}(2) = \{0,1\}$ identifiziert wird. Addition und Multiplikation sind auf $\mathrm{GF}(2)$ wie folgt definiert:

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Die definierenden Bedingungen des Hamming-Codes

$$\begin{array}{lllllll} (1) & & x_2 + x_3 + x_4 + x_5 & & & \text{ist} & \text{gerade} \\ (2) & x_1 & + x_3 + x_4 & + x_6 & & \text{ist} & \text{gerade} \\ (3) & x_1 + x_2 & & + x_4 & + x_7 & \text{ist} & \text{gerade} \end{array}$$

können dann einfach formuliert werden als lineares Gleichungssystem:

$$\begin{array}{lllllll} (1) & & x_2 + x_3 + x_4 + x_5 & & & = & 0 \\ (2) & x_1 & + x_3 + x_4 & + x_6 & & = & 0 \\ (3) & x_1 + x_2 & & + x_4 & + x_7 & = & 0 \end{array}$$

Oder in Matrixschreibweise: $H \cdot \vec{x} = \vec{0}$, mit

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

Die Matrix H ist die Kontrollmatrix des $(7,4)$-Hamming-Codes.

Definition 2.3.4. Es sei C ein linearer $(n, q^k, d; q)$-Code. Wenn $\vec{a}_1, \ldots, \vec{a}_k$ eine Basis von C ist, so heißt die Matrix G mit Zeilen $\vec{a}_1, \ldots, \vec{a}_k$ eine Generatormatrix für C.

Die Kontrollmatrix eines Codes C hängt mit dessen Generatormatrix mittels der Konzepte der Orthogonalität und des zu C dualen Codes $C^\perp$ zusammen.

Definition 2.3.5. Es seien $\vec{x} = (x_1, \ldots, x_n), \vec{y} = (y_1, \ldots, y_n) \in \mathrm{GF}(q)^n$. Unter dem inneren Produkt $\langle \vec{x} | \vec{y} \rangle$ verstehen wir

$$\langle \vec{x} | \vec{y} \rangle \; = \; \sum_{i=1}^{n} x_i y_i \; = \; x_1 y_1 + \cdots + x_n y_n \, .$$

Die Vektoren $\vec{x}$ und $\vec{y}$ heißen orthogonal, falls $\langle \vec{x} | \vec{y} \rangle = 0$ ist.

Lemma 2.3.6. *Es sei C ein linearer Code. Die Menge $C^\perp = \{\vec{y} \in \mathrm{GF}(q)^n \mid \langle \vec{x} | \vec{y} \rangle = 0 \; \forall \vec{x} \in C\}$ ist ebenfalls ein linearer Code, d. h. ein Unterraum von $\mathrm{GF}(q)^n$.*

Beweis.
Das innere Produkt ist offenbar linear in beiden Komponenten, d. h. es gilt insbesondere

$$\langle \vec{x} | \lambda_1 \vec{y}_1 + \lambda_2 \vec{y}_2 \rangle \; = \; \lambda_1 \langle \vec{x} | \vec{y}_1 \rangle + \lambda_2 \langle \vec{x} | \vec{y}_2 \rangle \quad , \quad \lambda_i \in \mathrm{GF}(q) \, .$$

Sind daher $\vec{y}_1, \vec{y}_2 \in C^\perp$, so ist jede Linearkombination $\lambda_1 \vec{y}_1 + \lambda_2 \vec{y}_2$ ebenfalls in $C^\perp$. Wie aus der Linearen Algebra bekannt ist, erfüllt $C^\perp$ damit die Kriterien für einen Untervektorraum. □

Definition 2.3.7. Der Code $C^\perp$ heißt der zu C duale Code. Eine Generatormatrix von $C^\perp$ heißt Kontrollmatrix von C.

Lemma 2.3.8. *Es sei C ein linearer Code der Länge n mit der Generatormatrix G über $\mathrm{GF}(q)$. Dann gilt für $\vec{y} \in \mathrm{GF}(q)^n$ die Äquivalenz*

$$\vec{y} \in C^\perp \;\Leftrightarrow\; G \cdot \vec{y} = \vec{0}\ .$$

Wir beweisen dieses Lemma zusammen mit

Satz 2.3.9. *Es sei C ein linearer Code der Länge n über $\mathrm{GF}(q)$, dann ist $(C^\perp)^\perp = C$ und $\dim(C) + \dim(C^\perp) = n$.*

Beweis.
Eine Generatormatrix von C sei

$$G = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ \vdots & & & \vdots \\ x_{k1} & x_{k2} & \cdots & x_{kn} \end{pmatrix}$$

mit den Zeilenvektoren $\vec{x}_1 = (x_{11}, \ldots, x_{1n}), \ldots, \vec{x}_k = (x_{k1}, \ldots, x_{kn})$. Es sei $\vec{y} = (y_1, \ldots, y_n)$:

$$(1):\ \vec{y} \in C^\perp \;\Leftrightarrow\; \forall \vec{x} \in C:\ \langle \vec{x} | \vec{y} \rangle = \vec{0} \;\Rightarrow\; \langle \vec{x}_1 | \vec{y} \rangle = \cdots = \langle \vec{x}_k | \vec{y} \rangle = 0\,,$$

da jedes $\vec{x} \in C$ eine Linearkombination $\vec{x} = \lambda_1 \vec{x}_1 + \cdots + \lambda_k \vec{x}_k$ der $\vec{x}_i$ ist. Die letztere Bedingung in (1) ist äquivalent mit

$$G \cdot \vec{y} = \begin{pmatrix} x_{11}y_1 & + & x_{12}y_2 & + & \cdots & + & x_{1n}y_n \\ \vdots & & & & & & \vdots \\ x_{k1}y_1 & + & x_{k2}y_2 & + & \cdots & + & x_{kn}y_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}\ .$$

Aus der Linearen Algebra ist bekannt, dass $\dim(C^\perp) = n - \mathrm{rg}(G) = n - \dim(C)$ ist, also $\dim(C) + \dim(C^\perp) = n$. Ferner gilt für ein beliebiges $\vec{x} \in C$ auch $\langle \vec{x} | \vec{y} \rangle = 0$ für alle $\vec{y} \in C^\perp$, also $\vec{x} \in (C^\perp)^\perp$ nach Definition. Das impliziert $C \subseteq (C^\perp)^\perp$. Wegen $\dim(C^\perp) + \dim((C^\perp)^\perp) = n$ ist $\dim(C) = \dim((C^\perp)^\perp)$ und damit $C = (C^\perp)^\perp$. □

Der Minimalabstand von linearen Codes lässt sich besonders einfach bestimmen. Wir beschreiben im Folgenden das Verfahren.

Definition 2.3.10. Es sei $\vec{x} \in \mathrm{GF}(q)^n$. Das Gewicht $w(\vec{x})$ ist die Anzahl der Koordinaten ungleich Null in $\vec{x}$. Das Minimalgewicht von C ist

$$\min\{w(\vec{x}) \mid \vec{x} \in C - \{0\}\}\,.$$

Satz 2.3.11. *Für beliebige $\vec{x}, \vec{y} \in \mathrm{GF}(q)^n$ gilt $d(\vec{x}, \vec{y}) = w(\vec{x} - \vec{y})$. Insbesondere ist für einen linearen Code der Minimalabstand gleich dem Minimalgewicht.*

Beweis.
Sei $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ und $x - y = (x_1 - y_1, \ldots, x_n - y_n)$. Dann gilt offensichtlich genau dann $x_i \neq y_i$, wenn $x_i - y_i \neq 0$ ist. Dementsprechend gilt $d(x, y) = w(x - y)$.

Nun sei C ein linearer Code mit Minimalabstand d_0 und Minimalgewicht w_0. Wähle nun $x, y \in C$ so, daß $d(x, y) = d_0$ gilt. Dann ist auch $x - y \in C \backslash \{0\}$, und es gilt:

$$d_0 = d(x, y) = w(x - y) \geq w_0.$$

Es sei x_0 das Codewort mit minimalem Gewicht, also $w(x_0) = w_0$. Dann gilt

$$w_0 = w(x_0 - 0) = d(x_0, 0) \geq d_0.$$

Insgesamt erhalten wir

$$d_0 \geq w_0 \geq d_0, \quad \text{also} \quad d_0 = w_0.$$

□

Satz 2.3.12. *Es sei H eine Kontrollmatrix eines linearen Codes C der Länge n über* $\mathrm{GF}(q)$*. Dann gilt für $\vec{x} \in \mathrm{GF}(q)^n$:*

$$\vec{x} \in C \;\Leftrightarrow\; H \cdot \vec{x} = \vec{0}\,.$$

Beweis.
Nach Satz 2.3.9 ist $(C^\perp)^\perp = C$. Nach Definition ist H also eine Generatormatrix von $\mathbb{C}^\perp$. Nach Lemma 2.3.8 gilt dann

$$\vec{x} \in (C^\perp)^\perp = C \;\Leftrightarrow\; H \cdot \vec{x} = \vec{0}\,.$$

□

Satz 2.3.13. *Es sei H die Kontrollmatrix des linearen Codes C. Dann ist das Minimalgewicht (und damit der Minimalabstand) von C die kleinste Anzahl d, für die es d linear abhängige Spalten in H gibt.*

Beweis.
Es seien $\vec{s}_1, \ldots, \vec{s}_n$ die Spalten von H, $\vec{x} = (x_1, \ldots, x_n) \in \mathrm{GF}(q)^n$ und $I = \{i \mid x_i \neq 0\}$, also $|I| = w(\vec{x})$. Es gilt

$$\vec{x} \in C \;\Leftrightarrow\; H \cdot \vec{x} = \vec{0} \;\Leftrightarrow\; \sum_{i=1}^{n} x_i \vec{s}_i = \vec{0} \;\Leftrightarrow\; \sum_{i \in I} x_i \vec{s}_i = \vec{0} \;\Leftrightarrow\; \{\vec{s}_i \mid i \in I\} \text{ lin. abh. }.$$

Es gibt daher genau dann d linear abhängige Spalten, wenn es ein Codewort vom Gewicht d gibt, woraus die Behauptung des Satzes folgt. □

Beispiel 2.3.14 ($(7, 4)$-Hamming-Code)**.** Die obigen Überlegungen liefern nun eine wesentlich einfachere Methode, den Minimalabstand des $(7, 4)$-Hamming-Codes zu bestimmen. In der Kontrollmatrix

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

dieses Codes sind je zwei Spalten linear unabhängig (da sie verschieden und damit keine Vielfachen voneinander sind). Hingegen ist $\vec{s}_1 + \vec{s}_2 + \vec{s}_3 = \vec{0}$. Damit gibt es 3 linear abhängige Spalten, aber keine zwei, d. h. der Minimalabstand ist 3.

Die Fragen nach der Qualität von Codes, die wir im vorigen Abschnitt gestellt haben, können natürlich auch für lineare Codes gestellt werden. Wie viele Codeworte kann ein linearer Code bei vorgegebener Länge und Mindestabstand besitzen? Die oberen Schranken nach Satz 2.2.13 (Singleton) und 2.2.14 (Kugelpackung) gelten natürlich auch für lineare Codes. Wir zeigen nun eine untere Schranke:

Satz 2.3.15 (Gilbert-Varshamov-Schranke)**.** *Wenn die Ungleichung*

$$\sum_{i=0}^{d-2} \binom{n-1}{i} (q-1)^i \;<\; q^{n-k}$$

gilt, dann gibt es einen $(n, q^k, d; q)$*-Code.*

Beweis.
Wir konstruieren den Code mittels seiner Kontrollmatrix H. Nach den Sätzen 2.3.11-2.3.13 genügt es, eine $((n-k), n)$-Matrix über GF(q) zu konstruieren, für die je $(d-1)$ Spalten linear unabhängig sind. Als erste Spalte können wir jeden Vektor $\vec{x} \neq \vec{0}$ in $V = \mathrm{GF}(q)^{n-k}$ wählen. Angenommen wir haben bereits m Spalten von H konstruiert, von denen je $d-1$ linear unabhängig sind. Offensichtlich kann man aus diesen m Spalten höchstens

$$f(m) \;=\; 1 + (q-1)\binom{m}{1} + (q-1)^2\binom{m}{2} + \cdots + (q-1)^{d-2}\binom{m}{d-2}$$

verschiedene Linearkombinationen bilden, die sich aus höchstens $d-2$ Spalten kombinieren lassen. Wenn $f(m) < q^{n-k}$ ist, kann man einen von diesen Linearkombinationen verschiedenen Vektor $\vec{v} \in V$ auswählen. Man sieht leicht, dass $\vec{v}$ als weitere Spalte von H gewählt werden kann. Aufgrund unserer Voraussetzung gilt aber $f(m) < q^{n-k}$ für alle $m \leq n-1$, womit die Existenz der Matrix H gezeigt ist. □

Beispiel 2.3.16 (Konstruktion eines linearen $(7, 4, 4; 2)$-Codes)**.** Die Kontrollmatrix muss die Dimension $(5, 7)$ besitzen, und je 3 Spalten müssen linear unabhängig sein.

Wahl von $\vec{s}_1$:
Die Bedingung ist hier nur $\vec{s}_1 \neq \vec{0}$. Wir wählen $\vec{s}_1 = (1, 0, 0, 0, 0)^T$.

Wahl von $\vec{s}_2$:
Die Bedingung ist nun, dass $\vec{s}_1$ und $\vec{s}_2$ linear unabhängig sind: $\vec{s}_2 \neq \lambda\vec{s}_1 \;\Leftrightarrow\; \vec{s}_2 \notin \{\vec{0}, \vec{s}_1\}$.
Wir wählen $\vec{s}_2 = (0, 1, 0, 0, 0)^T$.

Wahl von $\vec{s}_3$:
Bedingung: $\vec{s}_1, \vec{s}_2, \vec{s}_3$ linear unabhängig, also $\vec{s}_3 \neq \lambda_1\vec{s}_1 + \lambda_2\vec{s}_2$, d. h. $\vec{s}_3 \notin \{\vec{0}, \vec{s}_1, \vec{s}_2, \vec{s}_1 + \vec{s}_2\}$.
Wir wählen $\vec{s}_3 = (0, 0, 1, 0, 0)^T$.

Wahl von $\vec{s}_4$:
Die Bedingung lautet nun, dass die Mengen

$$\{\vec{s}_1, \vec{s}_2, \vec{s}_3\}\ ,\ \{\vec{s}_1, \vec{s}_2, \vec{s}_4\}\ ,\ \{\vec{s}_1, \vec{s}_3, \vec{s}_4\}\ ,\ \{\vec{s}_2, \vec{s}_3, \vec{s}_4\}$$

jeweils linear unabhängig sind. Das ist gleichbedeutend mit $\vec{s}_4 \notin \{\vec{0}, \vec{s}_1, \vec{s}_2, \vec{s}_3, \vec{s}_1 + \vec{s}_2, \vec{s}_1 + \vec{s}_3, \vec{s}_2 + \vec{s}_3\}$.
Wir wählen $\vec{s}_4 = (0, 0, 0, 1, 0)^T$.

Wahl von $\vec{s}_5$:
Nun müssen je drei der Vektoren $\{\vec{s}_1, \ldots, \vec{s}_5\}$ linear unabhängig sein. Das ist äquivalent zu

$$\vec{s}_5 \;\notin\; \{\vec{0}, \vec{s}_1, \vec{s}_2, \vec{s}_3, \vec{s}_4, \vec{s}_1 + \vec{s}_2, \ldots, \vec{s}_3 + \vec{s}_4\}\ .$$

Wir wählen $\vec{s}_5 = (1, 1, 0, 0, 1)^T$.

Wahl von $\vec{s}_6$:
Je drei der Vektoren $\{\vec{s}_1, \ldots, \vec{s}_6\}$ müssen linear unabhängig sein. Das ist äquivalent zu

$$\vec{s}_6 \notin \{\vec{0}, \vec{s}_1, \vec{s}_2, \vec{s}_3, \vec{s}_4, \vec{s}_5, \vec{s}_1 + \vec{s}_2, \ldots, \vec{s}_4 + \vec{s}_5\} .$$

Beobachtung: alle diese Linearkombinationen haben entweder eine 1 in der 5-ten Zeile, oder eine 0 in der 5-ten Zeile und höchstens 3 Einsen in den ersten vier Zeilen.
Wir wählen $\vec{s}_6 = (1, 1, 1, 1, 0)^T$.

Wahl von $\vec{s}_7$:
Je drei der Vektoren $\{\vec{s}_1, \ldots, \vec{s}_7\}$ müssen linear unabhängig sein. Das ist äquivalent zu

$$\vec{s}_7 \notin \{\vec{0}, \vec{s}_1, \vec{s}_2, \vec{s}_3, \vec{s}_4, \vec{s}_5, \vec{s}_6, \vec{s}_1 + \vec{s}_2, \ldots, \vec{s}_5 + \vec{s}_6\} .$$

Beobachtung: alle diese Linearkombinationen haben entweder eine 0 in der 5-ten Zeile, oder eine 1 in der 5-ten Zeile und höchstens 3 Einsen in den ersten vier Zeilen. Wir wählen $\vec{s}_7 = (1, 1, 1, 1, 1)^T$.

Damit haben wir die Kontrollmatrix

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

konstruiert.
Der Code hat nur 4 Codeworte, die wir bestimmen wollen. Dazu müssen wir das lineare Gleichungssystem $H \cdot \vec{x} = \vec{0}$ lösen, was im Prinzip durch Gaußsche Elimination geschieht. Im vorliegenden Fall hat das System schon die gewünschte Endform: durch die Wahl von x_1 und x_7 sind alle übrigen Koordinaten bestimmt. Die zwei Wahlen $(x_1, x_7) = (1, 0)$ und $(x_1, x_7) = (0, 1)$ liefern zwei linear unabhängige Codeworte, die als Zeilen einer Generatormatrix dienen können: Ist $x_1 = 1$ und $x_7 = 0$, so ist die Gleichung $H \cdot \vec{x} = \vec{0}$ wegen der letzten Zeile von H nur mit $x_5 = 0$ lösbar. Aus $(x_1, x_5, x_7) = (1, 0, 0)$ folgt $x_6 = 1$ wegen der ersten Zeile von H. Ebenso folgt $x_2 = x_3 = x_4 = 1$. Analog kann $(x_1, x_7) = (0, 1)$ behandelt werden.
Insgesamt gilt

$$(x_1, x_7) = (1, 0) \Rightarrow x_2 = 1, x_3 = 1, x_4 = 1, x_5 = 0, x_6 = 1 .$$

$$(x_1, x_7) = (0, 1) \Rightarrow x_2 = 0, x_3 = 1, x_4 = 1, x_5 = 1, x_6 = 0 .$$

Also $\vec{x}_1 = (1, 1, 1, 1, 0, 1, 0)^T$ und $\vec{x}_2 = (0, 0, 1, 1, 1, 0, 1)^T$.
Damit haben wir die Generatormatrix

$$G = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix} ,$$

und es ist $C = \{\vec{0}, \vec{x}_1, \vec{x}_2, \vec{x}_1 + \vec{x}_2\}$.

2.4 Fehlerkorrektur bei linearen Codes und Syndrome

Wir kehren zur Diskussion der Fehlerkorrektur aus dem vorigen Abschnitt zurück. Wird bei der Übertragung eine Zeichenfolge $\vec{x}'$ empfangen, die kein Codewort ist, so wird diese durch dasjenige Codewort $\vec{x}$ ersetzt, für das $d(\vec{x}, \vec{x}')$ minimal ist.

Es müssen also zwei Aufgaben gelöst werden:

Fehlererkennung: Es muß erkannt werden, dass $\vec{x}'$ kein Codewort ist.
Fehlerkorrektur: Es muß ein $\vec{x} \in C$ gefunden werden, für das $d(\vec{x}, \vec{x}')$ minimal ist.

Die gröbste Lösung dieser Aufgabe besteht in der Durchmusterung aller Codeworte, was wegen des großen Aufwands unbefriedigend ist. Bei linearen Codes bieten sich bessere Lösungen an: Die Fehlererkennung kann mittels der Überprüfung der Kontrollgleichung

$$H \cdot \vec{x} \;=\; 0$$

geschehen. Die Fehlerkorrektur geschieht mittels des so genannten Syndroms

$$\vec{y} \;=\; H \cdot \vec{x}\,.$$

Der Begriff des Syndroms ist aus der Medizin übernommen. Ein Syndrom ist eine Sammlung von Symptomen, die eine Krankheit anzeigen. In der Codierungstheorie ist die Krankheit der Fehler bei der Übertragung des Codeworts. Wir erläutern das Verfahren am Beispiel eines 1-Fehler-korrigierenden Codes. Dazu schreiben wir $\vec{x}' = \vec{x} + \vec{e}$ und nennen $\vec{e}$ den Fehlervektor. Wir gehen davon aus, dass bei der Übertragung genau ein Fehler gemacht wurde, also ist

$$\vec{e} = (0 \ldots 010 \ldots 0)$$

mit genau einer Eins an der j-ten Stelle für ein j. Für das Syndrom $\vec{y} = H\vec{x}'$ erhalten wir

$$H\vec{x}' \;=\; H \cdot (\vec{x} + \vec{e}) \;=\; H\vec{x} + H\vec{e} \;=\; H\vec{e}$$

da $H\vec{x} = \vec{0}$ ist. Hat H die Spaltenvektoren $\vec{s}_j$

$$\vec{s}_j \;=\; \begin{pmatrix} s_{1j} \\ s_{2j} \\ \vdots \\ s_{kj} \end{pmatrix}\,,$$

so ist $H\vec{e} = \vec{s}_j$ genau dann, wenn $\vec{e}$ die Eins an der j-ten Stelle enthält. Das Syndrom $H \cdot \vec{x}' = \vec{s}_j$ zeigt also unmittelbar an, dass der Fehler an der j-ten Stelle gemacht wurde. Bei der Korrektur ist daher die j-te Stelle zu ändern. Allgemein gilt folgender

Satz 2.4.1. *Es sei C ein linearer $(n, q^k, d; q)$-Code über $\mathrm{GF}(q)$ mit Minimalgewicht $d = 2e+1$ ($e \in \mathbb{N}$) und der Kontrollmatrix H. Zu $\vec{y} \in \mathrm{GF}(q)^k$ gibt es genau ein $\vec{e} \in \mathrm{GF}(q)^n$ mit $w(\vec{e}) \leq e$ und $H \cdot \vec{e} = \vec{y}$.*

Beweis.
Mit $\mathrm{rg}(H) = k$ ist auch $\dim(\{H\vec{x} \mid \vec{x} \in \mathrm{GF}(q)^n\}) = k$. Es seien $\vec{e}_1, \vec{e}_2 \in \mathrm{GF}(q)^n$ mit $w(\vec{e}_i) \leq e_i$ für $i = 1, 2$ und $H\vec{e}_1 = H\vec{e}_2 = \vec{y}$. Dann ist $H \cdot (\vec{e}_1 - \vec{e}_2) = \vec{0}$, also $\vec{e}_1 - \vec{e}_2 \in C$, und $w(\vec{e}_1 - \vec{e}_2) \leq w(\vec{e}_1) + w(\vec{e}_2) \leq 2e < d$. Das ist nur für $\vec{e}_1 = \vec{e}_2$ möglich. □

Aufgrund von Satz 2.4.1 kann die Fehlerkorrektur wie folgt vorgenommen werden: Man legt einen Speicher an, der jedem möglichen Syndrom $\vec{y}$ das eindeutig bestimmte $\vec{e}$ mit $H\vec{e} = \vec{y}$ mit minimalem Gewicht $w(\vec{e})$ zuordnet. Wird dann die Zeichenfolge $\vec{x}'$ mit $H\vec{x}' = \vec{y}$ empfangen, so wird $\vec{x}'$ zu $\vec{x} = \vec{x}' - \vec{e}$ korrigiert.

Beispiel 2.4.2. Beim $(7,4)$-Hamming-Code wird statt des Codeworts $\vec{x}$ das Wort $\vec{x}' = (0010001)^T$ empfangen, wobei ein Fehler gemacht wurde. Man nehme die Fehlerkorrektur vor.

Lösung: Es ist

$$H \cdot \vec{x}' = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \vec{s}_4 \,.$$

Der Fehler wurde also an der 4. Stelle begangen, d. h. die Zeichenfolge $\vec{x}'$ ist zum Codewort $\vec{x} = (0011001)^T$ zu korrigieren.

Kapitel 3

Mathematisches Modellieren

3.1 Ein mathematisches Modell der schwingenden Saite

Unter Verwendung physikalischer Annahmen wird ein mathematisches Modell der schwingenden Saite hergeleitet. Dieses ist eine partielle Differentialgleichung mit zwei Randbedingungen und zwei Anfangsbedingungen, welche mit mathematischen Methoden gelöst wird.

Wir beginnen mit dem physikalischen Modell:

1. Eine Drahtsaite ist an zwei festen Punkten der xy-Ebene eingespannt, etwa den Punkten $(0,0)$ und $(L,0)$, und schwingt in der xy-Ebene.

2. Jeder Massenpunkt der Drahtsaite schwingt parallel zur y-Achse.

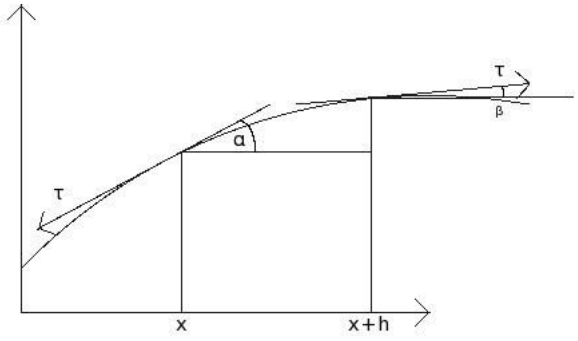
3. Der Massenpunkt, der mit Abszisse x parallel zur y-Achse schwingt, hat zum Zeitpunkt t eine Ordinate zur x-Achse, die wir mit $u(x,t)$ bezeichnen. Wir nehmen an, die Funktion u sei jeweils zweimal nach x und t differenzierbar. Die partiellen Ableitungen schreiben wir als u_x, u_{xx}, u_t und u_{tt}.

4. Der Massenpunkt zur Abszisse x erfährt zum Zeitpunkt t die Beschleunigung

$$u_{tt}(x,t) \;=\; \frac{\partial^2}{\partial t^2} u(x,t)\,.$$

   Wir ersetzen diesen Massenpunkt durch ein kleines Saitenstück S zwischen x und $x+h$, und nehmen vereinfachend an, dass dort die Beschleunigung konstant ist. Diese Beschleunigung ist nach dem Newtonschen Gesetz (Kraft=Masse $\cdot$ Beschleunigung) mit der auf das Saitenstück wirkenden Kraft verbunden.

5. Wir betrachten die Kräfte, die zum Zeitpunkt t an dem Saitenstück S zerren. Eine Kraft zerrt tangential nach links, und eine zerrt tangential nach rechts:



Die beiden Kräfte haben den gleichen (und von x unabhängigen) Betrag, etwa τ. Die Winkel, den die Tangenten in x bzw. $x+h$ mit der x-Achse einschließen, seien mit α bzw. β bezeichnet. Als Kraftkomponente senkrecht zur x-Achse ergibt sich

$$(1):\ \tau \cdot (\sin\beta - \sin\alpha)\,.$$

Wir machen zudem folgende vereinfachende Annahme: Da die Schwingungsfunktion $u(x,t)$ klein ist, ist $\tan\alpha \approx \sin\alpha$ sowie $\tan\beta \approx \sin\beta$. Damit ist die Kraft in (1) ungefähr

$$\tau \cdot (\tan\beta - \tan\alpha) \ = \ \tau \cdot (u_x(x+h,t) - u_x(x,t))\ .$$

Wegen

$$\lim_{h\to 0} \frac{u_x(x+h,t) - u_x(x,t)}{h} \ = \ u_{xx}(x,t)$$

gilt für kleine $h > 0$: $u_x(x+h,t) - u_x(x,t) \approx u_{xx}(x,t)$. Die Masse des kleinen Drahtstücks zwischen x und $x+h$ ist $\varrho \cdot h$, wobei ϱ die spezifische Massendichte bezeichnet. Das Newton'sche Gesetz liefert also

$$\tau \cdot (u_x(x+h,t) - u_x(x,t)) \ = \ \varrho \cdot h \cdot u_{tt}(x,t)$$

und damit die partielle Differentialgleichung

$$u_{tt}(x,t) \ = \ \frac{\tau}{\varrho} u_{xx}(x,t) \ =: \ c^2 \cdot u_{xx}(x,t)\,.$$

Die Einspannung der Saite in die beiden Punkte $(0,0)$ und $(L,0)$ liefert die beiden Randbedingungen

$$\forall t:\ u(0,t) \ = \ u(L,t) \ = \ 0\,.$$

Bezeichnet $f(x)$ die Lage der Saite zum Zeitpunkt $t=0$, und entspricht $g(x)$ der Geschwindigkeit (bei $t=0$) der zur Abszisse x zur y-Achse parallel schwingenden Massenpunkte, so folgen noch die beiden Anfangsbedingungen $\forall 0 \le x \le L:\ u(x,0) = f(x)$ und $u_t(x,0) = g(x)$.

Gesucht ist also eine Lösung u von:

- $u_{tt}(x,t) = c^2 \cdot u_{xx}(x,t)$,
- $\forall t:\ u(0,t) = u(L,t) = 0$,
- $\forall 0 \le x \le L:\ u(x,0) = f(x)$,
- $\forall 0 \le x \le L:\ u_t(x,0) = g(x)$.

3.2 Spezielle Lösungen eines vereinfachten Problems

Wir betrachten zunächst nur das vereinfachte Gleichungssystem, das durch Weglassen der Anfangsbedingungen entsteht:

$$(1):\ u_{tt}(x,t) = c^2 \cdot u_{xx}(x,t)\ ,\ \forall t:\ u(0,t) = u(L,t) = 0\,.$$

Eine wichtige Methode zur Lösung von solchen Gleichungen besteht in der Trennung der Variablen. Man schreibt die gesuchte Funktion als ein Produkt einer Funktion, die nur von x abhängt und einer

solchen, die nur von t abhängt: $u(x,t) = X(x) \cdot T(t)$. Einsetzen in (1) liefert $X(x)T''(t) = c^2 \cdot X''(x) \cdot T(t)$. Es folgt also an den Stellen (x,t) mit $X(x)T(t) \neq 0$ die Gleichung

$$\frac{T''(t)}{T(t)} \;=\; c^2 \cdot \frac{X''(x)}{X(x)}\,.$$

Da die linke Seite unabhängig von x ist, ist es auch die rechte Seite, und somit ist die rechte Seite (und damit wiederum die linke) konstant. Es gibt also eine Zahl λ mit

$$(2):\; \frac{T''(t)}{T(t)} \;=\; c^2 \cdot \frac{X''(x)}{X(x)} \;=\; \lambda$$

für alle (x,t) mit $X(x)T(t) \neq 0$. Wegen der Stetigkeit in t folgt die Aussage auch für alle t:

$$(3):\; \forall t:\; T''(t) \;=\; \lambda T(t)$$

und analog

$$(4):\; \forall 0 \leq x \leq L:\; X''(x) \;=\; \frac{\lambda}{c^2} \cdot X(x)$$

mit den Randbedingungen $X(0) = X(L) = 0$. Die Funktionen X und T sind somit Eigenfunktionen des Differentialoperators $K : f \mapsto f''$. Die Theorie der linearen Systeme von Differentialgleichungen 1. Ordnung ergibt, dass sämtliche Lösungen von (4) (ohne die Randbedingung) gegeben sind durch

$$(a)\quad \text{Falls } \lambda < 0: \qquad \alpha \cdot \cos\left(\frac{\sqrt{|\lambda|}}{c}x\right) \;+\; \beta \cdot \sin\left(\frac{\sqrt{|\lambda|}}{c}x\right)\;,$$

$$(b)\quad \text{falls } \lambda = 0: \qquad \alpha x + \beta\,,$$

$$(c)\quad \text{falls } \lambda > 0: \qquad \alpha \cdot e^{\frac{\sqrt{\lambda}}{c}x} \qquad + \qquad \beta \cdot e^{-\frac{\sqrt{\lambda}}{c}x}$$

mit beliebigen Koeffizienten $\alpha, \beta \in \mathbb{R}$. Einsetzen der Randbedingungen $X(0) = X(L) = 0$ ergibt das Folgende: Falls $\lambda < 0$ ist, bleibt nur $\alpha = 0$ (wegen $X(0) = 0$), sowie

$$\frac{\sqrt{|\lambda|} \cdot L}{\pi c} \;=:\; n \;\in \mathbb{Z}$$

wegen $X(L) = 0$, dafür ist $\beta \in \mathbb{R}$ beliebig. Zusammengefasst also

$$(a):\; X(x) \;=\; \beta \sin(\frac{n\pi}{L}x)\;,\; n \in \mathbb{Z}\,.$$

Dies entspricht den durch $n \in \mathbb{Z}$ parametrisierten Eigenwerten $\lambda_n = -(n\pi c/L)^2$ sowie den zugehörigen Eigenfunktionen $\sin(n\pi x/L)$. Im Falle $\lambda = 0$ folgt $\beta = 0$ aus $X(0) = 0$, sowie $\alpha = 0$ aus $X(L) = \alpha L = 0$. Es gibt also nur die triviale Lösung $X = 0$. Im Falle $\lambda > 0$ ist $\beta = -\alpha$ wegen $X(0) = 0$. Die Randbedingung $X(L) = 0$ impliziert

$$\alpha \cdot \left(e^{\frac{\sqrt{\lambda}}{c}L} - e^{-\frac{\sqrt{\lambda}}{c}L}\right) \;=\; 0\,,$$

und damit $\alpha = \beta = 0$, da $e^v - e^{-v} > 0$ ist für alle $v > 0$. Auch in diesem Fall gibt es also nur die triviale Lösung $X = 0$. Die einzigen von der Nullfunktion verschiedenen Lösungen ergeben sich daher für $\lambda < 0$, nämlich die reellen Vielfachen der Funktionen

$$X_n(x) \;=\; \sin\left(\frac{n\pi}{L} \cdot x\right)$$

für $n \in \mathbb{Z}$. Einsetzen der Eigenwerte λ_n in (3) liefert die Lösungen

$$T_n(t) \;=\; a_n \cdot \cos\Big(\frac{c\pi n}{L} \cdot t\Big) + b_n \cdot \sin\Big(\frac{c\pi n}{L} \cdot t\Big)$$

für den zweiten Faktor der Funktion u. Somit sind alle Lösungen des Randwertproblems

$$\begin{array}{rcl} u_{tt}(x,t) & = & c^2 \cdot u_{xx}(x,t) \\ u(0,t) & = & u(L,t) \end{array}$$

mit der speziellen Forderung $u(x,t) = X(x) \cdot T(t)$ von der Form

$$u(x,t) \;=\; \sin\Big(\frac{n\pi}{L} \cdot x\Big) \cdot \Big(a_n \cdot \cos\Big(\frac{c\pi n}{L} \cdot t\Big) + b_n \cdot \sin\Big(\frac{c\pi n}{L} \cdot t\Big)\Big) \;.$$

3.3 Superposition

Um einfachere Formeln zu erhalten, setzen wir von jetzt an $c = 1$ und $L = \pi$. Dies könnte auch durch die Substitution

$$v(x,t) \;=\; u(\frac{L}{\pi}x, \frac{Lc}{\pi}t)$$

erreicht werden. Das vereinfachte Modell (Weglassen der Anfangsbedingungen) lautet jetzt:

$$(1):\; u_{tt}(x,t) \;=\; u_{xx}(x,t) \;,\; \forall t:\; u(0,t) \;=\; u(\pi,t) \;=\; 0\,,$$

und es besitzt die speziellen Lösungen

$$(2):\; u^{(n)}(x,t) \;=\; \sin(nx) \cdot (a_n \cdot \cos(nt) + b_n \cdot \sin(nt))$$

mit $n \in \mathbb{N}$ und $a_n, b_n \in \mathbb{R}$. Die Lösungsmenge des Randwertproblems (1) hat nun die Eigenschaft der Superposition: Für n Lösungen $u^{(1)}, \ldots, u^{(n)}$ ist auch jede Linearkombination $\lambda_1 u^{(1)} + \cdots + \lambda_n u^{(n)}$ eine Lösung. Die Lösungsmenge bildet einen Untervektorraum des Vektorraums aller auf $[0,\pi] \times [0,\infty)$ stetigen Funktionen. Es stellt sich nun die Frage: können unendliche Linearkombinationen der speziellen Lösungen (2) so gewählt werden, dass auch die Anfangsbedingungen

$$u(x,0) \;=\; f(x) \quad,\quad u_t(x,0) \;=\; g(x)$$

erfüllt sind? Dies führt zu

$$u\;(x,t) \;=\; \sum_{n=1}^{\infty} \sin(nx) \cdot (\quad a_n \cos(nt) + b_n \sin(nt)) \qquad ,$$

$$u_t(x,t) \;=\; \sum_{n=1}^{\infty} \sin(nx) \cdot (-a_n \sin(nt) + b_n \cos(nt)) \cdot n\,.$$

$$(3):\; f(x) \;=\; u\;(x,0) \;=\; \sum_{n=1}^{\infty} a_n \sin(nx) \quad ,$$

$$(4):\; g(x) \;=\; u_t(x,0) \;=\; \sum_{n=1}^{\infty} n b_n \sin(nx)\,.$$

Die Darstellungen (3) und (4) sind spezielle Fourierreihen. Es ergibt sich also das Problem der Fourierreihenentwicklung von (weitgehend beliebigen) Funktionen $f(x)$ und $g(x)$. Dieses Problem werden wir im nächsten Abschnitt behandeln.

3.4 Innere Produkträume und Orthogonalsysteme von Funktionen

Die Theorie der Fourierreihen lässt sich in einen weiteren Rahmen stellen, wenn man die trigonometrischen Funktionen als Spezialfall eines Orthogonalsystems von Funktionen ansieht. Zunächst machen wir der einfacheren Rechnung halber von der Beziehung zwischen trigonometrischen Funktionen und der Exponentialfunktion Gebrauch:

$$\begin{aligned} (1): \quad e^{ix} &= \cos(x) + i \cdot \sin(x) \\ (2): \quad \cos(x) &= \tfrac{1}{2}(e^{ix} + e^{-ix}) \\ \sin(x) &= \tfrac{1}{2i}(e^{ix} - e^{-ix}) \quad . \end{aligned}$$

Definition 3.4.1. Eine Fourierreihe ist eine unendliche Reihe der Form

$$(3): \sum_{n=-\infty}^{\infty} c_n e^{2\pi i n L x} \ , \ L > 0, \ c_n \in \mathbb{C} \, .$$

Wegen (1) und (2) lässt sich jede Reihe des vorigen Abschnitts

$$\sum_{n=0}^{\infty} a_n \cos(2\pi n L x) \ + \ \sum_{n=0}^{\infty} b_n \sin(2\pi n L x)$$

in der Form (3) schreiben. Damit ist auch das Problem der schwingenden Saite in der Behandlung eingeschlossen. Der Einfachheit halber behandeln wir den Spezialfall $L = 1$ und setzen:

Definition 3.4.2. $E(x) \ = \ e^{2\pi i x}$.

Die Familie der Funktionen $(E(mx))_{m\in\mathbb{Z}}$ erfüllt die Orthogonalitätsrelation

$$\int_0^1 E(mx)E(-nx)dx \ = \ \begin{cases} 1 & \text{falls} \quad m = n \\ 0 & \text{sonst} \end{cases} \, .$$

Dieser Begriff ist von der Geometrie her motiviert: Zwei Vektoren $\vec{e}_1 \ = \ (x_1, \ldots, x_n)$ und $\vec{e}_2 \ = \ (y_1, \ldots, y_n)$ des $\mathbb{R}^n$ heißen orthogonal, falls ihr inneres Produkt $\langle \vec{e}_1 | \vec{e}_2 \rangle \ = \ 0$ ist. Da wir komplexwertige Funktionen betrachten, empfiehlt es sich, in Vektorräumen über $\mathbb{C}$ zu arbeiten. Führen wir im Vektorraum V der auf $[0, 1]$ stetigen komplexwertigen Funktionen das innere Produkt durch

$$\langle f|g \rangle \ = \ \int_0^1 f(x)\overline{g(x)}dx$$

ein, so ist das Integral in der Orthogonalitätsrelation der $E(mx)$ gerade das innere Produkt.

Definition 3.4.3. Ein Vektorraum H über $\mathbb{C}$ heißt innerer Produktraum (oder unitärer Raum, falls jedem Paar von Vektoren $\vec{x}, \vec{y} \in H$ eine komplexe Zahl $\langle \vec{x} | \vec{y} \rangle$, das innere Produkt, zugeordnet wird, so dass gilt:

$$\begin{array}{lrcll} (a): & \langle \vec{y} | \vec{x} \rangle & = & \overline{\langle \vec{x} | \vec{y} \rangle} & \\ (b): & \langle \vec{x} + \vec{y} | \vec{z} \rangle & = & \langle \vec{x} | \vec{z} \rangle \quad + & \langle \vec{y} | \vec{z} \rangle \\ (c): & \langle \alpha \vec{x} | \vec{y} \rangle & = & \alpha \langle \vec{x} | \vec{y} \rangle \quad \forall \alpha \in \mathbb{C} & \\ (d): & \langle \vec{x} | \vec{x} \rangle & \geq & 0 \quad \forall \vec{x} \in H & \\ (e): & \langle \vec{x} | \vec{x} \rangle & = & 0 \quad \Leftrightarrow & \vec{x} = \vec{0} \, . \end{array}$$

Bemerkung 3.4.4. Diese Eigenschaften haben folgende Konsequenzen: Aus (c) folgt, dass $\left\langle \vec{0}|\vec{y}\right\rangle = 0$ ist für alle $\vec{y} \in H$. Aus (a) und (c) folgt, dass $\langle \vec{x}|\alpha\vec{y}\rangle = \overline{\alpha}\,\langle \vec{x}|\vec{y}\rangle$ ist für $\alpha \in \mathbb{C}$. Aus (a) und (b) folgt das zweite Distributivgesetz $\langle \vec{x}|\vec{y}+\vec{z}\rangle = \langle \vec{x}|\vec{y}\rangle + \langle \vec{x}|\vec{z}\rangle$.

Definition 3.4.5. Die Norm (oder Länge) eines Vektors $\vec{x} \in H$ ist $\|\vec{x}\| = \sqrt{\langle \vec{x}|\vec{x}\rangle}$, bzw. $\langle \vec{x}|\vec{x}\rangle = \|\vec{x}\|^2$.

Im Folgenden sei H stets ein innerer Produktraum.

Definition 3.4.6. Eine endliche Folge $\vec{e}_1, \ldots, \vec{e}_n$ oder abzählbar unendliche Folge $(\vec{e}_j)$ von Vektoren $\vec{e}_j \in H$ heißt Orthogonalsystem, falls $\langle \vec{e}_i|\vec{e}_j\rangle = 0$ ist für alle $i \neq j$. Sie heißt Orthonormalsystem, falls zusätzlich $\langle \vec{e}_i|\vec{e}_i\rangle = 1$ ist für alle i. Bildet die Folge eine Basis eines endlichdimensionalen Unterraums V von H, so nennt man sie auch Orthonormalbasis von V.

Wir interessieren uns im Folgenden für den Abstand eines Vektors $\vec{y} \in H$ zu einem endlichdimensionalen Vektorraum $V \subseteq H$. Wir setzen dabei voraus, dass V von einer Orthonormalbasis (ONB) $\{\vec{e}_1, \ldots, \vec{e}_n\}$ erzeugt wird.

Definition 3.4.7. Es sei $V = \{\sum a_j\vec{e}_j \mid a_j \in \mathbb{C}\}$ ein endlichdimensionaler Unterraum von H mit Orthonormalbasis $B = \{\vec{e}_1, \ldots, \vec{e}_n\}$. Für $\vec{x} \in H$ nennen wir

$$P_V(\vec{x}) \;=\; \sum_{j=1}^{n} a_j\vec{e}_j \quad , \quad a_j = \langle \vec{x}|\vec{e}_j\rangle$$

die Projektion von $\vec{x}$ auf V. Die a_j heißen auch Fourierkoeffizienten von $\vec{x}$ bzgl. B.

Definition 3.4.8. Es sei $\vec{x} \in H$ und $M \subseteq H$. Der Abstand $d(\vec{x}, M)$ des Vektors $\vec{x}$ von der Menge M ist definiert durch

$$d(\vec{x}, M) \;=\; \inf_{\vec{u}\in M} \|\vec{u} - \vec{x}\| \,.$$

Satz 3.4.9. *Es sei V ein endlichdimensionaler Unterraum von H mit ONB $B = \{\vec{e}_1, \ldots, \vec{e}_n\}$, und $\vec{x} \in H$ beliebig. Es ist*

$$\langle \vec{x} - P_V(\vec{x})|\vec{u}\rangle \;=\; 0 \;\; \forall \vec{u} \in V \,.$$

Für $\vec{u} \in V$ gilt

$$d(\vec{x}, V) = \|\vec{x} - \vec{u}\| \;\Leftrightarrow\; \vec{u} = P_V(\vec{x}) \,,$$

d. h. unter allen Vektoren $\vec{u} \in V$ hat die Projektion $\vec{u} = P_V(\vec{x})$ den geringsten Abstand zu $\vec{x}$, und ist durch diese Eigenschaft eindeutig bestimmt.

Beweis.
Es sei $P_V(\vec{x}) = \sum a_j\vec{e}_j$ mit $a_j = \langle \vec{x}|\vec{e}_j\rangle$ und $\vec{u} = \sum b_j\vec{e}_j$ mit Koeffizienten $b_j \in \mathbb{C}$. Dann ist

$$\langle \vec{x} - P_V(\vec{x})|\vec{u}\rangle \;=\; \Big\langle \vec{x} - \sum a_j\vec{e}_j \Big| \sum b_j\vec{e}_j \Big\rangle$$

$$= \sum_{j=1}^{n} \bar{b}_j \,\langle \vec{x}|\vec{e}_j\rangle \;-\; \sum_{i=1}^{n}\sum_{j=1}^{n} \langle a_i\vec{e}_i|b_j\vec{e}_j\rangle \;=\; \sum_{j=1}^{n} a_j\bar{b}_j \;-\; \sum_{j=1}^{n} a_j\bar{b}_j \;=\; 0 \,.$$

Es sei $\vec{u} = P_V(\vec{x}) + \vec{w}$ für ein $\vec{w} \in V$, dann ist

$$\begin{aligned} \|\vec{u} - \vec{x}\|^2 &= \langle P_V(\vec{x}) + \vec{w} - \vec{x}|P_V(\vec{x}) + \vec{w} - \vec{x}\rangle \\ &= \langle P_V(\vec{x}) - \vec{x}|P_V(\vec{x}) - \vec{x}\rangle + \langle P_V(\vec{x}) - \vec{x}|\vec{w}\rangle + \overline{\langle P_V(\vec{x}) - \vec{x}|\vec{w}\rangle} + \langle \vec{w}|\vec{w}\rangle \\ &= \|P_V(\vec{x}) - \vec{x}\|^2 + \|w\|^2 \,. \end{aligned}$$

Also ist $\|\vec{u} - \vec{x}\| = \|P_V(\vec{x}) - \vec{x}\|$ genau dann, wenn $\vec{w} = \vec{0}$ ist bzw. $\vec{u} = P_V(\vec{x})$. □

Unser Hauptinteresse gilt unendlichdimensionalen inneren Produkträumen H, in denen unendliche ONB $B = (\vec{e}_j)$ gegeben sind, so dass sich jedes $\vec{x} \in H$ beliebig gut durch Linearkombinationen von endlich vielen $\vec{e}_j$ approximieren lässt.

Definition 3.4.10. Es sei $B = (\vec{e}_j)$ ein abzählbar unendliches Orthonormalsystem. Für $n \in \mathbb{N}$ sei $H_n = \langle \vec{e}_1, \ldots, \vec{e}_n \rangle = \{\lambda_1 \vec{e}_1 + \cdots + \lambda_n \vec{e}_n \mid \lambda_i \in \mathbb{C}\}$. Die Basis B heißt vollständig, falls es für alle $\vec{x} \in H$ und jedes $\varepsilon > 0$ ein $n = n(\varepsilon) \in \mathbb{N}$ und ein $\vec{x}_n \in H_n$ gibt mit $\|\vec{x} - \vec{x}_n\| < \varepsilon$.

Satz 3.4.11. *Es sei $B = (\vec{e}_j)$ ein abzählbar unendliches volllständiges Orthonormalsystem von H. Für alle $\vec{x} \in H$ gilt dann:*

$$\lim_{n\to\infty} \|\vec{x} - P_{H_n}(\vec{x})\| \;=\; 0\,.$$

Beweis.
Es sei $\varepsilon > 0$ gegeben. Nach Definition der Vollständigkeit gibt es $n_0 = n(\varepsilon)$ und $\vec{x}_{n_0} \in H_{n_0}$, so dass $\|\vec{x} - \vec{x}_{n_0}\| < \varepsilon$ ist. Für $n \geq n_0$ gilt nach Satz 3.4.9:

$$\|\vec{x} - P_{H_n}(\vec{x})\| \;\leq\; \|\vec{x} - P_{H_{n_0}}(\vec{x})\| \;\leq\; \|\vec{x} - \vec{x}_{n_0}\| \;<\; \varepsilon\,.$$

Also ist konvergiert die (reelle) Folge $(\|\vec{x} - P_{H_n}(\vec{x})\|)$ gegen Null für $n \to \infty$. □

3.5 Vollständigkeit des Systems der trigonometrischen Funktionen, Fourierreihen

Wir betrachten nun den inneren Produktraum $\mathcal{C} = \{f : [0,1] \to \mathbb{C} \mid f \text{ stetig}\}$ mit dem inneren Produkt

$$(1) : \langle f|g\rangle \;=\; \int_0^1 f(x)\overline{g(x)}dx$$

und dem Orthonormalsystem

$$l \;=\; (E_n)_{n=-\infty}^{\infty} \;,\; E_n : x \mapsto E(nx)$$

Die Funktionen E_n (mit $n \in \mathbb{Z}$) können zu einer Folge l_n mit $n \in \mathbb{N}$ gemacht werden, wenn $l_{2n} = E_n$ und $l_{2n+1} = E_{-n}$ gesetzt wird. Wie in der allgemeinen Diskussion des vorigen Abschnitts können wir endlichdimensionale Unterräume

$$C_N \;=\; \langle E_{-N}, E_{-N+1}, \ldots, 1 = E_0, E_1, \ldots, E_N\rangle \;\subset\; \mathcal{C}$$

betrachten und folgendes Problem behandeln: Es sei $f \in \mathcal{C}$ gegeben. Für welche $S_N \in C_N$ ist das durch (1) definierte Quadrat des Abstands

$$\int_0^1 |f(x) - S_N(x)|^2 dx$$

von f zu C_N minimal? Nach Satz 3.4.9 ist als S_N die Projektion von f auf C_N zu wählen:

$$S_N(x) \;=\; \sum_{n=-N}^{N} a_n E_n(x) \;,\; a_n = \int_0^1 f(x)E_{-n}(x)dx\,.$$

Man kann nun erwarten, dass in vielen Fällen

$$\lim_{N\to\infty} S_N(x) \;=\; f(x)$$

für alle $x \in [0,1]$ ist, d. h. dass f durch seine Fourierreihe

$$f(x) \;=\; \sum_{n=-\infty}^{\infty} a_n E_n(x)$$

dargestellt wird. Dies gibt Anlass zu folgender

Definition 3.5.1. Es sei $f:\ [0,1] \to \mathbb{C}$ stetig. Unter dem Fourierkoeffizienten $\hat{f}(n)$ von f verstehen wir

$$\hat{f}(n) \;=\; \int_0^1 f(x) E_{-n} dx\,,$$

unter der N-ten Partialsumme

$$S_N(x) \;=\; \sum_{n=-N}^{N} \hat{f}(n) E_n(x)$$

und unter der Fourierreihe von f

$$\sum_{n=-\infty}^{\infty} \hat{f}(n) E_n(x)$$

die Folge der Partialsummen S_N.

Wir zeigen zunächst die Vollständigkeit des Orthonormalsystems l. Hierfür muss gezeigt werden, dass jedes $f \in \mathcal{C}$ im quadratischen Mittel beliebig gut durch Funktionen aus der Folge der Vektorräume C_N approximiert werden kann: Sind $f \in \mathcal{C}$ und $\epsilon > 0$ gegeben, so gibt es N und $f_N \in C_N$ mit $\int_0^1 |f(x) - f_N(x)|^2\, dx < \epsilon$. Dafür wäre es hinreichend zu zeigen: sind $f \in \mathcal{C}$ und $\tilde{\epsilon} > 0$ gegeben, so gibt es N und $f_N \in C_N$ mit $\sup_{x\in[0,1]} |f(x) - f_N(x)| < \tilde{\epsilon}$. Dies ist nicht immer mit der Wahl $f_N = S_N$ möglich, aber wie Féjèr gezeigt hat stets mit $f_N = T_N$, wobei die T_N die arithmetischen Mittel der S_N bedeuten. Um die Partialsummen S_N bzw. deren arithmetisches Mittel T_N auszudrücken, benötigt man den Dirichletkern und den Féjèrkern, die wir im folgenden definieren wollen:

Definition 3.5.2. Der N-te Dirichletkern $D_n(x)$ ist gegeben durch

$$D_N(x) \;=\; \sum_{n=-N}^{N} E_n(x)\,.$$

Der N-te Féjèrkern $F_n(x)$ ist gegeben durch

$$F_N(x) \;=\; \frac{1}{N+1} \sum_{n=0}^{N} D_n(x)\,.$$

Ferner setzen wir

$$T_N(x) \;=\; \frac{1}{N+1} \sum_{n=0}^{N} S_n(x)\,.$$

Ein entscheidender Vorzug des Féjèrkerns gegenüber dem Dirichletkern ist, dass F_N stets nichtnegativ ist.

Satz 3.5.3. *Es gilt:*

$$\begin{aligned}
\text{Kerndarstellung}: \quad D_N(x) &= \tfrac{\sin(2\pi(N+\frac{1}{2})x)}{\sin(\pi x)} \\
F_N(x) &= \tfrac{1}{N+1} \cdot \tfrac{\sin^2(\pi(N+1)x)}{\sin^2(\pi x)} \\
\text{Faltungseigenschaft}: \quad S_N(x) &= \int_0^1 f(t) D_N(x-t)dt \\
T_N(x) &= \int_0^1 f(t) F_N(x-t)dt
\end{aligned}$$

$$\text{Normierung}: \int_0^1 D_N(x-t)dt \;=\; \int_0^1 F_N(x-t)dt \;=\; 1 \quad \forall x \in [0,1]\,.$$

Beweis.
Aus der Formel für die endliche geometrische Reihe ergibt sich

$$D_N(x) \;=\; E(-Nx) \cdot \frac{E((2N+1)x)-1}{E(x)-1} \;=\; \frac{E((N+\frac{1}{2})x) - E(-(N+\frac{1}{2})x)}{E(\frac{1}{2}x) - E(-\frac{1}{2}x)} \;=\; \frac{\sin(2\pi(N+\frac{1}{2})x)}{\sin(\pi x)}\,.$$

Die Formel für F_N ergibt sich aus der für D_N durch nochmalige Anwendung der Summenformel für die geometrische Reihe. Es ist zudem

$$\int_0^1 f(t) D_N(x-t)dt \;=\; \int_0^1 f(t) \sum_{n=-N}^{N} E_n(x-t)dt \;=\; \sum_{n=-N}^{N} E_n(x) \int_0^1 f(t) E_n(-t)dt \;=\; \sum_{n=-N}^{N} \hat{f}(n) E_n(x)\,,$$

wobei der letzte Ausdruck nach Definition gerade $S_N(x)$ ist. Analog ist

$$\int_0^1 f(t) F_N(x-t)dt \;=\; \frac{1}{N+1} \sum_{n=0}^{N} \int_0^1 f(t) D_n(x-t)dt \;=\; \frac{1}{N+1} \sum_{n=0}^{N} S_N(x) \;=\; T_N(x)\,.$$

Zudem ist

$$\int_0^1 D_N(x-t)dt \;=\; \sum_{n=-N}^{N} E_n(x) \int_0^1 E_{-n}(t)dt \;=\; 1\;,\quad \int_0^1 F_N(x-t)dt \;=\; \frac{1}{N+1} \sum_{n=0}^{N} \int_0^1 D_N(x-t)dt \;=\; 1\,.$$

□

Satz 3.5.4. *Sei $f : [0,1] \to \mathbb{C}$ stetig. Dann konvergiert die Folge der arithmetischen Mittel T_N der Partialsummen gleichmäßig auf $[0,1]$ gegen f.*

Beweis.
Es sei $\varepsilon > 0$ gegeben. Die stetige Funktion f ist auf dem kompakten Intervall $[0,1]$ beschränkt, und gleichmäßig stetig. Daher gibt es ein $M > 0$, so dass $|f(t)| \leq M$ ist für alle $t \in [0,1]$, sowie ein $\delta = \delta(\varepsilon)$, so dass für alle $x \in [0,1]$ gilt:

$$|f(t) - f(x)| < \varepsilon \text{ falls } |t-x| < \delta\,.$$

Wegen

$$|F_N(x-t)| \;\leq\; \frac{1}{\sin^2(\pi\delta)} \cdot \frac{1}{N+1} \text{ falls } |x-t| \geq \delta$$

folgt

$$(1) \quad : \quad \lim_{N\to\infty} \left(\int\limits_{|x-t|\geq\delta} f(t)F_N(x-t)dt \right) \;=\; 0 \quad \text{und}$$

$$(2) \quad : \quad \lim_{N\to\infty} \left(\int\limits_{|x-t|\geq\delta} f(x)F_N(x-t)dt \right) \;=\; 0$$

und aus der Normierung von F_N folgt

$$(3) : \left| \int\limits_{|x-t|\leq\delta} (f(t)-f(x))F_N(x-t)dt \right| \;<\; \varepsilon \,.$$

Es ist

$$\int\limits_0^1 f(t)F_N(x-t)dt \;=\; \int\limits_{|x-t|\leq\delta} f(t)F_N(x-t)dt + \int\limits_{|x-t|\geq\delta} f(t)F_N(x-t)dt \;=\; \int_0^1 f(x)F_N(x-t)\,dt$$

$$- \int\limits_{|x-t|>\delta} f(x)F_N(x-t)dt + \int\limits_{|x-t|\leq\delta} (f(t)-f(x))F_N(x-t)dt + \int\limits_{|x-t|\geq\delta} f(t)F_N(x-t)dt \,.$$

Mit (1), (2) und (3) folgt dann

$$\lim_{N\to\infty} \int\limits_0^1 f(t)F_N(x-t)dt \;=\; \int\limits_0^1 f(x)F_N(x-t)dt \;=\; f(x) \,.$$

□

Satz 3.5.5 (Konvergenz im quadratischen Mittel)**.** *Es sei* $f : [0,1] \to \mathbb{C}$ *stetig. Dann ist*

$$(\text{i}) : \lim_{N\to\infty} \int\limits_0^1 |f(x) - S_N(x)|^2 dx \;=\; 0 \,,$$

und es gilt die <u>*Parsevalsche Gleichung*</u>

$$(\text{ii}) : \int\limits_0^1 |f(x)|^2 dx \;=\; \sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2 \,.$$

Beweis.
Es sei $\varepsilon > 0$ gegeben. Nach Satz 3.5.4 gibt es $N_0 = N_0(\varepsilon)$, so dass für $N \geq N_0$ gilt: $|T_N(x) - f(x)| \leq \varepsilon$ für alle $x \in [0,1]$. Daraus folgt

$$\int\limits_0^1 |T_N(x) - f(x)|^2 dx \;\leq\; \varepsilon^2 \,.$$

Wegen $T_N \in C_N$ besitzt das System der E_n nach Definition 3.4.10 die Eigenschaft der Vollständigkeit. Aussage (i) folgt somit aus Satz 3.4.11. Nach der Cauchy-Schwarzschen Ungleichung folgt nun:

$$\left| \int_0^1 f(x)\overline{(f(x) - S_N(x))}dx \right| \leq \left(\int_0^1 |f(x)|^2 dx \right)^{\frac{1}{2}} \cdot \left(\int_0^1 |f(x) - S_N(x)|^2 dx \right)^{\frac{1}{2}},$$

also

$$\lim_{N\to\infty} \int_0^1 f(x)\overline{(f(x) - S_N(x))}dx = 0$$

und somit

$$\int_0^1 |f(x)|^2 dx = \lim_{N\to\infty} \int_0^1 f(x)\overline{S_N(x)}dx\,.$$

Ähnlich folgt

$$\lim_{N\to\infty} \int_0^1 \overline{S_N(x)}(f(x) - S_N(x))dx = 0$$

und

$$\lim_{N\to\infty} \int_0^1 f(x)\overline{S_N(x)}dx = \lim_{N\to\infty} \int_0^1 |S_N(x)|^2 dx = \sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2\,.$$

□

Die Beziehung (i) heißt auch Konvergenz im quadratischen Mittel. Aus ihr folgt nicht die punktweise Konvergenz. Es gibt Beispiele von stetigen Funktionen f, deren Fourierreihe nicht überall gegen f konvergiert. Wir schließen mit ein paar Sätzen aus der umfangreichen Theorie der Konvergenz der Fourierreihen:

Satz 3.5.6. *Es sei $f : [0,1] \to \mathbb{C}$ stetig. Gilt*

$$\lim_{N\to\infty} S_N(x) = S(x)$$

für ein $x \in [0,1]$, so folgt $S(x) = f(x)$. Wenn also die Fourierreihe einer stetigen Funktion f in einem Punkt x konvergiert, so konvergiert sie gegen den „richtigen Wert“ $f(x)$.

Beweis.
Aus dem Cauchyschen Grenzwertsatz folgt mit der Voraussetzung dann

$$\lim_{N\to\infty} T_N(x) = S(x)\,.$$

Nach Satz 3.5.4 ist dann $S(x) = f(x)$. □

Satz 3.5.7. *Es sei $f : [0,1] \to \mathbb{C}$ stetig differenzierbar und $f(0) = f(1)$. Dann gilt*

$$\lim_{N\to\infty} S_N(x) = f(x)$$

für alle $x \in [0,1]$.

Beweis.
Durch partielle Integration folgt

$$\int_0^1 f'(x)E(-nx)dx \;=\; [f(x)\cdot E(-nx)]_0^1 + 2\pi in\int_0^1 f(x)E(-nx)dx \;=\; 2\pi in\hat{f}(n)\,.$$

Nach der Parsevalschen Gleichung für $f'(x)$ ist

$$\sum_{n=-\infty}^{\infty} n^2|\hat{f}(n)|^2 \;<\; \infty\,,$$

damit gibt es für jedes $\varepsilon > 0$ ein $N_0 = N_0(\varepsilon)$, so dass für alle $N \geq N_0$ gilt:

$$\sum_{|n|\geq N} |\hat{f}(n)|^2 \;\leq\; \frac{\varepsilon^2}{N^2}\,.$$

Es folgt

$$\left|\sum_{2^kN_0\leq|n|\leq 2^{k+1}N_0} \hat{f}(n)E(nx)\right| \;\leq\; \left(\sum_{2^kN_0\leq|n|\leq 2^{k+1}N_0} |\hat{f}(n)|^2\right)^{\frac{1}{2}} \cdot \left(\sum_{2^kN_0\leq|n|\leq 2^{k+1}N_0} 1\right)^{\frac{1}{2}} \;\leq\; \varepsilon\cdot 2^{-\frac{k}{2}+1}\,.$$

Damit ist die Folge $S_N(x)$ gleichmäßig konvergent auf $[0,1]$ und konvergiert damit gegen eine stetige Funktion $S(x)$. Nach Satz 3.5.6 ist $S(x) = f(x)$ für alle $x \in [0,1]$. □

Bemerkung 3.5.8. Die Behauptung von Satz 3.5.7 gilt auch, wenn die Funktion f nur als stückweise stetig differenzierbar vorausgesetzt wird.

Kapitel 4

Fraktale Mengen und Dimensionen

4.1 Einleitung

Das Maß von Objekten war schon immer ein zentraler Begriff in Mathematik und Naturwissenschaften. Je nach der Dimension des Objekts war dieses Maß unter dem Namen Länge, Fläche oder Volumen bekannt. Der Zusammenhang zwischen Maß und Dimension ist durch die Skalierungseigenschaft gegeben: Ist $F \subset \mathbb{R}^n$ ein n-dimensionales Objekt, so kann man sein Maß mit dem des ähnlichen Objekts λF vergleichen. Die Dimension ist aus dem Exponenten des Skalierungsgesetzes

$$(1) \quad m(\lambda F) \;=\; \lambda^n \cdot m(F) \quad (\lambda > 0)$$

ersichtlich. Dieses Gesetz ist eine Folge von mehreren Eigenschaften des Maßes:
(2) Additivität: Ist $(F_i)_{i=1\ldots\infty}$ eine Folge von paarweise disjunkten Mengen, sie alle ein Maß besitzen (d. h. messbar sind), so ist

$$m\left(\bigcup_{i=1}^{\infty} F_i\right) \;=\; \sum_{i=1}^{\infty} m(F_i)\,,$$

sowie
(3) Translationsinvarianz: Ist $F \subseteq \mathbb{R}^n$ und $\vec{c} \in \mathbb{R}^n$, so ist $m(F + \vec{c}) = m(F)$, und
(4) die Überdeckung der n-dimensionalen Grundfigur des n-dimensionalen Würfels durch ähnliche Würfel (Selbstähnlichkeit). Eine Ähnlichkeitsabbildung ist eine Abbildung $S:\ D \to D$ für $D \subseteq \mathbb{R}^n$ mit

$$|S(\vec{x}) - S(\vec{y})| \;=\; \lambda|\vec{x} - \vec{y}| \;\; \forall \vec{x}, \vec{y} \in \mathbb{R}^n\,,\; \lambda > 0\,.$$

Sind $S_1, \ldots, S_m$ Ähnlichkeitsabbildungen, so heißt eine Menge F invariant unter $S_1, \ldots, S_m$, falls

$$F \;=\; \bigcup_{i=1}^{m} S_i(F)\,.$$

Ein n-dimensionaler Würfel W kann für jedes $N \in \mathbb{N}$ in N^n Teilwürfel W_i unterteilt werden. Jeder Teilwürfel ist das Bild des gesamten Würfels unter einer Ähnlichkeitsabbildung S_i von der Form

$$S_i:\ \vec{x} \;\mapsto\; \frac{1}{N}\vec{x} + \vec{y}_i \quad (1 \le i \le N^n)\,.$$

Nimmt man zusätzlich zu den Eigenschaften (2) und (3) noch an, dass die Ränder der Teilwürfel das Maß 0 besitzen, so folgt

$$m(W) \;=\; \sum_{i=1}^{N^n} m(W_i) \;\text{ und }\; m(N^{-1}W) \;=\; N^{-n} \cdot m(W)\,.$$

Ein Skalierungsgesetz von der Form (1) $m(\lambda F) = \lambda^s m(F)$ kann also nur für $s = n$ richtig sein. In diesen Beispielen ist die Dimension n stets eine positive ganze Zahl. In diesem Kapitel sollen nun Gebilde mit nichtganzer Dimension zur Sprache kommen, so genannte fraktale Mengen. Das möglicherweise wichtigste Beispiel gibt

Definition 4.1.1. Die Cantor-Menge C ist definiert als der Durchschnitt einer absteigenden Kette von Mengen:

$$C \;=\; \bigcap_{i=0}^{\infty} C_i \;\text{ mit }\; C_0 \;\supseteq\; C_1 \;\supseteq\; C_2 \;\supseteq\; \cdots\,.$$

Diese Kette wird rekursiv definiert. $C_0 = [0,1]$, und C_{i+1} entsteht aus C_i wie folgt: Ist

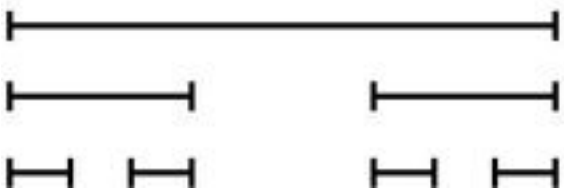
$$C_i \;=\; \bigcup_{j=0}^{l(i)} I_i^{(j)}$$

für abgeschlossene Intervalle $I_i^{(j)}$, so ist

$$C_{i+1} \;:=\; \bigcup_{j=1}^{l(i)} \left(I_{i,1}^{(j)} \cup I_{i,2}^{(j)} \right)\,,$$

wobei $I_{i,1}^{(j)}$ und $I_{i,2}^{(j)}$ aus $I_i^{(j)}$ durch Entfernung des offenen mittleren Drittels entstehen (C_{i+1} ist wieder eine Vereinigung abgeschlossener Intervalle).

Es gilt also $C_0 = [0,1]$, $C_1 = [0,\frac{1}{3}] \cup [\frac{2}{3},1]$, $C_2 = [0,\frac{1}{9}] \cup [\frac{2}{9},\frac{1}{3}] \cup [\frac{2}{3},\frac{7}{9}] \cup [\frac{8}{9},1]$, usw.



Wir haben die Erscheinung der Selbstähnlichkeit: Die Teilmengen $C \cap [0,\frac{1}{3}]$ und $C \cap [\frac{2}{3},1]$ sind Bilder von C unter der Ähnlichkeitsabbildung $S_1 :\; x \mapsto \frac{1}{3}x$ bzw. $S_2 :\; x \mapsto \frac{1}{3}x + \frac{2}{3}$, es ist also

$$C \;=\; S_1(C) \,\cup\, S_2(C)\,.$$

Ist nun $m(\cdot)$ irgend ein Maß, das die Skalierungseigenschaft (1) $m(\lambda F) = \lambda^s \cdot m(F)$ und die Eigenschaften (2) und (3) (Additivität und Translationsinvarianz) besitzt, so ist

$$m(C) \;=\; m(S_1(C)) \,+\, m(S_2(C)) \;=\; 2\left(\frac{1}{3}\right)^s \cdot m(C)\,.$$

Soll $m(C) \neq 0$ gelten, so folgt

$$s \;=\; \frac{\log(2)}{\log(3)}\,.$$

In Analogie zu dem Beispiel der Würfel hätte also C die (fraktale) Dimension $\frac{\log(2)}{\log(3)}$. Hausdorff hat ein solches Maß konstruiert. Es besitzt die Eigenschaft (2) der Additivität, was schwierig zu beweisen ist, und wir nicht zeigen wollen. Es ist schwierig zu berechnen. Ein einfacher zu bestimmendes Maß hat Minkowski konstruiert. Im Gegensatz zu Hausdorffs Maß besitzt es die Eigenschaft (2) der Additivität nur für endliche Folgen von disjunkten Mengen. Während früher fraktale Mengen als Kuriositäten galten, haben sie in der jüngeren Vergangenheit in vielen Wissenschaftszweigen an Bedeutung gewonnen.

4.2 Hausdorff-Maß und -Dimension

Definition 4.2.1. Der Durchmesser einer nichtleeren Teilmenge $U \subseteq \mathbb{R}^n$ ist durch

$$|U| \;=\; \sup\{\|\vec{x}-\vec{y}\| \mid \vec{x},\vec{y} \in U\}$$

definiert. Ist $\{U_i\}$ eine abzählbare (oder endliche) Auswahl von Mengen, deren Durchmesser jeweils höchstens δ beträgt, und die eine Menge $F \subseteq \mathbb{R}^n$ überdecken (d. h. $F \subseteq U_1 \cup U_2 \cup \cdots$ mit $0 < |U_i| \leq \delta$), so heißt $\{U_i\}$ eine δ-Überdeckung von F.

Definition 4.2.2. Es sei $F \subseteq \mathbb{R}^n$ und $s \geq 0$. Für jedes $\delta > 0$ definieren wir

$$H^s_\delta(F) \;=\; \inf\left\{\sum_{i=1}^{\infty} |U_i|^s \mid \{U_i\} \text{ ist } \delta\text{-Überdeckung von } F\right\}.$$

Bemerkung 4.2.3. Ist $\sum |U_i|^s = \infty$ für alle δ-Überdeckungen von F, so setzen wir $H^s_\delta(F) = \infty$.

Satz 4.2.4. *Es sei $F \subseteq \mathbb{R}^n$ und $s \geq 0$. $H^s_\delta(F)$ ist monoton nichtwachsend in δ und der Grenzwert*

$$\lim_{\delta \to 0} H^s_\delta(F)$$

existiert.

Beweis.
Für $\delta_1 < \delta_2$ ist jede δ_1-Überdeckung von F auch eine δ_2-Überdeckung, und es ist

$$\lim_{\delta \to 0} H^s_\delta(F) \;=\; \inf_{\delta > 0} H^s_\delta(F)\,.$$

□

Definition 4.2.5. Es sei $F \subseteq \mathbb{R}^n$ und $s \geq 0$. Der Wert

$$H^s(F) \;:=\; \lim_{\delta \to 0} H^s_\delta(F)$$

heißt s-dimensionales Hausdorff-Maß von F.

Satz 4.2.6 (Skalierungseigenschaft)**.** *Wenn $F \subseteq \mathbb{R}^n$ und $\lambda > 0$ ist, dann gilt*

$$H^s(\lambda F) \;=\; \lambda^s H^s(F)\,.$$

Beweis.
$\{U_i\}$ ist eine δ-Überdeckung von F genau dann, wenn $\{\lambda U_i\}$ eine $(\lambda\delta)$-Überdeckung von λF ist. Also gilt

$$H^s_{\lambda\delta}(\lambda F) \;=\; \inf\left\{\sum_{i=1}^{\infty} |\lambda U_i|^s \mid \{\lambda U_i\} \text{ ist } (\lambda\delta)\text{-Überdeckung von } \lambda F\right\}$$

$$=\; \lambda^s \cdot \inf\left\{\sum_{i=1}^{\infty} |U_i|^s \mid \{U_i\} \text{ ist } \delta\text{-Überdeckung von } F\right\} \;=\; \lambda^s H^s_\delta(F)\,.$$

Daraus folgt

$$H^s(\lambda F) \;=\; \lim_{\delta \to 0} H^s_{\lambda\delta}(\lambda F) \;=\; \lambda^s \lim_{\delta \to 0} H^s_\delta(F) \;=\; \lambda^s H^s(F)\,.$$

□

Satz 4.2.7 (Monotonieeigenschaft)**.** *Es sei $F \subseteq G \subseteq \mathbb{R}^n$, $n \in \mathbb{N}$ und $s \geq 0$. Dann ist $H^s(F) \leq H^s(G)$.*

Beweis.
Die Behauptung ist offensichtlich, da jede δ-Überdeckung von G auch eine von F ist. □

Das Hausdorff-Maß erfüllt die abzählbare Additivität: Ist $\{F_i\}_{i=1}^{\infty}$ ein abzählbare Folge von paarweise disjunkten Mengen, so ist

$$H^s\left(\bigcup_{i=1}^{\infty} F_i\right) \;=\; \sum_{i=1}^{\infty} H^s(F_i)\,.$$

Dies ist schwierig zu beweisen. Wir beschränken uns hier auf den Fall endlich vieler kompakter und paarweise disjunkter Mengen:

Satz 4.2.8. *Es seien $F_1, \ldots, F_m \subset \mathbb{R}^n$ kompakt und paarweise disjunkt und $s \geq 0$. Dann ist*

$$H^s\left(\bigcup_{i=1}^{m} F_i\right) \;=\; \sum_{i=1}^{m} H^s(F_i)\,.$$

Beweis.
Es genügt offenbar, die Behauptung für $m = 2$ zu beweisen. Es sei $\{U_i\}$ eine δ-Überdeckung von $F_1 \cup F_2$. Ist $\delta > 0$ genügend klein, so gilt wegen der Kompaktheit der F_j für jedes i

$$U_i \cap F_1 \;=\; \emptyset \quad \text{oder} \quad U_i \cap F_2 \;=\; \emptyset\,.$$

Dann ist $\mathcal{U}^1 = \{U_i \mid U_i \cap F_1 \neq \emptyset\}$ eine δ-Überdeckung von F_1 und $\mathcal{U}^2 = \{U_i \mid U_i \cap F_2 \neq \emptyset\}$ eine von F_2, d. h.

$$\sum_{U \in \mathcal{U}^1} |U|^s \;+\; \sum_{U \in \mathcal{U}^2} |U|^s \;\leq\; \sum_{i=1}^{\infty} |U_i|^s\,.$$

Daraus folgt

$$H^s(F_1) \;+\; H^s(F_2) \;\leq\; H^s(F_1 \cup F_2)\,.$$

Die umgekehrte Ungleichung gilt trivialerweise, woraus der Satz folgt. □

Die Definition der Hausdorff-Dimension beruht auf der folgenden Eigenschaft des Hausdorff-Maßes:

Satz 4.2.9. *Es sei $F \subseteq \mathbb{R}^n$. $H^s(F)$ ist monoton nichtwachsend in s, und es ist $H^s(F) = 0$ für $s > n$. Ist $0 < H^d(F) < \infty$, so ist $H^s(F) = 0$ für alle $s > d$ und $H^s(F) = \infty$ für $s < d$.*

Beweis.
Da $|U_i|^s$ monoton nichtwachsend in s ist gilt das auch für

$$H^s_\delta(F) \;=\; \inf\left\{\sum_{i=1}^{\infty} |U_i|^s \mid \{U_i\} \text{ ist } \delta\text{-Überdeckung von } F\right\} \text{ und für } H^s(F) \;=\; \lim_{\delta \to 0} H^s_\delta(F)\,.$$

Wir beweisen zunächst, dass $H^s(\mathbb{R}^n) = 0$ für $s > n$ ist. Für beliebige $F \subseteq \mathbb{R}^n$ folgt die Behauptung dann aus der Monotonie. Es sei

$$\mathbb{N} \;\longrightarrow\; \mathbb{Z}^n \;\;,\;\; i \;\longmapsto\; \left(m_1^{(i)}, \ldots, m_n^{(i)}\right)$$

irgend ein bijektive Abbildung, und W_i der Würfel

$$W_i \;=\; \left\{\vec{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n \mid m_j^{(i)} \;\leq\; x_j \;\leq\; m_j^{(i)} + 1\right\}\,.$$

Das ist eine Überdeckung des $\mathbb{R}^n$ mit Würfeln der Kantenlänge 1. Es sei (N_i) eine Folge natürlicher Zahlen. Jedes W_i kann durch N_i^n achsenparallele Würfel $W_{i,j}$ der Kantenlänge N_i^{-1} überdeckt werden. Der Durchmesser dieser ergibt sich zu $c_n N_i^{-1}$ mit $c_n = \sqrt{n}$. Für $s > n$ ist

$$\sum_{1 \le j \le N_i^n} |W_{i,j}|^s \;\le\; c_n N_i^{n-s} \,.$$

Setzen wir jetzt $N_i = 2^i N_0$, so folgt

$$\sum_{i=1}^{\infty} \sum_{1 \le j \le N_i^n} |W_{i,j}|^s \;\le\; c_n N_0^{n-s} \sum_{i=1}^{\infty} 2^{-i(s-n)} \;\underset{N_0 \to \infty}{\longrightarrow}\; 0 \,,$$

da die letzte Summe wegen $-(s-n) < 0$ eine geometrische Reihe endlichen Werts ist. Daraus folgt $H^s(\mathbb{R}^n) = 0$. Nun sei $H^d(F) = a$ für $0 < a < \infty$. Dann gibt es $\delta_0 > 0$, so dass für $\delta < \delta_0$ gilt:
Es gibt eine δ- Überdeckung $\{U_i\}$ von F mit

$$(1) \quad \sum_{i=1}^{\infty} |U_i|^d \;<\; 2a \,.$$

Für alle δ-Überdeckungen $\{V_i\}$ von F gilt

$$(2) \quad \sum_{i=1}^{\infty} |V_i|^d \;>\; \frac{1}{2} a \,.$$

Es sei $s > d$ und $\{U_i\}$ eine δ-Überdeckung, für die (1) gilt. Wegen $|U_i| < \delta$ und $s - d > 0$ folgt $|U_i|^s \le |U_i|^d \cdot \delta^{s-d}$, also

$$\sum_{i=1}^{\infty} |U_i|^s \;<\; 2a \cdot \delta^{s-d} \;\Rightarrow\; H^s_\delta(F) \;<\; 2 \cdot \delta^{s-d} \;\Rightarrow\; H^s(F) \;=\; \lim_{\delta \to 0} H^s_\delta(F) \;=\; 0 \,.$$

Es sei nun $s < d$, und $\{V_i\}$ eine δ-Umgebung für die (2) gilt. Wegen $|V_i| < \delta$ und $s - d < 0$ gilt $|V_i|^s \ge |V_i|^d \cdot \delta^{s-d}$, also jetzt

$$\sum_{i=1}^{\infty} |V_i|^s \;\ge\; \frac{1}{2} a \cdot \delta^{s-d} \;\Rightarrow\; H^s_\delta(F) \;\ge\; \frac{1}{2} a \cdot \delta^{s-d} \;\Rightarrow\; H^s(F) \;=\; \lim_{\delta \to 0} H^s_\delta(F) \;=\; \infty \,.$$

□

Bemerkung 4.2.10. Es gibt also höchstens einen „richtigen“ Exponenten d, für welchen das Hausdorff-Maß einen endlichen Wert annimmt. Insbesondere folgt

$$\inf \{s \mid H^s(F) = 0\} \;=\; \sup \{s \mid H^s(F) = \infty\} \,.$$

Definition 4.2.11. Für $n \in \mathbb{N}$ und $F \subseteq \mathbb{R}^n$ ist die Hausdorff-Dimension definiert durch

$$\dim_H(F) \;=\; \inf \{s \mid H^s(F) = 0\} \;=\; \sup \{s \mid H^s(F) = \infty\} \,.$$

Satz 4.2.12. *Es sei C die Cantor-Menge, dann ist*

$$\dim_H(C) \;\le\; \frac{\log(2)}{\log(3)} \,.$$

Beweis.
Nach Definition ist

$$C \;=\; \bigcap_{k=1}^{\infty} \;\text{ mit }\; C_1 \;\supset\; C_2 \;\supset\; \cdots \;\supset\; C_k \;\supset\; \cdots \;,$$

wobei C_k jeweils eine Vereinigung von 2^k Intervallen der Länge 3^{-k} ist. Es existiert damit zu jedem k eine δ_k-Überdeckung mit $\delta_k = 3^{-k}$ durch 2^k Mengen U_i. Es folgt:

$$d \;=\; \frac{\log(2)}{\log(3)} \;\;\Rightarrow\;\; H^d_\delta(C) \;\leq\; \sum_{i=1}^{2^k} |U_i|^d \;\leq\; 2^k (3^{-k})^{\frac{\log(2)}{\log(3)}} \;=\; 1$$

und somit $H^d(F) \leq 1 < \infty$. □

Es ist häufig sehr viel einfacher, obere Schranken für die Hausdorff-Dimension einer Menge anzugeben, als untere Schranken, oder die Hausdorff-Dimension exakt zu berechnen. Der Grund dafür ist der folgende: Für den Beweis, dass $\dim_H(F) < s_0$ ist genügt es zu zeigen, dass $H^s_\delta(F) \to 0$ strebt für $\delta \to 0$ und ein $s \leq s_0$. Dazu genügt eine obere Schranke der Form

$$(*) \quad H^s_\delta(F) \;\leq\; g(\delta)$$

für irgend eine Funktion g mit $g(\delta) \to 0$ für $\delta \to 0$. Nun ist aber

$$H^s_\delta(F) \;=\; \inf \left\{ \sum_{i=1}^{\infty} |U_i|^s \mid \{U_i\} \text{ ist } \delta\text{-Überdeckung von } F \right\}$$

als Infimum definiert. Für den Nachweis einer oberen Schranke für die Hausdorff-Dimension einer Menge genügt es daher, eine einzige Familie von Überdeckungen geeignet auszuwählen. Die analogen Argumente machen deutlich, dass es für den Nachweis einer unteren Schranke nötig ist, sämtliche Überdeckungen zu kontrollieren. Wir geben noch ein Beispiel eines fraktalen Objekts: die Koch-Kurve (nach von Koch). Unter einer Kurve stellt man sich im Allgemeinen ein Objekt der Dimension 1 vor. Es kann tatsächlich gezeigt werden, dass die Trägermenge einer glatten Kurve (d. h. das Bild einer stetig differenzierbaren Abbildung auf einem kompakten Intervall) die Hausdorff-Dimension 1 besitzt. Die Kochkurve ist das Bild des Einheitsintervalls $I = [0,1]$ unter einer stetigen aber nicht differenzierbaren Abbildung $f : [0,1] \to \mathbb{R}^2$. Wir erhalten f als Grenzwert einer Folge (f_n) von stückweise linearen Funktionen. Die f_k werden rekursiv definiert:

$\underline{k = 0}$
$f_0(t) = (t_0)^T$ für alle $t \in I$.

$\underline{k \to k+1}$
Das Intervall I wird in 4^k Teilintervalle

$$I_{j,k} \;=\; [a_{j,k}, b_{j,k}] \text{ mit } a_{j,k} \;=\; \frac{j}{4^k} \;,\; b_{j,k} \;=\; \frac{j+1}{4^k} \;,\; 0 \leq j \leq 4^k - 1$$

unterteilt. Auf $I_{j,k}$ ist f_k durch eine lineare Funktion $f_{j,k}$ definiert, die ihrerseits durch den Anfangswert $f_k(a_{j,k})$ und den Einheitsvektor $\vec{v}_{j,k}$ bestimmt ist. Es ist $f_k(a_{0,k}) = (0,0)^T$. Für $t = a_{j,k} + 4^{-k}u \in I_{j,k}$ mit $0 \leq u \leq 1$ ist

$$(*) \quad f_k(t) \;=\; f_k(a_{j,k}) \;+\; u\vec{v}_{j,k} \cdot 3^{-k} \,.$$

Dabei ergeben sich die $\vec{v}_{j,k}$ rekursiv wie folgt:

$\underline{k = 0}$
Es ist $\vec{v}_{1,0} = (1,0)^T$.

$\underline{k-1 \to k}$
Es sei $j = 4m+l$ mit $0 \le l \le 3$. Für $l = 0$ ist $\vec{v}_{j,k} = \vec{v}_{j,k-1}$. Für $l = 1$ geht $\vec{v}_{j,k}$ aus $\vec{v}_{j,k-1}$ durch Drehung um $\frac{1}{3}\pi$ in positivem Sinn hervor. Für $l = 2$ durch Drehung um $\frac{1}{3}\pi$ in negativem Sinn. Insgesamt wird das mittlere Drittel der Verbindungsstrecke durch ein gleichseitiges Dreieck ersetzt.

Aus $(*)$ ergibt sich $|f_k(t) - f_{k-1}(t)| \le 3^{-k}$. Damit ist (f_k) sogar gleichmäßig konvergent gegen ein $f = \lim f_k$, das dann wie die f_k auch stetig ist. Wir setzen $K_k = \{f_k(t) \mid 0 \le t \le 1\}$ und $K = \{f(t) \mid 0 \le t \le 1\}$. Man sieht leicht, dass K durch ein Quadrat der Seitenlänge 1 überdeckt werden kann.
Jedes K_k besteht aus 4^k Teilstrecken der Länge 3^{-k}. Da die von den Endpunkten jeder Teilstrecke begrenzte Teilmenge zur Gesamtmenge K ähnlich ist folgt, dass K_k von 4^k Quadraten der Länge 3^{-k} überdeckt werden kann. Also gilt mit $\delta_k = 3^{-k}$ und $d = \frac{\log(4)}{\log(3)}$

$$H^d_{\delta_k}(K) \;\le\; \sum_{j=1}^{4^k} |U_i|^d \;=\; \sqrt{2} \cdot 4^k (3^{-k})^{\frac{\log(4)}{\log(3)}} \;=\; \sqrt{2}$$

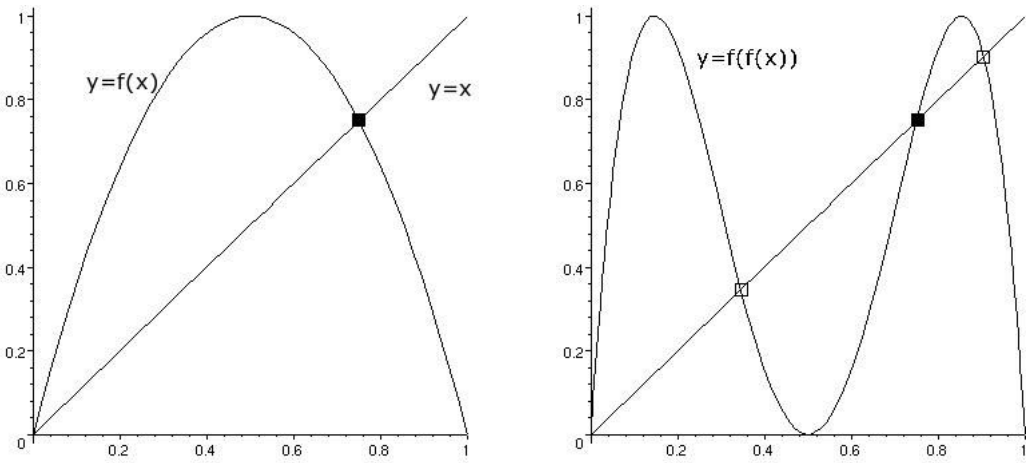
und damit $\dim_H(K) \le \frac{\log(4)}{\log(3)}$.

4.3 Dynamische Systeme

Dynamische Systeme treten in den verschiedensten Bereichen der Wissenschaft auf. Der Name ist ursprünglich von den Beispielen in der Physik motiviert. Wir beginnen mit einem Modell aus der Biologie, mit der die Populationsentwicklung einer bestimmten Tierart modelliert werden soll. Wenn die Population zum Ende eines Jahres x beträgt, so wird angenommen, dass sie am Ende des folgenden Jahres $f(x)$ ist. Wir erhalten dadurch eine Folge (x_n) mit $x_{n+1} = f(x_n)$. Wir haben besonderes Interesse an $\underline{\text{Fixpunkten}}$ $f(x) = x$ und $\underline{\text{periodischen Orbits}}$. Ein periodischer Orbit der Länge 2 ist zum Beispiel durch ein Paar (x_1, x_2) gegeben mit $f(x_1) = x_2$ und $f(x_2) = x_1$. Die Populationsentwicklung hat dann die Periode 2 und es ist

$$x_n \;=\; \begin{cases} x_1 & \text{für} \quad n \text{ ungerade} \\ x_2 & \text{für} \quad n \text{ gerade} \end{cases} .$$

Beispiel 4.3.1.



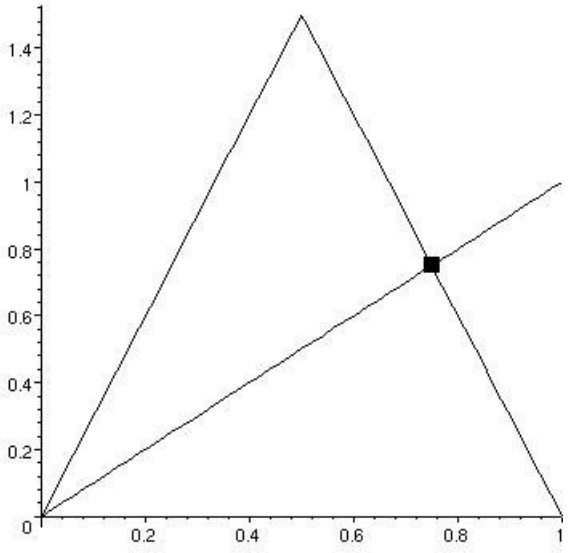
$f(x) = 4(x - x^2)$ besitzt den Fixpunkt $x = \frac{3}{4}$ und die Periode $x_1 = \frac{5}{8} + \frac{\sqrt{5}}{8}$, $x_2 = \frac{5}{8} - \frac{\sqrt{5}}{8}$.

Von Interesse ist auch die Frage, wie sich die Folge $x, f(x), f(f(x)), \ldots$ der Iterierten $f^{(n)}(x)$ für verschiedene Startwerte x verhält.

Definition 4.3.2. Es sei $D \subseteq \mathbb{R}^n$ und $f : D \to D$ stetig. Die Folge der Iterierten $\{f^{(n)}\}$ ist definiert durch $f^{(1)} = f$ und $f^{(n+1)} = f \circ f^{(n)}$. $\{f^{(n)}\}$ heißt diskretes dynamisches System. Eine Teilmenge $F \subseteq D$ heißt Attraktor von f, falls F eine abgeschlossene Menge ist, die invariant bzgl. f ist, d. h. $f(F) = F$, so dass es eine offene Menge V gibt, welche F enthält, so dass der Abstand von $f^{(n)}(x)$ zu F für $n \to \infty$ für alle $x \in V$ gegen Null konvergiert. Die Menge V heißt dann ein Attraktionsgebiet von f. Entsprechend heißt ein abgeschlossenes $F \subseteq D$ ein Repeller, wenn F invariant ist und es eine offene Menge $V \supset F$ gibt, so dass der Abstand $f^{(n)}(x)$ zu F für $n \to \infty$ gegen ∞ strebt für alle $x \in V \backslash F$

Attraktoren und Repeller bilden oft selbst fraktale Mengen, wie folgendes Beispiel zeigt:

Definition 4.3.3. Die Abbildung $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \frac{3}{2}(1 - |2x - 1|)$ heißt (wegen der Form des Graphen) Zelt-Abbildung.



Die Zelt-Abbildung.

Satz 4.3.4. *Die Cantor-Menge C ist ein Repeller der Zelt-Abbildung.*

Beweis.
Es sei $C = C_0 \cap C_1 \cap C_2 \cap \cdots$ die Cantor-Menge. Für die Ähnlichkeitsabbildungen $S_1(x) = \frac{1}{3}x$ und $S_2(x) = 1 - \frac{1}{3}x$ gilt offensichtlich $f(S_1(x)) = x$ und $f(S_2(x)) = x$ auf $[0, 1]$. Für diese Abbildungen gilt aber $C_{k+1} = S_1(C_k) \cup S_2(C_k)$, d. h. jedes $x \in C_{k+1}$ wird durch f in C_k abgebildet. Da ein $x \in C$ aber in allen C_k liegt, ist auch $f(x) \in C$, d. h. C ist invariant. Ist $x < 0$, so strebt $f^{(n)}(x) \to -\infty$ für $n \to \infty$, ebenso $f^{(n)}(x) \to -\infty$ für $n \to \infty$ falls $x > 1$. Ist $x \in [0, 1] \backslash C$, so gibt es mindestens ein k mit $x \notin C_k$. Da C_{k-1} sogar das surjektive Bild von C_k unter f ist folgt $f(x) \notin C_{k-1}$, ebenso $f^{(2)}(x) \notin C_{k-2}$ usw. Schließlich erhalten wir $f^{(n)}(x) \notin C_0 = [0, 1]$ für n groß genug, also $f^{(n)} \to \pm\infty$ für $n \to \infty$. Damit ist C ein Repeller für die Zelt-Abbildung, und der Abstoßungsbereich ist ganz $\mathbb{R}$. □

Index