

Prof. Silbermann

**Mathematik
für Informatiker**

Vorlesungsskript

Inhaltsverzeichnis

1	Elemente der Mengenlehre	2
1.1	Der Begriff der Menge	2
1.2	Funktionen	3
1.3	Mengenoperationen	9
1.4	Zahlenbereiche	12
1.4.1	Natürliche Zahlen	12
1.4.2	Reelle Zahlen	19
1.4.3	Komplexe Zahlen	25
2	Vektoren	35
2.1	Der Begriff des Vektors	35
2.2	Norm eines Vektors	42
2.3	Der Begriff des Skalarproduktes von Vektoren	44
2.4	Das Vektorprodukt (nur im Raum \mathbb{R}^3)	50
2.5	Allgemeine Vektorräume (lineare Räume)	54
2.6	Teilräume	58
2.7	Lineare Abbildungen	60
2.8	Lineare Gleichungssysteme und der Gaußsche Algorithmus . .	85

Kapitel 1

Elemente der Mengenlehre

Die „Mengenlehre“ - ein großes und relativ junges Teilgebiet der Mathematik - nimmt eine zentrale Stellung innerhalb der gesamten Mathematik ein. Ihr Begründer ist Georg Cantor (1845 - 1918).

1.1 Der Begriff der Menge

Eine Menge ist eine gedankliche Zusammenfassung von irgendwelchen wohlunterschiedenen Objekten zu einem Ganzen. Diese Objekte heißen Elemente der Menge.

Mengenbildung ist im üblichen Leben eine normale Erscheinung:

- Einwohner der Stadt Chemnitz
- Weltbevölkerung
- usw.

Wenn A eine Menge bezeichnet, x ein Element dieser Menge ist, so schreibt man $x \in A$ (x gehört zu A). Ist x kein Element von A , so schreibt man $x \notin A$ (oder $x \notin A$). Standardbezeichnungen einiger Mengen:

- \mathbb{N} - Menge der natürlichen Zahlen
- \mathbb{Z} - Menge der ganzen Zahlen

- \mathbb{Q} - Menge der rationalen Zahlen
- \mathbb{R} - Menge der reellen Zahlen
- \mathbb{C} - Menge der komplexen Zahlen

Zu diesen Mengen werden im weiteren noch einige Ausführungen vorgetragen.

Oftmals läßt sich eine Menge dadurch beschreiben, daß man eine Eigenschaft angibt, die sämtlichen Elementen dieser Menge und nur ihnen zukommt.

Beispiel: $\mathbb{N} := \{x \in \mathbb{Z} : x > 0\}$.

Leere Menge: Menge, die kein einziges Element enthält. Man bezeichnet diese mit \emptyset .

Frage: Warum führt man diese ein? \rightarrow Um Ausdrücke der Gestalt

$$\{x \in \mathbb{R} : x^2 = -1\}$$

zuzulassen. Dies ist nämlich gerade die leere Menge. Wenn eine Menge mindestens ein Element enthält, dann heißt sie nicht leer.

Definition 1.1.1 *Eine Menge A heißt Untermenge (oder Teilmenge) einer Menge B , wenn jedes Element der Menge A auch Element der Menge B ist: $A \subset B$ (oder $B \supset A$). Wenn $A \subset B$ und ein Element $x \in B$ existiert mit $x \notin A$, so heißt A echte Teilmenge von B .*

Im Besonderen gilt also $A \subset A$ für jede beliebige Menge.

Wenn $A \subset B$ und $B \subset A$, so heißen A und B gleich: $A = B$.

Festlegung: Die leere Menge \emptyset ist Teilmenge einer jeden Menge. Eine Menge heißt endlich, wenn sie endlich viele Elemente enthält, sonst heißt sie unendlich.

1.2 Funktionen

Definition 1.2.1 *Seien A, B zwei Mengen, deren Elemente beliebige Objekte sein dürfen. Wir nehmen an, daß auf Grund einer bestimmten Vorschrift*

f jedem Element $x \in A$ ein und nur ein Element der Menge B zugeordnet ist, welches wir mit $f(x)$ bezeichnen. Dann heißt f Funktion oder Abbildung der Menge A in die Menge B . Man schreibt dann $f : A \rightarrow B$ oder $f : x \mapsto y$. A heißt Definitionsbereich und B Zielmenge. Die Elemente $f(x)$ heißen Werte der Funktion und die Menge aller Werte $f(x)$, nämlich $f(A) := \{f(x) : x \in A\}$, heißt das Bild von A unter der Abbildung f .

Beachte: Eine Funktion ist durch ein Tripel, nämlich durch Definitionsbereich, Abbildungsvorschrift und Zielmenge festgelegt. So sind die Funktionen

$$\begin{aligned} f : \mathbb{Z} &\mapsto \mathbb{R}, & z &\mapsto z^2 \\ g : \mathbb{R} &\mapsto \mathbb{R}, & z &\mapsto z^2 \end{aligned}$$

nicht gleich.

Beispiel 1.2.2 $A \subset \mathbb{R}, B = \mathbb{R}, f : A \rightarrow B$ - sind Funktionen einer Veränderlichen (klassische Sprechweise), wie etwa

$$\begin{aligned} \sin : \mathbb{R} &\mapsto \mathbb{R}, & x &\mapsto \sin x \\ \cos : \mathbb{R} &\mapsto \mathbb{R}, & x &\mapsto \cos x \end{aligned}$$

Definition 1.2.3 Eine Abbildung $f : A \rightarrow B$ heißt Abbildung auf B , wenn $f(A) = B$. Für solche Funktionen ist auch der Begriff Surjektion gebräuchlich.

In der Mathematik ist es (aus ökonomischen Gründen) üblich, Sachverhalte (zumindest in einem vernünftigen Umfang) symbolisch auszudrücken. So schreibt man etwa: $f : A \rightarrow B$ ist eine Surjektion $\Leftrightarrow \forall y \in B \exists x \in A : y = f(x)$. Das Symbol \Leftrightarrow ist den beiden Symbolen \Rightarrow und \Leftarrow gleichwertig. Ihre Bedeutung wird wie folgt festgelegt:

Seien P_1 und P_2 zwei Aussagen. Wenn aus dem Fakt, daß P_1 wahr ist, folgt, daß die Aussage P_2 wahr ist, so schreibt man

$$P_1 \Rightarrow P_2 \text{ oder } P_2 \Leftarrow P_1$$

Man sagt auch, daß P_1 hinreichend für P_2 ist, und daß P_2 notwendig für P_1 ist.

Beispiel 1.2.4 Zwei Aussagen sind gegeben:

Aussage P_1 : Im gesamten Raum Chemnitz regnet es seit einer Stunde stark.

Aussage P_2 : Die Brückenstraße ist naß.

Offenbar: $P_1 \Rightarrow P_2$

P_1 ist also hinreichend für P_2 .

$P_2 \Rightarrow P_1$ gilt jedoch nicht, denn: Die Brückenstraße kann naß sein, ohne daß es geregnet hat (Wasserrohrbruch, ...).

P_2 ist also nicht notwendig für P_1 .

Das Symbol \Leftrightarrow steht also für: P_2 gilt genau dann, wenn P_1 gilt, oder P_1 ist zu P_2 äquivalent, oder dafür, daß P_1 gilt, ist notwendig und hinreichend, daß P_2 gilt.

Das Symbol $\forall y \in B$ wird gelesen: für jedes y aus B ...

Das Symbol $\exists x \in A$ wird gelesen: es gibt ein x aus A ...

Definition 1.2.5 Sei $E \subset A$ und $f : A \rightarrow B$ eine Funktion. Die Menge $f(E) := \{f(x) \in B : x \in E\}$ heißt Bild der Menge E unter der Abbildung f . Wenn $F \subset B$, so heißt $f^{-1}(F) := \{x \in A : f(x) \in F\}$ das (volle) Urbild der Menge F unter der Abbildung F . Besteht im besonderen F nur aus einem Element y , d.h. $F = \{y\}$, so heißt $f^{-1}(y) := f^{-1}(\{y\})$ ($= \{x \in A : f(x) = y\}$) das Urbild des Elements y .

Definition 1.2.6 Eine Funktion $f : A \rightarrow B$ heißt eineindeutig oder Injektion, wenn für beliebiges $y \in B$ die Menge $f^{-1}(y)$ aus höchstens einem Element besteht ($f^{-1}(y) = \emptyset$ ist also zugelassen).

Offenbar: f ist injektiv \Leftrightarrow aus $x_1 \neq x_2$ folgt $f(x_1) \neq f(x_2)$.

Definition 1.2.7 Die Abbildung $f : A \rightarrow B$ heißt Bijektion, wenn sie eine Injektion und eine Surjektion ist.

Im Zusammenhang mit Funktionen tritt häufig das Problem der Lösung bzw. der Lösbarkeit von Gleichungen auf, d.h. zu gegebenem $y \in B$ wird eine Lösung $x \in A$ der Gleichung $f(x) = y$ gesucht, wobei die Abbildung $f : A \rightarrow B$ gegeben ist. Dabei ist x Lösung, wenn $f(x)$ gleich y ist. „Lösen

von Gleichungen“ ist bis heute eine der wichtigsten Aufgaben der Mathematik (Differentialgleichungen, Integralgleichungen, lineare Gleichungssysteme, usw.).

Offenbar gelten folgende Aussagen:

1. $f(x) = y$ besitzt stets wenigstens eine Lösung, wenn f surjektiv ist (Existenz der Lösung).
2. $f(x) = y$ besitzt höchstens eine Lösung, wenn f injektiv ist. (Eindeutigkeit).
3. $f(x) = y$ besitzt zu jedem $y \in B$ genau eine Lösung $x \in A$, wenn f bijektiv ist.

Ein wichtiger Begriff ist die Verknüpfung (Hintereinanderausführung, bzw. Komposition) von Funktionen. Dazu seien $f : A \rightarrow B$, $g : B \rightarrow C$ Funktionen. Dann definiert man die Komposition $g \circ f$ durch

$$g \circ f : A \rightarrow C, \quad x \mapsto g(f(x))$$

$$\begin{array}{ccccc} A & \xrightarrow{f} & B & \xrightarrow{g} & C \\ & \searrow & & \nearrow & \\ & & g \circ f & & \end{array}$$

Beispiel 1.2.8 $f : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto x^2, \quad \sin : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \sin x.$

Komposition der Funktion f und \sin lautet: $\sin \circ f : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \sin x^2.$

Definition 1.2.9 Sei $f : A \rightarrow B$ eine Bijektion. Dann besteht für beliebiges y das volle Urbild $f^{-1}(y)$ aus genau einem Element $x_y \in A : f^{-1}(y) = x_y$. Die Abbildung $g : A \rightarrow B, \quad y \mapsto x_y$, heißt Umkehrfunktion zur Funktion f . Diese wird häufig mit f^{-1} bezeichnet.

Offenbar ist f^{-1} ebenfalls eine Bijektion. Außerdem gilt $f^{-1}(f(x)) = x$ für alle $x \in A$ und $f(f^{-1}(y)) = y$ für alle $y \in B$

Führt man für eine beliebige Menge C die identische Abbildung

$I_C : C \rightarrow C, \quad z \mapsto z$, ein, so kann dies auch wie folgt ausgedrückt werden: $f^{-1}f = I_A$ und $ff^{-1} = I_B$.

Definition 1.2.10 Sei $E \subset A$ und $f : A \rightarrow B$ eine Funktion. Die Funktion $F : E \rightarrow B$, $x \mapsto f(x)$, heißt *Einschränkung von f auf E* . Man schreibt auch $f|_E$ statt F . Die Funktion f heißt auch *Fortsetzung von F* .

Wir kehren noch einmal zum Begriff der Komposition zurück. Die Operation der Komposition ist eine assoziative Operation, d.h. es gilt

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

Dies ist leicht einzusehen: Für $\forall x \in A$ gilt nämlich

$$\begin{aligned} (h \circ (g \circ f))(x) &= h(g(f(x))) \\ ((h \circ g) \circ f)(x) &= (h \circ g)(f(x)) = h(g(f(x))). \end{aligned}$$

Das Assoziativgesetz ist von grundlegender Bedeutung. Bildet man nämlich die Verknüpfung von n Funktionen f_1, \dots, f_n , so ist es, wie man zeigen kann, völlig gleichgültig, wie man die Klammern setzt. Beispielsweise gilt:

$$f_1 \circ (\dots (f_{n-2} \circ (f_{n-1} \circ f_n)) \dots) = (\dots ((f_1 \circ f_2) \circ f_3) \circ \dots) \circ f_n$$

Das Kommutativgesetz, d.h. $f \circ g = g \circ f$, gilt im allgemeinen nicht:

$$f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto -x$$

$$g : \mathbb{R} \rightarrow \mathbb{R}, \begin{cases} x \mapsto 1, & x \geq 0 \\ x \mapsto 0, & x < 0 \end{cases}$$

$$g \circ f : \mathbb{R} \rightarrow \mathbb{R}, \begin{cases} x \mapsto 1, & x < 0 \\ x \mapsto 0, & x \geq 0 \end{cases}$$

$$f \circ g : \mathbb{R} \rightarrow \mathbb{R}, \begin{cases} x \mapsto -1, & x \geq 0 \\ x \mapsto 0, & x < 0 \end{cases}$$

Bemerkung 1.2.11 Die Menge der bijektiven Funktionen einer Menge A auf sich,

$$S(A) := \{f : A \rightarrow A : f \text{ ist bijektiv}\}$$

bildet bezüglich der Komposition eine sogenannte Gruppe, nämlich die symmetrische Gruppe von A , d. h. es gelten

1. $h \circ (g \circ f) = (h \circ g) \circ f$ (*Assoziativgesetz*)
2. $f \circ I_A = I_A \circ f = f$ (*neutrales Element*)
3. $f \circ f^{-1} = f^{-1} \circ f = I_A$ (*inverses Element*)

Der Begriff der Gruppe ist fundamental!

Vermerken schließlich noch folgende wichtige Begriffsbildung: Zwei Mengen A und B heißen äquivalent (oder gleichmächtig), wenn es eine Bijektion $f : A \rightarrow B$ gibt. Man schreibt auch $A \sim B$. Es gelten folgende Eigenschaften:

1. $A \sim A$ (Reflexivität)
2. $A \sim B \Rightarrow B \sim A$ (Symmetrie)
3. $A \sim B$ und $B \sim C \Rightarrow A \sim C$ (Transitivität)

Eine Relation in der Mathematik, die die Eigenschaften 1. bis 3. besitzt, heißt Äquivalenzrelation.

Mittels der von uns eingeführten Begriffe läßt sich der Terminus endliche Menge exakt definieren: Eine Menge A heißt endlich, wenn sie zu keiner ihrer echten Teilmengen äquivalent ist. Wir werden diese Definition jedoch nicht weiter benutzen.

1.3 Mengenoperationen

Seien A und Ω ($\neq \emptyset$) zwei Mengen. Wir nehmen an, daß jedem Element $\omega \in \Omega$ eine gewisse Untermenge der Menge A zugeordnet wird, die wir mit A_ω bezeichnen. Wir haben also ein System von Mengen $\{A_\omega\}_{\omega \in \Omega}$, wobei der Index ω die Menge Ω durchläuft. Die Menge Ω heißt auch Indexmenge.

Definition 1.3.1 Vereinigung der Mengen des Systems $\{A_\omega\}$ heißt die Menge C ($\subset A$), für die $x \in C$ dann und nur dann gilt, wenn $x \in A_\omega$ gilt für wenigstens ein $\omega \in \Omega$. Man schreibt

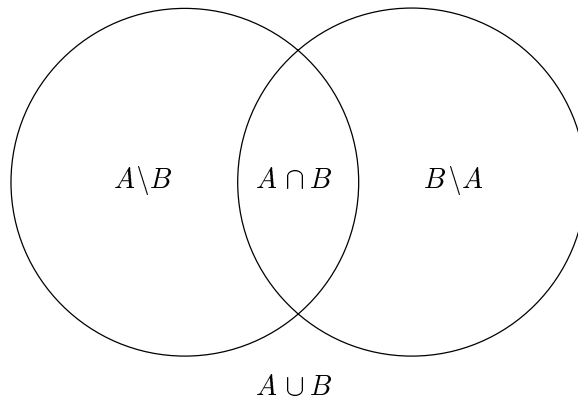
$$C = \bigcup_{\omega \in \Omega} A_\omega.$$

Spezialfälle:

$$\Omega = \{1, \dots, n\}, \quad C = \bigcup_{m=1}^n A_m, \quad \Omega = \mathbb{N} = \{1, 2, \dots\}, \quad C = \bigcup_{m=1}^{\infty} A_m.$$

Definition 1.3.2 Durchschnitt der Mengen des Systems $\{A_\omega\}$ heißt die Menge D , für die $x \in D$ genau dann gilt, wenn $x \in A_\omega$ für alle $\omega \in \Omega$ gilt. Man schreibt $D = \bigcap_{\omega \in \Omega} A_\omega$ und entsprechend $D = \bigcap_{m=1}^n A_m$ bzw. $D = \bigcap_{m=1}^{\infty} A_m$. Wenn $A \cap B = \emptyset$, so heißen A und B durchschnittsfremd oder disjunkt.

Definition 1.3.3 Unter der Differenz $A \setminus B$ (d.h. $A - B$) der Mengen A und B versteht man die Teilmenge der Menge A , die aus all jenen Elementen von A besteht, die nicht zu B gehören. ($A \supset B$ wird dabei nicht vorausgesetzt)



Beispiel 1.3.4

1. $A_1 = \{1, 2, 3\}, \quad A_2 = \{2, 3, 4\}$
 $\Rightarrow A_1 \cup A_2 = \{1, 2, 3, 4\}, \quad A_1 \cap A_2 = \{2, 3\}, \quad A_1 \setminus A_2 = \{1\}$
2. $\Omega = \{w \in \mathbb{R} : 0 < x < 1\}, \quad A_\omega = \{x \in \mathbb{R} : 0 < x < \omega\}$
 $\Rightarrow \bigcup_{\omega \in \Omega} A_\omega = \{x \in \mathbb{R} : 0 < x < 1\}, \quad \bigcap_{\omega \in \Omega} A_\omega = \emptyset$

Definition 1.3.5 Sei $A \subset M$. Die Differenz $M \setminus A$ heißt Komplement von A bzw. M . Man schreibt dafür A^c (oder auch $C_M A$).

Einige Eigenschaften (deren Beweise elementar ist):

$A \cup A^c = M, \quad A \cap A^c = \emptyset, \quad (A \cup B)^c = A^c \cap B^c, \quad (A \cap B)^c = A^c \cup B^c$ Wählen wir eine Grundmenge M und erklären für ihre Teilmengen $A \subset M, B \subset M, \dots$ die Mengenverknüpfungen \cup, \cap, \setminus , so erhalten wir eine sogenannte algebraische Struktur (System) mit den folgenden Eigenschaften:

1. $A \cup B = B \cup A$ und $A \cap B = B \cap A$ (Kommutativgesetze)
2. $A \cup (B \cap C) = (A \cup B) \cap C, \quad A \cap (B \cup C) = (A \cap B) \cup C$ (Assoziativgesetz)
3. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C), \quad A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (Distributivgesetze)
4. Es existieren für \cup und \cap die neutralen Elemente M und \emptyset :
 $A \cap M = A$ und $A \cup \emptyset = A$
5. Es existiert zu jeder Teilmenge $A \subset M$ ein Komplement $A^c \subset M$ mit
 $A \cap A^c = \emptyset$ und $A \cup A^c = M$

Die sogenannte Potenzmenge von M (Menge aller Teilmengen von M), versehen mit den Verknüpfungen \cap, \cup, \setminus bildet aufgrund dieser Eigenschaften eine sogenannte Mengenalgebra. Diese stellt ein Modell der abstrakten Booleschen Algebra dar.

Definition 1.3.6 Seien A und B zwei Mengen. Aus je zwei Elementen $a \in A$ und $b \in B$ bilden wir ein neues Objekt, ihr geordnetes Paar (a, b) . Dabei sei $(a, b) = (a', b') \Leftrightarrow a = a', \quad b = b'$. Die Menge aller geordneten Paare (a, b) mit $a \in A, \quad b \in B$ heißt direktes Produkt der Mengen A und $B : A \times B$.

Analog dazu bildet man das Produkt einer beliebigen endlichen Anzahl von Mengen: $A_1 \times A_2 \times \cdots \times A_n = \text{Menge aller geordneten } n\text{-Tupel } (x_1, \dots, x_n) \text{ mit } x_i \in A_i$. Im Spezialfall $A_1 = \cdots = A_n = A$ setzt man $A \times \cdots \times A = A^n$.

Beispiel 1.3.7 Sei:

1. $A = \mathbb{R}$

$$\mathbb{R}^2 = \{(a, b) : a, b \in \mathbb{R}\} \text{ (Gauß'sche Zahlenebene)}$$

$$\mathbb{R}^n = \{(a_1, \dots, a_n) : a_i \in \mathbb{R}, i = 1, \dots, n\}.$$

2. $f : A \rightarrow B$ eine Funktion. Graph der Funktion f heißt die Teilmenge $\{(a, b) \in A \times B : f(a) = b\} \subset A \times B$.

Vermerken: die Zuordnung Funktion $f \mapsto \text{Graph } f$ ist eineindeutig. Hieraus erwächst die Möglichkeit, Funktionen mit gewissen Mengen zu identifizieren. Frage: Wann ist eine Teilmenge von $A \times B$ Graph einer Funktion?

Definition 1.3.8 Sei X eine beliebige Menge. Relation auf X wird jede Teilmenge R von $X \times X$ genannt. Man sagt auch, daß zwei Elemente x und $y \in X$ in R -Relation zueinander stehen (xRy), wenn $(x, y) \in R$.

Beispiel 1.3.9 (Beispiele wichtiger Relationen:) Eine Relation R heißt

1. reflexiv, wenn $xRx, \forall x \in X$
2. symmetrisch, wenn aus $xRy \Rightarrow yRx$
3. antisymmetrisch, wenn aus xRy und $yRx \Rightarrow x = y$
4. transitiv, wenn aus $xRy, yRz \Rightarrow xRz$.

R heißt Äquivalenzrelation, wenn R reflexiv, symmetrisch und transitiv ist. Statt xRy wird dann häufig $x \sim y$ geschrieben.

$[x] := \{y \in X : y \sim x\}$ heißt Äquivalenzklasse. Es gilt: zwei Äquivalenzklassen $[x_1]$ und $[x_2]$ stimmen entweder überein oder sind durchschnittsfremd. Wenn also $y \in [x] \Rightarrow [y] = [x]$. Jedes Element $y \in [x]$ heißt auch Repräsentant. Jede Äquivalenzrelation R erzeugt eine Zerlegung von X in paarweise

disjunkte Mengen (Äquivalenzklassen) und jede Zerlegung einer Menge X in paarweise disjunkte (nichtleere) Teilmengen erzeugt eine Äquivalenzrelation: $x \sim y \Leftrightarrow x, y$ gehören zu ein und derselben Teilmenge.

R heißt Ordnungsrelation, wenn R reflexiv, antisymmetrisch und transitiv ist. Statt xRy schreibt man häufig $x \leq y$.

Reflexivität: $x \leq x$ für alle $x \in X$
 Antisymmetrie: $x \leq y$ und $y \leq x \Rightarrow x = y$
 Transitivität: $x \leq y$ und $y \leq z \Rightarrow x \leq z$.

1.4 Zahlenbereiche

1.4.1 Natürliche Zahlen

Die Menge der natürlichen Zahlen

$$\mathbb{N} := \{1, 2, 3, 4, \dots\} \subset \mathbb{R}$$

läßt sich durch die Peano-Axiome kennzeichnen. Sie lauten:

1. $1 \in \mathbb{N}$
2. $n \in \mathbb{N} \Rightarrow (n + 1) \in \mathbb{N}$
3. $n \neq m \Rightarrow (n + 1) \neq (m + 1)$
4. $n \in \mathbb{N} \Rightarrow n + 1 \neq 1$
5. Wenn für $A \subset \mathbb{N}$ gilt
 - (a) $1 \in A$
 - (b) $\forall n \in A \Rightarrow n + 1 \in A$

dann ist $A = \mathbb{N}$

Man kann die Menge der natürlichen Zahlen axiomatisch einführen, d.h. wir nehmen an, daß es eine Menge \mathbb{N} und eine Funktion $\delta : \mathbb{N} \rightarrow \mathbb{N}$ gibt, für die die folgenden Axiome gelten

- (i) $\delta : \mathbb{N} \rightarrow \mathbb{N}$ ist injektiv

(ii) Es gibt ein Element (wir bezeichnen es mit dem Symbol 1), das nicht zu δ gehört

(iii) Wenn für $A \subset \mathbb{N}$ gilt

(a) $1 \in A$

(b) $\forall n \in A \Rightarrow \delta(n) \in A$,

dann ist $A = \mathbb{N}$.

Nach längeren Überlegungen sieht man ein, daß allein auf Grundlage der Axiome (i) - (iii) eine Operation der Addition mit den uns bekannten Eigenschaften eingeführt werden kann ($z(n) =: n + 1$ -sogenannte Nachfolgefunktion) und die Eigenschaften 1. - 5. gelten. Wir werden dies hier unterlassen und sofort mit 1. bis 5. arbeiten. Das Axiom 5. nimmt eine besondere Stellung ein. Es ist die Grundlage für das Beweisprinzip der vollständigen Induktion. Wir erläutern dies nun.

Jedem Element $n \in \mathbb{N}$ sei eine Aussage $P(n)$ zugeordnet. Wir möchten wissen, ob diese Aussage für jedes $n \in \mathbb{N}$ wahr ist. Es gelte nun:

(a) $P(1)$ ist wahr (Induktionsanfang)

(b) für beliebiges $n \in \mathbb{N}$ gelte $P(n) \Rightarrow P(n + 1)$ (Induktionsschluß).

Dann ist die Aussage $A(n)$ für alle $n \in \mathbb{N}$ wahr.

Beweis. Sei $A \subset \mathbb{N}$ wie folgt erklärt: $A := \{n \in \mathbb{N} : P(n) \text{ ist wahr} \}$. Dann gilt $1 \in A$ (in Folge von Voraussetzung (a) und aus $n \in A \Rightarrow n + 1 \in A$ (infolge von (b).) . Also ist nach Axiom 5. $A = \mathbb{N}$. \square

Man beachte, daß der Induktionsschluß für ein beliebiges, jedoch festes $n \in \mathbb{N}$ zu beweisen ist. Dabei heißt $P(n)$ auch Induktionsannahme und $P(n + 1)$ Induktionsbehauptung. Wir geben nun noch ein wichtiges Beispiel an.

Wieviele Vertauschungen (Permutationen) des n -Tupels $(1, 2, 3, \dots, n)$ gibt es?

$n = 1$: (1)	–	eine Permutation
$n = 2$: (1, 2), (2, 1)	–	2 Permutationen
$n = 3$: (1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1)	–	6 Permutationen

Wir vermuten daher, daß es genau $n! := 1 \cdot 2 \cdot 3 \cdots n$ Permutationen des n -Tupels $(1, \dots, n)$ gibt. Diese Aussage bezeichnen wir mit $P(n)$ und behaupten, daß sie für jedes $n \in \mathbb{N}$ wahr ist.

Beweis. (vollständige Induktion): $n = 1$: $1! = 1$ wurde bereits gezeigt.

$P(n) \Rightarrow P(n+1)$: Wir unterteilen die Menge der Permutationen von $(1, 2, \dots, n+1)$ in $(n+1)$ Klassen. Die Klasse Nummer k ($k \in \{1, \dots, n+1\}$) enthalte dabei gerade alle Permutationen der Form (k, i_1, \dots, i_n) , wobei (i_1, \dots, i_n) eine beliebige Permutation von $(1, 2, \dots, k-1, k+1, \dots, n+1)$ ist. Nach Induktionsvoraussetzung enthält die Klasse Nummer k genau $n!$ Permutationen. Insgesamt erhält man dann also $(n+1)! = (n+1)n!$ Permutationen. \square

Wir haben offenbar folgendes Resultat gezeigt (andere Interpretation): Die symmetrische Gruppe $S(A)$ einer n -elementigen Menge A besteht aus genau $n!$ Elementen. Wenn nämlich $A = \{a_1, \dots, a_n\}$ und $\{a_{i_1}, \dots, a_{i_n}\}$ eine Permutation ist, dann wird durch die Zuordnung $a_1 \mapsto a_{i_1}, a_2 \mapsto a_{i_2}, \dots, a_n \mapsto a_{i_n}$ eine Bijektion der Menge A erzeugt. Umgekehrt erzeugt offenbar jede Bijektion der Menge A eine Permutation. Eine weitere Folgerung liefert uns die Anzahl der m -elementigen Teilmengen einer n -elementigen Menge $\{a_1, \dots, a_n\}$. Die Beweisidee demonstrieren wir anhand folgenden Beispiels:

Wieviele verschiedene Belegungen von zwei Zimmern mit jeweils m und $n-m$ ($n \geq m$) Betten durch n Bewohner ist möglich? Offenbar ist diese Zahl gleich $\frac{n!}{(n-m)!m!}$ (wie die Betten innerhalb eines Zimmers belegt werden, interessiert uns nicht). Somit entspricht jede m -elementige Teilmenge genau einer Belegung unserer Zimmer im genannten Sinn. Damit haben wir gezeigt, daß eine n -elementige Menge genau

$$\binom{n}{m} := \frac{n!}{(n-m)!m!} \quad (1.1)$$

Teilmengen besitzt, wobei $\binom{n}{n}$ per Definition gleich 1 gesetzt wird. Setzt man per Definition $0! = 1$, so ergibt sich $\binom{n}{n} = 1$ sofort aus 1.1. Gleichmaßen ergibt sich dann aus 1.1 auch $\binom{n}{0} = 1$. Diese Vereinbarung sehen wir von nun an als gültig an. Die Zahlen $\binom{n}{m}$ ($0 \leq m \leq n$) heißen Binomialkoeffizienten. Diese Bezeichnung rührt daher, daß sie im binomischen Lehrsatz auftreten.

Dieser lautet:

Für reelle (wie auch für komplexe) a, b und $n \in \mathbb{Z}_+ := \mathbb{N} \cup \{0\}$ gilt

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k} \quad \text{binomischer Lehrsatz.}$$

Für den Beweis benötigen wir die für $0 < m \leq n$ gültige Rekursion

$$\binom{n+1}{m} = \binom{n}{m} + \binom{n}{m-1}$$

die wir zunächst beweisen werden. Es gilt nämlich

$$\begin{aligned} \binom{n}{m} + \binom{n}{m-1} &= \frac{n!}{(n-m)!m!} + \frac{n!}{(m-1)!(n+1-m)!} \\ &= \frac{n!(n+1-m) + n!m}{(n+1-m)!m!} = \frac{n!(n+1)}{(n+1-m)!m!} = \frac{(n+1)!}{(n+1-m)!m!} = \binom{n+1}{m} \end{aligned}$$

Mittels der angegebenen Rekursion lassen sich die Binomialkoeffizienten leicht berechnen:

$n = 0$						1
$n = 1$				1		1
$n = 2$			1	2	1	
$n = 3$		1	3	3	1	
$n = 4$		1	4	6	4	1
$n = 5$	1	5	10	10	5	1

\Rightarrow die Binomialkoeffizienten sind somit in dem Pascalschen Dreieck angeordnet.

Beweis des binomischen Lehrsatzes (vollständige Induktion): Für $n = 1$ gilt

$$(a+b)^1 = a+b = \binom{1}{0} a^0 b^1 + \binom{1}{1} a^1 b^0.$$

Nehmen an, daß er für fixiertes n gilt und zeigen, daß er dann auch für $n+1$ gültig ist.

$$(a+b)^{n+1} = (a+b)(a+b)^n = (a+b) \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$$

$$= \sum_{k=0}^n \binom{n}{k} a^{k+1} b^{n-k} + \sum_{k=0}^n \binom{n}{k} a^k b^{n-k+1}$$

Für $j = k+1$ gilt:

$$\begin{aligned} &= \sum_{j=1}^{n+1} \binom{n}{j-1} a^j b^{n-j+1} + \sum_{k=0}^n \binom{n}{k} a^k b^{n-k+1} \\ &= \binom{n}{0} a^0 b^{n+1} + \sum_{k=1}^n \left\{ \binom{n}{k-1} + \binom{n}{k} \right\} a^k b^{n-k+1} + \binom{n}{n} a^{n+1} b^0 \end{aligned}$$

Mittels Rekursionsformel:

$$\begin{aligned} &= \binom{n+1}{0} a^0 b^{n+1} + \sum_{k=1}^n \binom{n+1}{k} a^k b^{n+1-k} + \binom{n+1}{n+1} a^{n+1} b^0 \\ &= \sum_{k=0}^{n+1} \binom{n+1}{k} a^k b^{n+1-k} \end{aligned}$$

Schließlich wollen wir kurz auf die multiplikativen Elementarbausteine der natürlichen Zahlen, die Primzahlen, eingehen.

Definition 1.4.1 Eine natürliche Zahl $m \in \mathbb{N}$ heißt Teiler von $n \in \mathbb{N}$, falls es ein $k \in \mathbb{N}$ gibt, mit $n = m \cdot k$. Wir schreiben dann $m|n$. Jede natürliche Zahl n besitzt also zumindest die Teiler 1 und n . Ist $n > 1$ und besitzt keine weiteren Teiler, so heißt n Primzahl.

Es gilt nun folgender

Satz 1.4.2 Jede Zahl $n \in \mathbb{N}, n \neq 1$ läßt sich eindeutig als Produkt von Primzahlen schreiben:

$$n = p_1^{r_1} \cdot p_2^{r_2} \cdots p_k^{r_k} \tag{1.2}$$

wobei p_j Primzahlen und $r_j \in \mathbb{N}$.

Der Beweis der Existenz einer solchen Darstellung kann über vollständige Induktion geführt werden. Dabei tritt eine Feinheit auf: Zum Nachweis von

$P(n+1)$ wird als Induktionsvoraussetzung $P(k), k = 1, \dots, n$ (und nicht lediglich $P(n)$) benötigt.

Aufgabe: Wie kann man dies auf den uns bekannten Fall zurückführen?

Definition 1.4.3 Zu $n, m \in \mathbb{N}$ heißt

$$ggT(n, m) := \max\{k \in \mathbb{N} : k \text{ teilt } n \text{ und } m\}$$

der größte gemeinsame Teiler von n und m , und

$$kgV(n, m) := \min\{k \in \mathbb{N} : n \text{ und } m \text{ teilen } k\}$$

das kleinste gemeinsame Vielfache von n und m . Aus praktischen Erwägungen wird in der Darstellung 1.2 das Auftreten von Faktoren der Gestalt $p_i^0 (= 1)$ zugelassen. Dann kann man für beliebige natürliche Zahlen m und n etwa

$$n = p_1^{r_1} \cdots p_k^{r_k}, \quad m = p_1^{s_1} \cdots p_k^{s_k}$$

schreiben, denn wenn in der ursprünglichen Darstellung z.B. für n der Faktor $p_i^{r_i}$ nicht auftritt, so kann $p_i^{r_i}$ mit $r_i = 0$ als Faktor hinzugefügt werden, ohne daß sich am Produkt etwas ändert. Sind (die erweiterten) Primzahlzerlegungen von n und m bekannt und etwa

$$n = p_1^{r_1} \cdots p_k^{r_k}$$

$$m = p_1^{s_1} \cdots p_k^{s_k}$$

so lassen sich $kgV(n, m)$ und $ggT(n, m)$ leicht berechnen. Es gilt nämlich:

$$ggT(n, m) = p_1^{\min(r_1, s_1)} \cdots p_k^{\min(r_k, s_k)}$$

$$kgV(n, m) = p_1^{\max(r_1, s_1)} \cdots p_k^{\max(r_k, s_k)}$$

Wenn die erweiterten Primzahldarstellungen nicht bekannt sind, kann man den Euklidischen Algorithmus zum Aufsuchen des größten gemeinsamen Teilers einsetzen. Dieser stützt sich auf folgenden Fakt (Division mit Rest): Sei

$q \in \mathbb{N}$ und $p \in \mathbb{Z}_+$. Dann gibt es eindeutig bestimmte Zahlen $m, r \in \mathbb{Z}_+$ mit $r < q$ und

$$p = mq + r.$$

Die Existenz kann über vollständige Induktion bewiesen werden.

Eindeutigkeit: Sei noch eine weitere Darstellung $p = m_1q + r_1$ gegeben.

$$\begin{aligned} \Rightarrow 0 &= (m - m_1)q + r - r_1 \\ \Rightarrow |(m - m_1)q| &= |r - r_1| < q \\ \Rightarrow |(m - m_1)| < 1 &\Rightarrow |m - m_1| = 0, \text{ da } m - m_1 \text{ eine ganze Zahl ist} \\ \Rightarrow m - m_1 &= 0 \text{ und } r - r_1 = 0 \end{aligned}$$

Wir geben nun einen Satz an, der nicht nur die Existenz eines größten gemeinsamen Teilers, sondern noch einen Algorithmus (Euklidischer Algorithmus) zu dessen Berechnung liefert.

Satz 1.4.4 (Satz vom größten gemeinsamen Teiler) *Es seien n, m natürliche Zahlen. Dann existiert genau ein $ggT(m, n)$. Dabei gilt*

$$ggT(m, n) = s_1n + s_2m \quad (\triangle)$$

mit gewissen ganzen Zahlen s_1 und s_2 . Umgekehrt ist jeder gemeinsame Teiler d von n und m , der die Gestalt (\triangle) besitzt, größter gemeinsamer Teiler.

Beweis. Sind d und d' größte gemeinsame Teiler von n und m , dann gilt nach Definition $d|d'$ und $d'|d$. Daraus folgt $d = d'$. Ferner: besitzt ein Teiler von n und m die Gestalt (\triangle) , dann ist er $ggT(n, m)$, denn jeder gemeinsame Teiler von n und m ist dann auch Teiler der rechten Seite von (\triangle) .

Existenz: Falls $n = m \Rightarrow$ trivial. Sei nun $n > m$. Falls m n teilt, dann ist $ggT(n, m) = m$. Andernfalls erhalten wir durch wiederholte Anwendung der Division mit Rest folgendes:

$$n = mq_1 + r_1, \quad 0 < r_1 < m$$

$$m = r_1q_2 + r_2, \quad 0 < r_2 < r_1 \text{ oder } r_2 = 0.$$

Wenn $r_2 \neq 0$ fahren wir fort:

$$r_1 = r_2 q_3 + r_3, \quad 0 < r_3 < r_2 < r_1 \text{ oder } r_3 = 0 \text{ usw.}$$

Die Reste r_1, r_2, r_3, \dots bilden eine streng abnehmende, monotone Folge nichtnegativer ganzer Zahlen. Daher müssen wir nach endlich vielen Schritten zum Rest 0 gelangen. Für eine gewisse natürliche Zahl k ergibt sich also:

$$r_{k-4} = r_{k-3} q_{k-2} + r_{k-2}, \quad 0 < r_{k-2} < r_{k-3} \quad (\triangle.k-2)$$

$$r_{k-3} = r_{k-2} q_{k-1} + r_{k-1}, \quad 0 < r_{k-1} < r_{k-2} \quad (\triangle.k-1)$$

$$r_{k-2} = r_{k-1} q_k + r_k, \quad 0 < r_k < r_{k-1} \quad (\triangle.k)$$

$$r_{k-1} = r_k q_{k+1}, \quad (\triangle.k+1)$$

Wir zeigen nun, daß sich r_k - der letzte von Null verschiedene Rest - in der Form $r_k = s_1 n + s_2 m$ mit geeigneten Zahlen $s_1, s_2 \in \mathbb{Z}$ darstellen läßt, und r_k sowohl n als auch m teilt.

Beginnen wir mit der zweiten Behauptung: Setzen $(\triangle.k+1)$ in $(\triangle.k)$ ein $\Rightarrow r_{k-2} = r_k l_{k-2}$; setzen dies in $(\triangle.k-1)$ ein $\Rightarrow r_{k-3} = r_k l_{k-3}$ usw. $\Rightarrow r_k$ teilt sowohl n als auch m .

Weiter: aus $(\triangle.k) \Rightarrow r_k = r_{k-2} - r_{k-1} q_k$.

Mittels $(\triangle.k-1)$ kann hierin r_{k-1} durch eine Linearkombination von r_{k-2} und r_{k-3} ersetzt werden usw. Schließlich erhalten wir, daß $r_k = s_1 n + s_2 m$ mit $s_1, s_2 \in \mathbb{Z}$. \square

Definition 1.4.5 Folge in A wird jede Funktion $f : \mathbb{N} \rightarrow A$ genannt. Statt $f(n)$ wird häufig f_n geschrieben. Eine Folge kann somit auch als Familie $\{f_n\}_{n \in \mathbb{N}}$ aufgefaßt werden.

1.4.2 Reelle Zahlen

$$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$$

$$\mathbb{Q} = \left\{ \frac{m}{n} \mid m \in \mathbb{Z}, n \in \mathbb{Z} \setminus \{0\} \right\}$$

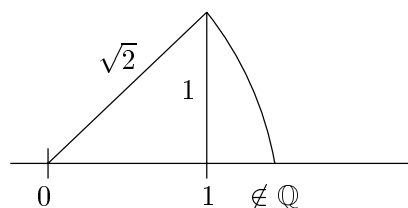
werden eingeführt, um Gleichungen der Gestalt $n + x = m$ ($n, m \in \mathbb{N}$ lösen zu können bzw. solche für $n \cdot x = m$ $n, m \in \mathbb{N}, n \neq 0$). \mathbb{Q} ist ebenfalls noch nicht ausreichend: es gibt zum Beispiel keine rationale Zahl, deren Quadrat gleich 2 ist.

Beweis. Nehmen an, daß es eine solche rationale Zahl a gibt, $a = \frac{n}{m}$, $n, m \in \mathbb{N}, n \neq 0$. Wir können annehmen, daß n und m teilerfremd sind, d.h. der Bruch gekürzt ist. ($\Leftrightarrow \text{ggT}(n, m) = 1$)

Es folgt nun $2 = \frac{n^2}{m^2} \Rightarrow 2m^2 = n^2 \Rightarrow n^2$ gerade $\Rightarrow n$ gerade, etwa $n = 2k, k \in \mathbb{N}$

(n ungerade $\Rightarrow n = 2t + 1 \Rightarrow n^2 = 4t^2 + 4t + 1$ ist ungerade)

$\Rightarrow m^2$ gerade ($2m^2 = 4k^2$) $\Rightarrow m$ gerade. Dies ist aber ein Widerspruch zur Teilerfremdheit von n und m . \square



Die rationalen Zahlen lassen „Lücken“ auf der Zahlengeraden. Idee: Lücken füllen. Diese Prozedur ist aufwendig und wird hier nicht erklärt. Wir geben stattdessen ein definierendes Axiomensystem für die reellen Zahlen an. Wir nehmen also an, daß es eine unendliche Menge gibt, die mit zwei binären Operationen versehen werden kann, und für die folgendes gilt: (binären Operator nennen wir Addition: $+$, und Multiplikation: \cdot).

(I) Regeln der Addition

- (a) $x + (y + z) = (x + y) + z$
- (b) $x + y = y + x$
- (c) $x + 0 = 0 + x = x$
- (d) $x + (-x) = (-x) + x = 0$

(II) Regeln der Multiplikation

- (a) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$

- (b) $x \cdot y = y \cdot x$
- (c) $x \cdot 1 = 1 \cdot x = x$
- (d) $x \cdot \frac{1}{x} = \frac{1}{x} \cdot x = 1 \quad (x \neq 0)$

(III) Distributivgesetz

$x \cdot (y + z) = x \cdot y + x \cdot z$ (gibt an, daß die Multiplikation mit der Addition verträglich ist)

(IV) Ordnungseigenschaften (es ist eine Ordnungsrelation \leq gegeben mit:)

- (a) für beliebige $x, y \in \mathbb{R}$ gilt $x \leq y$ oder $y \leq x$
- (b) $x \leq y \Rightarrow x + z \leq y + z$
- (c) $x \leq y$ und $z \geq 0 \Rightarrow x \cdot z \leq y \cdot z$

(b) und (c) gibt an, daß die Ordnungsrelation mit den algebraischen Operationen verträglich ist)

Schreiben $x < y$, wenn $x \leq y$ und $x \neq y$ gilt.

Vermerken: diese Axiome werden auch von den rationalen Zahlen erfüllt. Somit benötigen wir wenigstens noch ein Axiom, welches die Menge \mathbb{Q} von \mathbb{R} unterscheidet, dies ist das

(V) Vollständigkeits- oder Stetigkeitsaxiom

Zerlegt man die Menge \mathbb{R} in zwei Mengen N und M , d.h. $\mathbb{R} = N \cup M$, wobei $N \neq \emptyset, M \neq \emptyset$ und für $\forall x \in N, \forall y \in M : x < y$. Dann gibt es genau ein Element $s \in \mathbb{R}$ (Schnittzahl genannt) mit $x \leq s \leq y$ für alle $x \in N$ und für alle $y \in M$.

Bemerkung 1.4.6 zu vorangegangenen Regeln:

- Eine Menge G mit einer binären Operation „+“, die die Regeln der Addition erfüllt, heißt Abelsche Gruppe.
- Eine Menge G mit den binären Operationen „+“ und „ \cdot “, die die Regeln der Addition, der Multiplikation und das Distributivgesetz erfüllt, heißt Körper (vermerken, daß $G \setminus \{0\}$ bezüglich der Multiplikation eine Abelsche Gruppe ist)

- \mathbb{Q} erfüllt nicht das Vollständigkeitsaxiom. Zu

$$N = \{x \in \mathbb{Q} : x^2 < 2\} \cup \{x \in \mathbb{Q} : x < 0\},$$

$$M = \{x \in \mathbb{Q} : x^2 > 2\} \cap \{x \in \mathbb{Q} : x > 0\}$$

gibt es nämlich keine Schnitzzahl in \mathbb{Q} . Dies wäre ja $s = \sqrt{2} \notin \mathbb{Q}$.

Geht man von dieser Axiomatik aus, so entsteht die Frage, wie die natürlichen Zahlen eingeführt werden können. Man definiert dazu den Begriff der induktiven Menge: Eine Teilmenge $K \subset \mathbb{R}$ heißt induktiv, wenn $1 \in K$ und aus $k \in K$ auch $k + 1 \in K$ folgt. Der Durchschnitt aller induktiven Mengen ist wieder eine induktive Menge (Beweis!); diese wird mit \mathbb{N} bezeichnet und erfüllt gerade die ersten fünf Axiome des Punktes 1.4.1 (Beweis).

Rechnen mit Ungleichungen und Beträgen

Aus den Grundregeln für die Ordnungsrelation \leq folgen weitere Regeln für das Rechnen mit Ungleichungen

1. $x \leq y \Rightarrow -x \geq -y$
2. $x \leq y$ und $z \leq 0 \Rightarrow x \cdot z \geq y \cdot z$
3. $x^2 \geq 0$
4. $x \leq y$ und $u \leq v \Rightarrow x + u \leq y + v$
5. $0 \leq x \leq y$ und $0 \leq u \leq v \Rightarrow x \cdot u \leq y \cdot v$

Beweisen z.B. 1.:

$$x \leq y \Rightarrow x + (-x - y) \leq y + (-x - y) \Rightarrow -y \leq -x$$

Definition 1.4.7 Zu $a \in \mathbb{R}$ heißt $|a| := \begin{cases} a & \text{falls } a \geq 0 \\ -a & \text{falls } a < 0 \end{cases}$, Betrag von a .

Eigenschaften des Betrages:

1. $|a| \geq 0$ und $|a| = 0 \Leftrightarrow a = 0$

2. $|a \cdot b| = |a| \cdot |b|$
3. $|a + b| \leq |a| + |b|$ Dreiecksungleichung

Wir werden später einsehen (in der Grenzwerttheorie), daß diese Eigenschaften fundamental sind.

Vermerken: Wenn $a, b \in \mathbb{R}$, so wird mit $|a - b|$ der Abstand von a zu b bezeichnet.

Definition 1.4.8 Sei M eine Teilmenge von \mathbb{R} .

- (i) $x \in \mathbb{R}$ heißt obere Schranke von M , falls $m \leq x$ gilt für alle $m \in M$. Analog wird der Begriff „untere Schranke“ eingeführt.
- (ii) M heißt nach oben (bzw. nach unten) beschränkt, falls es eine obere (untere) Schranke von M gibt.
- (iii) $s \in \mathbb{R}$ heißt Supremum von M , falls s eine obere Schranke von M ist und für jede obere Schranke x von M die Beziehung $s \leq x$ gilt ($\Rightarrow s$ ist im Falle seiner Existenz eindeutig bestimmt). Analog wird der Begriff „Infimum von M “ als größte untere Schranke definiert.

Beispiel 1.4.9 $[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}$, $\sup[a, b] = b$, $\inf[a, b] = a$.
oder: $(a, b) := \{x \in \mathbb{R} : a < x < b\}$, $\sup(a, b) = b$, $\inf(a, b) = a$.

Satz 1.4.10 Jede nichtleere und nach oben (bzw. nach unten) beschränkte Menge $M \subset \mathbb{R}$ besitzt ein Supremum (bzw. ein Infimum).

Beweis. Man zerlege \mathbb{R} wie folgt:

$$L := \{x \in \mathbb{R} : \exists m \in M \text{ mit } x < m\}$$

$$K := \{y \in \mathbb{R} : \text{für } \forall m \in M \text{ gilt } y \geq m\}.$$

$$\Rightarrow \mathbb{R} = L \cup K; \quad L \neq \emptyset \text{ (da } M \neq \emptyset); \quad K \neq \emptyset$$

(da M nach oben beschränkt ist); $x < y$ für $\forall x \in L$ und $\forall y \in K$.

Nach dem Vollständigkeitsaxiom gibt es eine Schnitzzahl s . Diese ist gerade das Supremum von M . □

Folgerung 1.4.11 *Es gilt:*

1. \mathbb{N} ist nach oben unbeschränkt.
2. Für jedes $x > 0$ gibt es ein $n \in \mathbb{N}$ mit $0 < \frac{1}{n} < x$ (Axiom des Archimedes).
3. Zwischen zwei reellen Zahlen $x < y$ gibt es immer (unendlich viele) rationale Zahlen.

Beweis.

1. Nehmen an, daß \mathbb{N} beschränkt ist $\Rightarrow \exists s = \sup \mathbb{N}$. Betrachten $\mathbb{Z}_+ := \{n - 1 : n \in \mathbb{N}\}$ Für $x = n - 1 \in \mathbb{Z}_+$ gilt nun $x = n - 1 \leq s - 1$. Damit ist $s - 1$ eine obere Schranke von \mathbb{Z}_+ . Wegen $\mathbb{N} \subset \mathbb{Z}_+$ ist dann auch $s - 1$ eine obere Schranke für \mathbb{N} . \Rightarrow Widerspruch zu $s = \sup \mathbb{N}$.
2. Nach 1. kann $\frac{1}{x}$ keine obere Schranke für \mathbb{N} sein.

$$\Rightarrow \quad \exists n \in \mathbb{N} \text{ mit } 0 < \frac{1}{x} < n \Rightarrow 0 < 1 < n \cdot x$$

$$\Rightarrow \quad 0 < \frac{1}{n} < x.$$

3. Nach 2. $\exists n \in \mathbb{N}$ mit $0 < \frac{1}{n} < y - x \Rightarrow x < x + \frac{1}{n} < y$
Man wähle nun $m \in \mathbb{Z}$ mit $\frac{m-1}{n} \leq x < \frac{m}{n}$

$$\Rightarrow \quad x < \frac{m}{n} = \frac{m-1}{n} + \frac{1}{n} \leq x + \frac{1}{n} < y \Rightarrow \frac{m}{n} \in (x, y)$$

□

Zifferndarstellung reeller Zahlen

Seien n und g natürliche Zahlen, $g > 1$. Eine Darstellung der Form

$$n = \sum_{j=0}^k r_j g^j, \quad r_j \in \{0, 1, \dots, g-1\}$$

heißt g -adische Zifferndarstellung von n . Dabei heißen
 r_j – Ziffern
 g – Basis
 k – Stellenzahl (falls $r_k \neq 0$)

Man schreibt $n = (r_k \ r_{k-1} \cdots r_0)_g$

Häufig verwendete Basen sind

$g = 2$ (Dualdarstellung, Ziffern: 0, 1)
 $g = 10$ (Dezimaldarstellung, Ziffern: 0, ..., 9)
 $g = 16$ (Hexadezimaldarstellung, Ziffern: 0, ..., 9, A, ..., F)

Wir kommen nun zur g -adischen Darstellung reeller Zahlen. Sei dazu $x \in \mathbb{R}$ und es gelte ohne Einschränkung der Allgemeinheit $x > 0$. Wir spalten dazu den ganzen Anteil $[x] =: n$ (entier) ab: $x = n + x_0$, $n \in \mathbb{Z}_+$, $0 \leq x_0 < 1$. Der ganze Teil wird wie üblich umgerechnet. Für $x_0 \neq 0$ sucht man eine Darstellung nach Potenzen von $\frac{1}{g} = g^{-1}$ der Form

$$x_0 = \sum_{j=1}^{\infty} s_j g^{-j} = s_1 g^{-1} + s_2 g^{-1} + s_2 g^{-1} + s_3 g^{-3} + s_4 g^{-4} + \dots$$

($s_j \in \{0, 1, \dots, g-1\}$). Man schreibt dafür $x_0 = (0 \ s_1 \ s_2 \dots)$. Die Ziffern s_j lassen sich durch iterierte Multiplikation darstellen:

$$\begin{aligned} g x_0 &= s_1 + x_1, \quad 0 < x_1 < 1 \\ g x_1 &= s_2 + x_2, \quad 0 < x_2 < 1 \\ g x_2 &= s_3 + x_3, \quad 0 < x_3 < 1 \\ &\dots \end{aligned}$$

Dieser Prozeß bricht möglicherweise mit $x_k = 0$ ab.

Bemerkung 1.4.12 *Man kann anhand des Verfahrens unschwer erkennen, daß die Dezimaldarstellung jeder rationalen Zahl $\frac{m}{n}$ endlich oder periodisch ist. Die Umkehrung gilt auch: Jede periodische Dezimalzahl stellt eine rationale Zahl dar.*

1.4.3 Komplexe Zahlen

Die Gleichung $x^2 + 1 = 0$ hat in \mathbb{R} offenbar keine Lösung. Frage: Kann die Menge der reellen Zahlen (unter Beibehaltung der arithmetischen Operationen) so erweitert werden, daß obige Gleichung (und allgemeiner)

lösbar ist?

Eine solche Erweiterung existiert: Menge der komplexen Zahlen \mathbb{C} . Wir beschreiben diese Erweiterung wie folgt: Vermerken zunächst, daß $\mathbb{R}^2 := \mathbb{R} \times \mathbb{R} := \{(x, y) : x, y \in \mathbb{R}\}$

Geometrische Interpretation: Ebene.

Wir führen in die Menge \mathbb{R}^2 zwei binäre Operationen ein, die wir wieder Addition und Multiplikation nennen:

$$\text{Addition} \quad : \quad (a_1, b_1) + (a_2, b_2) := (a_1 + a_2, b_1 + b_2)$$

$$\text{Multiplikation} \quad : \quad (a_1, b_1) \cdot (a_2, b_2) := (a_1 \cdot a_2 - b_1 \cdot b_2, a_1 \cdot b_2 + a_2 \cdot b_1)$$

Das algebraische System $(\mathbb{R}^2, +, \cdot)$ bezeichnen wir mit \mathbb{C} . Es gilt $\mathbb{R} \subset \mathbb{C}$ im folgenden Sinne: Wir identifizieren \mathbb{R} mit $\mathbb{R} \times \{0\} := \{(x, 0) : x \in \mathbb{R}\}$. Dabei gilt $(x, 0) + (y, 0) = (x + y, 0)$ und $(y, 0) \cdot (x, 0) = (xy, 0)$. Auf \mathbb{R} stimmen also die „neuen“ Operationen der Addition und Multiplikation mit den „alten“ überein.

Vermerken: Die Operationen der Addition und Multiplikation sind assoziativ und kommutativ. Außerdem gelten die Distributivgesetze. (Nachweis elementar).

Wenn $(a, b) \neq (0, 0)$, dann $\exists (c, d) \in \mathbb{C}$ mit $(a, b)(c, d) = (1, 0)$, nämlich $(c, d) := (\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2})$. \mathbb{C} erfüllt die Regeln der Addition, die Regeln der Multiplikation und die Distributivgesetze des Punktes 1.4.2 und ist somit ein Körper.

Man überprüft direkt, daß $(a, b) = (a, 0) + (0, 1)(b, 0)$ gilt. Man schreibt dies auch als $a + ib$ ($= a + bi$) mit $i = (0, 1)$. Und es gilt $i^2 = -1$.

Wenn $c = a + ib$, $a, b \in \mathbb{R}$, so heißt a Realteil und b Imaginärteil der komplexen Zahl c ; $a = \text{Re}c$, $b = \text{Im}c$.

Absolutbetrag einer komplexen Zahl $c = a + ib$ ($a, b \in \mathbb{R}$) ist nach Definition die Zahl $|c| := \sqrt{a^2 + b^2}$

Beachte: Nach Definition ist die Quadratwurzel \sqrt{d} aus einer reellen Zahl $d \geq 0$ die kleinste reelle Zahl $x \geq 0$ mit $x^2 = d$!

Eigenschaften:

1. $|c| \geq 0, c = 0 \Leftrightarrow |c| = 0$
2. $|cd| = |c||d|$
3. $|c + d| \leq |c| + |d|$ (Dreiecksungleichung)

Die Eigenschaften 1. und 2. werden durch direktes Nachrechnen überprüft. Der Beweis der Eigenschaft 3. ist schwieriger. Er stützt sich auf

Hilfssatz 1.4.13 (Schwarzsche oder auch Cauchysche Ungleichung)

Seien a_i, b_i reelle Zahlen, $i = 1, \dots, n$. Dann gilt

$$\left(\sum_{i=1}^n a_i b_i \right)^2 \leq \left(\sum_{i=1}^n a_i^2 \right) \left(\sum_{i=1}^n b_i^2 \right) \quad (1.3)$$

Beweis. Wir setzen

$$A = \sum_{i=1}^n a_i^2, \quad B = \sum_{i=1}^n b_i^2, \quad C = \sum_{i=1}^n a_i b_i$$

$$\Rightarrow 0 \leq \sum (Ba_i - Cb_i)^2 = B^2 \sum a_i^2 - 2BC \sum a_i b_i + C^2 \sum b_i^2$$

$$= B^2 A - 2BC^2 + C^2 B = B^2 A - C^2 B = B(BA - C^2)$$

Wenn $B = 0 \Rightarrow$ 1.3 gilt offenbar; wenn $B \neq 0$, d.h. $B > 0$ gilt $\Rightarrow BA \geq C^2$ und damit die Behauptung. \square

Nun können wir die Eigenschaft 3. beweisen:

Beweis. Sei $c = c_1 + ic_2$, $d = d_1 + id_2$. Daraus folgt:

$$\begin{aligned} |c + d|^2 &= (c_1 + d_1)^2 + (c_2 + d_2)^2 \\ &= c_1^2 + c_2^2 + 2(c_1 d_1 + c_2 d_2) + d_1^2 + d_2^2 \\ &\leq c_1^2 + c_2^2 + d_1^2 + d_2^2 + 2(c_1^2 + c_2^2)^{\frac{1}{2}} (d_1^2 + d_2^2)^{\frac{1}{2}} \quad (\text{nach 1.3}) \\ &= ((c_1^2 + c_2^2)^{\frac{1}{2}} + (d_1^2 + d_2^2)^{\frac{1}{2}})^2 \\ &= (|c| + |d|)^2. \end{aligned}$$

\square

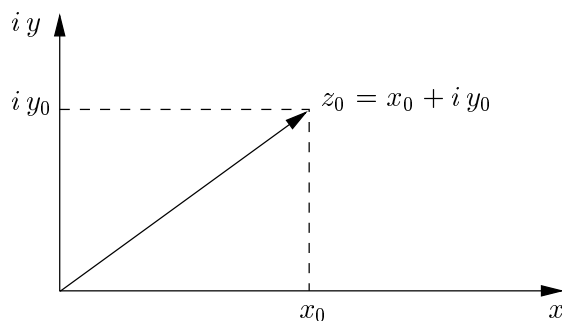
Die konjugiert komplexe Zahl

Wenn $z = x + iy$, ($x, y \in \mathbb{R}$) komplexe Zahl ist, dann heißt $\bar{z} := x - iy$ die zu z konjugiert komplexe Zahl.

Eigenschaften:

1. $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$
2. $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$
3. $z \bar{z} = |z|^2$
4. $z + \bar{z}$ – reelle Zahl, $\bar{\bar{z}} = z$ wenn z reell $\Rightarrow z = \bar{z}$,

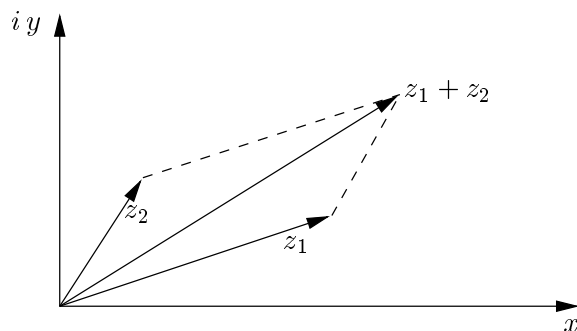
Geometrisch lassen sich die komplexen Zahlen als Punkte (Vektoren) einer Ebene, der komplexen Ebene bzw. der Gaußschen Zahlenebene darstellen.



Geometrische Interpretation der binären Operationen

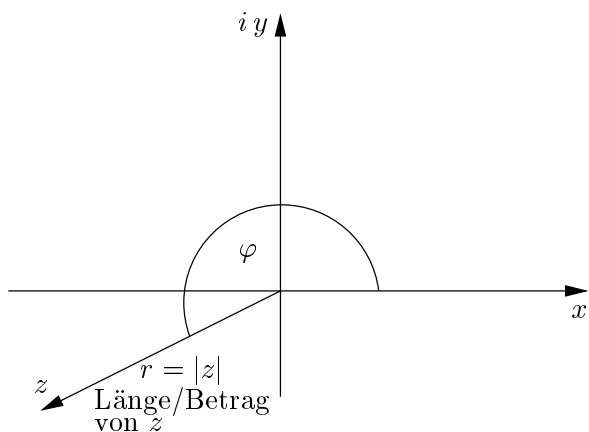
1. Addition

Vektoraddition mittels „Parallelogrammregel“.



2. Multiplikation

Diese wird transparent, wenn man die trigonometrische Darstellung einer komplexen Zahl einführt. Das heißt wir arbeiten mit Polarkoordinaten.



$$z = r(\cos \varphi + i \sin \varphi)$$

Wenn $z \neq 0$ ist, dann ist φ bis auf ganzzahlige Vielfache von 2π bestimmt und wird Argument der komplexen Zahl genannt. Das Argument einer komplexen Zahl $z \neq 0$ ist also eine Menge. Wenn jedoch $\varphi_1, \varphi_2 \in \arg z \Rightarrow \cos \varphi_1 = \cos \varphi_2, \sin \varphi_1 = \sin \varphi_2$. (Das heißt, daß die Periodizität der Sinus- und Cosinusfunktion ausgenutzt wird.) Bisweilen bezeichnet man mit $\arg z$ irgendein Element aus dem Argument von z .

Vermerken: wenn $z = 0$, dann ist das Argument nicht erklärt.

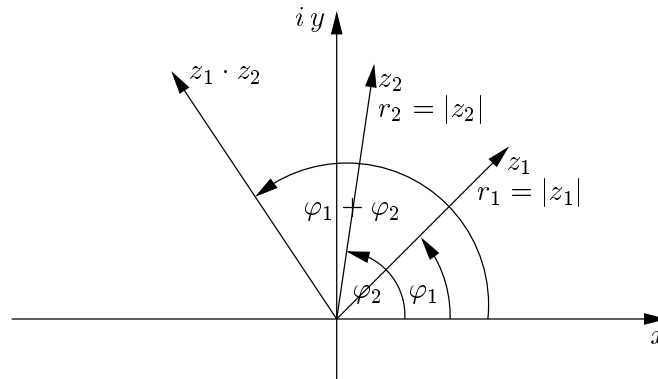
Es gelte nun:

$$z_1 = r_1(\cos \varphi_1 + i \sin \varphi_1)$$

$$z_2 = r_2(\cos \varphi_2 + i \sin \varphi_2)$$

$$\begin{aligned} \Rightarrow z_1 z_2 &= r_1 r_2 (\cos \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2 + i(\cos \varphi_1 \sin \varphi_2 + \sin \varphi_1 \cos \varphi_2)) \\ &= r_1 r_2 (\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)) \end{aligned}$$

(Additionstheorem)



Beträge werden multipliziert, Argumente addiert!

Wenn $z = r(\cos \varphi + i \sin \varphi)$, so zeigt man mittels Induktion, daß $z^n = r^n(\cos(n\varphi) + i \sin(n\varphi))$.

Kehren zur Ausgangsfrage zurück: Wie löst man Gleichungen der Gestalt $z^n = a$, $a \in \mathbb{C}$ und $n \in \mathbb{N}$ gegeben?

Betrachten zunächst den Fall $a = 1$ und bestimmen die n -ten Einheitswurzeln ε_k , $k = 0, \dots, n-1$, nämlich die Lösungen von $z^n = 1$. Sei nun z irgendeine komplexe Zahl mit $z^n = 1$. Wegen $|z|^n = |z^n| = 1 \Rightarrow |z| = 1$, das heißt z liegt auf dem Einheitskreis $\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$

$$\Rightarrow z = \cos \varphi + i \sin \varphi$$

$$\Rightarrow z^n = \cos(n\varphi) + i \sin(n\varphi) = 1$$

$$\Rightarrow \sin(n\varphi) = 0 \text{ (und } \cos(n\varphi) = 1)$$

$$\Rightarrow n\varphi = 2\pi p, \quad p \in \mathbb{Z}$$

$$\Rightarrow \varphi = \frac{2\pi p}{n}, \quad p \in \mathbb{Z}$$

Ferner gilt $p = ln + r$, $l \in \mathbb{Z}$ und $r \in \mathbb{Z}_+$, $0 \leq r < n$

$$\Rightarrow \varphi = 2\pi(l + \frac{r}{n}) \text{ und } \sin \varphi = \sin \frac{2\pi r}{n}, \cos \varphi = \cos \frac{2\pi r}{n}, \quad 0 \leq r < n$$

Also gibt es höchstens n Wurzeln der Gleichung $z^n = 1$ und diese befinden sich unter den Zahlen

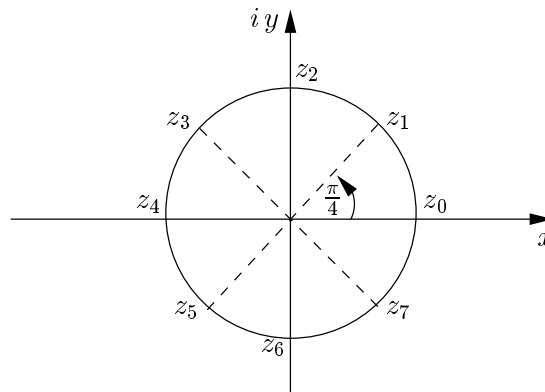
$$\varepsilon_k = \cos \frac{2\pi k}{n} + i \sin \frac{2\pi k}{n}, \quad k = 0, 1, \dots, n-1 \quad (1.4)$$

Nun überprüft man unmittelbar, daß jede der Zahlen 1.4 Wurzel der Gleichung $z^n = 1$ ist. Wir haben also folgenden Satz bewiesen:

Satz 1.4.14 *Es gibt genau n verschiedene komplexe Zahlen $\varepsilon_0, \dots, \varepsilon_{n-1}$, die die Gleichung $z^n = 1$ erfüllen, diese sind gegeben durch*

$$\varepsilon_k = \cos \frac{2\pi k}{n} + i \sin \frac{2\pi k}{n} \quad k = 0, 1, \dots, n-1.$$

Beispiel 1.4.15 *Lösungen von $z^8 = 1$*



Wir betrachten nun $z^n = a$ mit $a \neq 0$.

Zunächst schreibt man $a = r(\cos \varphi + i \sin \varphi)$, direktes Nachrechnen liefert, daß $z_k = \sqrt[n]{r}(\cos \frac{\varphi + 2\pi k}{n} + i \sin \frac{\varphi + 2\pi k}{n})$ der Gleichung $z^n = a$ genügt. Wir haben zu zeigen, daß es keine weiteren Lösungen gibt. Der Beweis dafür stützt sich auf folgende Überlegung: Seien z_1 und z_2 zwei Lösungen von $z^n = a$. Dann gilt $(\frac{z_2}{z_1})^n = 1$, das heißt es gibt ein $0 \leq k < n$ mit $\frac{z_2}{z_1} = \varepsilon_k$,

das heißt $z_2 = \varepsilon_k z_1$. Hieraus folgt, daß jede Lösung von der Gestalt $\varepsilon_k z_1$ ist, wobei $0 \leq k < n$.

Hieraus folgt: Jedes Polynom der Gestalt $z^n - a$ besitzt n Wurzeln ($a \neq 0$). Dieser Fakt kann verallgemeinert werden.

Hauptsatz der Algebra: In der Menge der komplexen Zahlen besitzt jedes Polynom n -ten Grades genau n Nullstellen. Die Aussage des Hauptsatzes der Algebra kann auch wie folgt gefaßt werden: Jedes (komplexe) Polynom P vom Grade n , d.h. $P(z) = a_n z^n + \dots + a_0$ mit $a_n \neq 0$ kann als

$$P(z) = a_n(z - z_1) \dots (z - z_n)$$

geschrieben werden, wobei z_1, \dots, z_n fixierte komplexe Zahlen sind; diese heißen Nullstellen oder Wurzeln des Polynoms P . Tritt unter den Wurzeln von P die Wurzel mit dem Index j genau k_j mal auf, so heißt die Zahl k_j die Vielfachheit der Wurzel k_j . Damit kann P auch als

$$P(z) = a_n(z - \lambda_1)^{k_1} \dots (z - \lambda_l)^{k_l}$$

geschrieben werden, wobei die Zahlen $\lambda_1, \dots, \lambda_l$ paarweise verschieden und $k_1 + \dots + k_l = n$ gilt. Die multiplikativen Elementarbausteine komplexer Polynome sind somit Faktoren der Gestalt $(z - \lambda)$ (bis auf den Faktor a_n , der ein Polynom vom Grade 0 ist). Für reelle Polynome gilt im Reellen diese Aussage nicht. Beispiel: $x^2 + 1$. Der Hauptsatz der Algebra kann aber benutzt werden, um auch die elementaren (multiplikativen Bausteine reeller Polynome im Reellen zu bestimmen.

Folgende Überlegung zeigt dies sofort: Jedes reelle Polynom P kann auch als komplexes Polynom betrachtet werden. Wenn nun λ eine komplexe Nullstelle von P ist (und eine solche existiert nach dem Hauptsatz der Algebra, dann gilt $P(\lambda) = \overline{P(\lambda)} = 0$. Nun ist aber

$$\overline{P(\lambda)} = a_n \overline{\lambda}^n + \dots + a_0;$$

somit ist auch $\overline{\lambda}$ Nullstelle von P . Ist nun λ nicht reell, so tritt ein Faktor der Gestalt $(z - \lambda)(z - \overline{\lambda}) = z^2 + az + b$ mit reellen Zahlen a, b auf. Somit kann ein reelles Polynom vom Grade n als

$$d(x - \lambda_1)^{k_1} \dots (x - \lambda_m)^{k_m} (x^2 + a_1 x + b_1)^{s_1} \dots (x^2 + a_p x + b_p)^{s_p}$$

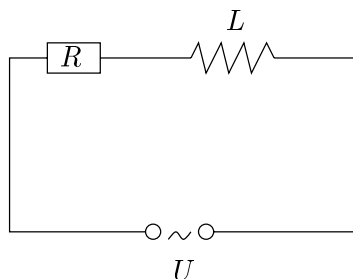
geschrieben werden, wobei $d \neq 0$ eine reelle Zahl, x_1, \dots, λ_m die reellen Nullstellen von P und $x^2 + a_j x + b_j$ paarweise verschiedene Polynome sind, die keine reellen Nullstellen besitzen. Außerdem gilt

$$k_1 + \dots + k_m + 2s_1 + \dots + 2s_p = n.$$

Folgerung. Jedes reelle Polynom P , dessen Grad ungerade ist, besitzt wenigstens eine reelle Nullstelle.

Die elementaren (multiplikativen) Bausteine reeller Polynome (im Reellen) sind somit Polynome vom Grade ≤ 2 (mit den oben angegebenen Eigenschaften). Dieser Fakt spielt später bei der Integration rationaler Funktionen eine wichtige Rolle. Komplexe Zahlen finden mannigfaltige Anwendungen. In der Elektrotechnik werden sie beispielsweise zur Untersuchung von Wechselstromkreisen herangezogen.

Beispiel 1.4.16 (Widerstand und Spule in Reihe geschaltet) Bei diesem Beispiel betrachten wir einen Wechselstromkreis.



R : Ohmscher Widerstand

L : Induktivität

$U(t) = U_0 \cos \omega t$: Spannung mit $U_0 > 0$

Kirchhoffsche Gesetze: Die Stromstärke $I(t)$ genügt der Differentialgleichung

$$L \frac{dI}{dt} + RI = U.$$

Diese Gleichung ist zu lösen. Setzen $e^{i\omega t} := \cos \omega t + i \sin \omega t$ und sehen die Größen $U(t)$ und $I(t)$ als Realteile gewisser komplexwertiger Funktionen

$U^*(t)$ und $I^*(t)$ an. Weil wir es mit periodischen Vorgängen zu tun haben, setzt man

$$U^*(t) = U_0 e^{i\omega t}, \quad I^*(t) = I_0^* e^{i\omega t}, \quad I_0^* \in \mathbb{C}$$

Setzen dies in obige Differentialgleichung ein:

$$L I_0^* i\omega e^{i\omega t} + R I_0^* e^{i\omega t} = U_0 e^{i\omega t} \Rightarrow I_0^* = \frac{U_0}{R + i\omega L}.$$

Polardarstellung von I_0^* : $I_0^* = I_0 e^{-i\alpha}$ mit $I_0 = \frac{U_0}{\sqrt{R^2 + \omega^2 L^2}}$, $\tan \alpha = \frac{\omega L}{R}$.
Mit der komplexen Stromstärke $I^* = I_0 e^{i(\omega t - \alpha)}$ folgt für die physikalische Stromstärke $I(t) = I_0 \cos(\omega t - \alpha)$. Dies ist tatsächlich (die einzige) Lösung obiger Differentialgleichung.

Kapitel 2

Vektoren

2.1 Der Begriff des Vektors

In der Physik, besonders in der Mechanik, unterscheidet man zwischen skalaren und vektoriellen Größen.

Skalare Größen: Temperatur, Luftfeuchtigkeit, ...

Vektorielle Größen: Kraft, Geschwindigkeit, Beschleunigung, ...

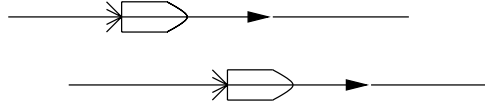
Intuitiv weiß man, daß vektorielle Größen durch **Länge** (Betrag, Norm) und durch **Richtung** gekennzeichnet sind.

Eine strenge mathematische Fassung ist nicht ganz einfach und beinhaltet sowohl die korrekte Definition der Länge als auch der Richtung. Wir werden diese Begriffe schrittweise herausarbeiten. Kehren wir zunächst wieder zur Physik zurück. Dort werden Vektoren als gerichtete Strecken dargestellt: Wenn wir uns den 3-dimensionalen physikalischen Raum denken, und uns dort zwei Punkte P und Q gegeben sind, so wird mit $v = \overrightarrow{PQ}$ der Vektor bezeichnet, der P als Anfangspunkt und Q als Endpunkt besitzt, d. h.



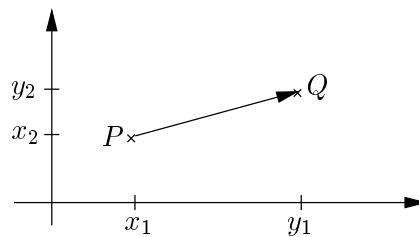
Diese Konzeption ist jedoch zu eng gefaßt (sie führt auf logische Schwierigkeiten). Beispielsweise sei uns eine Rakete gegeben, deren Triebwerk arbeitet

(also eine Kraft wirkt) und die sich geradlinig bewegt. Denken wir uns die gleiche Rakete mit dem selben Triebwerk bestückt, nun allerdings auf einer Bahn (Geraden), die zur ersten parallel ist:



In beiden Fällen besitzt die wirkende Kraft gleiche Länge und gleiche Richtung. Also sollte der Begriff des Vektors nicht von den Begriffen Anfangspunkt und Endpunkt abhängen. Daher werden Vektoren, die durch Parallelverschiebung ineinander überführt werden können, als gleich angesehen (freie Vektoren). Mathematisch bedeutet dies, daß in die Menge der gerichteten Strecken eine Äquivalenzrelation eingeführt werden kann: Zwei gerichtete Strecken heißen äquivalent, wenn sie durch eine Parallelverschiebung ineinander überführt werden können. Ein Vektor ist dann eine **Äquivalenzklasse**. Der Mathematiker ist damit immer noch nicht zufrieden; im 3-dimensionalen Raum können wir uns gerichtete Strecken vorstellen. Wie sieht das aber im Höherdimensionalen aus?

Um die Frage beantworten zu können, gehen wir wieder zum 3-dimensionalen physikalischen Raum zurück. Wir stellen uns vor, daß in ihm ein kartesisches Koordinatensystem gegeben ist. Dann kann dieser mit dem Raum \mathbb{R}^3 identifiziert werden; das Tripel (x_1, x_2, x_3) bedeutet dann gerade, daß x_1, x_2 und x_3 die Koordinaten eines gegebenen Punktes P sind: $P = P(x_1, x_2, x_3)$. Seien nun $P = P(x_1, x_2, x_3)$ und $Q = Q(y_1, y_2, y_3) \in \mathbb{R}^3$ fixierte Punkte. Wir betrachten \overrightarrow{PQ} , also einen Repräsentanten eines Vektors v (man sagt auch, v sei an P angelegt).



Wir betrachten den Punkt $S = S(y_1 - x_1, y_2 - x_2, y_3 - x_3)$. Dann ist, wie man leicht sehen kann, auch \overrightarrow{OS} ein Repräsentant von v . Dieser ist eindeutig

durch den Punkt

$$(y_1 - x_1, y_2 - x_2, y_3 - x_3) \in \mathbb{R}^3$$

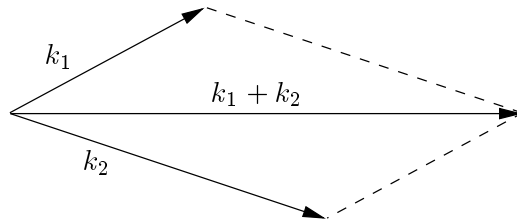
festgelegt. Also können wir sagen, daß zwischen allen (freien) Vektoren, die im Raume \mathbb{R}^3 „wirken“, und allen Punkten des Raumes \mathbb{R}^3 eine eindeutige Beziehung besteht.

Diese Zuordnung wird wie folgt gegeben. Ist v ein (freier) Vektor im Raume \mathbb{R}^3 , so ordnen wir ihm den Repräsentanten zu, dessen Anfangspunkt gerade der Koordinatenursprung ist. Der Endpunkt dieses Repräsentanten ist dann gerade der Punkt, den wir v zuordnen. Wollen wir diesen Vektor an den Punkt $R = (z_1, z_2, z_3) \in \mathbb{R}^3$ anlegen, so haben wir den Vektor (die gerichtete Strecke) \overrightarrow{RS} zu bilden mit $S = (z_1 + (y_1 - x_1), z_2 + (y_2 - x_2), z_3 + (y_3 - x_3))$. Mit anderen Worten, wenn wir den Vektor v an einen Punkt $R \in \mathbb{R}^3$ anzulegen wollen, haben wir den Punkt $S = (z_1 + (y_1 - x_1), \dots, z_3 + (y_3 - x_3))$ zu bilden, d. h. durch v wird eine Abbildung durch $(z_1, z_2, z_3) \mapsto (z_1 + (y_1 - x_1), \dots, z_3 + (y_3 - x_3))$ definiert. Diese Abbildung heißt Translations- oder Verschiebungsabbildung, die durch v erzeugt wird. Bezeichnung: τ_v oder τ_a , $a = (y_1 - x_1, \dots, y_3 - x_3)$. Da Vektoren in einer eindeutigen Beziehung zu den eben definierten Translationen stehen, werden diese Translationen auch freie Vektoren genannt. In dieser Terminologie ist ein freier Vektor eine Translation τ_a (die auch Translationsoperator genannt wird). Dieser Begriff ist nun völlig korrekt definiert. Die Formulierung, an einen Punkt $P = (x_1, \dots, x_n)$ den Vektor τ_a , $a = (a_1, \dots, a_3)$ anlegen, bedeutet nun das Bild $\tau_a(P)$ zu bestimmen (d. h. in unserer früheren Terminologie $\overrightarrow{P\tau_a(P)}$ zu betrachten).

Aus der Physik ist bekannt, daß zwei Kräfte $\overrightarrow{K_1}, \overrightarrow{K_2}$, die auf einen punktförmigen Körper wirken, äquivalent durch eine dritte Kraft $\overrightarrow{K_3}$ ersetzt werden können, die als Summe dieser Vektoren definiert ist (Superpositionsprinzip):

$$\overrightarrow{K_3} = \overrightarrow{K_1} + \overrightarrow{K_2} = \overrightarrow{K_2} + \overrightarrow{K_1};$$

die Summenbildung wird dabei nach der Parallelogrammregel vollzogen, d. h.



Auch hier entsteht die Frage, wie wir diese Operation mathematisch exakt fassen können. Wir denken uns die Vektoren $\vec{K_1}$ und $\vec{K_2}$ an den Koordinatenursprung angelegt. Dann sind $\vec{K_1}$ und $\vec{K_2}$ entsprechend durch ihre Endpunkte $(x_1, \dots, x_3), (y_1, \dots, y_3)$ festgelegt. Dies gilt auch für $\vec{K_3}$, dessen Endpunkt durch (z_1, \dots, z_3) gekennzeichnet sei. Man kann leicht einsehen, daß dann gerade

$$z_i = x_i + y_i \quad (i = 1, 2, 3)$$

gelten muß. Dies führt zunächst auf die Idee, daß der Raum \mathbb{R}^3 (allgemeiner: der Raum \mathbb{R}^n) mit der binären Operation einer Addition versehen werden kann:

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n)$$

Vektoren können mit Skalaren (d. h. mit reellen Zahlen) multipliziert werden (Streckung des Vektors und „Richtung“ bleibt erhalten, wenn $x > 0$, wird umgekehrt, wenn $x < 0$).

Dies führt auf eine Abbildung $\mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, $(\lambda, (x_1, x_2, \dots, x_n)) \mapsto (\lambda x_1, \lambda x_2, \dots, \lambda x_n)$, die sogenannte Skalarmultiplikation; diese wird als λx geschrieben ($\lambda \in \mathbb{R}, x \in \mathbb{R}^n$). Entsprechend heißen die Elemente $\lambda \in \mathbb{R}$ auch Skalare. Wenn $(x_1, \dots, x_n) \in \mathbb{R}$ gilt, dann heißen x_1, x_2, \dots, x_n die Komponenten oder Koordinaten des Punktes (x_1, \dots, x_n) . Die oben eingeführten algebraischen Operationen werden also komponentenweise (koordinatenweise) ausgeführt. Wir setzen noch $O = (0, 0, \dots, 0)$. Dann gilt $(v, w, z \in \mathbb{R}^n, \alpha, \beta \in \mathbb{R})$

1. $v + w = w + v$
2. $v + (w + z) = (v + w) + z$

3. $v + 0 = 0 + v = v$
4. $v + (-v) = (-v) + v = 0$
- und
5. $1 \cdot v = v$
6. $\alpha(\beta v) = (\alpha\beta)v$
7. $(\alpha + \beta)v = \alpha v + \beta v$
8. $\alpha(v + w) = \alpha v + \alpha w$

Vermerken: Die Eigenschaften 1. bis 4. bedeuten, daß $(\mathbb{R}^n, +)$ eine Abelsche Gruppe ist.

Die Eigenschaften 5. bis 8. bedeuten, daß für alle Elemente aus \mathbb{R}^n die Multiplikation mit allen Elementen (Skalaren) des Körpers \mathbb{R} definiert ist, wobei die genannten Gesetze gelten. Man sagt auch, daß \mathbb{R}^n ein Vektorraum (linearer Raum ist auch gebräuchlich) über dem Körper \mathbb{R} ist. Dies gibt Anlaß zu folgender Definition.

Definition 2.1.1 *Sei V eine beliebige Menge, für die eine binäre Operation „+“ definiert ist, d. h. $(x, y) \mapsto x + y$. Sei ferner K ein beliebiger Körper und eine Operation $K \times V \rightarrow V$, $(\alpha, v) \mapsto \alpha v$ gegeben. Wenn für diese Operationen die Eigenschaften 1. bis 8. gelten, dann heißt V linearer Raum über K .*

Alternative Definition:

Eine Abelsche Gruppe $(V, +)$ heißt linearer Raum über K , wenn eine Abbildung $K \times V \rightarrow V$, $(\alpha, v) \mapsto \alpha v$ gegeben ist, die die Eigenschaften 5. bis 8. besitzt.

Lineare Algebra = Theorie der linearen Räume

Wir werden vorrangig lineare Räume über \mathbb{R} oder \mathbb{C} betrachten.

Definition 2.1.2 *Sei V ein linearer Raum über dem Körper K ($= \mathbb{R}$ oder \mathbb{C}). Für fixiertes $a \in V$ heißt die Abbildung $v \mapsto a + v$ freier Vektor in V . Wir bezeichnen diese wieder mit \mathcal{T}_a und nennen die Abbildung auch Translation (um den Vektor a).*

Bemerkung 2.1.3 Den Vektor \mathcal{T}_a an den Punkt $P \in V$ „anlegen“ bedeutet nun das Element $\mathcal{T}_a(P)$ zu bilden; das geordnete Paar $(P, \mathcal{T}_a(P))$ bezeichnen wir auch mit $\overrightarrow{P\mathcal{T}_a(P)}$.

Offenbar sind die Abbildungen $\mathcal{T}_a : V \rightarrow V$ Bijektionen; dabei gilt

$$\mathcal{T}_a^{-1} = \mathcal{T}_{-a}, \quad \mathcal{T}_{a+b} = \mathcal{T}_a \circ \mathcal{T}_b, \quad \mathcal{T}_0 \mathcal{T}_a = \mathcal{T}_a \mathcal{T}_0 = \mathcal{T}_a.$$

\Rightarrow Die Menge $T(V)$ aller Translationen bildet bezüglich der binären Operation der Komposition eine Abelsche Gruppe; definiert man noch $\lambda \mathcal{T}_a$ durch $\mathcal{T}_{\lambda a}$ ($\lambda \in K$), so wird $T(V)$ sogar zu einem linearen Raum; dabei ist die Abbildung $\varphi : V \rightarrow T(V), a \mapsto \mathcal{T}_a$ offenbar eine Bijektion, für die noch

$$\begin{aligned} \varphi(\lambda a) &= \lambda \varphi(a) \\ \varphi(a + b) &= \varphi(a) \circ \varphi(b) \end{aligned}$$

gelten (z. B. $\varphi(a + b) = \mathcal{T}_{a+b} = \mathcal{T}_a \circ \mathcal{T}_b = \varphi(a) \circ \varphi(b)$).

Die Abbildung φ ist also mit den linearen Strukturen von V und $T(V)$ verträglich. Sie wird daher lineare Abbildung genannt.

Definition 2.1.4 Seien V_1 und V_2 lineare Räume über K . Eine Abbildung $\varphi : V_1 \rightarrow V_2$ heißt linear (auch linearer Operator genannt), wenn

$$\begin{aligned} \varphi(\lambda v) &= \lambda \varphi(v) \\ \varphi(a + b) &= \varphi(a) + \varphi(b) \end{aligned}$$

für alle $\lambda \in K$ und alle $a, b \in V_1$ gelten.

Eine lineare Abbildung $\varphi : V_1 \rightarrow V_2$ heißt Isomorphismus, wenn φ eine Bijektion ist. Die Räume V_1 und V_2 werden dann auch isomorph genannt.

Bemerkung 2.1.5

1. Die binären Operationen in V_1 und V_2 ebenso wie die Multiplikation mit Skalaren haben wir einheitlich mit „+“ und „ \cdot “ bezeichnet!
2. Isomorphe Räume können sich zwar durch die Natur ihrer Elemente unterscheiden, sind aber in ihrer linearen Struktur gleich. Daher werden solche Räume häufig identifiziert.

Beispielsweise können wir V mit $T(V)$ identifizieren. Das Studium der linearen Struktur von $T(V)$ (und einiger weiteren Eigenschaften) ist somit das Studium dieser Eigenschaften in V ! Wir werden uns daher auf die Untersuchung von V beschränken!

Definition 2.1.6 Sei V ein Vektorraum über K ($= \mathbb{R}$ oder \mathbb{C}). Für jedes Element $v \in V$, $v \neq 0$, heißt die Menge

$$B = \{\lambda v : \lambda \in \mathbb{R} \text{ und } \lambda > 0\}$$

Richtung. Zwei Vektoren v_1 und v_2 besitzen gleiche Richtung, wenn sie beide zu einer Menge der Gestalt B gehören (\Leftrightarrow es gibt einen Skalar $\mu > 0$ mit $v_1 = \mu v_2$).

Weitere Beispiele von linearen Räumen

1. \mathbb{C}^n ist ein linearer Raum über \mathbb{C} , wenn man die Addition durch

$$(\xi_1, \dots, \xi_n) + (\mu_1, \dots, \mu_n) := (\xi_1 + \mu_1, \dots, \xi_n + \mu_n)$$

und die Multiplikation mit Skalaren durch

$$\lambda(\xi_1, \dots, \xi_n) := (\lambda\xi_1, \dots, \lambda\xi_n)$$

einführt.

2. Sei P_n die Menge aller reellen Polynome, d. h. die Menge aller Funktionen

$$P : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \sum_{i=0}^n a_i x^i,$$

wobei a_0, \dots, a_n fixierte reelle Zahlen sind (analog führt man den Begriff komplexes Polynom ein).

Addition: $(\sum_{i=0}^n a_i x^i) + (\sum_{i=0}^n b_i x^i) := \sum_{i=0}^n (a_i + b_i) x^i$

Multiplikation mit Skalaren: $\lambda \sum_{i=0}^n a_i x^i := \sum_{i=0}^n (\lambda a_i) x^i$

3. Sei A die Menge aller Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$.

Addition: $(f_1 + f_2)(x) := f_1(x) + f_2(x)$

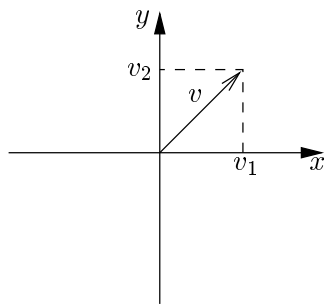
Multiplikation mit Skalaren: $(\lambda f)(x) := \lambda f(x)$

4. $L = \{x \in \mathbb{R} : x > 0\}$ ist Abelsche Gruppe (L, \cdot) .
Multiplikation mit Skalaren α aus $\mathbb{R} : \alpha * l := l^\alpha$

Fazit: Lineare Räume begegnen uns auf Schritt und Tritt!

2.2 Norm eines Vektors

Wir wenden uns nun dem Begriff Norm oder Länge eines Vektors zu. Dazu betrachten wir wieder den Fall eines konkreten linearen Raumes, nämlich des linearen Raumes \mathbb{R}^2 .



$$v = (v_1 + v_2)$$

Nach dem Lehrsatz des Pythagoras macht es Sinn, die Länge des Vektors v durch

$$\|v\| := \sqrt{v_1^2 + v_2^2}$$

zu erklären.

Allgemeiner: Sei $n \in \mathbb{N}$ beliebig. Die Länge oder Norm eines Vektors $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ wird durch

$$\|v\| := \sqrt{\sum_{i=1}^n v_i^2}$$

erklärt.

Die so erklärte Norm besitzt folgende Eigenschaften:

- (i) $\|v\| \geq 0$ und $\|v\| = 0 \Leftrightarrow v = 0$
- (ii) $\|xv\| = |\lambda| \cdot \|v\| \quad \forall \lambda \in \mathbb{R}$
- (iii) $\|v + w\| \leq \|v\| + \|w\|$ (Dreiecksungleichung)

Die Eigenschaften (i) und (ii) sind offensichtlich. Die Eigenschaft (iii) ist schwieriger zu beweisen. Sie ist aber eine direkte Folgerung aus der schon bewiesenen Cauchy-Schwarzschen Ungleichung (Punkt 1.4.3.).

Wir definieren nun allgemein:

Definition 2.2.1 *Gibt es für einen Vektorraum über K ($= \mathbb{R}$ oder \mathbb{C}) eine Abbildung $\|\cdot\| : V \rightarrow \mathbb{R}$, $v \mapsto \|v\|$, die die Eigenschaften (i) bis (iii) erfüllt, so heißt $\|\cdot\|$ eine Norm für V und $(V, \|\cdot\|)$ heißt normierter Raum, $\|v\|$ Norm (oder Länge) des Vektors v .*

Weitere Beispiele für normierte Räume

1. \mathbb{R}^n :

- $\|v\|_\infty := \max\{|v_1|, \dots, |v_n|\}$
- $\|v\|_2 := \sqrt{\sum_{i=1}^n v_i^2}$ (Euklidische Norm)
- $\|v\|_1 := \sum_{i=1}^n |v_i|$

2. \mathbb{C}^n :

- $\|v\|_\infty := \max\{|v_1|, \dots, |v_n|\}$
- $\|v\|_2 := \sqrt{\sum_{i=1}^n |v_i|^2}$
- $\|v\|_1 := \sum_{i=1}^n |v_i|$

Man kann zeigen, daß beliebige Normen auf \mathbb{R}^n oder \mathbb{C}^n im folgenden Sinne äquivalent sind: Wenn $\|\cdot\|_\alpha$ und $\|\cdot\|_\beta$ Normen auf \mathbb{R}^n oder auf \mathbb{C}^n sind, dann gibt es zwei Zahlen $c_1, c_2 > 0$ mit

$$c_1 \|x\|_\alpha \leq \|x\|_\beta \leq c_2 \|x\|_\alpha \quad \forall x \in \mathbb{R}^n \text{ (oder } \mathbb{C}^n \text{)}.$$

Das bedeutet:

Die Wahl der **konkreten** Norm ist zunächst nicht bedeutsam, wenn es um „Konvergenz“ geht. Dies ist für allgemeine lineare Räume nicht mehr richtig (und zwar für solche, für die $\dim V = \infty$ gilt).

Mit dem Begriff der Norm ist ein weiterer wichtiger Begriff verknüpft: Sei $(V, \|\cdot\|)$ ein normierter Raum. Dann heißt

$$d(v, w) := \|v - w\| \quad (v, w \in V)$$

Abstand der Vektoren v und w .

Eigenschaften

1. $d(v, w) \geq 0$ und $d(v, w) = 0 \Leftrightarrow v = w$
2. $d(v, w) = d(w, v)$
3. $d(v, z) \leq d(v, w) + d(w, z)$ (Dreiecksungleichung)

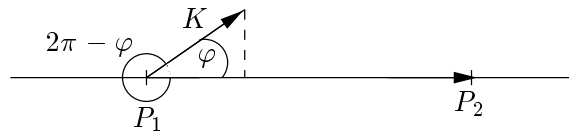
Man definiert allgemeiner (Unter X verstehen wir irgendeine Menge.): Sei $d : X \times X \rightarrow \mathbb{R}$ eine Funktion, die die eben formulierten Eigenschaften 1. bis 3. besitzt. Dann heißt d Metrik auf X , und das Paar (X, d) metrischer Raum. Die Zahl $d(x_1, x_2)$ wird als Abstand der Punkte x_1 und x_2 interpretiert. (Äußerst wichtige Begriffsbildung \rightarrow Analysis)

2.3 Der Begriff des Skalarproduktes von Vektoren

Wir beginnen wieder mit einem Beispiel:

Ein Massepunkt sei auf einer Geraden unter Wirkung der konstanten Kraft K verschiebbar.

Fragen: Welche Arbeit A wird geleistet, wenn dieser Punkt unter der Wirkung von K aus dem Punkt P_1 in den Punkt P_2 verschoben wird?



Antwort (Physik):

$A = \|k\| \cdot \|\overrightarrow{P_1 P_2}\| \cdot \cos \varphi$, wobei φ der Winkel ist, den K mit $\overrightarrow{P_1 P_2}$ bildet (keine Zweideutigkeit möglich, da $\cos \varphi = \cos(2\pi - \varphi)$ gilt!).

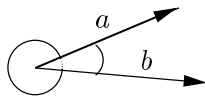
Daher definiert man: Das Skalarprodukt zweier Vektoren $a, b \in \mathbb{R}^3$ ist nach Definition die Zahl

$$\langle a, b \rangle := \|a\| \cdot \|b\| \cdot \cos \varphi,$$

wobei φ der Winkel ist, den diese beiden Vektoren miteinander bilden.

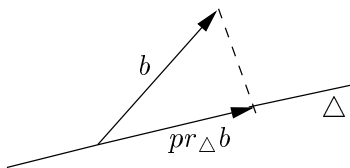
Bemerkung 2.3.1

1. Es gilt $\cos(2\pi - \varphi) = \cos \varphi$, daher ist es gleich, ob man den größeren oder kleineren der beiden Winkel wählt, den a und b miteinander bilden:



2. Sei $a \neq 0$ und $\Delta = \{\lambda a : \lambda \in \mathbb{R}\}$ die durch a gegebene (gerichtete) Gerade. Als orthogonale Projektion des Vektors b auf Δ bezeichnet man den Vektor

$$\begin{aligned} pr_{\Delta} b &:= (\|b\| \cos \varphi) \frac{a}{\|a\|} \\ &= \frac{\langle a, b \rangle}{\|a\|^2} a \quad (a \neq 0) \end{aligned}$$



Es gilt:

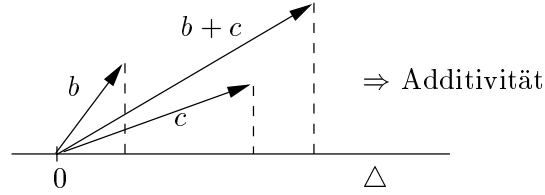
a)

$$\begin{aligned} pr_{\Delta}(b + c) &= pr_{\Delta} b + pr_{\Delta} c \\ pr_{\Delta} \lambda b &= \lambda pr_{\Delta} b \end{aligned}$$

\Rightarrow Die Abbildung $pr_{\Delta} : \mathbb{R}^3 \rightarrow \Delta$ ist eine lineare Abbildung. (Δ ist offenbar ein linearer Teilraum von \mathbb{R}^3 .)

$$b) \quad \|pr_{\Delta}(b + c)\| \leq \|pr_{\Delta} b\| + \|pr_{\Delta} c\|$$

Zeigen z. B. a)



weiter:

Wenn $\lambda \geq 0$, dann gilt Homogenität.

Wenn $\lambda \leq 0 \Rightarrow pr_{\Delta} \lambda b = |\lambda| \cdot \|b\| \cdot \cos(\pi - \varphi) = \lambda \cdot \|b\| \cdot \cos \varphi. \quad \square$

Diese Aussagen können nun benutzt werden, um folgende Eigenschaften des Skalarproduktes $\langle \cdot, \cdot \rangle$ zu zeigen:

$$\langle a, b \rangle = \langle b, a \rangle \quad (2.1)$$

$$\langle \lambda a, b \rangle = \lambda \langle a, b \rangle \quad (2.2)$$

$$\langle a + b, c \rangle = \langle a, c \rangle + \langle b, c \rangle \quad (2.3)$$

$$\langle a, a \rangle \geq 0, \quad \langle a, a \rangle = 0 \Leftrightarrow a = 0 \quad (2.4)$$

Wir beweisen z. B. (2.3):

Wenn $c = 0$, dann ist die Behauptung offensichtlich;

$c \neq 0$: Aus (2.1) folgt $\langle a + b, c \rangle = \langle c, a + b \rangle$. Aus 2.a) folgt nun

$$\begin{aligned} pr_{\Delta_c}(a + b) &= pr_{\Delta_c} a + pr_{\Delta_c} b \\ \frac{\langle c, a + b \rangle}{\|c\|} \frac{c}{\|c\|} &= \frac{\langle c, a \rangle}{\|c\|} \frac{c}{\|c\|} + \frac{\langle c, b \rangle}{\|c\|} \frac{c}{\|c\|} \end{aligned}$$

$$\Rightarrow \quad \langle c, a + b \rangle = \langle c, a \rangle + \langle c, b \rangle \text{ und damit die Behauptung.}$$

\square

Mittels (2.1) bis (2.4) läßt sich $\langle a, b \rangle$ durch die Koordinaten von a und b ausdrücken. Um dies einzusehen, führen wir die Einheitsvektoren $e_1 := (1, 0, 0)$, $e_2 := (0, 1, 0)$, $e_3 := (0, 0, 1)$ ein und vermerken

$$\mathbb{R}^3 \ni C = (C_1, C_2, C_3) = C_1 e_1 + C_2 e_2 + C_3 e_3.$$

$$\text{Aus } \langle e_i, e_j \rangle = \delta_{ij} := \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad (\text{Kroneckersymbol})$$

folgt nun unmittelbar ($a = (a_1, a_2, a_3)$, $b = (b_1, b_2, b_3)$)

$$\boxed{\langle a, b \rangle = a_1 b_1 + a_2 b_2 + a_3 b_3}$$

Diese Überlegungen nimmt man als Anlaß für folgende Definition

Definition 2.3.2 Sei V ein Vektorraum über \mathbb{R} . Sei

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$$

eine Abbildung, die die Eigenschaften 2.1 bis 2.4 aufweist. Dann heißt diese Abbildung Skalarprodukt und das Paar $(V, \langle \cdot, \cdot \rangle)$ Vektorraum mit Euklidischem Skalarprodukt (= Euklidischer Raum).

Wichtige Begriffsbildung für die Geometrie!

Beispiel 2.3.3

1. \mathbb{R}^n ist ein Euklidischer Vektorraum mit dem Skalarprodukt

$$\langle a, b \rangle = \sum_{k=1}^n a_k b_k \quad (2.5)$$

(Verallgemeinerung des Euklidischen Raumes \mathbb{R}^3)

2. Es bezeichne $C[0, 1]$ die Menge aller stetigen und reellwertigen Funktionen auf $[0, 1] := \{x \in \mathbb{R} : 0 \leq x \leq 1\}$. \Rightarrow reeller Vektorraum.
Skalarprodukt: $\langle f, g \rangle = \int_0^1 f g \, dx$

Satz 2.3.4 (Schwarzsche Ungleichung)

- a) Ist V ein Euklidischer Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$, so gilt

$$|\langle u, v \rangle| \leq \|u\| \cdot \|v\| \quad \text{mit } \|v\| := \sqrt{\langle v, v \rangle}.$$

b) Für jeden Euklidischen Vektorraum ist durch

$$\|v\| := \sqrt{\langle v, v \rangle}$$

eine Norm definiert.

Beweisidee.

zu a): Man betrachte den Ausdruck $\langle \alpha u + \beta v, \alpha u + \beta v \rangle \geq 0$ mit zunächst beliebig vorgegebenen Skalaren α und β . Man forme diese unter Berücksichtigung der Bilinearität um und spezifiziere $\alpha := \|v\|$ und $\beta := \pm \|u\|$.

zu b): Man beweise die Dreiecksungleichung mit Hilfe der Schwarzschen Ungleichung (Modell: komplexe Zahlen!).

□

Bemerkung 2.3.5 Für den Euklidischen Raum \mathbb{R}^n mit dem Standard-Skalarprodukt (2.5) haben wir die Schwarzsche Ungleichung bereits im Punkt komplexe Zahlen bewiesen.

$$\left(\left| \sum_{i=1}^n x_i y_i \right| \leq \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^n y_i^2 \right)^{\frac{1}{2}} \right)$$

Definition 2.3.6 Zwei Vektoren $u, v \in V$ heißen zueinander senkrecht oder orthogonal, wenn $\langle u, v \rangle = 0$ gilt. Man schreibt dies auch $u \perp v$ ($\Leftrightarrow v \perp u$).

Beispiel 2.3.7 Die Einheitsvektoren $e_i \in \mathbb{R}^3$ sind für $i \neq j$ zueinander orthogonal:

$$e_i \perp e_j, i \neq j.$$

Definition 2.3.8 Sei V ein Euklidischer Raum, $v, u \in V$. Die Zahl

$$\frac{\langle u, v \rangle}{\|u\| \cdot \|v\|}$$

wird als Kosinus des Winkels φ interpretiert, den die Vektoren u, v miteinander „bilden“.

Eine Reihe elementargeometrischer Probleme lassen sich mittels des Skalarproduktes lösen.

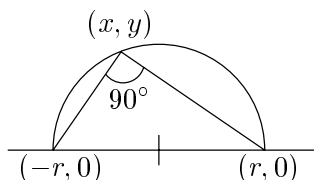
Beispiel 2.3.9

1. Wenn $a \perp b$ gilt, dann ist $\|a+b\|^2 = \|a\|^2 + \|b\|^2$ (Satz des Pythagoras)

$$\begin{aligned} \|a+b\|^2 &= \langle a+b, a+b \rangle \\ &= \langle a, a \rangle + \underbrace{\langle a, b \rangle}_0 + \underbrace{\langle b, a \rangle}_0 + \langle b, b \rangle \\ &= \|a\|^2 + \|b\|^2 \end{aligned}$$

2. (Satz des Thales)

Die Winkel über dem Durchmesser eines Kreises sind rechte.



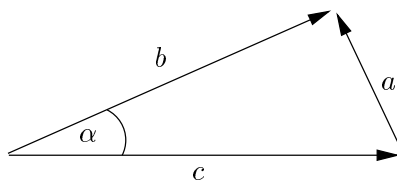
Wir bilden das Skalarprodukt der Vektoren $\overrightarrow{(-r, 0)(x, y)}$, $\overrightarrow{(r, 0)(x, y)}$, d. h. der Vektoren $(x+r, y)$ und $(x-r, y)$: dieses ist gleich

$$(x+r)(x-r) + y^2 = x^2 - r^2 + y^2 = 0$$

\Rightarrow Winkel ist ein rechter.

3. Kosinussatz:

$$\|a\|^2 = \|b\|^2 + \|c\|^2 - 2\|b\|\|c\|\cos \alpha$$

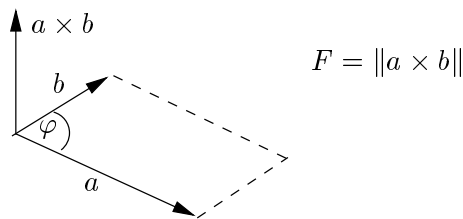


$$\begin{aligned}
a = b - c \quad \Rightarrow \quad \|a\|^2 &= \langle b - c, b - c \rangle \\
&= \|b\|^2 + \|c\|^2 - 2\|b\|\|c\| \frac{\langle b, c \rangle}{\|b\|\|c\|} \\
&= \|b\|^2 + \|c\|^2 - 2\|b\|\|c\| \cos \alpha
\end{aligned}$$

2.4 Das Vektorprodukt (nur im Raum \mathbb{R}^3)

Das Vektorprodukt (äußeres Produkt oder Kreuzprodukt) zweier Vektoren $a, b \in \mathbb{R}^3$ wird durch folgende Eigenschaften (eindeutig) bestimmt:

- (i) $\|a \times b\| = \|a\|\|b\|\sin \varphi$, wobei φ einer der beiden Winkel zwischen a und b ist ($-\sin \varphi = \sin(2\pi - \varphi)$!). Der Betrag des Vektorproduktes $a \times b$ ist somit gleich dem Flächeninhalt des von a und b aufgespannten Parallelogramms. (Rechnen Sie dies nach!)
- (ii) $a \times b$ steht senkrecht auf a und auf b . Somit liegt $a \times b$ bis auf das Vorzeichen (Richtung) fest.
- (iii) Das Tripel $(a, b, a \times b)$ der Vektoren $a, b, a \times b$ bildet (in dieser Reihenfolge) ein sogenanntes Rechtssystem (Rechte-Hand-Regel). Dies bedeutet: Dreht man den Vektor a auf kürzestem Weg in die Richtung von b , so zeigt $a \times b$ in die Richtung, in die sich eine Rechtsschraube bei dieser Drehung bewegen würde.



Bemerkung 2.4.1 Das Vektorprodukt findet in der Physik mannigfaltige Anwendungen (Drehmoment, ...).

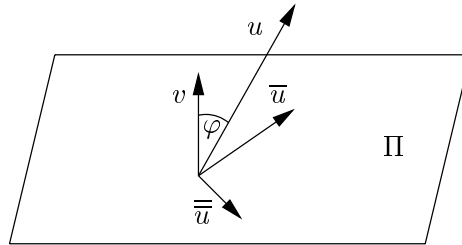
Eigenschaften des Vektorproduktes

$$(1) \quad a \times b = -(b \times a)$$

- (2) $(\lambda a) \times b = \lambda(a \times b)$
 (3) $a \times (b + c) = a \times b + a \times c$

Beweis.

- (1): Vertauschung der Reihenfolge von a und b ändert lediglich den Drehsinn und damit die Richtung von $a \times b$.
 (2): Multiplikation von a mit λ ändert den Flächeninhalt um das $|\lambda|$ -fache; die Richtung bleibt erhalten, falls $\lambda > 0$ ist und ändert sich für $\lambda < 0$.
 (3): Dieser Beweis ist etwas schwieriger und stützt sich auf folgende Feststellung:
 Ist $\|v\| = 1$, so kann $u \times v$ wie folgt konstruiert werden.



Dabei bezeichnet \bar{u} die orthogonale Projektion von u auf Π , und es gilt $\|\bar{u}\| = \|u\| \sin \varphi$. $\bar{\bar{u}} = u \times v$ wird nun durch eine Drehung von \bar{u} um $\frac{\pi}{2}$ (im Uhrzeigersinn) konstruiert.

Wegen (2) können wir $\|a\| = 1$ annehmen (falls $a \neq 0$ gilt). Sei nun Π die Ebene, auf der a senkrecht steht. Ferner gilt $\overline{b+c} = \bar{b} + \bar{c}$. Unter Berücksichtigung obiger Feststellung kann man nun die Behauptung unschwer einsehen.

□

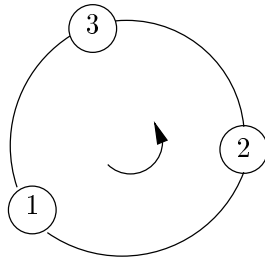
Mittels der Eigenschaften des Vektorproduktes kann nunmehr $a \times b$ mittels der Koordinaten von a und b ausgedrückt werden:

Es bezeichnen e_1, e_2, e_3 wiederum die Einheitsvektoren im Raume \mathbb{R}^3 . Aus der Definition des Vektorproduktes folgt dann

$$e_i \times e_i = 0 \quad , i = 1, 2, 3$$

$$e_1 \times e_2 = e_3, \quad e_2 \times e_3 = e_1, \quad e_3 \times e_1 = e_2$$

Merkregel:



zyklisches Durchlaufen
der Indizes.

Damit erhält man

$$\begin{aligned} a \times b &= \left(\sum_{i=1}^3 a_i e_i \right) \times \left(\sum_{i=1}^3 b_i e_i \right) \\ &= \sum_{i,j=1}^3 a_i b_j (e_i \times e_j) \\ &= (a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1) \end{aligned}$$

Der Physiker liebt es, diesen Sachverhalt „formal“ wie folgt darzustellen. Seien uns zwei Punkte $(a_1, a_2), (b_1, b_2)$ des Raumes \mathbb{R}^2 gegeben. Wir schreiben die Koordinaten als Zahlenschema $A = \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}$ auf und ordnen diesem Schema (welches auch Matrix genannt wird) eine Zahl $\det A$ zu, die Determinante von A genannt wird:

$$\det A := a_1 b_2 - b_1 a_2 \quad \left(\left(\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} \right) \right).$$

Seien uns nun n Vektoren $(a_{ij})_{j=1}^n$, $i = 1, \dots, n$, gegeben. Wir schreiben die Koordinaten wieder als Zahlenschema auf:

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

welches wiederum Matrix, genauer (n, n) -Matrix, genannt wird. Dieser Matrix kann wieder eine Zahl $\det A$, die Determinante von A zugeordnet werden. Wir geben diese zunächst noch für $n = 3$ an: (Entwicklung nach der ersten Zeile)

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13} \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

Der Physiker schreibt nun formal

$$a \times b = \det \begin{pmatrix} e_1 & e_2 & e_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}$$

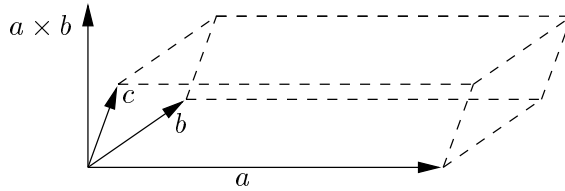
Bemerkung 2.4.2

1. Das Vektorprodukt ist nicht assoziativ, d. h. es gilt im allgemeinen $a \times (b \times c) \neq (a \times b) \times c$. Man wähle z. B. $a = e_1, b = e_2, c = e_2$
 $\Rightarrow e_1 \times (e_2 \times e_2) = e_1 \times 0 = 0$ und $(e_1 \times e_2) \times e_2 = e_3 \times e_2 = -e_1$.
2. (Jacobiidentität)
Für beliebige Vektoren a, b und $c \in \mathbb{R}^3$ gilt
 $a \times (b \times c) + b \times (c \times a) + c \times (a \times b) = 0$
(Man zeige dazu $a \times (b \times c) = \langle a, c \rangle b - \langle a, b \rangle c$)

Definition 2.4.3 Seien $a, b \in \mathbb{R}^3$ beliebigen Vektoren. Den Ausdruck $\langle a \times b, c \rangle$ nennt man Spatprodukt der Vektoren a, b und c .

Man überprüft durch elementares Nachrechnen, daß $\langle a \times b, c \rangle = \det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$ gilt.

Bemerkung 2.4.4 Das Spatprodukt kann als orientiertes, d. h. vorzeichenbehaftetes Volumen des von den Vektoren a, b und c aufgespannten Spates betrachtet werden:



$$\text{Grundfläche } A = \|a \times b\|, h = \|pr_{a \times b} c\| = \left| \left\langle \frac{a \times b}{\|a \times b\|}, c \right\rangle \right| \Rightarrow A \cdot h = V = |\langle a \times b, c \rangle|.$$

2.5 Allgemeine Vektorräume (lineare Räume)

Sei V ein linearer Raum (Vektorraum) über dem Körper K (K ist hier ein beliebiger Körper).

Beispiel 2.5.1

1. K ist über K ein linearer Raum.
2. Ist K' ein Teilkörper von K , so ist K ein linearer Raum über K' .
($\Rightarrow \mathbb{C}$ ist ein linearer Raum über \mathbb{R})
3. Wir betrachten den Koordinatenraum K^n , der wie folgt erklärt ist: Sei n eine natürliche Zahl und $K^n = K \times K \times \cdots \times K$ die Menge aller n -Tupel (a_1, \dots, a_n) mit $a_i \in K$, $i = 1, \dots, n$.
Addition: $(a_1, \dots, a_n) + (b_1, \dots, b_n) := (a_1 + b_1, \dots, a_n + b_n)$
Multiplikation mit Skalaren: $t(a_1, \dots, a_n) = (ta_1, \dots, ta_n)$
Für $n = 2, 3$ und $K = \mathbb{R}$ erhalten wir gerade \mathbb{R}^2 und \mathbb{R}^3 mit den schon früher definierten Operationen.
4. $L := \{\lambda \in \mathbb{R} : \lambda > 0\}$ ist eine abelsche Gruppe bezüglich der Multiplikation reeller Zahlen. Wir erklären für $l \in L$ und $\alpha \in \mathbb{R}$ das Produkt $\alpha \star l$ wie folgt: $a \star l = l^\alpha$. Die Axiome des Vektorraumes bedeuten dann gerade

$$(ab) \star l = a \star (b \star l) \quad \Leftrightarrow l^{ab} = (l^b)^a$$

und

$$(a + b) \star l = a \star l + b \star l \quad \Leftrightarrow l^{a+b} = l^a l^b.$$

5. Sei $F_{\mathbb{R}}(x)$ die Menge aller Funktionen $f : X \rightarrow \mathbb{R}$. Diese bildet einen linearen Raum, wenn die algebraischen Operationen punktweise eingeführt werden.
6. Wenn L ein linearer Raum über \mathbb{C} ist, dann kann ein neuer Vektorraum wie folgt gebildet werden: Die Abelsche Gruppe bleibt unverändert, die Multiplikation mit einer komplexen Zahl wird durch $z \star l := \bar{z}l$ erklärt.

Folgende Aussagen gelten in einem linearen Raum L :

Satz 2.5.2

1. $0 \cdot l = a \cdot 0 = 0 \quad \forall l \in L, \forall a \in K$.
2. $(-1)l = -l \quad \forall l \in L$. (Mit 1 ist hier das Einselement des Körpers K bezeichnet.)
3. Gilt mit $a \in K$ und $l \in L$ $al = 0 \Rightarrow a = 0$ oder $l = 0$.
4. Für beliebige $a_1, \dots, a_n \in K$ und $l_1, \dots, l_n \in L$ ist die Summe $\sum_{k=1}^n a_k l_k = a_1 l_1 + \dots + a_n l_n$ eindeutig erklärt. Diese Summe nennt man auch Linearkombination der Vektoren l_1, \dots, l_n . Analog ist das Produkt $(\prod_{k=1}^n a_k)l = a_1 \cdots a_n l$ eindeutig bestimmt.

Beweis. Zeigen z. B. 1., 2. und 3.:

- 1.: $0l = (0 + 0)l = 0l + 0l \Rightarrow 0 = 0l$ und analog gilt $a \cdot 0 = a(0 + 0) = a \cdot 0 + a \cdot 0$, d. h. $a \cdot 0 = 0$.
- 2.: $l + (-1)l = 1 \cdot l + (-1) \cdot l = (1 - 1)l = 0l = 0$. Also muß $(-1)l$ das zu l inverse Element sein, d. h. $(-1)l = -l$.
- 3.: Sei $a \neq 0 \Rightarrow (a^{-1}a)l = l = 0$; ist $l \neq 0$ und $a \neq 0$ so folgt mit voriger Überlegung $l = 0$.

□

Definition 2.5.3 Seien l_1, \dots, l_k eine endliche Anzahl von Vektoren des linearen Raumes L . Sie heißen linear unabhängig, wenn aus $\lambda_1 l_1 + \dots + \lambda_k l_k = 0$ ($\lambda_i \in K$) immer $\lambda_i = 0$ ($i = 1, \dots, k$) folgt. Andernfalls heißen sie linear abhängig.

Beispiel 2.5.4

1. Ein einzelner Vektor l ist linear unabhängig $\Leftrightarrow l \neq 0$.
2. Ist unter den Vektoren l_1, \dots, l_k der Nullvektor, so sind sie linear abhängig.
3. Tritt ein Vektor mehrfach in l_1, \dots, l_k auf, so sind diese ebenfalls linear abhängig.
4. Wir betrachten den Koordinatenraum K^n über K und setzen $e_i = (0, 0, \dots, 0, e, 0, \dots)$ (e - Einselement von K ; steht an i -ter Stelle). Es gelte $(\lambda_1, \dots, \lambda_n) = \sum_{k=1}^n \lambda_k e_k = 0$ ($= (0, \dots, 0)$). Daraus folgt $\lambda_k = 0$,
 $k = 1, \dots, n$. Folglich sind die Vektoren e_k linear unabhängig.

Satz 2.5.5 Eine endliche Anzahl von Vektoren l_1, \dots, l_n ($n \geq 2$) ist genau dann linear abhängig, wenn einer dieser Vektoren Linearkombination der anderen ist.

Beweis. Es seien die Vektoren l_1, \dots, l_n linear abhängig. Dann existieren Skalare λ_i , welche nicht alle identisch Null sind, für die $\sum_{k=1}^n \lambda_k l_k = 0$ gilt. Sei $\lambda_j \neq 0$. Dann gilt $l_j = -\lambda_j^{-1} \sum_{k=1, k \neq j}^n \lambda_k l_k$, d. h. l_j ist Linearkombination der anderen Vektoren. Sei nun z. B. l_n Linearkombination von l_1, \dots, l_{n-1} , d. h. $l_n = \sum_{k=1}^{n-1} \lambda_k l_k \Rightarrow \sum_{k=1}^{n-1} \lambda_k l_k - l_n = 0$ und $\lambda_n = -1 \Rightarrow$ Diese Vektoren sind linear abhängig. \square

Definition 2.5.6 Sei n eine natürliche Zahl. Ein linearer Raum L heißt n -dimensional, wenn es in ihm n linear unabhängige Vektoren, jedoch nicht $n+1$ solcher gibt ($\dim L = n$).

Definition 2.5.7 Sei L ein n -dimensionaler linearer Raum. Jede Menge $\{e_1, \dots, e_n\}$ linear unabhängiger Vektoren heißt Basis des Raumes L .

Bemerkung 2.5.8 Der Raum $L = \{0\}$ besitzt im obigen Sinne keine Basis. Seine Dimension wird per Definition gleich Null gesetzt.

Satz 2.5.9 Sei L ein n -dimensionaler Raum und $\{e_1, \dots, e_n\}$ eine Basis von L . Dann läßt sich jeder Vektor $l \in L$ eindeutig als Linearkombination

der Vektoren e_1, \dots, e_n darstellen, d. h. es gilt $l = \lambda_1 e_1 + \dots + \lambda_n e_n$, wobei die Skalare eindeutig bestimmt sind (Diese heißen auch Koordinaten des Vektors l bezüglich der Basis $\{e_1, \dots, e_n\}$).

Beweis. Aus der Definition der Basis folgt, daß die Vektoren l, e_1, \dots, e_n linear abhängig sind. Es existieren somit Skalare μ_i mit $\mu_0 l + \mu_1 e_1 + \dots + \mu_n e_n = 0$, wobei nicht alle dieser Skalare gleich Null sind. Offenbar gilt $\mu_0 \neq 0$, da anderenfalls auch die Skalare μ_1, \dots, μ_n gleich Null sein müßten (wegen der linearen Unabhängigkeit der Vektoren e_1, \dots, e_n). $\Rightarrow l = \lambda_1 e_1 + \dots + \lambda_n e_n$ mit $\lambda_i = -\mu_0^{-1} \mu_i$ ($i = 1, \dots, n$).

Eindeutigkeit: Es sei $l = \mu_1 e_1 + \dots + \mu_n e_n$ eine weitere Darstellung von l . Dann folgt $(\lambda_1 - \mu_1) e_1 + \dots + (\lambda_n - \mu_n) e_n = 0$. Aus der linearen Unabhängigkeit der Vektoren e_1, \dots, e_n folgt nun $\mu_k - \lambda_k = 0$ ($k = 1, \dots, n$). \square

Definition 2.5.10 Sei $\{e_1, \dots, e_n\}$ eine Basis von L und sei $l = \lambda_1 e_1 + \dots + \lambda_n e_n$. Dann heißen die Skalare λ_i ($\in K$) Koordinaten des Vektors l bezüglich der Basis $\{e_1, \dots, e_n\}$.

Satz 2.5.11 Sei L ein n -dimensionaler Raum über K . Dann ist L dem (n -dimensionalen) Koordinatenraum K^n isomorph.

Beweisidee. Fixieren eine Basis $\{e_1, \dots, e_n\}$ des Raumes L . Dann gilt für einen beliebigen Vektor $l \in L$ $l = \lambda_1 e_1 + \dots + \lambda_n e_n$, wobei der Vektor $(\lambda_1, \dots, \lambda_n) \in K^n$ eindeutig bestimmt ist. Die Abbildung $A : L \rightarrow K^n$, $l \mapsto (\lambda_1, \dots, \lambda_n)$, ist ein Isomorphismus (d. h. ist eine bijektive und lineare Abbildung). \square

Folgerung 2.5.12 Sind L_1 und L_2 zwei n -dimensionale Vektorräume über K , dann sind sie isomorph.

Satz 2.5.13 Sei L ein linearer Raum und e_1, \dots, e_n linear unabhängige Elemente von L . Wenn jeder Vektor aus L als Linearkombination der Elemente e_1, \dots, e_n dargestellt werden kann, dann gilt $\dim L = n$.

Beweisidee. Man zeige mittels vollständiger Induktion folgenden Fakt: Wenn g_1, \dots, g_k irgendein System von linear unabhängigen Vektoren in L ist und jeder der Vektoren g_j Linearkombination der linear unabhängigen Vektoren t_1, \dots, t_m ist, dann gilt $k \leq m$. \square

Beispiel 2.5.14 Sei \mathcal{P}_n der Raum aller reellen Polynome vom Grade $\leq n$. Dann folgt $\dim \mathcal{P}_n = n + 1$. Wir zeigen, daß die Polynome $1, \dots, t^n$ eine Basis bilden.

Lineare Unabhängigkeit: $\lambda_0 \cdot 1 + \lambda_1 \cdot t^1 + \dots + \lambda_n \cdot t^n = 0$.

Differentiation: $\lambda_1 \cdot 1 + 2 \cdot \lambda_2 \cdot t^1 + \dots + n \cdot \lambda_n \cdot t^{n-1} = 0$.

Fahren so fort und erhalten nach n Schritten $n! \lambda_n = 0$ usw. $\Rightarrow \lambda_0 = \dots = \lambda_n = 0$.

Da jedes Polynom aus \mathcal{P}_n als Linearkombination der Polynome $1, t, \dots, t^n$ darstellbar ist, folgt nach Satz 2.5.16 $\dim \mathcal{P}_n = n + 1$. Gleichzeitig haben wir die Methode des Koeffizientenvergleiches begründet: Wenn $\lambda_0 \cdot 1 + \lambda_1 t^1 + \dots + \lambda_n t^n = \mu_0 \cdot 1 + \mu_1 t^1 + \dots + \mu_n t^n$ gilt, folgt $\lambda_i = \mu_i \ \forall i = 0, \dots, n$.

Diese Aussagen gelten auch für komplexe Polynome.

2.6 Teilräume

Definition 2.6.1 Eine nichtleere Menge M eines linearen Raumes L heißt Teilraum oder auch Unterraum, wenn für alle $l_1, l_2 \in L$ und beliebige Skalare λ_1, λ_2 immer $\lambda_1 l_1 + \lambda_2 l_2 \in M$ gilt.

Bemerkung 2.6.2 Ein Teilraum $M \subset L$ ist selbst auch ein linearer Raum. (Beweisen Sie dies!)

Beispiel 2.6.3

1. $M = \{0\}$ und $M = L$ sind offenbar Teilräume (die auch trivial genannt werden).
2. Die geordneten n -Tupel (a_1, \dots, a_n) mit $a_{m+1} = a_{m+2} = \dots = a_n = 0$ ($m < n$) bilden einen Teilraum von K^n .
3. Die Menge aller geraden Polynome P mit $P(t) = P(-t)$ bildet einen Unterraum von \mathcal{P} .

Wir geben nun einen schönen und wichtigen Satz an, der zum Teil überraschende Folgerungen hat.

Satz 2.6.4 (Steinitzscher Austauschsatz) Sei L ein n -dimensionaler Raum und

$B = \{\tilde{b}_1, \dots, \tilde{b}_n\}$ eine Basis. Ferner sei $C_k = \{c_1, \dots, c_k\}$ ($k \leq n$) eine Menge linear unabhängiger Vektoren in L . Dann gibt es eine Teilmenge $B_k = \{b_1, \dots, b_k\} \subset B$ derart, daß $(B \setminus B_k) \cup C_k$ wieder eine Basis von L ist.

Der Beweis (, der sich auf das Prinzip der vollständigen Induktion stützt,) wird hier nicht gebracht.

Folgerung 2.6.5 Sei L ein n -dimensionaler Raum und die Vektoren $e_1, \dots, e_k \in L$ seien linear unabhängig. Dann kann die Menge $\{e_1, \dots, e_k\}$ zu einer Basis ergänzt werden, d. h. es existieren Vektoren l_1, \dots, l_{n-k} derart, daß $\{e_1, \dots, e_k, l_1, \dots, l_{n-k}\}$ eine Basis von L ist.

Beweis. Sei B eine Basis von L . Dann gibt es eine Teilmenge $B_k = \{b_1, \dots, b_k\} \subset B$, die gegen $\{e_1, \dots, e_k\}$ getauscht werden kann. \square

Definition 2.6.6 Ist $S \subset L$ eine Menge von Vektoren, so bezeichnet man mit $\mathcal{L}(S)$ die Menge aller Linearkombinationen von Elementen aus S (d. h. $l \in \mathcal{L}(S) \Leftrightarrow l = \lambda_1 l_1 + \dots + \lambda_k l_k$ mit $l_1, \dots, l_k \in S$, $k \in \mathbb{N}$.) und nennt dies lineare Hülle von S . Dafür werden auch die Bezeichnungen $\text{lin}S$ oder $\text{span}S$ verwendet. $\mathcal{L}(\{e_1, \dots, e_k\})$ heißt der auf die Vektoren e_1, \dots, e_k aufgespannte Raum.

Bemerkung 2.6.7 $\mathcal{L}(S)$ ist ein Teilraum von L .

Satz 2.6.8 Sei L ein n -dimensionaler Raum und e_1, e_2, \dots, e_k ein System von linear unabhängigen Vektoren des Raumes L . Dann gilt $\dim \text{span}\{e_1, \dots, e_k\} = k$.

Beweis. Unter Berücksichtigung von Folgerung 2.6.5 gelte $\dim \text{span}\{e_1, \dots, e_k\} = m > k$. Dann kann $\{e_1, \dots, e_k\}$ zu einer Basis von $\text{span}\{e_1, \dots, e_k\}$ erweitert werden: $\{e_1, \dots, e_k, l_1, \dots, l_{m-k}\}$. Dann gilt (bezüglich dieser Basis) z. B. für l_1

$$\begin{aligned} l_1 &= \lambda_1 e_1 + \dots + \lambda_k e_k + 0 \cdot l_1 \\ l_1 &= 0 \cdot e_1 + \dots + 0 \cdot e_k + 1 \cdot l_1 \end{aligned}$$

\square

Hilfssatz 2.6.9 Wenn L n -dimensional ist, so ist jeder Teilraum M von L der Dimension $\dim M \leq n$.

Beweis. Sei $\dim M > n$ (d. h. $\dim M = \infty$ oder m mit $m > n$). Dann gibt es in M mindestens $n+1$ linear unabhängige Elemente. Diese sind aber auch in L linear unabhängig. Das widerspricht sich mit $\dim L = n$. \square

2.7 Lineare Abbildungen

Zur Erinnerung: Seien V_1, V_2 lineare Räume über K . Eine Abbildung $f : V_1 \rightarrow V_2$ heißt linearer Operator, wenn

$$\begin{aligned} f(v_1 + v_2) &= f(v_1) + f(v_2) & \forall v_1, v_2 \in V_1, \\ f(\lambda v) &= \lambda f(v) & \forall \lambda \in K, \forall v \in V \end{aligned}$$

gilt.

Bemerkung 2.7.1

1. Es gilt $f(0) = 0$ sowie $f(-l) = -f(l)$. Mittels Induktion zeigt man $f(\sum \lambda_i v_i) = \sum \lambda_i f(v_i)$. ($f(0) = f(0 + 0) = f(0) + f(0)$)
2. Ein linearer Operator $f : V \rightarrow K$ heißt auch lineares Funktional.

Satz 2.7.2 Seien L und M lineare Räume über K und $\{l_1, \dots, l_n\}$ sowie $\{m_1, \dots, m_n\}$ Familien von Elementen aus L und M ($n \in \mathbb{N}$).

1. Wenn $\mathcal{L}(\{l_1, \dots, l_n\}) = L$ gilt, dann gibt es höchstens eine lineare Abbildung $f : L \rightarrow M$ mit $f(l_i) = m_i$, $i = 1, \dots, n$.
2. Sind die Elemente l_1, \dots, l_n linear unabhängig, so existiert eine lineare Abbildung mit $f(l_i) = m_i$.

Beweis.

- 1.: Seien f und \tilde{f} zwei lineare Abbildungen mit $f(l_i) = m_i$ und $\tilde{f}(l_i) = m_i$ ($i = 1, \dots, n$). Sei $l = \lambda_1 l_1 + \dots + \lambda_n l_n \Rightarrow$

$$\begin{aligned} f(l) &= \sum \lambda_i f(l_i) = \sum \lambda_i m_i \\ \tilde{f}(l) &= \sum \lambda_i \tilde{f}(l_i) = \sum \lambda_i m_i \end{aligned}$$

$\Rightarrow f(l) = \tilde{f}(l)$ für alle $l \in L$, d. h. $f = \tilde{f}$.

2.: Setzen $l = \lambda_1 l_1 + \dots + \lambda_n l_n$. (Diese Darstellung ist eindeutig, weil $\{l_1, \dots, l_n\}$ eine Basis ist.) Definieren $f : L \rightarrow M, l \mapsto l = \lambda_1 m_1 + \dots + \lambda_n m_n$. Diese Funktion ist linear und erfüllt die Gleichung $f(l_i) = m_i$.

□

Bemerkung 2.7.3

1. Lineare Operatoren werden häufig mit A, B, \dots bezeichnet. Statt $A(x)$ schreibt man auch Ax .
2. Sind A und B zwei lineare Operatoren und ist ihre Komposition $A \circ B$ erklärt, dann ist $A \circ B$ ebenfalls ein linearer Operator.

Definition 2.7.4 Sei $A : L \rightarrow M$ ein linearer Operator.

Die Menge $\ker A := \{x \in L : Ax = 0\}$ heißt Kern oder Nullraum des Operators A .

Bemerkung 2.7.5 $\ker A$ ist ein linearer Raum. (Beweisen Sie dies!)

Satz 2.7.6 Ein linearer Operator $A : L \rightarrow M$ ist genau dann injektiv, wenn $\ker A = \{0\}$ gilt.

Beweis. Sei A injektiv, d. h. aus $Ax = Ay$ folge $x = y$.

Damit gilt $Ax - Ay = A(x - y) = 0 \Leftrightarrow x = y$, d. h. $\ker A = \{0\}$. ($A^{-1}(0) = 0$)

Nun gelte $\ker A = \{0\}$ sowie $Ax = Ay \Leftrightarrow A(x - y) = 0 \Rightarrow x = y$. □

Satz 2.7.7 Sei $A : L \rightarrow M$ ein linearer Operator und L ein endlichdimensionaler Raum. Dann gilt $\dim \ker A + \dim \operatorname{im} A = \dim L$. ($\operatorname{im} A$ ist ein linearer Raum!)

Beweis. Offenbar ist $\ker A$ als Teilraum von L mit $\dim L < \infty$ ebenfalls endlichdimensional. Nun sei $\{e_1, \dots, e_k\}$ eine Basis von $\ker A$. Diese kann durch Vektoren e_{k+1}, \dots, e_n zu einer Basis von L ergänzt werden. Jedes Element aus $\operatorname{im} A$ besitzt folgende Gestalt:

$$A \left(\sum_{i=1}^n \lambda_i e_i \right) = \sum_{i=1}^n \lambda_i A e_i = \sum_{i=k+1}^n \lambda_i A e_i.$$

Das Bild von A wird somit durch die Vektoren Ae_{k+1}, \dots, Ae_n aufgespannt. Diese Vektoren sind aber linear unabhängig:

$$0 = \sum_{i=k+1}^n \lambda_i Ae_i = A \left(\sum_{i=k+1}^n \lambda_i e_i \right) \Rightarrow k = \sum_{i=k+1}^n \lambda_i e_i \in \ker A.$$

Dann gilt aber $k = \sum_{i=1}^k \lambda_i e_i$ mit gewissen Skalaren λ_i . Wir haben somit zwei Darstellungen des Vektors k bezüglich der Basis $\{e_1, \dots, e_n\}$ erhalten. $\Rightarrow \lambda_i = 0, i = 1, \dots, n$. Damit gilt nun $\dim \operatorname{im} A = n - k$. \square

Dieser Satz ist für die Lösungstheorie von linearen Gleichungen in endlich-dimensionalen Räumen von enormer Bedeutung.

Folgerung 2.7.8 *Sei $\dim L < \infty$ und $A : L \rightarrow L$ ein linearer Operator. Dann gilt: Entweder ist die Gleichung $Ax = b$ für alle $b \in L$ lösbar, und dann besitzt $Ax = 0$ nur die Nulllösung, oder sie ist nicht für jedes b lösbar, und dann besitzt $Ax = 0$ nichttriviale Lösungen. $Ax = b$ ist nämlich genau dann für alle $b \in L$ lösbar, wenn A eine Surjektion ist. Dann muß aber $\dim \operatorname{im} A = \dim L$ gelten und nach dem Dimensionssatz $\dim \ker A = 0$. Dies ist nach Satz 2.7.6 zur Injektivität von A äquivalent.*

Folgerung 2.7.9 *Sei $A : L \rightarrow L$ ein linearer Operator und $\dim L < \infty$. Dann ist L genau dann injektiv, wenn L surjektiv ist (Injektivität \Leftrightarrow Bijektivität).*

Bemerkung 2.7.10 *In unendlichdimensionalen Räumen gilt diese Aussage nicht. Im Raum \mathcal{P} der reellen Polynome ist z. B. der lineare Operator $A : \mathcal{P} \rightarrow \mathcal{P}, f(x) \mapsto xf(x)$ injektiv, aber nicht surjektiv.*

Die Matrixdarstellung eines linearen Operators

Seien N und M endlichdimensionale Räume über K und entsprechend $\{e_1, \dots, e_n\}$ und $\{g_1, \dots, g_m\}$ Basen in N und M . Sei ferner $A : N \rightarrow M$ ein linearer Operator. Wir versuchen die Wirkungsweise von A auf einen Vektor $l \in N$ in Koordinaten darzustellen:

$$l = l_1 e_1 + \dots + l_n e_n \Rightarrow Al = \sum_{k=1}^n l_k Ae_k = \sum_{k=1}^n l_k \left(\sum_{i=1}^m a_{ik} g_i \right) = \sum_{i=1}^m \left(\sum_{k=1}^n a_{ik} l_k \right) g_i.$$

Wir haben ermittelt: Wenn $Al = \sum_{i=1}^m b_i g_i$ gilt, dann ist $b_i = \sum_{k=1}^n a_{ik} l_k$. Die Wirkungsweise von A wird somit durch die $m \times n$ Zahlen a_{ik} , $i = 1, \dots, m$ und $k = 1, \dots, n$ vollständig beschrieben, wenn die Basen $\{e_1, \dots, e_n\}$ und $\{g_1, \dots, g_m\}$ in den entsprechenden Koordinatendarstellungen fixiert bleiben. Diese Zahlen schreiben wir als ein Zahlenschema auf $((m \times n)$ -Matrix):

$$[A] = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Wir schreiben nun die Koordinaten l_1, \dots, l_n unseres Vektors l als $(n \times 1)$ -Matrix nieder:

$$\begin{pmatrix} l_1 \\ \vdots \\ l_n \end{pmatrix}.$$

Der Skalar b_i ist nun das „Skalarprodukt“ der i -ten Zeile unserer Matrix $[A]$ mit dem Vektor (l_1, \dots, l_n) . Deshalb wird das „Produkt“ von $[A]$ mit

$$\begin{pmatrix} l_1 \\ \vdots \\ l_n \end{pmatrix} \text{ wie folgt eingeführt:}$$

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} l_1 \\ \vdots \\ l_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^n a_{1k} l_k \\ \vdots \\ \sum_{k=1}^n a_{mk} l_k \end{pmatrix} \quad (2.6)$$

Dies ist auch der Grund, warum Vektoren aus K^n häufig als Spaltenvektoren geschrieben werden. Es ist leicht einzusehen, daß durch $[A]$ eine lineare Abbildung von K^n in K^m erzeugt wird, nämlich durch (2.6). Die lineare Abbildung wird ebenfalls mit $[A]$ bezeichnet. Es gilt nun folgendes Diagramm:

$$\begin{array}{ccc} & A & \\ N & \rightarrow & M \\ E_1 \downarrow & & \uparrow E_2 \\ K^n & \rightarrow & K^m \\ & [A] & \end{array}$$

$E_1 : N \rightarrow K^n, l \mapsto (l_1, \dots, l_n)$ (Koordinaten bezüglich $\{e_1, \dots, e_n\}$)

$E_2 : K^m \rightarrow M, (b_1, \dots, b_m) \mapsto (b_1 g_1, \dots, b_m g_m)$

Weil E_1 und E_2 Bijektionen sind, können die Eigenschaften von A mittels derer von $[A]$ studiert werden ($A = E_2[A]E_1, [A] = E_2^{-1}AE_1^{-1}$).

Definition 2.7.11 Die Matrix $[A]$ heißt Matrixdarstellung des Operators A bezüglich der Basen $\{e_1, \dots, e_n\}$ und $\{g_1, \dots, g_m\}$.

Bemerkung 2.7.12 Die Matrixdarstellung hängt von den Basen ab; ändern sich die Basen, so ändert sich die Matrixdarstellung!

Wir geben nun ein Beispiel an, wie obige Überlegungen genutzt werden können.

Wir betrachten den Raum \mathcal{P}_n aller reellen Polynome vom Grade höchstens n . Eine Basis dieses Raumes ist $\{1, t, t^2, \dots, t^n\}$. Wir definieren einen Operator D (Differentialoperator) durch $Dt^k := k \cdot t^{k-1}$ und $D \cdot 1 := 0$. (Nach Satz 2.7.2 existiert ein eindeutig bestimmter linearer Operator mit diesen Eigenschaften). Damit gilt

$$\begin{aligned} D \cdot 1 &= 0 \cdot 1 + 0 \cdot t + \dots + 0 \cdot t^{n-1} + 0 \cdot t^n \\ D \cdot t^1 &= 1 \cdot 1 + 0 \cdot t + \dots + 0 \cdot t^{n-1} + 0 \cdot t^n \\ D \cdot t^2 &= 0 \cdot 1 + 2 \cdot t + \dots + 0 \cdot t^{n-1} + 0 \cdot t^n \\ &\vdots \\ D \cdot t^n &= 0 \cdot 1 + 0 \cdot t + \dots + n \cdot t^{n-1} + 0 \cdot t^n \end{aligned}$$

Trägt man die gewonnenen Koeffizienten spaltenweise ein, erhält man

$$[D] = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 2 & 0 & \dots & 0 \\ 0 & 0 & 0 & 3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \ddots & n \\ 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Sei nun speziell $n = 2$ und $A : \mathcal{P}_2 \rightarrow \mathcal{P}_2$ durch $f \mapsto (t+1)D^2f - 2Df + f$ ($D^2 := D \circ D$!) gegeben. Andere Schreibweise: $Af = (t+1)\frac{d^2f}{dt^2} - 2\frac{df}{dt} + f$. Wir betrachten die Differentialgleichung

$$(Af)(t) = 3t^2 - 3t + 1, \quad f \in \mathcal{P}_2. \quad (2.7)$$

Zunächst suchen wir $[A]$ bezüglich der Standardbasis $\{1, t, t^2\}$:

$$\begin{aligned} A1 &= 1 \cdot 1 + 0 \cdot t + 0 \cdot t^2 \\ At &= -2 \cdot 1 + 1 \cdot t + 0 \cdot t^2 \\ At^2 &= 2 \cdot 1 - 2 \cdot t + 1 \cdot t^2 \end{aligned}$$

Somit lautet die Matrixdarstellung (spaltenweise eingetragen!)

$$[A] = \begin{pmatrix} 1 & -2 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Die Gleichung (2.7) ist nun der Gleichung

$$[A]x = \begin{pmatrix} 1 \\ -3 \\ 3 \end{pmatrix} \quad (2.8)$$

äquivalent, d. h. jede Lösung $x = (x_1, x_2, x_3)$ erzeugt vermöge $x_0 \cdot 1 + x_1 \cdot t + x_2 \cdot t^2$ eine Lösung von (2.7) und jede Lösung $a_0 \cdot 1 + a_1 \cdot t + a_2 \cdot t^2$ erzeugt eine Lösung (a_0, a_1, a_2) von (2.8).

Wir schreiben nun $[A]x$ aus:

$$\begin{aligned} 1x_1 - 2x_2 + 2x_3 &= 1 \\ x_2 - 2x_3 &= -3 \\ x_3 &= 3 \end{aligned}$$

$$\Rightarrow x_3 = 3, x_2 = 3, x_1 = 1$$

Somit ist $1 + 3t + 3t^2$ die einzige Lösung der Differentialgleichung (2.7). (Beachte: (2.8) ist für jede rechte Seite lösbar und somit ist $[A] : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ eine Bijektion nach Folgerung 2.7.12)

Dieses Beispiel kann modifiziert werden:

Sei $\mathcal{P}_n(e^{i\omega t})$ der lineare Raum aller Funktionen der Gestalt $f(t) = (a_0 + a_1 t + \dots + a_n t^n)e^{i\omega t}$, $a_i \in \mathbb{C}$, $(f : \mathbb{R} \rightarrow \mathbb{C})$.

Basis: $\{e^{i\omega t}, te^{i\omega t}, \dots, t^n e^{i\omega t}\}$ (Man benutze $e^{i\omega t}e^{-i\omega t} = 1$.)

Wir führen den Operator D der Differentiation durch

$$\begin{aligned} De^{i\omega t} &:= i\omega e^{i\omega t} \\ Dt^k e^{i\omega t} &:= kt^{k-1}e^{i\omega t} + i\omega t^k e^{i\omega t} = (kt^{k-1} + i\omega)t^k e^{i\omega t} \end{aligned}$$

ein. Man bestimme seine Matrixdarstellung in obiger Basis!

Die bislang aufgebaute Theorie liefert die Möglichkeit, Differentialgleichungen der Gestalt

$$a_n D^n f + a_{n-1} D^{n-1} f + \dots + f = b_0 e^{i\omega t} + \dots + b_n t^n e^{i\omega t}$$

auf das Studium linearer Gleichungssysteme zurückzuführen (*).

Diese Betrachtungen zeigen eindringlich: Wir müssen lernen, mit linearen Gleichungssystemen umzugehen.

(*) In der Praxis geschieht dies wie folgt:

Wir wählen als Ansatz $f = x_0 e^{i\omega t} + x_1 t e^{i\omega t} + \dots + x_n t^n e^{i\omega t}$ und setzen dies in unsere Differentialgleichung ein. Danach werden die Koeffizienten bei gleichen Basiselementen gleichgesetzt. Es entsteht ein System von n linearen Gleichungen in $n + 1$ Unbekannten x_0, \dots, x_n . Aus der allgemeinen Theorie der Differentialgleichungen folgt, daß dieser Ansatz gegebenenfalls modifiziert werden muß. Dazu führt man das charakteristische Polynom p , $p(\lambda) := a_n \lambda^n + \dots + a_0$ ein; wenn $p(i\omega) \neq 0$ ist, dann wählt man obigen Ansatz; ist $i\omega$ eine k -fache Nullstelle von p , dann setzt man $f(t) := t^k (x_0 + \dots + x_n t^n) e^{i\omega t}$. (Man muß also den Raum der Ansatzfunktionen modifizieren.)

Einschub

Lineare Differentialgleichungen mit konstanten Koeffizienten

Modell: Ein Massepunkt M mit der Masse m sei an einer horizontalen Feder befestigt und liege auf einer ebenfalls horizontalen x -Achse. Bei ungespannter Feder befinde sich M im Nullpunkt. (Gleichgewichtslage - man nimmt also an, daß in diesem Fall keinerlei Kräfte auf M wirken!) Verschiebt man M , so übt die (ausgedehnte oder zusammengedrückte) Feder eine sogenannte Rückstellkraft K aus, die M in die Gleichgewichtslage zurückzubringen versucht. Bei kleinem x (d.h. bei kleiner Verschiebung) ist häufig in guter Näherung $K = -kx$, wobei k eine Proportionalitätskonstante ist, die von der Feder abhängt und Federsteifigkeit genannt wird. (Man nimmt also die Gültigkeit des Hookeschen Gesetzes an.) Befindet sich M zur Zeit t im Punkt $x(t)$ und sehen wir von dämpfenden Reibungskräften ab, so gilt nach dem Newtonschen Kraftgesetz („Kraft ist Masse mal Beschleunigung“)

$$m\ddot{x} = -kx \tag{2.9}$$

(2.4) nennt man die Gleichung des (ungedämpften) harmonischen Oszillators. In der Regel wird man jedoch Reibungs- oder Dämpfungskräfte in Rechnung stellen müssen, z. B. wenn M sich in einem zähen Medium bewegt. Diese hemmenden Kräfte werden in vielen Fällen proportional zur Geschwindigkeit von M sein, also die Form $-r\dot{x}$ mit festem r haben. Die Differentialgleichung (2.9) muß dann zur Gleichung

$$m\ddot{x} = -kx - r\dot{x} \quad \text{oder} \quad \ddot{x} + \frac{r}{m}\dot{x} + \frac{k}{m}x = 0 \quad (2.10)$$

erweitert werden. Die Rückstellkraft und die Dämpfungskraft sind gewissermaßen die „inneren“ Kräfte unserer Masse-Feder-Medium-Vorrichtung. In vielen Fällen wirkt aber auf M noch zusätzlich eine zeitabhängige „äußere“ Kraft, eine sogenannte Zwangskraft oder erregende Kraft $K(t)$ (z. B. Erdanziehung oder Mondanziehung). In diesem Fall tritt an die Stelle von (2.10) die Differentialgleichung

$$m\ddot{x} = -kx - r\dot{x} + K(t) \quad \text{oder} \quad \ddot{x} + \frac{r}{m}\dot{x} + \frac{k}{m}x = \frac{1}{m}K. \quad (2.11)$$

Wir interessieren uns also für Differentialgleichungen der Gestalt

$$\ddot{u} + a\dot{u} + bu = s \quad (a, b \in \mathbb{R}). \quad (2.12)$$

Wir werden im folgenden kurz auf die Theorie solcher Gleichungen eingehen.

Einführende Feststellungen

Unter J verstehen wir im folgenden eine der Mengen $(a, b) := \{x \in \mathbb{R} : a < x < b\}$, $(-\infty, a) := \{x \in \mathbb{R} : x < a\}$, $(a, \infty) := \{x \in \mathbb{R} : x > a\}$ und \mathbb{R} (werden häufig auch offene Intervalle genannt). Sei $f : J \rightarrow \mathbb{C}$ gegeben. Wir definieren $\operatorname{Re} f : J \rightarrow \mathbb{R}$ durch $(\operatorname{Re} f)(x) = \operatorname{Re} f(x)$, $\operatorname{Im} f : J \rightarrow \mathbb{R}$ durch $(\operatorname{Im} f)(x) := \operatorname{Im} f(x) \Rightarrow f = \operatorname{Re} f + i \operatorname{Im} f$. Eine Funktion $f : J \rightarrow \mathbb{C}$ heißt im Punkt $x_0 \in J$ differenzierbar, wenn $\operatorname{Re} f$ und $\operatorname{Im} f$ in diesem Punkt differenzierbar sind. Im diesem Fall heißt $f'(x_0) := (\operatorname{Re} f)'(x_0) + i(\operatorname{Im} f)'(x_0)$ Ableitung der Funktion f im Punkt x_0 . Analog werden höhere Ableitungen definiert: $f^{(n)}(x_0) := (\operatorname{Re} f)^{(n)}(x_0) + i(\operatorname{Im} f)^{(n)}(x_0)$.

Beispiel 2.7.13 Wir setzen $e^{ix} := \cos x + i \sin x$, $x \in \mathbb{R}$. Diese Funktion bildet \mathbb{R} auf $\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$ ab. Diese Funktion ist differenzierbar; Ableitung ist gleich $-\sin x + i \cos x = i(\cos x + i \sin x)$, d.h. $(e^{ix})' = ie^{ix}$ (Ableitung ergibt sich durch formales Anwenden der Kettenregel).

Sei nun $z \in \mathbb{C}$ beliebig. Wir definieren eine äußerst wichtige Funktion, die sogenannte Exponentialfunktion $\exp : \mathbb{C} \rightarrow \mathbb{C}$ durch $z \mapsto \exp z = e^z := e^x e^{iy} = e^x (\cos y + i \sin y)$, ($z = x + iy$ mit $x, y \in \mathbb{R}$). Unter Berücksichtigung der Eigenschaften der Funktion $x \mapsto e^x$, $x \in \mathbb{R}$, und der Additionstheoreme erhält man sofort

$$\exp(z_1 + z_2) = \exp z_1 \exp z_2 \quad (z_1, z_2 \in \mathbb{C} \text{ beliebig}).$$

Wir werden es häufig mit Funktionen der Gestalt $x \mapsto e^{\lambda x}$ zu tun haben, wobei λ eine komplexe Zahl ist. Diese Funktion ist differenzierbar:

$$\begin{aligned} e^{\lambda x} &= e^{(\lambda_1 + i\lambda_2)x}, \lambda_1, \lambda_2 \in \mathbb{R} \\ &= e^{\lambda_1 x} e^{i\lambda_2 x} = e^{\lambda_1 x} (\cos \lambda_2 x + i \sin \lambda_2 x) \\ &= e^{\lambda_1 x} \cos \lambda_2 x + i(e^{\lambda_1 x} \sin \lambda_2 x) \\ \Rightarrow (e^{\lambda x})' &= (e^{\lambda_1 x} \cos \lambda_2 x)' + i(e^{\lambda_1 x} \sin \lambda_2 x)' \\ &= \lambda_1 e^{\lambda_1 x} \cos \lambda_2 x - \lambda_2 e^{\lambda_1 x} \sin \lambda_2 x + i(\lambda_1 e^{\lambda_1 x} \sin \lambda_2 x + \lambda_2 e^{\lambda_1 x} \cos \lambda_2 x) \\ &= (\lambda_1 + i\lambda_2)(e^{\lambda_1 x} \cos \lambda_2 x) + (\lambda_1 + i\lambda_2)e^{\lambda_1 x} (i \sin \lambda_2 x) \\ &= (\lambda_1 + i\lambda_2)e^{\lambda_1 x} (\cos \lambda_2 x + i \sin \lambda_2 x) \\ &= \lambda e^{\lambda x} \end{aligned}$$

(merke: Ergebnis ergibt sich durch formales Anwenden der Kettenregel)

Definition 2.7.14 *Unter einer linearen (inhomogenen) Differentialgleichung mit konstanten Koeffizienten versteht man eine Gleichung der Gestalt*

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y = q, \quad (\text{IH})$$

wobei $a_0, \dots, a_{n-1} \in \mathbb{C}$ gegebene Konstanten und $q : J \rightarrow \mathbb{C}$ eine gegebene Funktion sind. Unter einer Lösung von (IH) versteht man eine n -mal differenzierbare Funktion $y : J \rightarrow \mathbb{C}$, die die Bedingung (IH) in jedem Punkt von J erfüllt. Die Zahl n heißt Ordnung, die Funktion q Stör- oder Steuerfunktion der Gleichung. Falls $q \equiv 0$, spricht man von der homogenen Gleichung; in diesem Fall gilt also

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_0y = 0. \quad (\text{H})$$

In der Theorie der linearen Differentialgleichungen benutzt man häufig die Operatorschreibweise: Wenn $f : J \rightarrow \mathbb{C}$ differenzierbar ist, dann versteht

man unter Df die Ableitung f' ; ist f n -mal differenzierbar, dann ist $D^n f := \underbrace{(D \circ D \circ \dots \circ D)}_{n\text{-mal}} f = f^{(n)}$.

Wenn P ein Polynom ist, $P(z) = z^n + a_{n-1}z^{n-1} + \dots + a_0z^0$, dann versteht man unter $P(D)$ den Operator $D^n + a_{n-1}D^{n-1} + \dots + a_1D + a_0 \text{ Id}$; dieser wird auch Differentialoperator genannt. Ist y n -mal differenzierbar, dann ist also

$$P(D)f = f^{(n)} + a_{n-1}f^{(n-1)} + \dots + a_1f' + a_0f.$$

Die Gleichung (IH) schreibt sich dann als

$$P(D)y = q$$

(kompakte und sehr nützliche Schreibweise).

Bezeichnen mit $\mathcal{F}^{(n)}(J)$ ($n = 0, 1, \dots$) die Menge aller auf J definierten komplexwertigen Funktionen, die n -mal differenzierbar sind. $\mathcal{F} := \mathcal{F}^{(0)}$ sei dabei die Menge aller auf J definierten komplexwertigen Funktionen. $\mathcal{F}^{(n)}$ ist ein linearer Raum über \mathbb{C} :

$$\begin{aligned} (f_1 + f_2)(x) &:= f_1(x) + f_2(x). \\ (\lambda f)(x) &:= \lambda \cdot f(x). \end{aligned}$$

Man überprüft nun sofort

$$P(D) : \mathcal{F}^{(n)} \longrightarrow \mathcal{F} \quad (\text{grad } P \leq n)$$

ist eine lineare Abbildung (linearer Operator).

\implies Um die Gleichung $P(0)y = q$ zu lösen, müssen wir eine partikuläre Lösung von (IH) finden und die Lösungen von (H) bestimmen.

Bemerkung 2.7.15 *Wir sind maßgeblich an reellen Differentialgleichungen interessiert. Es zeigt sich, dass ihr Studium vorteilhaft auf komplexe, d.h. auf (IH) zurückgeführt werden kann.*

Wichtige Bemerkung: Wenn $P = P_1P_2 (= P_2P_1)$ ist, dann gilt auch $P(D) = P_1(D)P_2(D) = P_2(D)P_1(D)$.

Der Eindeigkeitsatz: Wie die einföhrenden Beispiele belegen, involvieren durch (IH) beschreibbare Vorgänge aus Natur und Technik in der Regel noch die Vorgabe von n Anfangswerten

$$y(x_0), y'(x_0), \dots, y^{(n-1)}(x_0) \quad (2.13)$$

an einer Stelle $x_0 \in J$.

Bewegungsvorgang mit $n = 2$:

$$\begin{array}{ll} y(t_0) & - \text{Anfangsort} \\ y'(t_0) & - \text{Anfangsgeschwindigkeit} \end{array}$$

(zum Zeitpunkt t_0).

Wir zeigen, dass eine Lösung von (IH) (natürlich im Falle ihrer Existenz) eindeutig durch die Vorgabe der n Werte (2.8) festgelegt ist. Der Beweis stützt sich auf

Hilfssatz 2.7.16 *Sei $g : J \rightarrow \mathbb{C}$ eine differenzierbare Funktion. Genügt g auf J einer Ungleichung*

$$|g'| \leq C|g|$$

mit $C > 0$ und ist $g(x_0) = 0$ für ein $x_0 \in J \Rightarrow g \equiv 0$.

Beweisidee:

- a) Wir betrachten zunächst den Spezialfall $g \geq 0$. Trick: Föhren die Funktion $f := ge^{Cx}$ ein. Diese ist wegen $f'e^{-Cx}(g' - Cg) \leq 0$ monoton fallend, hat in x_0 eine Nullstelle und ist somit ≤ 0 für $x \geq x_0$. Zusammen mit $f \geq 0$ ergibt sich somit $f(x) = 0$ für $x \geq x_0$, d.h. $g(x) = 0$ für $x \geq x_0$. Mittels $f = ge^{Cx}$ zeigt man analog $g(x) = 0$ für $x < x_0$.

- b) Den allgemeinen Fall föhren wir auf a) zurück. Setzen $G = g\bar{g} (\geq 0)$

$$\Rightarrow |G'| = |g'\bar{g} + g\bar{g}'| \leq 2|g'\bar{g}| \leq 2CG.$$

Eindeutigkeitssatz. Seien $y_1, y_2 : J \rightarrow \mathbb{C}$ Lösungen von (IH) mit gleichen Anfangswerten an einer Stelle $x_0 \in J$:

$$y_1^{(k)}(x_0) = y_2^{(k)}(x_0), \quad k = 0, \dots, n-1.$$

Dann gilt $y_1 = y_2$ auf ganz J .

Beweis. Wir setzen $y := y_2 - y_1$ und wenden den Hilfssatz auf $g := \sum_{k=0}^{n-1} |y^{(k)}|^2$ an. Die Funktion g ist differenzierbar, da y n -mal differenzierbar ist, und es gilt

$$g' = \left(\sum_{k=0}^{n-2} y^{(k)} \overline{y^{(k+1)}} + y^{(k+1)} \overline{y^{(k)}} \right) + y^{(n-1)} \overline{y^{(n)}} + y^{(n)} \overline{y^{(n-1)}}.$$

Wegen $|y^{(k)}| \leq \sqrt{g}$ ($k = 0, \dots, n-1$) und

$$y^{(n)} = -a_{n-1}y^{(n-1)} - \dots - a_0y \quad (\text{beachte, dass } y \text{ Lösung von } (H) \text{ ist})$$

folgt nun

$$|g'| \leq C|g|, \quad \text{wobei } C := 2 \left(n-1 + \sum_{k=0}^{n-1} |a_k| \right).$$

Weiter ist

$$g(x_0) = 0 \xrightarrow{\text{HS}} g \equiv 0 \Rightarrow y \equiv 0, \quad \text{d.h. } y_1 = y_2 \quad \square$$

Folgerung 2.7.17

- a) Der \mathbb{C} -Vektorraum L aller komplexen Lösungen der homogenen Gleichung (H) n -ter Ordnung hat die Dimension $\leq n$.
- b) Sind y_1, \dots, y_n linear unabhängige Lösungen von (H), so bilden diese eine Basis von L (eine Basis von L heißt auch Fundamentalsystem für die homogene Gleichung).

Beweis.

- a) Sei $x_0 \in K$ beliebig. Wir betrachten die Abbildung $A : L \rightarrow \mathbb{C}^n$, $y \mapsto (y(x_0), \dots, y^{(n-1)}(x_0))$. Diese ist linear und injektiv. (Eindeutigkeitssatz ist zu berücksichtigen) $\Rightarrow \dim L \leq n$ (Surjektivität der Abbildung ist noch nicht bewiesen!)
- b) Folgt aus a). □

Ein Fundamentalsystem für die homogene Gleichung

Wir betrachten die homogene Gleichung

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y = 0.$$

Wenn y Lösung dieser Gleichung ist, dann ist $y \in \mathcal{F}^\infty(J) : y^{(n)} = -a_{n-1}y^{(n-1)} - \dots - a_0y$. $\Rightarrow y$ besitzt Ableitungen beliebiger Ordnung.

Mit dem Polynom $P(z) = z^n + a_{n-1}z^{n-1} + \dots + a_1z + a_0$ (das auch charakteristisches Polynom der Differentialgleichung genannt wird), können wir diese auch als

$$P(D)y = 0$$

schreiben. Wir wählen nun $y = e^{\lambda x}$, wobei λ eine noch näher zu bestimmende komplexe Zahl ist und berechnen

$$P(D)e^{\lambda x}.$$

Unter Berücksichtigung von $De^{\lambda x} = \lambda e^{\lambda x}$ erhalten wir sofort

$$P(D)e^{\lambda x} = (\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0)e^{\lambda x} = P(\lambda)e^{\lambda x}.$$

Die rechte Seite ist genau dann gleich Null, wenn λ eine (komplexe) Nullstelle des Polynoms P ist. Hat P n verschiedene Nullstellen $\lambda_1, \dots, \lambda_n$, so hat man $e^{\lambda_1 x}, \dots, e^{\lambda_n x}$ verschiedene Lösungen von (H).

Falls mehrfache Nullstellen auftreten, ist ihre Anzahl kleiner als n . Trotzdem gibt es auch in diesem Fall n linear unabhängige Lösungen. Um eine Vorstellung zu erhalten, was wir zu erwarten haben, betrachten wir

$$(D - \lambda \text{ Id})^2 e^{\lambda x} \text{ mit } \lambda \in \mathbb{C} \text{ fixiert.}$$

Das charakteristische Polynom $(z - \lambda)^2$ besitzt die zweifache Nullstelle λ . Nun ist bereits $(D - \lambda \text{Id})e^{\lambda x}$ gleich Null. Betrachten $(D - \lambda \text{Id})^2 \lambda e^{\lambda x} = (D - \lambda \text{Id})(D - \lambda \text{Id})xe^{\lambda x} = (D - \lambda \text{Id})e^{\lambda x} = 0 \Rightarrow e^{\lambda x}xe^{\lambda x}$ sind Lösungen von $(D - \lambda \text{Id})^2 = 0$. Die benutzte Idee zeigt, dass folgender Sachverhalt richtig ist: Ist f n -mal differenzierbar, dann gilt $(D - \lambda \text{Id})^n f e^{\lambda x} = f^{(n)} e^{\lambda x}$. (Beweis mittels Induktion.) Spezifiziert man f zu x^k , $k < m \Rightarrow (D - \lambda \text{Id})^m x^k e^{\lambda x} = 0$.

Satz 2.7.18 (*Fundamentalsystem*)

Sei P das charakteristische Polynom der Gleichung (H), seien $\lambda_1, \dots, \lambda_r$ die verschiedenen Nullstellen des Polynoms P und k_1, \dots, k_r deren jeweiligen Vielfachheiten. Dann hat (H) folgende linear unabhängige Lösungen:

$$\begin{aligned} \text{zu } \lambda_1 \text{ die } k_1 \text{ Lösungen: } & e^{\lambda_1 x}, xe^{\lambda_1 x}, \dots, x^{k_1-1}e^{\lambda_1 x} \\ \text{zu } \lambda_2 \text{ die } k_2 \text{ Lösungen: } & e^{\lambda_2 x}, xe^{\lambda_2 x}, \dots, x^{k_2-1}e^{\lambda_2 x} \\ & (*) \\ \text{zu } \lambda_r \text{ die } k_r \text{ Lösungen: } & e^{\lambda_r x}, xe^{\lambda_r x}, \dots, x^{k_r-1}e^{\lambda_r x}. \end{aligned}$$

Jede Lösung von (H) ist Linearkombination dieser n Lösungen.

Beweisidee: $P(D) = Q(d)(P - \lambda_j \text{Id})^{k_j} \Rightarrow P(D)x^m e^{\lambda_j x} = 0$ für $m \leq k_j - 1$. Lineare Unabhängigkeit werden wir nicht beweisen (aus Zeitgründen). Für $n = 2$ ist diese allerdings ganz einfach:

Zwei verschiedene Nullstellen: λ_1, λ_2 : $e^{\lambda_1 x}, e^{\lambda_2 x}$ sind linear unabhängige Lösungen von (H): $\mu_1 e^{\lambda_1 x} + \mu_2 e^{\lambda_2 x} = 0 \Rightarrow \mu_1 e^{(\lambda_1 - \lambda_2)x} + \mu_2 = 0 \Rightarrow \mu_1 = 0$ da anderenfalls $e^{(\lambda_1 - \lambda_2)x}$ konstant sein müsste $\frac{1}{2}$.

Doppelte Nullstelle λ_1 : $e^{\lambda_1 x}, xe^{\lambda_1 x}$ sind Lösungen von (H) (wurde gezeigt). Lineare Unabhängigkeit: $\mu_1 x e^{\lambda_1 x} + \mu_2 e^{\lambda_1 x} = 0 \Leftrightarrow \mu_1 x + \mu_2 = 0 \Rightarrow \mu_1 = \mu_2 = 0$.

Folgerung 2.7.19 Die Gleichung (H) besitzt zu beliebig vorgegebenen Anfangswerten (a_0, \dots, a_{n-1}) bei x_0 genau eine Lösung y mit $y^{(k)}(x_0) = a_k$, $k = 0, \dots, n - 1$.

Beweis. Sei $A : L \rightarrow \mathbb{C}^n$ die im vorigen Punkt definierte lineare Abbildung. Weil A injektiv ist und $\dim L = n \Rightarrow \text{im } A$ ist n dimensional (Dimensionssatz) $\Rightarrow \text{im } A = \mathbb{C}^n$.

Beispiel 2.7.20 $y^{(4)} - 3y^{(3)} + 3y^{(2)} - y' = 0$

$$\begin{array}{ll}
\text{charakteristisches Polynom :} & z^4 - 3z^3 + 3z^2 - z \\
\text{einfache Nullstelle:} & z = 0 \\
\text{dreifache Nullstelle:} & z = 1 \quad ((z-1)^3 = z^3 - 3z^2 + 3z - 1) \\
\text{Fundamentalsystem:} & e^0 = 1, e^x, xe^x, x^2e^x
\end{array}$$

Reelle Lösungen: Die Koeffizienten a_0, a_1, \dots, a_{n-1} der Differentialgleichung (H) seien jetzt reell. Man interessiert sich dann häufig nur für reelle Lösungen. Ungeachtet dessen geht man zu ihrer Berechnung zweckmäßigerweise durch das Komplexe. Der folgende Hilfssatz liefert den Schlüssel:

Hilfssatz 2.7.21 Sowohl der Realteil u wie auch der Imaginärteil v einer komplexen Lösung $z = u + iv$ der Gleichung (H) sind (reelle) Lösungen von (H).

Beweis. Wegen $z^{(k)} = u^{(k)} + iv^{(k)}$ liefert $P(D)z = 0$

$$(u^{(n)} + a_{n-1}u^{(n-1)} + \dots + a_0u) + i(v^{(n)} + a_{n-1}v^{(n-1)} + \dots + a_0v) = 0.$$

Die Summen in den Klammern sind reelle Funktionen und müssen somit gleich Null sein. \square

Eine reelle, k -fache Nullstelle des charakterischen Polynoms liefert die k reellen Lösungen $e^{\lambda x}, e^{\lambda x}, \dots, x^{k-1}e^{\lambda x}$. Die nichtreellen Nullstellen von P treten in Paaren konjugiert komplexer auf. Es seien $\lambda = \alpha + i\beta, \bar{\lambda} = \alpha - i\beta$ ($\beta \neq 0$) ein solches Paar und k sei die Vielfachheit von λ (und von $\bar{\lambda}$). Das Paar $(\lambda, \bar{\lambda})$ liefert die $2k$ komplexen Lösungen

$$\begin{array}{l}
e^{(\alpha+i\beta)x}, xe^{(\alpha+i\beta)x}, \dots, x^{k-1}e^{(\alpha+i\beta)x} \\
e^{(\alpha-i\beta)x}, xe^{(\alpha-i\beta)x}, \dots, x^{k-1}e^{(\alpha-i\beta)x} .
\end{array}$$

Der Hilfssatz zeigt nun, dass

$$\begin{array}{l}
e^{\alpha x} \cos \beta x, xe^{\alpha x} \cos \beta x, \dots, x^{k-1}e^{\alpha x} \cos \beta x \\
e^{\alpha x} \sin \beta x, xe^{\alpha x} \sin \beta x, \dots, x^{k-1}e^{\alpha x} \sin \beta x
\end{array} \tag{2.14}$$

$2k$ reelle Lösungen von (H) sind.

Nach diesem Muster erhält man n reelle Lösungen von (H). Diese sind linear unabhängig über \mathbb{R} , da sich die Funktionen (2.8) als Linearkombination der

ursprünglichen komplexen Lösungen ausdrücken lassen. Wie im komplexen Fall zeigt man nun

$$\dim L_{\mathbb{R}} = n \quad (L_{\mathbb{R}} = \text{lin. Raum aller reellen Lösungen von (H) über } \mathbb{R}).$$

Beispiel 2.7.22

$1^0.$	$y'' - 2y' + 5y = 0$
<i>Charakteristisches Polynom:</i>	$\lambda^2 - 2\lambda + 5$
<i>Nullstellen:</i>	$\lambda_1 = 1 + i, \lambda_2 = 1 - 2i$
<i>komplexes Fundamentalsystem:</i>	$e^{(1+2i)x}, e^{(1-2i)x}$
<i>reelles Fundamentalsystem:</i>	$e^x \cos 2x, e^x \sin 2x$
$2^0.$	$y^{(4)} + 2y^{(2)} + y = 0$
<i>Charakteristisches Polynom:</i>	$\lambda^4 + 2\lambda^2 + 1$
<i>Nullstellen:</i>	$i \text{ zweifach}, -i \text{ zweifach}$
<i>komplexes Fundamentalsystem:</i>	$e^{ix}, xe^{ix}, e^{-ix}, xe^{-ix}$
<i>reelles Fundamentalsystem:</i>	$\cos x, x \cos x, \sin x, x \sin x$

Berechnung einer partikulären Lösung bei speziellen Inhomogenitäten

Betrachten wir zunächst die inhomogene Differentialgleichung

$$y'' + ay' + by = Ae^{i\omega t}, \quad A, \omega \in \mathbb{R}.$$

Ansatz: $y_p(t) = Be^{i\omega t}$.

Setzen ein und erhalten

$$\underbrace{((i\omega)^2 + a(i\omega) + b)}_{P(i\omega)} Be^{i\omega t} = Ae^{i\omega t}.$$

Wenn $P(i\omega) \neq 0 \Rightarrow B = \frac{A}{P(i\omega)}$ und $\frac{A}{P(i\omega)}e^{i\omega t}$ ist eine partikuläre Lösung. Ist aber $i\omega$ eine Nullstelle von P , dann ist dieser Ansatz nicht verwendungsfähig.

Versuchen den Ansatz

$$U_p(t) = Bte^{i\omega t}.$$

Einsetzen: $P(i\omega)Bte^{i\omega t} + (2i\omega + a)Be^{i\omega t} = Ae^{i\omega t}$

$$\implies B = \frac{A}{2i\omega + a}$$

und $\frac{1}{2i\omega + a}te^{i\omega t}$ ist partielle Lösung obiger Differentialgleichung.

Diese Vorgehensweise lässt sich nun verallgemeinern (ohne Beweis). Betrachten die inhomogene Differentialgleichung (IH), d.h.

$$P(D)y = q \tag{*}$$

mit $q(x) = (b_0 + b_1x + \dots + b_mx^m)e^{\mu x}$ ($b_0, \dots, b_m, \mu \in \mathbb{C}$).

Betrachten von zwei Fällen:

a) $P(\mu) \neq 0$; dann ist eine Funktion der Gestalt

$$y = (c_0 + c_1x + \dots + c_mx^m)e^{\mu x}$$

Lösung von (*). Für $m = 0$ ist $y(x) = \frac{b_0}{P(\mu)}e^{\mu x}$ Lösung von (*).

b) Ist ω eine k -fache Nullstelle von P , dann besitzt $P(D)y = q$ eine Lösung der Gestalt

$$y(x) = (c_0 + c_1x + \dots + c_mx^m)x^k e^{\mu x},$$

und für $m = 0$ die Lösung

$$y(x) = \frac{b_0}{P^{(k)}(\mu)}x^k e^{\mu x}.$$

Die Koeffizienten c_0, \dots, c_m werden durch Koeffizientenvergleich ermittelt.

Diese Vorgehensweise wird benutzt, um auch reelle Differentialgleichungen zu lösen: Seien die Koeffizienten des charakteristischen Polynoms P reell und die Störfunktion q der Realteil einer komplexen Funktion $q + ir$. Ist nun z eine komplexe Lösung von $P(D)z = q + ir$, so ist der Realteil von z eine reelle Lösung von $P(D)y = q$.

Die allgemeine Lösung von $P(D)z = q + ir$ erhält man, indem man zu einer partikulären Lösung die Lösungen der homogenen Gleichung hinzufügt (siehe auch Punkt 2.8). Ist P reelles Polynom, so erhält man durch das Heraustrennen des Realteils aus der allgemeinen Lösung die allgemeine Lösung der reellen Gleichung $P(D)y = q$. Die in der allgemeinen Lösung auftretenden Konstanten (Koeffizienten in der Linearkombination der linear unabhängigen Lösungen der homogenen Gleichung) werden durch die Vorgabe von n Anfangswerten festgelegt oder "angepasst".

Wir diskutieren nun das Lösungsverhalten der Differentialgleichung

$$\ddot{u} + a \dot{u} + bu = 0$$

mit $a, b \in \mathbb{R}$.

Bestimmen die Wurzeln des charakteristischen Polynoms $z^2 + az + b$:

$$\begin{aligned} z^2 + az + b &= z^2 + \frac{a}{2}2z + b - \frac{a^2}{4} + \frac{a^2}{4} \\ &= \left(z + \frac{a}{2}\right)^2 + b - \frac{a^2}{4}. \end{aligned}$$

Es ist $z^2 + az + b = 0 \Leftrightarrow (z + \frac{a}{2})^2 = \frac{a^2}{4} - b = \frac{1}{4}(a^2 - 4b)$; $\Delta := a^2 - 4b$ wird auch Diskriminante der quadratischen Gleichung $z^2 + az + b = 0$ genannt.

Für die Wurzeln λ_1, λ_2 gilt nun

$$\lambda_{1/2} = \begin{cases} \frac{(-a \pm \sqrt{\Delta})}{2}, & \text{falls } \Delta > 0 \quad (\text{zwei verschiedene reelle Wurzeln}) \\ -\frac{a}{2}, & \text{falls } \Delta = 0 \quad (\text{eine reelle Doppelwurzel liegt vor}) \\ \frac{(-a \pm i\sqrt{-\Delta})}{2}, & \text{falls } \Delta < 0 \quad (\text{zwei konjugiert komplexe Wurzeln liegen vor}). \end{cases}$$

Diskutieren diese 3 Fälle:

1. $\Delta > 0$: komplexes Fundamentalsystem: $e^{\lambda_1 t}, e^{\lambda_2 t}$
reelles Fundamentalsystem: $e^{\lambda_1 t}, e^{\lambda_2 t}$
2. $\Delta = 0$: komplexes Fundamentalsystem: $e^{\lambda_1 t}, te^{\lambda_1 t}$
reelles Fundamentalsystem: $e^{\lambda_1 t}, te^{\lambda_1 t}$
3. $\Delta < 0$: komplexes Fundamentalsystem: $e^{\lambda_1 t}, e^{\overline{\lambda_1} t}$ ($\lambda_1 = \alpha + i\beta, \beta \neq 0$)
reelles Fundamentalsystem: $e^{\alpha t} \cos \beta t, e^{\alpha t} \sin \beta t$
($\alpha = -\frac{a}{2}, \beta = \frac{\sqrt{-\Delta}}{2}$).

Anwendungen

Freie Schwingungen eines Massenpunktes

Wir betrachten unser eingängliches Beispiel mit verschwindender Zwangskraft:

$$\ddot{x} + \frac{r}{m}\dot{x} + \frac{k}{m}x = 0 \quad (r \geq 0 \text{ konstant})$$

Mit $\rho := \frac{1}{2}\frac{r}{m}$, $\omega_0 := \sqrt{\frac{k}{m}}$ ergibt sich

$$\ddot{x} + \omega_0^2 x = 0 \quad \text{ungedämpfter harmonischer Oszillator} \quad (\rho = 0) \quad (2.15)$$

$$\text{bzw. } \ddot{x} + 2\rho\dot{x} + \omega_0^2 x = 0 \quad \text{gedämpfter harmonischer Oszillator} \quad (\rho > 0) \quad (2.16)$$

Ungedämpfter Oszillator

Alle Lösungen von (2.15) werden gegeben durch

$$x(t) = C_1 \cos \omega_0 t + C_2 \sin \omega_0 t \quad (2.17)$$

(Fall 3 mit $a = 0$, $b = \omega_0^2$, also $\Delta = -4\omega_0^2 < 0$, $\alpha = 0$, $\beta = \omega_0$)

mit den freien Konstanten C_1, C_2 ; diese machen es möglich, die Lösung an die Anfangsbedingung $x(t_0) = x_0$, $\dot{x}(t_0) = x_1$ anzupassen \Leftrightarrow die Konstanten werden festgelegt.

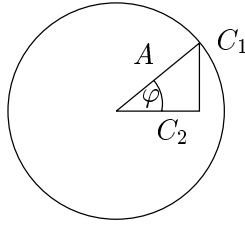
$C_1 = C_2 = 0 \Rightarrow M$ verharrt bewegungslos im Nullpunkt.

Gilt $C_1^2 + C_2^2 > 0$, so schreiben wir (2.17) in der Form

$$x(t) = A \left(\frac{C_1}{A} \cos \omega_0 t + \frac{C_2}{A} \sin \omega_0 t \right)$$

mit $A = \sqrt{C_1^2 + C_2^2}$. Wegen $\left(\frac{C_1}{A}\right)^2 + \left(\frac{C_2}{A}\right)^2 = 1$ gibt es genau ein $\varphi \in [0, 2\pi)$ mit

$$\frac{C_1}{A} = \sin \varphi, \quad \frac{C_2}{A} = \cos \varphi.$$



Also gilt $x(t) = A(\sin \varphi \cos \omega_0 t + \cos \varphi \sin \omega_0 t) = A \sin(\omega_0 t + \varphi)$. $\Rightarrow M$ schwingt zwischen den Maximalausschlägen $-A$ und A periodisch hin und her. Die Größe φ heißt auch Phasenkonstante: Wegen $x(0) = A \sin \varphi$ legt sie die Lage von M zur Zeit $t = 0$ fest.

Gedämpfter Oszillator

$$a = 2\rho, b = \omega_0^2 \Rightarrow \Delta = 4(\rho^2 - \omega_0^2)$$

Es gibt drei Fälle:

I. $\rho > \omega_0$, d. h. eine „starke Dämpfung“ liegt vor $\Rightarrow \Delta > 0$ und somit

$$x(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t} \text{ mit } \lambda_{1,2} := \rho \pm \sqrt{\rho^2 - \omega_0^2} < 0$$

Es gilt $\lim_{t \rightarrow \infty} x(t) = 0$ (unabhängig von den Anfangsbedingungen).
Geschwindigkeit: $\dot{x}(t) = C_1 \lambda_1 e^{\lambda_1 t} + C_2 \lambda_2 e^{\lambda_2 t}$; im nichttrivialen Fall (o.E.d.A. $C_1 \neq 0$) wird sie genau dann gleich Null sein, wenn $e^{(\lambda_1 - \lambda_2)t} = -\frac{C_2 \lambda_2}{C_1 \lambda_1}$ gilt. Das kann für höchstens einen Wert t vorkommen (Funktion exp ist injektiv!).

\Rightarrow aperiodische Bewegung, bei der M höchstens einmal seine anfängliche Bewegungsrichtung umkehrt, im übrigen aber in den Nullpunkt hineinkriecht.



II. $\rho = \omega_0$ (immernoch starke Dämpfung)

$$\Delta = 0 \Rightarrow x(t) = (C_1 + C_2 t) e^{-\rho t}$$

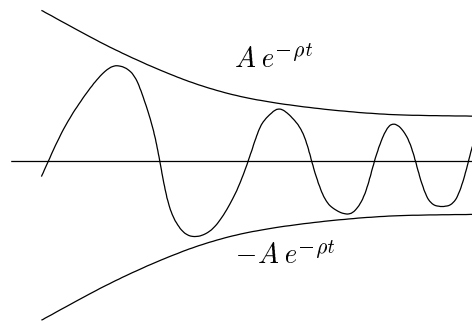
Hier liegt der gleiche Bewegungstyp wie unter I. vor.

III. Das ist der interessanteste Fall.

$\rho < \omega_0$ (schwache Dämpfung)

$$\Rightarrow \Delta < 0 \text{ und } x(t) = e^{-\rho t}(C_1 \cos \omega_1 t + C_2 \sin \omega_1 t) \text{ mit } \omega_1 := \sqrt{\omega_0^2 - \rho^2},$$

also $x(t) = Ae^{-\rho t} \sin(\omega_1 t + \varphi)$ im nichttrivialen Fall (selbe Verfahrensweise wie im Falle des ungedämpften Oszillators). Wegen $\rho > 0$ gilt auch diesmal $x(t) \rightarrow 0$ für $t \rightarrow \infty$. Der Sinusterm in der Lösung sorgt allerdings dafür, daß M unaufhörlich um den Nullpunkt herumpendelt.



Periodische Zwangskräfte

Wir kehren noch einmal zu unserem ursprünglichen Beispiel zurück. Wir hatten für dieses folgende Differentialgleichung ermittelt:

$$\ddot{x} + \frac{r}{m}\dot{x} + \frac{k}{m}x = \frac{1}{m}K \quad (r \geq 0)$$

Wir schreiben nun diese Gleichung in der Gestalt

$$\ddot{x} + 2\rho\dot{x} + \omega_0^2 x = f \text{ mit } \rho := \frac{r}{2m}, \omega_0 = \sqrt{\frac{k}{m}}, f = \frac{1}{m}K.$$

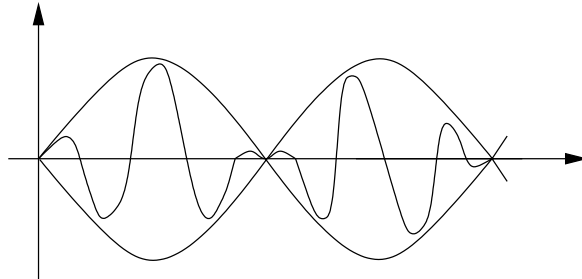
Wir betrachten nur den Fall schwacher oder sogar verschwindender Dämpfung: $0 \leq \rho < \omega_0$. Ferner sei uns eine „Kosinuserregung“, also die Differentialgleichung $\ddot{x} + 2\rho\dot{x} + \omega_0^2 x = a \cos \omega t$ ($a, \omega > 0$) gegeben.

a.) $\rho = 0$

Allgemeine Lösung ($\omega \neq \omega_0$): $x(t) = C_1 \cos \omega_0 t + C_2 \sin \omega_0 t + \frac{a}{\omega_0^2 - \omega^2} \cos \omega t$

da $p(i\omega) = \omega_0^2 - \omega^2$ gilt. Ruht der Massenpunkt zur Zeit $t = 0$, d. h. ist $x(0) = \dot{x}(0) = 0$, so ergibt sich für ihn durch Anpassung der Konstanten ($C_1 = -\frac{a}{\omega_0^2 - \omega^2}$, $C_2 = 0$)

$$\begin{aligned} x(t) &= \frac{a}{\omega_0^2 - \omega^2} (\cos \omega t - \cos \omega_0 t). \\ \Rightarrow x(t) &= \frac{2a}{\omega_0^2 - \omega^2} \sin \frac{\omega_0 - \omega}{2} t \sin \frac{\omega_0 + \omega}{2} t \end{aligned}$$



Der Massenpunkt vollführt eine Sinusschwingung mit der zeitlich veränderlichen und zwar auch ihrerseits „sinusförmigen“ Amplitude $A(\omega t) := \frac{2a}{\omega_0^2 - \omega^2} \sin \frac{\omega_0 - \omega}{2} t$.

Die allgemeine Lösung im Falle $\omega = \omega_0$ ist

$$x(t) = C_1 \cos \omega_0 t + C_2 \sin \omega_0 t + \frac{a}{2\omega_0} t \sin \omega_0 t.$$

Für die Resonanz besagt die Lösung, daß die Feder mit wachsendem t immer stärker ausschlägt; in der Praxis bedeutet dies, daß die Feder zu Bruch geht.

b.) $\rho > 0$ (gedämpfter harmonischer Oszillator)

Die allgemeine Theorie liefert uns $x_n(t) = e^{-\rho t} (C_1 \cos \omega_1 t + C_2 \sin \omega_1 t)$ mit $\omega_1 = \sqrt{\omega_0^2 - \rho^2}$. Eine partikuläre Lösung der inhomogenen Gleichung läßt sich am bequemsten finden, indem man zunächst die korrespondierende komplexe Differentialgleichung $\ddot{x} + 2\rho\dot{x} + \omega_0^2 x = ae^{i\omega t}$ durch den Ansatz $z_p(t) := Ae^{i\omega t}$ löst und dann $x_p(t) := \operatorname{Re} z_p(t)$ als Lösung der ursprünglichen Gleichung nimmt.

Wegen $p(i\omega) = \omega_0^2 - \omega^2 + 2\rho\omega i$ erhält man

$$A = \frac{a}{\omega_0^2 - \omega^2 - 2\rho\omega i} \quad \text{und}$$

$$z_p(t) = \frac{a}{\omega_0^2 - \omega^2 - 2\rho\omega i} e^{i\omega t} = a \frac{\omega_0^2 - \omega^2 - 2\rho\omega i}{(\omega_0^2 - \omega^2)^2 + 4\rho^2\omega^2} (\cos \omega t + i \sin \omega t), \text{ also}$$

$$x_p(t) = \underbrace{\frac{(\omega_0^2 - \omega^2)a}{(\omega_0^2 - \omega^2)^2 + 4\rho^2\omega^2}}_{\parallel C_1} \cos \omega t + \underbrace{\frac{2\rho\omega a}{(\omega_0^2 - \omega^2)^2 + 4\rho^2\omega^2}}_{\parallel C_2} \sin \omega t$$

Setzen $A = \sqrt{C_1^2 + C_2^2} = \frac{a}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\rho^2\omega^2}}$, wegen $\left(\frac{C_1}{A}\right)^2 + \left(\frac{C_2}{A}\right)^2 = 1$

gibt es wieder ein eindeutig bestimmtes $\varphi \in [0, 2\pi)$ mit $\sin \varphi = \frac{C_2}{A}$, $\cos \varphi = \frac{C_1}{A} \Rightarrow$

$$x_p(t) = \frac{a}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\rho^2\omega^2}} \sin(\omega t + \varphi).$$

Somit genügt der kosinuserregte und schwach gedämpfte harmonische Oszillator dem Bewegungsgesetz

$$x(t) = e^{-\rho t} (C_1 \cos \omega_1 t + C_2 \sin \omega_1 t) + \frac{a}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\rho^2\omega^2}} \sin(\omega t + \varphi).$$

Da der erste Summand für große t („nach einem Einschwingvorgang“) vernachlässigbar klein ist, wird die Bewegung im wesentlichen durch die partikuläre Lösung beschrieben, also durch eine Sinusschwingung mit der Erregerfrequenz ω und einer von ihr abhängenden **konstanten** Amplitude. Das Bewegungsgesetz des sinuserregten und schwachgedämpften Oszillators wird analog ermittelt. Das Resonanzphänomen kann auch im gedämpften Fall beobachtet werden; es ist theoretisch nicht ganz so kraß ausgeprägt, praktisch aber kaum weniger verheerend. Wir betrachten nun wieder den kosinuserregten harmonischen Oszillator, d. h.

$$\ddot{x} + 2\rho\dot{x} + \omega_0^2 x = a \cos \omega t \quad (a, \omega > 0; 0 < \rho < \omega_0)$$

Nach einem Einschwingvorgang stellt sich eine Sinusschwingung mit der Amplitude

$$A(\omega) := \frac{a}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\rho^2\omega^2}}$$

ein. Wir schauen uns die Abhängigkeit der Amplitude $A(\omega)$ von ω an: Dazu betrachten wir den Radikanden $f(\omega) := (\omega_0^2 - \omega^2)^2 + 4\rho^2\omega^2$.

$$\begin{aligned}\text{Ableitung: } f'(\omega) &= -2 \cdot 2(\omega_0^2 - \omega^2)\omega + 8\rho^2\omega \\ &= -4\omega_0^2\omega + 4\omega^3 + 8\rho^2\omega\end{aligned}$$

Für $\rho \geq \frac{\omega_0}{\sqrt{2}}$ folgt $f'(\omega) \geq 4\omega^3 \Rightarrow f(\omega)$ wächst streng monoton und unbeschränkt (gegen ∞). Für $\rho < \frac{\omega_0}{\sqrt{2}}$ gilt jedoch: Die Funktion f fällt zunächst streng bis zur Stelle $\omega_R := \sqrt{\omega_0^2 - 2\rho^2}$, und erst dann beginnt sie streng und unbeschränkt zu wachsen (lokales Minimum liegt vor: $f'(\omega) = -4\omega_0^2\omega + 4\omega^3 + 8\rho^2\omega = 0 \Rightarrow \omega^2 = -2\rho^2 + \omega_0^2$; wegen der Bedingung $\rho < \frac{\omega_0}{\sqrt{2}}$ ist die rechte Seite größer Null $\Rightarrow \omega_R = \sqrt{\omega_0^2 - 2\rho^2}$ usw.). Somit wächst die Amplitude $A(\omega)$ streng auf $(0, \omega_R)$, erreicht ihren Maximalwert

$$A_{max} = \frac{a}{2\rho\sqrt{\omega_R^2 - \rho^2}}$$

an der Stelle $\omega = \omega_R$ und strebt dann monoton fallend gegen Null.

Für den kosinuserregten harmonische Oszillator stellt sich das Resonanzphänomen wie folgt dar: Im Falle $\rho < \frac{\omega_0}{\sqrt{2}}$ - und nur in ihm - nimmt die Amplitude der erzwungenen Schwingung stark, aber nicht unbeschränkt zu, wenn die Erregerfrequenz ω in die Nähe der Eigenfrequenz ω_0 kommt.

Sei $A : C^2(\mathbb{R}) \rightarrow C(\mathbb{R})$ der lineare Differentialoperator $u \mapsto \ddot{u} + a\dot{u} + bu$ ($a, b \in \mathbb{R}$). Wir zeigen $\dim \ker A = 2$. (Die Methode, die beschrieben wird, bleibt auch im Fall $a, b \in \mathbb{C}$ gültig.) Es genügt, sich auf den komplexwertigen Fall zu beschränken. Es ist nun günstig, den Operator $u \mapsto \dot{u}$ mit D zu bezeichnen.
 \Rightarrow

$$\begin{aligned}Au &= D^2u + aDu + bu, \text{ d. h.} \\ A &= D^2 + aD + bI,\end{aligned}$$

wobei $I : C^2(\mathbb{R}) \rightarrow C^2(\mathbb{R})$ der identische Operator $u \mapsto u$ ist.

Das charakteristische Polynom p besitzt zwei Wurzeln: λ_1, λ_2 . Statt A schreibt man häufig $(p(z) = (z - \lambda_1)(z - \lambda_2))$

$$A = p(D) = (D - \lambda_1 I)(D - \lambda_2 I).$$

Offenbar gilt nun im Falle $\boxed{\lambda_1 \neq \lambda_2}$

$$I = \frac{1}{\lambda_1 - \lambda_2} ((\lambda_1 I - D) + (D - \lambda_2 I)).$$

Wir setzen nun

$$A_1 := \frac{1}{\lambda_1 - \lambda_2} (\lambda_1 I - D)$$

$$A_2 := \frac{1}{\lambda_1 - \lambda_2} (D - \lambda_2 I)$$

Offenbar gilt nun für jedes $x \in C^2(\mathbb{R})$ wegen $I = A_1 + A_2$

$$x = Ix = A_1 x + A_2 x,$$

also $x = x_1 + x_2$ mit $x_1 = A_1 x$ und $x_2 = A_2 x$. Sei nun $x \in \ker A$, d. h. $p(D)x = 0$. Für die entsprechenden x_1 und x_2 erhalten wir

$$\begin{aligned} (D - \lambda_1 I)x_2 &= \frac{1}{\lambda_1 - \lambda_2} (D - \lambda_1 I)(D - \lambda_2 I)x = 0 \\ (D - \lambda_2 I)x_1 &= \frac{1}{\lambda_1 - \lambda_2} (D - \lambda_2 I)(\lambda_1 I - D)x = 0 \\ (D - \lambda_2 I)(\lambda_1 I - D) &= (\lambda_1 I - D)(D - \lambda_2 I)! \end{aligned}$$

Umgekehrt ist jede Lösung von $(D - \lambda_1 I)y = 0$ und $(D - \lambda_2 I)y = 0$ auch eine Lösung von $p(D)y = 0$. Somit gilt: $x \in \ker p(D) \Leftrightarrow$ die entsprechenden x_1 und x_2 gehören zu

$\ker(D - \lambda_1 I) = \ker(\lambda_1 I - D)$ und entsprechend zu $\ker(D - \lambda_2 I)$ (und $x = x_1 + x_2$). Somit haben wir die Aufgabe auf das Beschreiben des Kernes von Operatoren der Gestalt $D - \lambda I$ ($\lambda \in \mathbb{C}$) zurückgeführt. Wir machen nun folgenden Trick: Für ε_α , $t \mapsto e^{\alpha t}$ und jede beliebige Funktion $x \in C^2(\mathbb{R})$ gilt nun

$$D(\varepsilon_\alpha x) = \varepsilon_\alpha D x + x D \varepsilon_\alpha = \varepsilon_\alpha D x + x \alpha \varepsilon_\alpha = \varepsilon_\alpha (D x + \alpha I x).$$

Wenn nun $x \in \ker(D - \lambda I)$, so wenden wir obige Überlegung auf $\varepsilon_{-\lambda} x$ an:

$$D(\varepsilon_{-\lambda} x) = \varepsilon_{-\lambda} (D - \lambda I)x = 0 \Rightarrow (\varepsilon_{-\lambda} x) \text{ ist eine Konstante } C, \text{ d. h.}$$

$$x(t) = C e^{\lambda t}. \Rightarrow \ker p(D) = \{C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}\}.$$

$\boxed{\lambda_1 = \lambda_2 =: \lambda}$: Dieser Fall ist ziemlich einfach. Sei $x \in C^2(\mathbb{R})$ beliebig. Dann gilt

$$D^2(\varepsilon_\alpha x) = \varepsilon_\alpha (D + \alpha I)^2 x.$$

Ist nun $x \in \ker p(D)$, so wenden wir obige Überlegung auf $\varepsilon_{-\lambda}x$ an:

$$D^2 \varepsilon_{-\lambda} x = \varepsilon_{-\lambda} (D - \lambda I)^2 x = 0$$

$\Rightarrow \varepsilon_{-\lambda}$ ist ein Polynom vom Grade höchstens 1, d. h. $\varepsilon_{-\lambda} x = C_1 + C_2 t \Rightarrow x(t) = (C_1 + C_2 t) e^{\lambda t}$.

2.8 Lineare Gleichungssysteme und der Gaußsche Algorithmus

Unter einem linearen Gleichungssystem aus m Gleichungen und n Unbekannten versteht man einen Ausdruck der Gestalt

$$\begin{aligned} a_{11}x_1 + \dots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n &= b_m, \end{aligned} \tag{2.18}$$

wobei a_{ij} und b_i fixierte Elemente in einem Körper K und $x_1, \dots, x_n \in K$ gesucht sind (besonders wichtig: $K = \mathbb{R}$ oder \mathbb{C}). Die Matrix $A := (a_{ij})_{i=1, j=1}^{m, n}$ heißt (Koeffizienten-) Matrix des Systems und

$$\begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix}$$

heißt erweiterte Matrix des gegebenen Gleichungssystems. Besitzt es wenigstens eine Lösung (x_1, \dots, x_n) , so heißt es lösbar, anderenfalls unlösbar.

Vermöge der Matrix A definieren wir eine lineare Abbildung von K^n in K^m (die wir wieder mit A bezeichnen), die einem Vektor (l_1, \dots, l_n) den Vektor $(c_1, \dots, c_m) \in K$ nach der Regel $c_i = \sum_{k=1}^n a_{ik} l_k$ zuordnet. (Die Linearität kann leicht gezeigt werden.) Mit anderen Worten

$$A(l_1, \dots, l_n) = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} l_1 \\ \vdots \\ l_n \end{pmatrix}$$

Bemerkung 2.8.1 Die Matrixdarstellung von A in den Standardbasen ist gerade A !

Somit ist folgendes gezeigt: Das Gleichungssystem (2.18) ist mit der Operatorgleichung

$A(x_1, \dots, x_n) = (b_1, \dots, b_m)$ identisch. Man schreibt dafür auch kurz und bündig $Ax = b$.

Wir betrachten im weiteren die möglichen drei Fälle:

$n > m$ (unterbestimmtes System)

Aus $\dim \operatorname{im} A \leq m$ folgt mit $n = \dim \ker A + \dim \operatorname{im} A$ unmittelbar $\dim \ker A > 0$. Somit besitzt die homogene Gleichung **nichttriviale** Lösungen. Ist \tilde{x} eine Lösung von $Ax = b$ und x_0 eine Lösung der homogenen Gleichung $Ax = 0$, so gilt $A(\tilde{x} + x_0) = A\tilde{x} + Ax_0 = A\tilde{x} + 0 = b$, und somit ist $\tilde{x} + x_0$ eine weitere Lösung. Es gelte nun $Ax = Ay = b \Rightarrow A(x - y) = 0$, d. h. $z = x - y \in \ker A$. Die allgemeine Lösung von $Ax = b$ ist somit von der Gestalt $x = x_0 + \tilde{x}$ und setzt sich additiv aus einer partikulären Lösung \tilde{x}_0 von $Ax = b$ und einer beliebigen Lösung des homogenen Systems $Ax = 0$ zusammen.

Vermerken: Diese Überlegung gilt für jede lineare Operatorgleichung $Ax = b$ ($A : V \rightarrow V$ ist linearer Operator).

$m > n$ (überbestimmtes System)

Aus $n = \dim \ker A + \dim \operatorname{im} A$ und $m > n$ folgt $\dim \operatorname{im} A \leq n < m$, d. h. $\operatorname{im} A \neq K^m$. Damit ist das Gleichungssystem $Ax = b$ nicht für alle rechten Seiten lösbar.

$m = n$

Es gilt die Fredholmsche Alternative, d. h. entweder besitzt $Ax = b$ für alle b genau eine Lösung und es gilt $\ker A = \{0\}$, oder jede Lösung x setzt sich wie folgt zusammen: $x = x_0 + \tilde{x}$ mit $x_0 \in \ker A$ und $A\tilde{x} = b$ (falls die Lösung x existiert).

Gauß-Elimination oder Gaußscher Algorithmus

Wir kehren zurück zum linearen Gleichungssystem $Ax = b$ mit gegebener Koeffizientenmatrix $A \in K^{(m,n)}$ ($K^{(m,n)}$ bezeichne die Menge aller $(m \times n)$ -Matrizen mit Einträgen aus K ; in der Praxis ist in der Regel $K = \mathbb{R}$ oder \mathbb{C}). Da A als linearer Operator von K^n in K^m interpretiert werden kann, hat die früher eingeführte Sprechweise Gültigkeit ($\ker A$, homogene Gleichung, partikuläre Lösung, . . .). Das im folgenden beschriebene Eliminationsverfahren nach Gauß ist das Standardverfahren zur Lösung linearer Gleichungssysteme. Das Prinzip besteht darin, mittels elementarer Umformungen die Koeffizientenmatrix A schrittweise in eine obere Dreiecksmatrix zu transformieren.

Wir notieren das Gleichungssystem $Ax = b$ in Form der erweiterten Matrix

$$(A|b) = \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array} \right)$$

Eliminationsprozeß (1. Teilschritt)

I. Ist $a_{11} \neq 0$?

- (a) Falls $a_{11} = 0$, so suche man in der ersten Spalte von A ein Element $a_{k1} \neq 0$ und vertausche die k -te Zeile mit der ersten (Vertauschung von Gleichungen; Lösungsmenge bleibt dabei unverändert).
- (b) Falls alle Elemente der ersten Spalte verschwinden, so suche man in der Restmatrix $(a_{ij})_{i=1, j=2}^{m,n}$ ein Element a_{ij} und vertausche die j -te Spalte mit der ersten (Vertauschung der Variablen; notieren!). Die Lösungsmenge bleibt dabei wieder unverändert, und man gehe nach (a).
- (c) Sind in der Restmatrix alle Elemente gleich 0, so ist der Eliminationsprozeß beendet.

II. Ist $a_{11} \neq 0$ erreicht, so heißt dieses Element das Pivotelement für den ersten Eliminationsschritt. Man führe folgende Operationen durch: Man addiere für $i = 2, 3, \dots, m$ das $\left(-\frac{a_{i1}}{a_{11}}\right)$ -fache der ersten Zeile zur Zeile Nummer i .

Als Ergebnis des ersten Eliminationsschritts erhalten wir eine erweiterte Matrix der Gestalt

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ 0 & a_{22}^{(2)} & \dots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & a_{m2}^{(2)} & \dots & a_{mn}^{(2)} & b_m^{(2)} \end{array} \right) \quad (2.19)$$

Offensichtlich ist der Eliminationsprozeß reversibel, d. h. durch entsprechende Additionen von Vielfachen der ersten Gleichung zu den folgenden erhält man aus (2.19) das Ausgangssystem zurück. Insbesondere bleibt dabei die Lösungsmenge unverändert.

2. Teilschritt

Der eben skizzierte Eliminationsprozeß wird nun auf die Restmatrix

$$\left(\begin{array}{ccc|c} a_{22}^{(2)} & \dots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & & \vdots & \vdots \\ a_{m2}^{(2)} & \dots & a_{mn}^{(2)} & b_m^{(2)} \end{array} \right)$$

angewendet. Das Verfahren bricht nach höchstens $m - 1$ Teilschritten ab, wenn kein Pivotelement mehr gefunden werden kann. In diesem Fall hat die erweiterte Matrix die Gestalt

$$\left(\begin{array}{cccccc|c} a_{11} & \dots & \dots & \dots & \dots & a_{1n} & b_1 \\ & a_{22}^{(2)} & & & & a_{2n}^{(2)} & b_2^{(2)} \\ & & \ddots & & & \vdots & \vdots \\ & & & a_{rr}^{(r)} & \dots & a_{rn}^{(r)} & b_r^{(r)} \\ & & & & & 0 & b_{r+1}^{(r)} \\ & & & & & & \vdots \\ 0 & & & & & & b_m^{(r)} \end{array} \right) \quad (2.20)$$

mit $a_{11} \cdot a_{22}^{(2)} \cdots a_{rr}^{(r)} \neq 0$. Wir erkennen nun sofort, daß (2.20) genau dann eine Lösung besitzt, wenn $b_{r+1}^{(r)} = b_{r+2}^{(r)} = \dots = b_m^{(r)} = 0$ gilt. Man setzt einfach $x_{r+1} = x_{r+2} = \dots = x_m = 0$ (Umnummerierung beachten!) und löst dann (2.20) durch „Rückwärtssubstitution“. Man erhält damit sofort

eine partikuläre Lösung. Um die Lösungen der homogenen Gleichung zu bestimmen, verfahren wir wie folgt. Wir suchen einen Vektor (x_1, \dots, x_n) in dem die Koordinaten x_{r+1}, \dots, x_n beliebig vorgegeben sind und bestimmen die Koordinaten x_1, \dots, x_r aus dem Gleichungssystem

$$\begin{aligned} a_{11}x_1 + \dots + a_{1r}x_r &= -(a_{1r+1}x_{r+1} + \dots + a_{1n}x_n) \\ a_{22}^{(2)}x_2 + \dots + a_{2r}^{(2)}x_r &= -(a_{2r+1}^{(2)}x_{r+1} + \dots + a_{2n}^{(2)}x_n) \\ &\vdots \\ a_{rr}^{(r)}x_r &= -(a_{rr+1}^{(r)}x_{r+1} + \dots + a_{rn}^{(r)}x_n) \end{aligned}$$

Dieses ist eindeutig lösbar (Rückwärtssubstitution). Jeder so bestimmte Vektor liegt im Kern von A . Umgekehrt erfüllt jeder Vektor, der im Kern liegt, gerade letzteres Gleichungssystem. Wegen der Dimensionsformel $n = \dim \operatorname{im} A + \dim \ker A$ folgt $\dim \ker A = n - r$ ($r := \dim \operatorname{im} A$).

Definition 2.8.2 Die maximale Anzahl linear unabhängiger Zeilenvektoren der Matrix A heißt der Zeilenrang von A .

Man kann sich davon überzeugen, daß beim Eliminationsprozeß (und nur für diesen muß das nachgewiesen werden) sich der Rang von A nicht ändert. Offenbar kann sich dieser nicht vergrößern, da durch den Eliminationsprozeß lineare Operationen im linearen Raum, der durch die Zeilen von A aufgespannt wird, durchgeführt werden. Wäre der Rang der transformierten Matrix streng kleiner, dann käme man sofort zu einem Widerspruch, da man von dem neuen System wieder durch lineare Operationen zu den alten gelangt.

Satz 2.8.3 Es gilt $\operatorname{rang} A = \dim \operatorname{im} A (= r)$.

Definition 2.8.4 Die maximale Anzahl linear unabhängiger Spaltenvektoren der Matrix A heißt Spaltenrang von A .

Satz 2.8.5 Der Spaltenrang von A ist gleich dem Zeilenrang von A .

Beweis. Sei $\{e_1, \dots, e_n\}$ die Standardbasis von K^n . Dann gilt gerade $Ae_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$. $\Rightarrow \operatorname{im} A$ ist auf die Vektoren Ae_j , d. h. auf die Spaltenvektoren

$\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}, j = 1, \dots, n$ aufgespannt. Dann muß aber die maximale Anzahl von linear unabhängigen Spaltenvektoren gleich $\dim \operatorname{im} A$ sein. Aus Satz 2.8.3 folgt dann die Behauptung. \square

Bemerkung 2.8.6 *Wir können somit einfach vom Rang einer Matrix sprechen, und dieser ist gleich $\dim \operatorname{im} A$.*

Satz 2.8.7 (Kronecker-Capelli)

Das lineare Gleichungssystem $Ax = b$ ist genau dann lösbar, wenn $\operatorname{rang}(A|b) = \operatorname{rang} A$ gilt.

Beweis. $Ax = b$ ist lösbar $\Leftrightarrow b \in \operatorname{im} A \Leftrightarrow b$ ist Linearkombination der Spaltenvektoren von $A \Leftrightarrow \operatorname{rang}(A|b) = \operatorname{rang} A$. \square

Beispiel 2.8.8

(1)

$$\begin{aligned}
 x_1 + x_2 + x_3 &= 3 \\
 x_1 - x_2 - x_3 &= 4 \\
 x_1 + 3x_2 + 3x_3 &= 1
 \end{aligned}$$

Erweiterte Matrix: $\left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 1 & -1 & -1 & 4 \\ 1 & 3 & 3 & 1 \end{array} \right)$

Gaußelimination: \Downarrow

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & -2 & -2 & 1 \\ 0 & 2 & 2 & -2 \end{array} \right)$$

\Downarrow

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & -2 & -2 & 1 \\ 0 & 0 & 0 & -1 \end{array} \right)$$

Das Gleichungssystem besitzt keine Lösung!

(2) Wir betrachten obiges Gleichungssystem mit neuer rechter Seite

$$\begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix} \Rightarrow$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & -2 & -2 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

Zur Berechnung einer partikulären Lösung setzt man $x_3 = 0 \Rightarrow x_2 = -\frac{1}{2}$, $x_1 = \frac{7}{2}$, d. h. $x_p = (-\frac{1}{2}, \frac{7}{2}, 0)$. Zur Berechnung der allgemeinen Lösung des homogenen Gleichungssystems setze man $x_3 = 1$ und findet als Lösung von

$$\begin{aligned} x_1 + x_2 &= -1 \\ -2x_2 &= 2 \end{aligned}$$

$x_2 = -1$, $x_1 = 0$; also gilt $(0, -1, 1) \in \ker A \Rightarrow$ Die allgemeine Lösung des linearen Gleichungssystems ist

$$x = \left(-\frac{1}{2}, \frac{7}{2}, 0 \right) + \lambda(0, -1, 1). \quad (\dim \ker A = 1!)$$

(3) Wir zeigen mittels eines Beispiels, daß das unbedachte Lösen von Gleichungssystemen mit dem Rechner zu völlig falschen Ergebnissen führen kann: Zur Lösung des linearen Gleichungssystems

$$\begin{pmatrix} 10^{-4} & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

simulieren wir eine dreistellige Rechnung (\doteq heißt gleich bis auf Rundung):

$$\left(\begin{array}{cc|c} 10^{-4} & 1 & 1 \\ 1 & 1 & 2 \end{array} \right) \Rightarrow \left(\begin{array}{cc|c} 10^{-4} & 1 & 1 \\ 0 & 1 - 10^4 & 2 - 10^4 \end{array} \right) \doteq \left(\begin{array}{cc|c} 10^{-4} & 1 & 1 \\ 0 & -10^4 & -10^4 \end{array} \right)$$

$\Rightarrow x_2 \doteq 1$, $x_1 \doteq 0$. Das Resultat ist völlig falsch:

$$\begin{aligned} x_1 &= 1 + \frac{1}{9999} = 1,0001\dots \\ x_2 &= 1 - \frac{1}{9999} = 0,999899\dots \end{aligned}$$