§ 2: Vektoren und Vektorräume

a) Vektoren in der Ebene und im Raum

Vektoren werden anschaulich dargestellt durch *Pfeile,* d.h. durch gerichtete Verbindungsstrecken zweier Punkte. Sie sind festgelegt durch die Angabe von Anfangs- und Endpunkt, aber auch beispielsweise durch die Angabe von Anfangspunkt, Länge und Richtung, wobei diese Richtung jedoch für Pfeile der Länge Null nicht definiert ist.

Pfeile dieser Art sind nützlich beispielsweise für die Darstellung von elektrischen oder magnetischen Feldern wie etwa den in Abbildung eins dargestellten: dem elektrischen Feld einer abstoßenden Punktladung und dem Magnetfeld eines stromdurchflossenen Leiters.

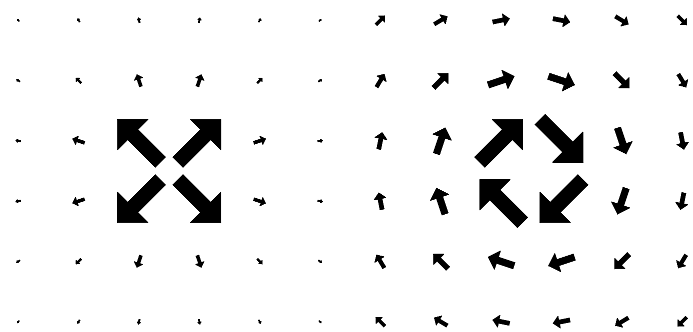

Abb. 1: Zwei elektromagnetische Felder

Zwei Pfeile lassen sich addieren, falls der Endpunkt des ersten gleich dem Anfangspunkt des zweiten ist; die Summe ist dann derjenige Pfeil, der den Anfangspunkt des ersten mit dem Endpunkt des zweiten Pfeils verbindet. Auch läßt sich ein Pfeil mit einer reellen Zahl multiplizieren, wenn wir vereinbaren, daß das Ergebnis jeder Pfeil sein soll, der denselben Anfangspunkt und dieselbe Richtung hat wie der ursprüngliche Pfeil, dessen Länge aber mit der reellen Zahl multipliziert wurde. (Eine

Multiplikation mit einer negativen Zahl soll dabei bedeuten, daß der Pfeil an seinem Anfangspunkt gespiegelt und dann mit dem Betrag der Zahl multipliziert wird.)

Sobald wir uns allerdings dafür interessieren, wie sich ein Teilchen im kombinierten Kraftfeld der Punktladung und des stromdurchflossenen Leiters bewegt, reichen Pfeile nicht mehr aus: Wir haben zwar für jeden Punkt der Ebene (außer dem Nullpunkt) einen Kraftpfeil für jedes Feld, aber natürlich müssen wir in jedem Punkt die beiden *dort* beginnenden Kraftpfeile addieren, was mit Pfeilen nicht geht.

Die Lösung dieses Problems ist wohl bekannt: Die beiden Pfeile werden gemäß dem „Parallelogramm der Kräfte“ kombiniert, d.h. der eine Pfeil wird so verschoben, daß sein Anfangspunkt gleich dem Endpunkt des anderen Pfeils ist. Wie Abbildung zwei zeigt, ist das Ergebnis unabhängig von der Reihenfolge der Summanden, d.h.

$$\vec{v} + \vec{w} = \vec{w} + \vec{v}\,.$$

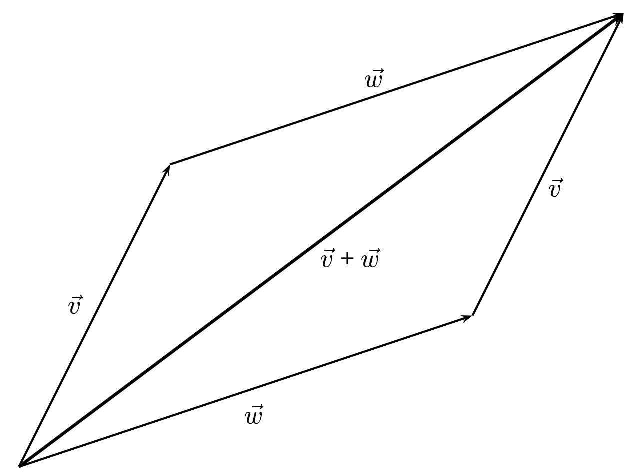

Abb. 2: Das „Parallelogramm der Kräfte“

Wir definieren daher als neuen Begriff einen *Vektor* als etwas, das zwar wie ein Pfeil eine Länge und eine Richtung haben soll, aber keinen

Anfangspunkt. Mathematisch exakt ausgedrückt ist also ein Vektor eine *Äquivalenzklasse* von Pfeilen, wobei zwei Pfeile genau dann äquivalent sind, wenn sie dieselbe Länge und (so die Länge von Null verschieden ist) dieselbe Richtung haben.

Vektoren werden in der Literatur meist durch Fraktur- oder Fettbuchstaben bezeichnet; da sich Fettdruck schlecht an der Tafel realisieren läßt und Frakturbuchstaben meist zu Hörerprotesten führen, verwenden wir hier stattdessen lateinische Buchstaben, die mit einem Pfeil überstrichen sind, also $\vec{u}, \vec{v}, \vec{w}, \dots$. Die Addition zweier Vektoren wird durch das gewöhnliche Pluszeichen ausgedrückt, wir schreiben also $\vec{v}+\vec{w}$. Aus dem „Parallelogramm der Kräfte“ in Abbildung zwei liest man sofort ab, daß $\vec{v}+\vec{w}=\vec{w}+\vec{v}$ ist. Abbildung drei zeigt die Summe der beiden Kraftfelder aus Abbildung eins.

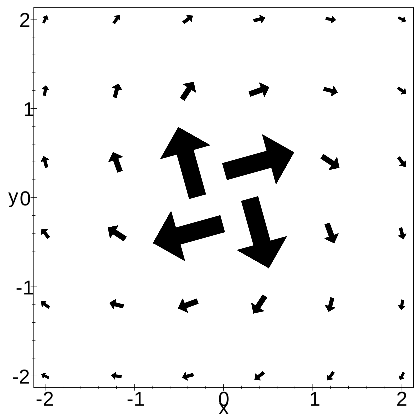

Abb. 3: Die Summe der beiden Felder aus Abbildung eins

Als weitere Eigenschaft der Vektoraddition wollen wir festhalten, daß es zu jedem Vektor $\vec{v}$ einen Vektor $\vec{w}$ gibt, so daß

$$\vec{v}+\vec{w}=\vec{0}$$

der Nullvektor ist; $\vec{w}$ ist einfach der entgegengesetzt orientierte Vektor $\vec{v}$. Wir bezeichnen diesen Vektor kurz als $-\vec{v}$.

Die Addition des Nullvektors ändert natürlich nichts am anderen Summanden, d.h.

$$\vec{v}+\vec{0}=\vec{0}+\vec{v}=\vec{v} \quad \text{für alle Vektoren } \vec{v}\,.$$

Schließlich gilt für die Vektoraddition auch noch das Assoziativgesetz

$$\big(\vec{u}+\vec{v}\big)+\vec{w}=\vec{u}+\big(\vec{v}+\vec{w}\big)\,,$$

wie man sich leicht überzeugt, indem man das Diagramm für die Konstruktion von $\vec{v}+\vec{w}$ an den Endpunkt des Vektors $\vec{u}$ verschiebt; siehe Abbildung vier.

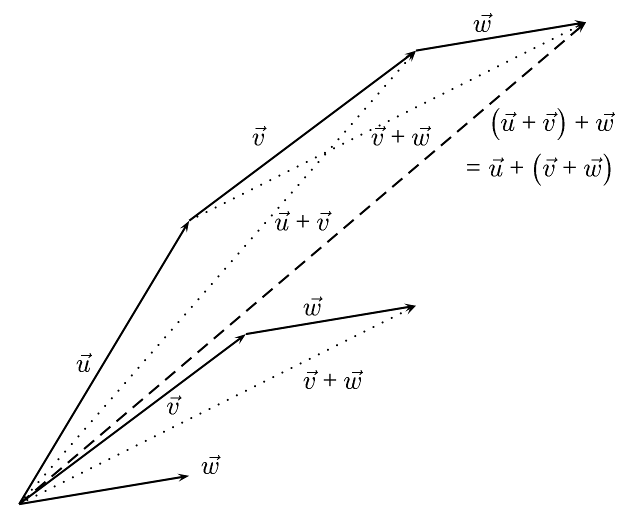

Abb. 4: Das Assoziativgesetz der Vektoraddition

Außer der Addition von Vektoren können wir auch ihre Streckung, d.h. ihre Multiplikation mit einer reellen Zahl, definieren: Ist $\vec{v}$ ein Vektor und $\lambda>0$ eine positive reelle Zahl, so soll $\lambda\vec{v}$ dieselbe Richtung haben wie $\vec{v}$ und die λ-fache Länge; für $\lambda<0$ soll $\lambda\vec{v}$ die entgegengesetzte Richtung haben und die $|\lambda|$-fache Länge. Für $\lambda=0$ schließlich ist $\lambda\vec{v}$ der Nullvektor.

Anwendung des Strahlensatzes auf das Dreieck in Abbildung fünf zeigt, daß für diese Multiplikation das Distributivgesetz

$$\lambda(\vec v + \vec w) = \lambda\vec v + \lambda\vec w$$

gilt: Als Strahlen betrachten wir von $\vec v$ und $\vec v + \vec w$ aufgespannten Halbgeraden, und wir schneiden mit den beiden parallelen Geraden durch die eingezeichneten Vektoren $\lambda\vec w$ und $\vec w$. Dabei sollen die fett eingezeichneten Vektoren die mit λ multiplizierten sein; im Bild ist $\lambda = 0{,}4$.

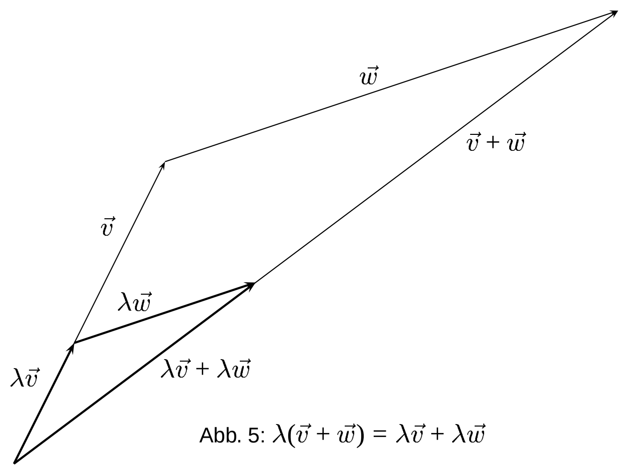

Abb. 5: $\lambda(\vec v + \vec w) = \lambda\vec v + \lambda\vec w$

Das andere Distributivgesetz $(\lambda + \mu)\vec v = \lambda\vec v + \mu\vec v$ ist ziemlich trivial: Da sich alles auf der von $\vec v$ aufgespannten Geraden abspielt und wir diese mit der reellen Zahlengeraden identifizieren können, läßt sich diese Regel auf das gewöhnliche Distributivgesetz in $\mathbb{R}$ zurückführen, genauso wie sich die Regel $(\lambda\mu)\vec v = \lambda(\mu\vec v)$ auf das gewöhnliche Assoziativgesetz der Multiplikation in $\mathbb{R}$ zurückführen läßt.

b) Definition des Vektorraums

Damit haben wir alle Rechenregeln zusammen, die wir für die Definition eines Vektorraums brauchen. Da wir Vektoren auch mit Zahlen multiplizieren wollen, müssen wir zwei Arten von Objekten betrachten:

Vektoren, die wir weiterhin mit $\vec v, \vec w$ *usw.* bezeichnen, sowie Skalare, für die wir griechische Buchstaben verwenden.

Für die Skalare lassen wir, wie bereits in §1*c)* diskutiert, Elemente eines beliebigen Körpers zu; für den Anfänger ist es wahrscheinlich am einfachsten, sich die Skalare zunächst als reelle Zahlen vorzustellen.

Definition: k sei ein Körper. Eine Menge V heißt *Vektorraum* über k oder k-Vektorraum, wenn es zwei Verknüpfungen

$$+\colon V \times V \to V \quad \text{und} \quad \cdot\colon k \times V \to V$$

gibt, so daß gilt:

I.1) Das Assoziativgesetz der Vektoraddition
$$(\vec u + \vec v) + \vec w = \vec u + (\vec v + \vec w) \qquad \text{für alle } \vec u, \vec v, \vec w \in V$$

I.2) Es gibt einen Vektor $\vec 0 \in V$, so daß für jeden Vektor $\vec v \in V$ gilt
$$\vec v + \vec 0 = \vec 0 + \vec v = \vec v\,.$$

I.3) Zu jedem Vektor $\vec v \in V$ gibt es einen Vektor $-\vec v \in V$, so daß
$$\vec v + (-\vec v) = (-\vec v) + \vec v = \vec 0\,.$$

I.4) Das Kommutativgesetz der Vektoraddition
$$\vec u + \vec v = \vec v + \vec u \qquad \text{für alle } \vec u, \vec v \in V$$

II.1) Das Distributivgesetz bei der Addition von Skalaren
$$(\lambda + \mu)\vec v = \lambda\vec v + \mu\vec v \quad \text{für alle } \lambda, \mu \in k \text{ und alle } \vec v \in V\,.$$

II.2) Das Distributivgesetz bei der Addition von Vektoren
$$\lambda(\vec v + \vec w) = \lambda\vec v + \lambda\vec w \quad \text{für alle } \lambda \in k \text{ und alle } \vec v, \vec w \in V\,.$$

II.3) Kompatibilität von Körper- und Skalarmultiplikation
$$(\lambda\mu)\vec v = \lambda(\mu\vec v) \quad \text{für alle } \lambda, \mu \in k \text{ und alle } \vec v \in V\,.$$

II.4) Multiplikation mit der Eins
$$1\vec v = \vec v \quad \text{für alle } \vec v \in V\,.$$

II.5) Multiplikation mit der Null *bzw.* mit dem Nullvektor
$$0\vec v = \vec 0 \quad \text{für alle } \vec v \in V \qquad \text{und} \qquad \lambda\vec 0 = \vec 0 \quad \text{für alle } \lambda \in k\,.$$

Bemerkung: Die Forderungen I.1–I.4 in der Definition des Körpers und des Vektorraums sowie die Forderungen II.1–II.4 in der Körperdefinition sind fast identisch, und in der Tat beschreiben sie eine gemeinsame mathematische Struktur, die sogenannte *abelsche Gruppe*. Da wir diese nicht weiter benötigen werden, sei auf Einzelheiten verzichtet.

Vektoren und Vektorräume sind als mathematische Begriffe recht jung: Rechnerische Methoden zur Lösung geometrischer Probleme wurden zwar schon ab etwa 1636 von RENÉ DESCARTES (1596–1650) eingesetzt (kartesische Koordinaten), aber erst gegen Mitte des 19. Jahrhunderts wurden Ansätze entwickelt, um geometrische Objekte *koordinatenfrei* algebraisch zu behandeln. Ein erster Durchbruch war das 1844 erschienene Buch *Die Ausdehnungslehre* von HERMANN GÜNTER GRASSMANN (1809–1877, oberes Bild): Er betrachtete abstrakte Objekte, die unter anderem alle Vektorraumaxiome erfüllten, die darüber hinaus allerdings auch miteinander multipliziert werden konnten, so daß er etwas komplizierteres als einen Vektorraum definiert hatte: eine sogenannte Algebra. Sie spielt noch heute eine große Rolle bei der Charakterisierung der Lage zweier Vektorräume ineinander. 1888 definierte GIUSEPPE PEANO (1858–1932, unteres Bild) in seinem Buch *Calcolo geometrico secondo l'Ausdehnungslehre di H. Grassmann preceduto dalle operazioni della logica deduttiva* Vektorräume (über $\mathbb{R}$) im obigen Sinne; in diesem Buch treten auch erstmalig mengentheoretische Symbole wie $\cap$, $\cap$ und $\in$ auf. Ab etwa 1920 wandte STEFAN BANACH (1892–1945) PEANOs Theorie an auf Funktionenräume und lineare Operatoren.

c) Erste Beispiele

Standardbeispiel sind natürlich die $\mathbb{R}$-Vektorräume $\mathbb{R}^n$. Wir schreiben ihre Elemente als Spaltenvektoren der Form

$$\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \quad \vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}, \quad \dots$$

und haben die beiden Rechenoperationen

$$\vec{v} + \vec{w} = \begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{pmatrix} \quad \text{und} \quad \lambda\vec{v} = \begin{pmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{pmatrix}.$$

Alle Rechenregeln folgen sofort aus den entsprechenden Regeln für reelle Zahlen, und genauso können wir auch für einen beliebigen Körper k die k-Vektorräume k^n definieren.

Auf den ersten Blick seltsam erscheint, daß $\mathbb{R}$ ein $\mathbb{Q}$-Vektorraum ist: Vektoraddition ist die gewöhnliche Addition reeller Zahlen und Multiplikation mit Skalaren die Multiplikation einer reellen Zahl mit einer rationalen. Auch hier folgen alle Vektorraumaxiome sofort aus den üblichen Rechenregeln für reelle Zahlen, für die es natürlich gleichgültig ist, daß hier einige der betrachteten Zahlen sogar rational sind.

Interessanter ist das folgende Beispiel: Für eine offene Teilmenge U von $\mathbb{R}$, als z.B. ein offenes Intervall $U = (a, b)$ und eine natürliche Zahl $n \in \mathbb{N}$ definieren wir die Menge

$$\mathcal{C}^n(U, \mathbb{R}) = \{ f: U \to \mathbb{R} \mid f \text{ ist } n\text{-mal stetig differenzierbar} \}.$$

Für deren Elemente seien Addition und Skalarmultiplikation punktweise definiert, d.h. für $f, g \in \mathcal{C}^n(U, \mathbb{R})$ und $\lambda \in \mathbb{R}$ setzen wir

$$f + g: \begin{cases} U \to \mathbb{R} \\ t \mapsto f(t) + g(t) \end{cases} \quad \text{und} \quad \lambda f: \begin{cases} U \to \mathbb{R} \\ t \mapsto \lambda f(t), \end{cases}.$$

Abbildung sechs zeigt für $(a, b) = (0, 3\pi)$ zu den beiden dünn eingezeichneten Funktionen $f(t) = \sin t$ und $g(t) = \sin 3t$ die dick eingezeichnete Funktion $f + g$, und Abbildung sieben zeigt f zusammen mit der Funktion $2f$.

Damit dies alles wohldefiniert ist, müssen wir uns noch überlegen, daß mit f und g die Funktionen $f + g$ und λf wieder in $\mathcal{C}^n(U, \mathbb{R})$ liegen. Das ist aber klar, denn die Summe zweier stetiger *bzw.* differenzierbarer Funktionen ist wieder stetig *bzw.* differenzierbar, und wegen der Rechenregel $(f + g)' = f' + g'$ gilt dies auch für die höheren Ableitungen. Genauso kann man für λf argumentieren. Da alle Rechenoperationen auf

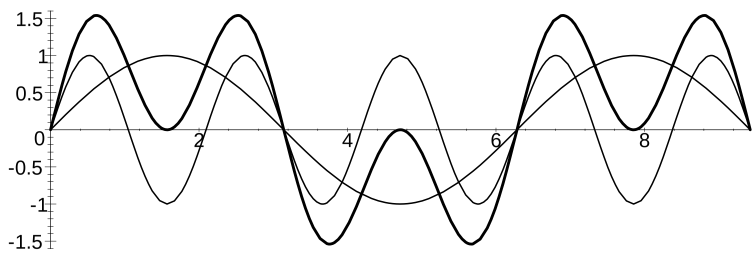

Abb. 6: Die Summe von $f(t) = \sin t$ und $g(t) = \sin 3t$

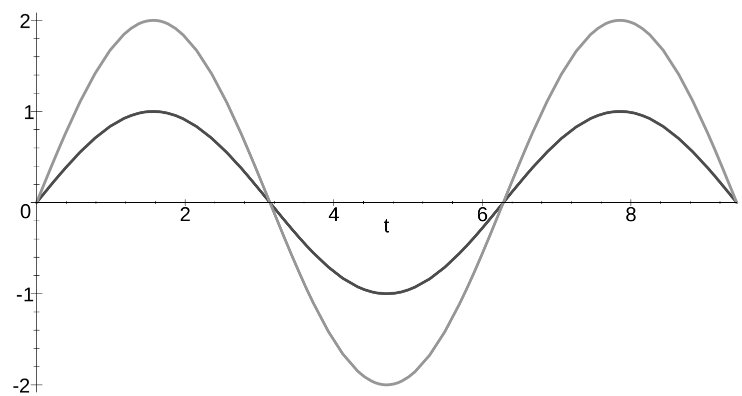

Abb. 7: $f(t) = \sin t$ zusammen mit $2f$

die gewöhnliche reelle Addition und Multiplikation für die Funktionswerte zurückgeführt ist, folgen die Vektorraumaxiome aus den üblichen Rechenregeln für reelle Zahlen: Um etwa das Assoziativgesetz

$$(f+g)+h = f+(g+h)$$

nachzuweisen, müssen wir zeigen, daß für jede reelle Zahl t die Funktionen auf beiden Seiten denselben Wert haben, d.h.

$$\big((f+g)+h\big)(t) = \big(f+(g+h)\big)(t) \quad \text{für alle } t \in \mathbb{R}\,.$$

Dazu rechnen wir beide Seiten aus:

$$\big((f+g)+h\big)(t) = (f+g)(t) + h(t) = \big(f(t)+g(t)\big) + h(t)$$

und

$$\big(f+(g+h)\big)(t) = f(t) + (g+h)(t) = f(t) + \big(g(t)+h(t)\big)\,,$$

und die beiden rechten Seiten stimmen in der Tat überein nach dem Assoziativgesetz für die Addition reeller Zahlen.

Die restlichen Axiome folgen genauso, nur etwas einfacher.

Ganz entsprechend lassen sich auch die Mengen

$$\mathcal{C}^0(U,\mathbb{R}) = \{ f\colon U \to \mathbb{R} \mid f \text{ ist stetig} \}$$

und

$$\mathcal{C}^\infty(U,\mathbb{R}) = \{ f\colon U \to \mathbb{R} \mid f \text{ ist beliebig oft differenzierbar} \}$$

sowie

$$\mathcal{C}^\omega(U,\mathbb{R}) = \left\{ f\colon U \to \mathbb{R} \;\middle|\; \begin{array}{l} f \text{ ist um jeden Punkt } x \in U \text{ durch} \\ \text{eine TAYLOR-Reihe darstellbar} \end{array} \right\}$$

zu $\mathbb{R}$-Vektorräumen machen. (Wer noch nicht weiß, was eine TAYLOR-Reihe ist, wird es im nächsten Kapitel lernen.)

Als trivialstes Beispiel überhaupt haben wir schließlich noch über jedem Körper k den Nullvektorraum, der nur aus dem Nullvektor $\vec{0}$ besteht.

d) Lineare Abbildungen

Vektorräume werden erst richtig interessant, wenn man mit ihren Elementen etwas mehr tun kann als sie nur zu addieren und mit Skalaren zu multiplizieren. In der Geometrie etwa möchte man Vektoren gelegentlich auch drehen, bei Vektorräumen von differenzierbaren Funktionen möchte man deren Elemente differenzieren und so weiter. Viele derartige Operationen lassen sich unter dem Begriff der linearen Abbildung einordnen:

Definition: *a)* Eine Abbildung $\varphi\colon V \to W$ heißt *linear,* wenn für alle Vektoren $\vec{u}, \vec{v} \in V$ und alle $\lambda, \mu \in k$ gilt:

$$\varphi(\lambda\vec{u} + \mu\vec{v}) = \lambda\varphi(\vec{u}) + \mu\varphi(\vec{v})\,.$$

b) Unter dem *Kern* von φ verstehen wir die Menge

$$\operatorname{Kern}\varphi \underset{\text{def}}{=} \left\{ \vec{v} \in V \;\middle|\; \varphi(\vec{v}) = \vec{0} \right\}.$$

c) Das *Bild* von φ ist die Menge

$$\operatorname{Bild}\varphi \underset{\text{def}}{=} \left\{ \vec{w} \in W \;\middle|\; \text{Es gibt } \vec{v} \in V \text{ mit } \varphi(\vec{v}) = \vec{w} \right\}.$$

Die beiden allereinfachsten Beispiele für lineare Abbildungen sind für jeden Vektorraum V die identische Abbildung $V \to V$ sowie die Nullabbildung, die jedem Vektor $\vec{v} \in V$ den Nullvektor zuordnet. Letztere kann man wahlweise als Abbildung $V \to V$ oder als Abbildung von V in den Nullvektorraum auffassen.

Ebenfalls völlig trivial ist die Linearität von *Projektionen* wie etwa der Projektion $\mathbb{R}^3 \to \mathbb{R}^2$, die jedem Vektor seine ersten beiden Komponenten zuordnet.

Ein etwas interessanteres Beispiel einer linearen Abbildung ist

$$\varphi\colon \begin{cases} \mathbb{R}^3 \to \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x - y \\ y - z \end{pmatrix} \end{cases};$$

sie ist linear, denn

$$\varphi\left(\lambda \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \mu \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}\right) = \varphi\left(\begin{pmatrix} \lambda x_1 + \mu x_2 \\ \lambda y_1 + \mu y_2 \\ \lambda z_1 + \mu z_2 \end{pmatrix}\right) = \begin{pmatrix} \lambda x_1 + \mu x_2 - \lambda y_1 - \mu y_2 \\ \lambda y_1 + \mu y_2 - \lambda z_1 - \mu z_2 \end{pmatrix}$$

und

$$\lambda \cdot \varphi\left(\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}\right) + \mu \cdot \varphi\left(\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}\right) = \lambda \begin{pmatrix} x_1 - y_1 \\ y_1 - z_1 \end{pmatrix} + \mu \begin{pmatrix} x_2 - y_2 \\ y_2 - z_2 \end{pmatrix} = \begin{pmatrix} \lambda x_1 - \lambda y_1 + \mu x_2 - \mu y_2 \\ \lambda y_1 - \lambda z_1 + \mu y_2 - \mu z_2 \end{pmatrix},$$

was offensichtlich dasselbe ist.

Bei Vektorräumen von Funktionen ist beispielsweise für jeden Punkt $t_0 \in (a,\, b)$ die Auswertungsabbildung

$$\mathcal{C}^n\big((a,\, b), \mathbb{R}\big) \to \mathbb{R}; \quad f \mapsto f(t_0)$$

nach Definition der Vektorraumoperationen von $\mathcal{C}^n\big((a,\, b), \mathbb{R}\big)$ linear, denn $\lambda f + \mu g$ wurde ja gerade so definiert, daß für t_0 wie auch für jeden anderen Punkt aus $(a,\, b)$ gilt

$$\big(\lambda f + \mu g\big)(t_0) = \lambda f(t_0) + \mu g(t_0).$$

Allgemeiner können wir auch die *Abtastung* einer Funktion betrachten: Für vorgegebene Punkte $t_1, \ldots, t_N \in (a,\, b)$ definieren wir die Abbildung

$$\varphi\colon \begin{cases} \mathcal{C}^n\big((a,\, b), \mathbb{R}\big) \to \mathbb{R}^N \\ f \mapsto \begin{pmatrix} f(t_1) \\ \vdots \\ f(t_N) \end{pmatrix} \end{cases},$$

die f an N Argumenten auswertet. Anwendung ist beispielsweise die Digitalisierung eines Signals, etwa eines Musikstücks für eine CD. Hier würde die Funktion f die zeitliche Variation des Schalldrucks beschreiben (die man zumindest als stetig annehmen kann, d.h. $n = 0$), und die Punkte t_i wären gleichmäßig über die Länge des Musikstücks verteilt, jeweils 44 100 Stück pro Sekunde.

Diese Abbildung ist linear, weil jede ihrer Komponentenabbildungen $f \mapsto f(t_i)$ linear ist. Wir können die Linearität dieser Digitalisierung eines Signals aber auch inhaltlich interpretieren: Die Eigenschaft

$$\varphi(\lambda f + \mu g) = \lambda\varphi(f) + \mu\varphi(g)$$

bedeutet für positive λ und μ, daß es gleichgültig ist, ob man zwei verschiedene Signale (z.B. Mikrophonkanäle) zunächst in einem analogen Mischpult vereinigt und dann digitalisiert oder zunächst digitalisiert und dann digital mischt. (Dieses setzt natürlich voraus, daß man sowohl analog als auch digital mit perfekter Genauigkeit arbeitet – keine sehr realistische Annahme.)

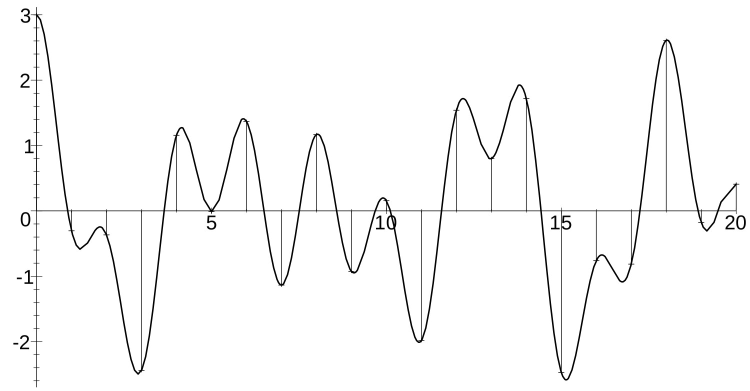

Abb. 8: Abtastung eines Signals

Ein anderes Beispiel einer linearen Abbildung zwischen Vektorräumen von Funktionen ist die Differentiation

$$\mathcal{C}^{n+1}\big((a,b),\mathbb{R}\big) \to \mathcal{C}^{n}\big((a,b),\mathbb{R}\big); \quad f \mapsto \dot{f}\,,$$

denn $(\lambda f + \mu g)' = \lambda f' + \mu g'$. Auch die Abbildung

$$\mathcal{C}^{2}\big((a,b),\mathbb{R}\big) \to \mathcal{C}^{0}\big((a,b),\mathbb{R}\big); \quad f \mapsto f'' + \omega^2 f$$

ist für jedes $\omega \in \mathbb{R}$ linear; ihr Kern besteht genau aus jenen Funktionen $f(t)$, die der Schwingungsdifferentialgleichung

$$f''(t) + \omega^2 f(t) = 0$$

genügen, enthält also beispielsweise die Funktionen $\cos \omega t$ und $\sin \omega t$.

e) Untervektorräume

Kern und Bild einer linearen Abbildung werden im allgemeinen unendliche Mengen sein, so daß selbst bei Vektorräumen wie dem $\mathbb{R}^3$ zunächst nicht ganz klar ist, wie man sie mit endlichem Aufwand beschreiben kann. Bei nichtlinearen Abbildungen kann so etwas in der Tat ein großes Problem sein, aber hier im Linearen reichen unsere vorhandenen Werkzeuge zumindest für Vektorräume wie einen $\mathbb{R}^n$ völlig aus. Zur Klärung der Begriffe beginnen wir mit einer

Definition: Eine Teilmenge $U \subseteq V$ eines k-Vektorraums V heißt *Untervektorraum,* in Zeichen $U \le V$, wenn U nicht leer ist und mit je zwei Vektoren $\vec{u}, \vec{v} \in U$ und Skalaren $\lambda, \mu \in k$ auch den Vektor $\lambda\vec{u} + \mu\vec{v}$ enthält.

Lemma: $\varphi: V \to W$ sei eine lineare Abbildung zwischen zwei k-Vektorräumen. Dann ist Kern φ ein Untervektorraum von V und Bild φ ein Untervektorraum von W.

Beweis: Sind $\vec{u}$ und $\vec{v}$ Elemente des Kerns von $\varphi: V \to W$ und $\lambda, \mu \in k$ Skalare, so ist

$$\varphi(\lambda\vec{u} + \mu\vec{v}) = \lambda\varphi(\vec{u}) + \mu\varphi(\vec{u}) = \lambda\vec{0} + \mu\vec{0} = \vec{0}\,,$$

also liegt auch $\lambda\vec{u} + \mu\vec{v}$ im Kern von φ. Außerdem ist dieser nicht leer, denn wegen

$$\varphi(\vec{0}) = \varphi(0 \cdot \vec{0}) = 0 \cdot \varphi(\vec{0}) = \vec{0}$$

liegt der Nullvektor in Kern φ. Also ist der Kern ein Untervektorraum.

Ähnlich ist die Situation für das Bild: Für zwei Vektoren $\vec{v}, \vec{w} \in$ Bild φ gibt es Vektoren $\vec{r}, \vec{s} \in V$, so daß $\varphi(\vec{r}) = \vec{v}$ und $\varphi(\vec{s}) = \vec{w}$ ist. Wegen der Linearität von φ liegt dann für zwei Skalare $\lambda, \mu \in k$ auch

$$\lambda\vec{v} + \mu\vec{w} = \lambda\varphi(\vec{r}) + \mu\varphi(\vec{s}) = \varphi(\lambda\vec{r} + \mu\vec{s})$$

im Bild von φ, das somit ein Untervektorraum von W ist. ■

Kern und Bild einer linearen Abbildung haben natürlich etwas mit deren Injektivität und Surjektivität zu tun; erinnern wir uns zunächst an die Definition dieser Begriffe:

Definition: *a)* Eine Abbildung $\varphi: M \to N$ zwischen zwei Mengen heißt *injektiv,* wenn keine zwei verschiedenen Elemente von M dasselbe Bild haben, d.h. aus der Gleichheit von $\varphi(m_1)$ und $\varphi(m_2)$ folgt für zwei Elemente $m_1, m_2 \in M$, daß $m_1 = m_2$ ist.
b) φ heißt *surjektiv,* wenn es zu jedem $n \in N$ (mindestens) ein $m \in M$ gibt, so daß $\varphi(m) = n$ ist.
c) φ heißt *bijektiv* oder auch „eins-zu-eins (1-1)“, wenn φ injektiv und surjektiv ist.

Lemma: $\varphi\colon V \to W$ ist genau dann injektiv, wenn Kern φ der Nullvektorraum ist; φ ist genau dann surjektiv, wenn Bild $\varphi = W$ ist.

Beweis: Die zweite Aussage ist zu trivial, als daß man etwas dazu sagen müßte, betrachten wir also die erste. Falls φ injektiv ist, hat insbesondere der Nullvektor nur ein einziges Urbild, d.h. der Kern besteht nur aus dem Nullvektor, der natürlich immer im Kern liegt. Ist umgekehrt Kern φ der Nullraum und haben zwei Vektoren $\vec{u}, \vec{v} \in V$ dasselbe Bild, so ist

$$\varphi(\vec{u} - \vec{v}) = \varphi(\vec{u}) - \varphi(\vec{v}) = \vec{0}\,,$$

d.h. $\vec{u} - \vec{v}$ liegt im Kern und muß daher gleich dem Nullvektor sein, so daß $\vec{u} = \vec{v}$ ist. Dies zeigt die Injektivität von φ. ■

Als Beispiel einer Anwendung dieses Lemmas betrachten wir noch einmal die Digitalisierung eines Signals: Aufgrund der hoch gelobten CD-Qualität erwarten wir, daß in diesem Fall die Abtastung eine „einigermaßen injektive“ lineare Abbildung ist. Das Wort „einigermaßen injektiv“ ist zwar kein wohldefinierter mathematischer Begriff, aber schon die Tatsache, daß bei einer CD die Abtastwerte nicht als reelle Zahlen gespeichert werden, sondern als 16bit-Zahlen, macht eine „echte“ Injektivität unmöglich. Überlegen wir uns, was sonst noch alles schiefgehen kann.

Nach dem gerade bewiesenen Lemma reicht es, wenn wir den Kern der Abbildung kennen. Dort liegt, bei einer Abtastung mit 44 100 Hz und in Sekunden gemessener Zeit, beispielsweise die Funktion

$$f(t) = \sin\bigl(44\,100\,\pi t\bigr) = \sin\bigl(22\,050 \cdot 2\pi t\bigr)\;;$$

denn für jedes ganzzahlige Vielfache von $1/44\,100$ ist das Argument des Sinus ein ganzzahliges Vielfaches von π, der Sinus also Null.

Die Funktion $f(t)$ entspricht einer reinen Schwingung mit einer Frequenz von 22,05 kHz. Solche Frequenzen sind zwar sehr wichtig für die Navigation von Fledermäusen, sie sind aber unhörbar für Käufer von CDs, so daß uns dieses Element des Kerns nicht weiter stört.

Betrachten wir aber beispielsweise die Funktionen

$$g(t) = \sin\bigl(66\,150\,\pi t\bigr) \quad\text{und}\quad h(t) = \sin\bigl(22\,050\,\pi t\bigr)\,.$$

Für $k \in \mathbb{Z}$ und $t = k/44\,100$ ist

$$g(t) = g\left(\frac{k}{44\,100}\right) = \sin\left(66\,150\pi \cdot \frac{k}{44\,100}\right) = \sin\left(\frac{3k\pi}{2}\right) = \begin{cases} 0 & \text{für gerades } k \\ -1 & \text{für } k \equiv 1 \bmod 4 \\ 1 & \text{für } k \equiv 3 \bmod 4 \end{cases}$$

und

$$h(t) = h\left(\frac{k}{44\,100}\right) = \sin\left(22\,050\pi \cdot \frac{k}{44\,100}\right) = \sin\left(\frac{k\pi}{2}\right) = \begin{cases} 0 & \text{für gerades } k \\ 1 & \text{für } k \equiv 1 \bmod 4 \\ -1 & \text{für } k \equiv 3 \bmod 4 \end{cases},$$

wobei $k \equiv a \bmod b$ bedeuten soll, daß $k - a$ durch b teilbar ist. Damit sind die beiden Funktionen g und h an allen Abtaststellen entgegengesetzt gleich, d.h. die Funktion $g + h$ liegt im Kern von φ.

Dieses Element des Kern stört uns erheblich mehr, denn es hat zur Folge, daß die beiden Funktion $g(t) = \sin\bigl(66\,150\,\pi t\bigr)$ und $-h(t) = -\sin\bigl(22\,050\,\pi t\bigr)$ auf dieselbe Weise digitalisiert werden. g beschreibt aber einen für Menschen unhörbaren Ton mit einer Frequenz von 33,075 kHz, während h mit nur 11,025 kHz durchaus hörbar ist. Abbildung neun zeigt die beiden Schwingungen; die Zeitachse ist der besseren Übersicht wegen in Millisekunden beschriftet, und die Abtastwerte sind durch Quadrate markiert.

Die Digitalisierungsabbildung kann also höchstens dann injektiv sein, wenn wir uns auf Funktionen beschränken, an deren Aufbau keine Schwingungen mit einer Frequenz von 22,05 kHz oder höher beteiligt sind. Was das bedeutet, und ob dann wirklich Injektivität gilt, werden wir in der *Höheren Mathematik II* im Kapitel über harmonische Analyse genauer untersuchen.

f) Lineare Abhängigkeit

Im $\mathbb{R}^3$ definieren zwei Vektoren eine Ebene – es sei denn, sie liegen, wenn man sie am gleichen Anfangspunkt beginnen läßt, auf einer Geraden, d.h einer der beiden Vektoren ist ein Vielfaches des anderen.

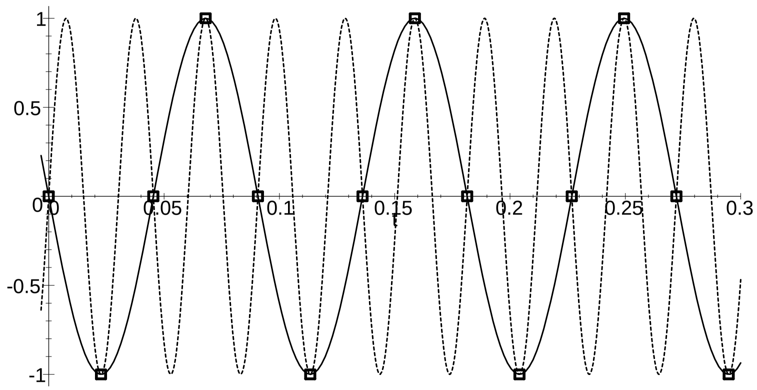

Abb. 9: Zwei verschiedene Signale, die gleich abgetastet werden

Entsprechend spannen drei Vektoren im allgemeinen den gesamten $\mathbb{R}^3$ auf – es sei denn, sie liegen, wenn man sie am gleichen Anfangspunkt beginnen läßt, in einer Ebenen, d.h. einer der drei ist als Summe von Vielfachen der anderen beiden darstellbar.

Der Begriff der *linearen Abhängigkeit* verallgemeinert diese Ausnahmebedingungen so, daß sie auf beliebige Vektorräume angewandt werden können:

Definition: $\vec{v}_1, \ldots, \vec{v}_n$ seien Elemente des k-Vektorraums V.

a) Eine *Linearkombination* von $\vec{v}_1, \ldots, \vec{v}_n$ ist eine Summe der Form

$$\lambda_1\vec{v}_1 + \cdots + \lambda_n\vec{v}_n$$

mit Skalaren $\lambda_i \in k$; ist diese Summe gleich dem Vektor $\vec{v} \in V$, so sagen wir, $\vec{v}$ sei als Linearkombination von Vektoren aus M darstellbar.

b) Die Menge aller Vektoren, die sich als Linearkombination der Vektoren $\vec{v}_1, \ldots, \vec{v}_n$ darstellen lassen, bezeichnen wir mit $[\vec{v}_1, \ldots, \vec{v}_n]$; wir nennen sie das *Erzeugnis* von $\vec{v}_1, \ldots, \vec{v}_n$.

c) Eine Linearkombination wie in *a)* heißt *nichttrivial*, falls mindestens ein λ_i von Null verschieden ist; ansonsten heißt sie *trivial*.

d) Die Vektoren $\vec{v}_1, \ldots, \vec{v}_n$ heißen *linear unabhängig*, wenn der Nullvektor nicht als nichttriviale Linearkombination von $\vec{v}_1, \ldots, \vec{v}_n$ darstellbar ist, d.h. eine Gleichung der Form

$$\lambda_1\vec{v}_1 + \cdots + \lambda_n\vec{v}_n = \vec{0}$$

kann nur gelten, wenn alle λ_i verschwinden.

e) Sind $\vec{v}_1, \ldots, \vec{v}_n$ *nicht* linear unabhängig, so bezeichnen wir sie als *linear abhängig*.

f) Eine *Teilmenge* $M \subseteq V$ eines Vektorraums V heißt *linear unabhängig*, wenn jede Auswahl endlich vieler verschiedener Vektoren $\vec{v}_1, \ldots \vec{v}_m$ (für beliebiges $m \in \mathbb{N}$) linear unabhängig ist.

g) Das *Erzeugnis* $[M]$ einer Teilmenge $M \subseteq V$ eines Vektorraums V ist die Menge aller Vektoren aus V, die als Linearkombination aus endlich vielen Vektoren aus V dargestellt werden können.

Beispielsweise sind also die Vektoren

$$\begin{pmatrix}1\\2\\3\end{pmatrix}, \quad \begin{pmatrix}2\\4\\6\end{pmatrix} \in \mathbb{R}^3$$

linear abhängig, da der zweite das Zweifache des ersten ist, und auch die Vektoren

$$\begin{pmatrix}1\\0\\0\end{pmatrix}, \quad \begin{pmatrix}2\\3\\0\end{pmatrix}, \quad \begin{pmatrix}1\\1\\0\end{pmatrix} \in \mathbb{R}^3$$

sind linear abhängig, denn

$$\lambda\begin{pmatrix}1\\0\\0\end{pmatrix} + \mu\begin{pmatrix}2\\3\\0\end{pmatrix} + \nu\begin{pmatrix}1\\1\\0\end{pmatrix} = \begin{pmatrix}\lambda + 2\mu + \nu\\3\mu + \nu\\0\end{pmatrix}$$

ist gleich dem Nullvektor wann immer $\nu = -3\mu$ und $\lambda = -2\mu - \nu = \mu$ ist. Eine nichttriviale Darstellung des Nullvektors ist beispielsweise

$$\begin{pmatrix}1\\0\\0\end{pmatrix} + \begin{pmatrix}2\\3\\0\end{pmatrix} - 3\begin{pmatrix}1\\1\\0\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}.$$

Die drei Vektoren

$$\begin{pmatrix}1\\0\\0\end{pmatrix},\quad \begin{pmatrix}0\\1\\0\end{pmatrix},\quad \begin{pmatrix}0\\0\\1\end{pmatrix} \in \mathbb{R}^3$$

dagegen sind linear unabhängig, denn

$$\lambda_1\begin{pmatrix}1\\0\\0\end{pmatrix} + \lambda_2\begin{pmatrix}0\\1\\0\end{pmatrix} + \lambda_3\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}\lambda_1\\\lambda_2\\\lambda_3\end{pmatrix}$$

ist genau dann gleich dem Nullvektor, wenn alle λ_i verschwinden.

Allgemein sind die Vektoren $\vec{v}_1, \dots, \vec{v}_n$ aus einem beliebigen Vektorraum V dann trivialerweise linear abhängig, wenn zwei Vektoren $\vec{v}_i$ und $\vec{v}_j$ (mit $j \neq i$) gleich sind, denn dann ist beispielsweise

$$1 \cdot \vec{v}_i + (-1) \cdot \vec{v}_j = \vec{0}$$

eine nichttriviale Darstellung des Nullvektors. Ebenfalls trivial ist die lineare Abhängigkeit, falls einer der Vektoren $\vec{v}_i$ gleich dem Nullvektor ist: Dann ist bereits

$$1 \cdot \vec{v}_i = \vec{0}$$

eine solche Darstellung. Eine Menge, die den Nullvektor enthält, ist also stets linear abhängig.

Auch in Vektorräumen von Funktionen können wir leicht Beispiele für lineare Abhängigkeit und Unabhängigkeit finden. In $\mathcal{C}^0(\mathbb{R}, \mathbb{R})$ sind etwa Sinus und Kosinus linear unabhängig, denn gäbe es $\lambda_{1/2} \in \mathbb{R}$ mit

$$\lambda_1 \sin x + \lambda_2 \cos x = 0 \quad \text{für alle } x \in \mathbb{R}$$

mit $\lambda_1 \neq 0$, so wäre

$$\tan x = \frac{\sin x}{\cos x} = -\frac{\lambda_2}{\lambda_1}$$

eine konstante Funktion; wäre $\lambda_2 \neq 0$, könnten wir entsprechend auf die Konstanz des Kotangens schließen.

Genauso sieht man, daß die Funktionen $\sin^2 x$ und $\cos^2 x$ linear unabhängig sind, denn auch die Quadrate von Tangens und Kotangens

sind nicht konstant. Dagegen sind die drei Funktionen $\sin^2 x$, $\cos^2 x$ und 1 (konstante Funktion) linear abhängig, denn

$$\sin^2 x + \cos^2 x - 1 = 0 \quad \text{für alle } x \in \mathbb{R}\,.$$

Elementare Beispiele von Linearkombinationen sind etwa die Zerlegung eines Vektors in seine Komponenten entlang der Achsen eines gegebenen Koordinatensystems, also etwa

$$\lambda\begin{pmatrix}1\\0\\0\end{pmatrix} + \mu\begin{pmatrix}0\\1\\0\end{pmatrix} + \nu\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}\lambda\\\mu\\\nu\end{pmatrix},$$

oder die „übliche“ Darstellung eines Polynoms durch Potenzen der Variable:

$$f(x) = \lambda_0 + \lambda_1 x + \lambda_2 x^2 + \lambda_3 x^3\,.$$

Im Vektorraum $V = \mathbb{R}[x]$ aller reeller Polynome in x ist demgemäß das Erzeugnis $[1, x, x^2, x^3]$ der Untervektorraum aller Polynome vom Grad höchstens drei.

Auch für Erzeugnisse unendlicher Mengen gibt es einfache Beispiele in $\mathbb{R}[x]$; beispielsweise ist das Erzeugnis

$$[1, x^2, x^4, x^6, x^8, \dots]$$

der Menge aller gerader Potenzen gleich die Menge aller Polynome, in denen nur gerade x-Potenzen vorkommen, also (wie man sich leicht überlegt) gleich der Menge aller gerader Polynome, d.h. der Polynome $f \in \mathbb{R}[x]$ mit $f(-x) = f(x)$ für alle $x \in \mathbb{R}$.

Da Konstanten und x-Potenzen stetige Funktionen sind, können wir auch im Vektorraum $\mathcal{C}^0(\mathbb{R}, \mathbb{R})$ aller stetiger Funktionen das Erzeugnis derselben Menge betrachten, und wieder erhalten wir die Menge aller gerader Polynome. Das mag auf den ersten Blick verwundern, da einige vielleicht erwartet hätten, daß auch die Funktion

$$\cos x = 1 - \frac{x^2}{2} + \frac{x^4}{24} - \frac{x^6}{720} + - \cdots = \sum_{i=0}^{\infty} (-1)^i \frac{x^{2i}}{(2i)!}$$

in $[1, x^2, x^4, x^6, x^8, \dots]$ liegt, aber dies ist eine *unendliche* Summe, und $[M]$ war ausdrücklich definiert als die Menge aller Linearkombinationen, in denen jeweils nur *endlich* viele Elemente aus M auftreten.

In der Musik (und in der Signalverarbeitung) spielen Linearkombinationen von Sinus- und Kosinusschwingungen eine große Rolle: Der Aufbau eines Tons aus Grundschwingung und Oberschwingungen ist mathematisch betrachtet einfach eine Linearkombination

$$f(t) = \sum_{i=1}^{n} \sin 2\pi i\nu t\,,$$

wobei ν die (Grund-)Frequenz des Tons ist. Bei einem Orchester, das den Kammerton a auf 440 Hz festlegt, sind also alle möglichen Klänge, die dieser Ton auf den verschiedenen Instrumenten annehmen kann, Funktionen aus dem Erzeugnis

$$[\sin 440 \cdot 2\pi t,\ \sin 880 \cdot 2\pi t,\ \sin 1320 \cdot 2\pi t,\ \dots] \le \mathcal{C}^0(\mathbb{R}, \mathbb{R})\,.$$

Abbildung zehn zeigt den Ton, den die g-Saite einer Geige produziert zusammen mit der (kaum sichtbaren) Grundschwingung von 196 Hz sowie den ersten acht Oberschwingungen; außerdem ist zum Vergleich gestrichelt eine reine Schwingung der Frequenz 196 Hz eingezeichnet. Wie man sieht, spielen in diesem Beispiel die Oberschwingungen mit der doppelten und der dreifachen Grundfrequenz die größte Rolle.

(Wer selbst Töne aus Grund- und Oberschwingungen synthetisieren möchte, findet unter `http://www.gac.edu/~huber/fourier/` ein Java-Applet, das die entsprechenden Summenkurven zeichnen und die dazugehörigen Töne hörbar machen kann.)

Linearkombinationen sind somit ein einfaches Mittel, um aus relativ wenigen einfachen Funktionen oder Vektoren kompliziertere aufzubauen. Insbesondere bieten sie auch die Möglichkeit, Untervektorräume mit endlichem Aufwand zu beschreiben: Im $\mathbb{R}^n$ etwa ist jeder Untervektorraum mit Ausnahme des Nullraums $\{\vec{0}\}$ eine unendliche Menge, aber wie wir bald sehen werden, läßt sich jeder dieser Untervektorräume als Erzeugnis von endlich vielen Vektoren darstellen.

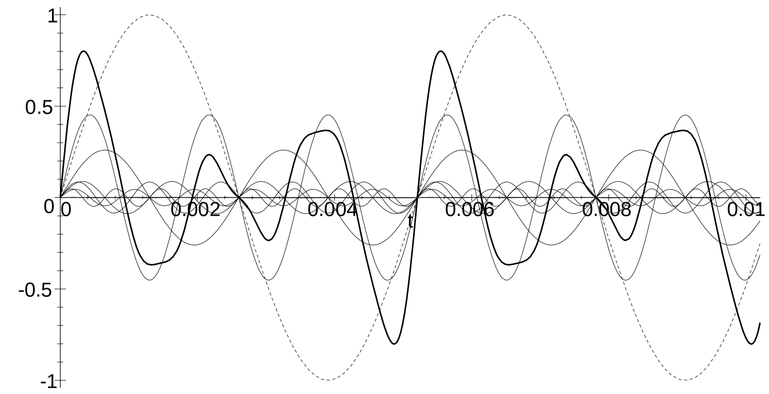

Abb. 10: Ton der g-Saite einer Geige und seine Komponenten

Als ersten Schritt dazu wollen wir uns überlegen, daß die Teilmenge $[M]$ stets ein Untervektorraum ist:

Lemma: Für jede Teilmenge M eines k-Vektorraums V ist $[M] \le V$ ein Untervektorraum von V; es ist der kleinste Untervektorraum von V, der M enthält.

Beweis: Nach dem Untervektorraumkriterium müssen wir zeigen, daß $[M]$ nicht leer ist und mit je zwei Vektoren $\vec{u}, \vec{v} \in [M]$ und zwei Skalaren $\lambda, \mu \in k$ auch den Vektor $\lambda\vec{u} + \mu\vec{v}$ enthält.

Die erste Eigenschaft ist (fast) trivial: Ist $\vec{v}$ irgendein Vektor aus M, so ist $1\vec{v} = \vec{v}$ eine Linearkombination von $\vec{v}$, liegt also in $[M]$, und insbesondere ist damit $M \subseteq [M]$. Die einzige kleine Schwierigkeit ergibt sich, wenn $M = \emptyset$ die leere Menge ist. Hier müssen wir uns auf die übliche Konvention berufen, daß leere Summen gleich Null sein sollen, eine „Linearkombination" aus null Vektoren als entsprechend gleich dem Nullvektor, der somit auch im Falle $M = \emptyset$ in $[M]$ liegt.

Nun seien

$$\vec{u} = \lambda_1\vec{u}_1 + \cdots + \lambda_n\vec{u}_n \quad \text{und} \quad \vec{v} = \lambda_1\vec{v}_1 + \cdots + \lambda_m\vec{v}_m$$

zwei Linearkombinationen von Vektoren aus M. Da wir zu jeder Linearkombination Summanden der Form $0\vec{w}$ hinzufügen können, ohne etwas an der Summe zu ändern, können wir die beiden Linearkombinationen auch in der Form

$$\vec{u} = \alpha_1\vec{w}_1 + \cdots + \alpha_\ell\vec{w}_\ell \quad \text{und} \quad \vec{v} = \beta_1\vec{w}_1 + \cdots + \beta_\ell\vec{w}_\ell$$

schreiben, wobei

$$\{\vec{w}_1, \dots, \vec{w}_\ell\} = \{\vec{u}_1, \dots, \vec{u}_n\} \cup \{\vec{v}_1, \dots, \vec{v}_m\}$$

ist mit irgendeiner beliebigen Nummerierung der Elemente. Dann ist aber klar, daß auch

$$\lambda\vec{u} + \mu\vec{w} = (\lambda\alpha_1 + \mu\beta_1)\vec{w}_1 + \cdots + (\lambda\alpha_\ell + \mu\beta_\ell)\vec{w}_\ell$$

eine Linearkombination von Vektoren aus M ist und somit in $[M]$ liegt.

Schließlich müssen wir noch zeigen, daß $[M]$ der *kleinste* Untervektorraum von V ist, der M enthält. Wir wissen bereits, daß $[M]$ ein Untervektorraum von V ist, der M enthält; um zu sehen, daß es der kleinste ist, betrachten wir irgendeinen Untervektorraum U von V ist, der M enthält. Dann ist U insbesondere ein Vektorraum, enthält also mit je zwei Vektoren auch deren sämtliche Linearkombinationen. Induktiv folgt, daß er mit jeder endlichen Anzahl von Vektoren auch deren sämtliche Linearkombinationen enthält, also enthält er mit M auch alle Vektoren aus $[M]$. Damit ist $[M] \subseteq U$ für jeden Untervektorraum U, der M enthält, und $[M]$ ist somit in der Tat der kleinste solche Untervektorraum von V. ■

Vektorräume wurden früher und werden auch gelegentlich noch heute als *lineare Räume* bezeichnet; da $[M]$ somit der kleinste lineare Raum ist, der M enthält, nennt man $[M]$ auch die *lineare Hülle* von M.

Am ökonomischsten ist die Darstellung eines Untervektorraums $U \leq V$ in der Form $U = [M]$ dann, wenn M möglichst wenig Elemente enthält. Wir wollen uns als nächstes überlegen, daß dies höchstens dann der Fall sein kann, wenn M linear unabhängig ist:

Lemma: Falls die Menge $M \subseteq V$ linear abhängig ist, gibt es ein Element $\vec{v} \in M$, das sich als Linearkombination der übrigen, d.h. von Vektoren aus $M \setminus \{\vec{v}\}$, schreiben läßt. Insbesondere ist dann auch

$$[M] = \big[M \setminus \{\vec{v}\}\big] .$$

Beweis: Wenn M linear abhängig ist, gibt es eine nichttriviale Linearkombination von Vektoren $\vec{v}_i \in M$, so daß

$$\lambda_1\vec{v}_1 + \cdots + \lambda_n\vec{v}_n = \vec{0}$$

ist mit Körperelementen λ_i, die nicht alle gleich Null sind. Sei zum Beispiel $\vec{\lambda}_j \neq 0$. Dann kann obige Gleichung nach $\vec{v}_j$ aufgelöst werden; für $1 < j < n$ etwa ist

$$\vec{v}_j = -\frac{\lambda_1}{\lambda_j}\vec{v}_1 - \cdots - \frac{\lambda_{j-1}}{\lambda_j}\vec{v}_{j-1} - \frac{\lambda_{j+1}}{\lambda_j}\vec{v}_{j+1} - \cdots - \frac{\lambda_n}{\lambda_j}\vec{v}_n ,$$

und entsprechend läßt sich $\vec{v}_j$ auch im Falle $j = 1$ oder $j = n$ als Linearkombination der übrigen $\vec{v}_i$ schreiben. ■

g) Die Dimension eines Vektorraums

Die Dimension eines Vektorraums soll natürlich so definiert werden, daß $\mathbb{R}^n$ die Dimension n hat; wir müssen die Zahl n also irgendwie als Eigenschaft von (Mengen von) Vektoren aus $\mathbb{R}^n$ rekonstruieren.

Offensichtlich kann jeder Vektor aus $\mathbb{R}^n$ als Linearkombination der n Einheitsvektoren geschrieben werden:

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + a_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} .$$

Es fällt aber schwer sich eine Menge aus weniger als n Vektoren vorzustellen, aus der sich ebenfalls *jeder* Vektor aus $\mathbb{R}^n$ als Linearkombination darstellen läßt.

Diese Eigenschaft machen wir uns zunutze, um allgemein Dimensionen zu definieren:

Definition: *a)* Eine Teilmenge $M \subseteq V$ eines k-Vektorraums V heißt *Erzeugendensystem*, wenn $[M] = V$ ist.
b) Wir sagen, der k-Vektorraum V sei *endlichdimensional*, wenn er ein endliches Erzeugendensystem hat; ansonsten bezeichnen wir V als *unendlichdimensional*.