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.)

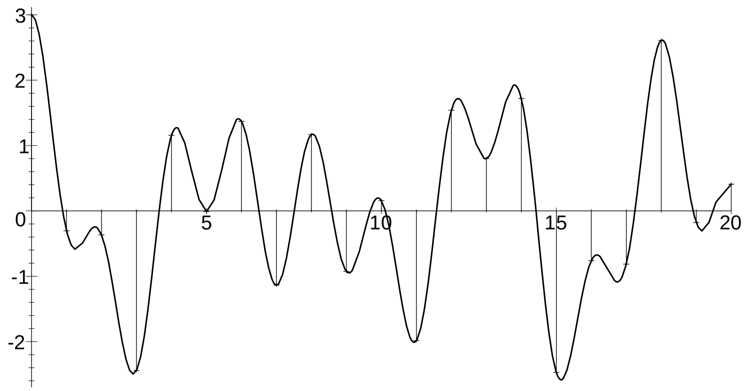

Abb. 8: Abtastung eines Signals

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: $\varphi\colon V \to W$ 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 $\varphi\colon V \to W$ 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 $\varphi\colon M \to N$ 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.

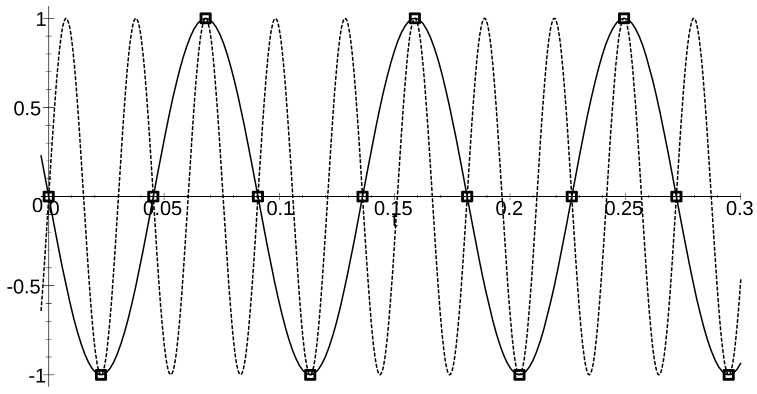

Abb. 9: Zwei verschiedene Signale, die gleich abgetastet werden

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.

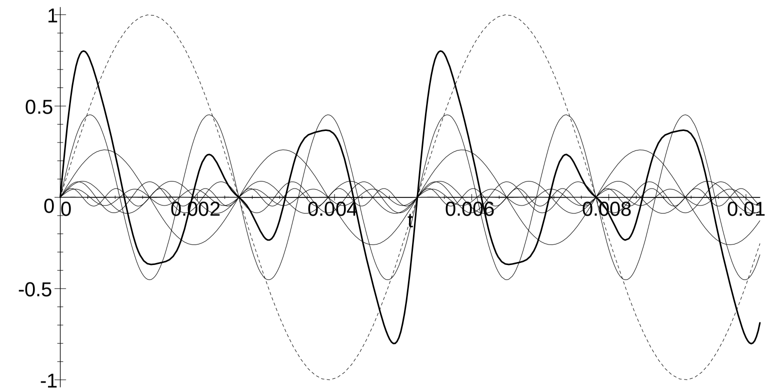

Abb. 10: Ton der g-Saite einer Geige und seine Komponenten

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

$$\{\vec w_1, \ldots, \vec w_\ell\} = \{\vec u_1, \ldots, \vec u_n\} \cup \{\vec v_1, \ldots, \vec v_m\}$$

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 $U \le V$ 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

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

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.*

c) Wir sagen, der endlichdimensionale k-Vektorraum V habe die Dimension n, in Zeichen $n = \dim_k V$ oder kurz $n = \dim V$, wenn er ein n-elementiges Erzeugendensystem enthält, aber kein Erzeugendensystem mit weniger als n Elementen.
d) Dem Nullvektorraum $\{\vec{0}\}$ ordnen wir (formal) die Dimension Null zu.

Als Beispiel eines unendlichdimensionalen Vektorraums haben wir den Vektorraum aller reeller Polynome. Hätte dieser nämlich ein endliches Erzeugendensystem, bestehend etwa aus den Polynomen f_1 bis f_n, so ließe sich sich jedes Polynom als Linearkombination

$$f = \lambda_1 f_1 + \cdots + \lambda_n f_n$$

schreiben. Auf diese Weise aber erhält man nur Polynome, deren Grad nicht größer ist als der größte Grad eines f_i. Damit sind auch alle Vektorräume $\mathcal{C}^k\big((a,b),\mathbb{R}\big)$ unendlichdimensional, denn sie enthalten insbesondere alle Polynome.

Endlichdimensional sind natürlich die reellen Vektorräume $\mathbb{R}^n$, denn $\mathbb{R}^n$ wird von seinen n Einheitsvektoren erzeugt. Da wir aber noch nicht sicher wissen, daß es kein Erzeugendensystem mit *weniger* als n Vektoren gibt, können wir im Augenblick nur sagen, daß die Dimension von $\mathbb{R}^n$ *höchstens* n ist.

Für $\mathbb{R}^2$ sieht man leicht, daß sie genau zwei ist: Ansonsten gäbe es nämlich ein Erzeugendensystem aus nur einem Vektor, d.h. alle Vektoren aus $\mathbb{R}^2$ wären proportional zueinander, was natürlich nicht der Fall ist. Für beliebiges n müssen wir ähnlich argumentieren mit linearer Abhängigkeit anstelle von Proportionalität; die Methoden dazu entwickelt der nächste Abschnitt.

h) Basen

Im $\mathbb{R}^3$ läßt sich jeder Vektor

$$\vec{v} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} = a\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + c\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

auf genau eine Weise als Linearkombination der drei Einheitsvektoren

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

schreiben; dies entspricht der Tatsache, daß wir im $\mathbb{R}^3$ drei Koordinaten haben.

Mit dem Begriff der *Basis* soll dieser Sachverhalt (soweit möglich) auf beliebige Vektorräume verallgemeinert werden. Da bei der Beschreibung eines Punkts durch seine Koordinaten deren Reihenfolge wesentlich ist, werden wir Basen meist nicht einfach als Mengen auffassen, sondern als geordnete Systeme, im endlichen Fall also als Tupel:

Definition: Ein System $\mathcal{B}$ von Vektoren $\vec{b}_1, \vec{b}_2, \ldots$ eines k-Vektorraums V heißt *Basis* von V, wenn gilt:
1.) Die Menge der $\vec{b}_i$ erzeugt den Vektorraum V, *und*
2.) $\mathcal{B}$ ist linear unabhängig.

Wenn es nicht auf die Reihenfolge ankommt, bezeichnen wir gelegentlich auch die Menge der Basisvektoren als Basis; in diesem Sinne ist also eine Basis einfach ein linear unabhängiges Erzeugendensystem.

Es ist klar, daß die Einheitsvektoren $\vec{e}_1, \vec{e}_2, \vec{e}_3$ eine Basis des $\mathbb{R}^3$ bilden. Ihre wesentliche Eigenschaft der eindeutigen Darstellbarkeit eines jeden Vektors als Linearkombination teilt sie mit jeder anderen Basis:

Lemma: Ist $\mathcal{B}$ eine Basis eines Vektorraums V, so läßt sich jeder Vektor $\vec{v} \in V$ auf genau eine Weise als Linearkombination

$$\vec{v} = \lambda_1 \vec{b}_1 + \cdots + \lambda_r \vec{b}_r$$

von Basisvektoren $\vec{b}_i$ aus $\mathcal{B}$ darstellen.

Beweis: Nach der ersten Eigenschaft aus der Definition einer Basis müssen die Basisvektoren V erzeugen, also läßt sich jeder Vektor $\vec{v} \in V$ als Linearkombination von endlich vielen Elementen aus $\mathcal{B}$ darstellen. Auch wenn wir von zwei solchen Darstellungen ausgehen, ist die Menge der daran beteiligten Vektoren aus $\mathcal{B}$ noch endlich; wir können also annehmen, daß es r Vektoren $\vec{b}_1, \dots, \vec{b}_r$ gibt, so daß

$$\begin{aligned}\vec{v} &= \lambda_1\vec{b}_1 + \cdots + \lambda_r\vec{b}_r \\ &= \mu_1\vec{b}_1 + \cdots + \mu_r\vec{b}_r\end{aligned}$$

ist, wobei wir einfach λ_i oder μ_i gleich Null setzen, wenn $\vec{b}_i$ in der entsprechenden Darstellung nicht vorkommt. Subtrahieren wir die beiden Darstellungen voneinander, erhalten wir eine Darstellung des Nullvektors als Linearkombination

$$\vec{0} = (\lambda_1 - \mu_1)\vec{b}_1 + \cdots + (\lambda_r - \mu_r)\vec{b}_r$$

von Basisvektoren. Da diese nach der zweiten definierenden Eigenschaft einer Basis linear unabhängig sind, müssen alle Koeffizienten $\lambda_i - \mu_i$ verschwinden. Damit sind die beiden betrachteten Darstellungen von $\vec{v}$ als Linearkombination der Vektoren aus $\mathcal{B}$ gleich, mit anderen Worten: Es gibt genau eine solche Darstellung. ■

Lemma: Ein Erzeugendensystem eines k-Vektorraum V ist genau dann eine Basis, wenn es minimal ist.

Beweis: Das Erzeugendensystem $\mathcal{B}$ sei eine Basis. Um zu zeigen, daß es minimal ist, müssen wir uns überlegen, daß jeder Basisvektor $\vec{v}$ aus $\mathcal{B}$ wirklich notwendig ist, daß also $\mathcal{B}$ ohne diesen Vektor $\vec{v}$ kein Erzeugendensystem mehr ist.

Falls es eines wäre, könnte insbesondere der Vektor $\vec{v}$ als Linearkombination der restlichen Vektoren aus $\mathcal{B}$ geschrieben werden. Gleichzeitig hat er aber die Darstellung $\vec{v} = \vec{v}$, deren rechte Seite man auch als Linearkombination von Elementen aus $\mathcal{B}$ auffassen kann. Somit ist seine Basisdarstellung nicht eindeutig, im Widerspruch zum gerade bewiesenen Lemma. Daher muß $\mathcal{B}$ minimal sein.

Umgekehrt sei $\mathcal{B}$ ein minimales Erzeugendensystem. Um zu zeigen, daß es eine Basis ist, reicht der Nachweis der linearen Unabhängigkeit von $\mathcal{B}$.

Sei also $\lambda_1\vec{b}_1 + \cdots + \lambda_n\vec{b}_n = \vec{0}$ eine Darstellung des Nullvektors als Linearkombination von Elementen aus $\mathcal{B}$. Falls darin einer der Koeffizienten λ_i nicht verschwindet, läßt sich der zugehörige Vektor $\vec{v}_i$ als Linearkombination der restlichen $\vec{v}_j$ schreiben. Dann reicht aber bereits $\mathcal{B}$ ohne $\vec{v}_i$ zur Erzeugung aus, $\mathcal{B}$ ist also nicht minimal. Somit müssen alle λ_i verschwinden, $\mathcal{B}$ ist also linear unabhängig und damit eine Basis. ■

Lemma: Eine System von linear unabhängigen Elementen eines Vektorraums ist genau dann eine Basis, wenn es maximal ist.

Beweis: $\mathcal{B}$ sei eine Basis. Dann läßt sich jeder Vektor $\vec{v} \in V$ als Linearkombination der Elemente von $\mathcal{B}$ schreiben, nimmt man also $\vec{v}$ zu $\mathcal{B}$ hinzu, ist das System nicht mehr linear unabhängig.

Umgekehrt sei $\mathcal{B}$ maximal linear unabhängig, und $\vec{v}$ sei ein beliebiger Vektor; wir müssen zeigen, daß er als Linearkombination der Vektoren aus $\mathcal{B}$ darstellbar ist. Das ist trivial, falls $\vec{v}$ bereits zu $\mathcal{B}$ gehört.

Andernfalls ist $\mathcal{B}$ zusammen mit $\vec{v}$ linear abhängig, da $\mathcal{B}$ ja als *maximal* linear unabhängig vorausgesetzt war. Somit gibt es ein nichttriviale Linearkombination

$$\lambda\vec{v} + \lambda_1\vec{b}_1 + \cdot + \lambda_n\vec{b}_n = \vec{0}$$

mit Vektoren $\vec{b}_i \in \mathcal{B}$. Darin muß $\lambda \neq 0$ sein, denn sonst wäre $\mathcal{B}$ linear abhängig. Also liegt

$$\vec{v} = -\frac{\lambda_1}{\lambda}\vec{b}_1 + \cdots + -\frac{\lambda_n}{\lambda}\vec{b}_n$$

in $[\mathcal{B}]$, und $\mathcal{B}$ ist ein Erzeugendensystem. ■

Basen lassen sich somit auch charakterisieren als minimale Erzeugendensysteme oder als maximale Systeme linear unabhängiger Vektoren.

Als nächstes stellt sich die Frage, wann es Basen gibt. Glücklicherweise hat *jeder* Vektorraum eine Basis; der Beweis ist allerdings für unendlichdimensionale Vektorräume logisch nicht ganz einfach. Für diese Vorlesung wollen wir uns daher mit einem Beweis für endlichdimensionale

Vektorräume begnügen. Wir beweisen dazu den etwas allgemeineren, tatsächlich ebenfalls für beliebige Vektorräume gültigen

Basisergänzungssatz: $M \subset V$ sei eine linear unabhängige Teilmenge des endlichdimensionalen Vektorraums V. Dann gibt es eine Basis $\mathcal{B}$ von V, die M enthält.

Beweis: Da V nach Voraussetzung endlichdimensional ist, gibt es zunächst einmal überhaupt eine endliche Menge $E \subset V$, die V erzeugt. Falls E einen oder mehrere der Vektoren aus M enthält, entfernen wir diese; was übrigbleibt, sei die Menge N, d.h. $N = E \setminus M$.

Damit sind M und N zwei disjunkte Teilmengen von V, deren Vereinigung die Menge E enthält und somit insbesondere ein Erzeugendensystem von V ist.

Konkret sei $M = \{\vec{b}_1, \dots, \vec{b}_r\}$ und $N = \{\vec{v}_1, \dots, \vec{v}_s\}$; dann wird V also erzeugt von

$$M \cup N = \{\vec{b}_1, \dots, \vec{b}_r, \vec{v}_1, \dots, \vec{v}_s\} .$$

Wir beweisen die Behauptung durch Induktion nach der Elementanzahl s von N.

Für $s = 0$ bilden die Elemente von M, in irgendeiner Weise angeordnet, bereits eine Basis, und wir sind fertig.

Für $s > 0$ sind wir fertig, falls $M \cup N$ linear unabhängig ist; denn dann ist $M \cup N$ eine Basis von V, die M enthält.

Andernfalls gibt es Elemente $\lambda_1, \dots, \lambda_r$ und $\mu_1, \dots, \mu_s$, die nicht alle gleichzeitig verschwinden, so daß

$$\lambda_1\vec{b}_1 + \cdots + \lambda_r\vec{b}_r + \mu_1\vec{v}_1 + \cdots + \mu_s\vec{v}_s = \vec{0}$$

ist. In dieser Gleichung können nicht alle μ_i verschwinden, denn sonst wären die $\vec{b}_i$ linear abhängig, im Widerspruch zur Voraussetzung. Also gibt es (mindestens) ein $\mu_i \neq 0$, und der zugehörige Vektor $\vec{v}_i$ läßt sich als Linearkombination der restlichen $\vec{v}_j$ und der $\vec{b}_\ell$ ausdrücken.

Damit wird $V = [M \cup (N \setminus \{\vec{v}_i\})]$ auch von der um $\vec{v}_i$ verminderten Menge erzeugt, und wir haben nur noch $s - 1$ Vektoren $\vec{v}_j$. Daher gibt es nach Induktionsannahme eine Basis von V, die M enthält. ∎

Korollar: Jeder endlichdimensionale Vektorraum $V \neq \{\vec{0}\}$ hat eine Basis.

Beweis: Man wende den obigen Satz an auf eine Menge M, die aus einem einzigen Vektor $\vec{v} \neq \vec{0}$ besteht. ∎

Um die Sonderrolle des Nullvektorraums zu eliminieren, vereinbaren wir, daß er die leere Menge als Basis haben soll; dies ist kompatibel mit der üblichen Interpretation von leeren Summen und leeren Aussagen.

Wie bereits erwähnt, gelten sowohl der Basisergänzungssatz als auch das obige Korollar für beliebige Vektorräume, d.h. also auch im Falle unendlicher Dimension. Für interessierte Leser sei kurz erwähnt, wie man hier vorgeht. Das wesentliche neue Hilfsmittel ist das ZORNsche Lemma, benannt nach dem deutschen Mathematiker MAX ZORN (1906–1993), der es, nachdem er Deutschland wegen der nationalsozialistischen Politik verlassen mußte, um 1935 an der amerikanischen Yale Universität bewies. Es besagt folgendes:

Gegeben sei eine nichtleere partiell geordnete Menge $\mathcal{M}$, d.h. für manche Paare von Elementen $A, B \in \mathcal{M}$ ist eine Relation $A < B$ erklärt mit der Eigenschaft, daß mit $A < B$ und $B < C$ auch $A < C$ gilt, wohingegen nie $A < A$ ist. Diese partiell geordnete Menge habe die zusätzliche Eigenschaft, daß es zu jeder Kette

$$A_1 < A_2 < A_3 < \cdots$$

von Elementen aus $\mathcal{M}$ ein Element A_∞ gebe mit der Eigenschaft, daß $A_i < A_\infty$ für alle i. Dann gibt es in $\mathcal{M}$ ein *maximales* Element, d.h. ein Element B, zu dem es kein $C \in \mathcal{M}$ gibt mit $B < C$.

Dieses Lemma kann nicht aus den üblichen Axiomen der Mengenlehre hergeleitet werden, sondern ist äquivalent zum sogenannten *Auswahlaxiom*. Für dieses bewies um 1940 der österreichische Mathematiker KURT GÖDEL (1906–1978), seit 1940 im amerikanischen Exil in Princeton, daß sowohl dieses Axiom als auch seine Negation mit den restlichen Axiomen der Mengenlehre kompatibel ist; das gleiche gilt demnach auch für das ZORNsche Lemma. Man kann daher wählen, ob man eine Mathematik mit oder ohne ZORNsches Lemma bevorzugt. Die meisten Mathematiker haben sich für „mit“ entschieden, es gibt aber auch welche, die das ZORNsche Lemma ablehnen.

Aus dem ZORNschen Lemma folgt der Basisergänzungssatz recht einfach: Als Menge $\mathcal{M}$ nehmen wir die Menge aller linear unabhängiger Teilmengen $A \subset V$, die M enthalten; die partielle Ordnungsrelation sei die gewöhnliche (echte) Teilmengenbeziehung. Die Kettenbedingung des ZORNschen Lemmas ist offensichtlich erfüllt, denn für eine Kette

$$M \subset A_1 \subset A_2 \subset A_3 \subset \cdots$$

aus linear unabhängigen Mengen A_i, die M enthalten, ist auch

$$A_\infty = \bigcup_{i \geq 1} A_i$$

eine linear unabhängige Teilmenge von V, die M enthält, da jede endliche Menge von Vektoren aus A_∞ bereits in einer der Mengen A_m liegt. Also gibt es nach dem ZORNschen Lemma ein maximales Element $\mathcal{B} \in \mathcal{M}$. Diese Menge $\mathcal{B}$ ist linear unabhängig, da sie in $\mathcal{M}$ liegt, und sie ist eine Basis, denn gäbe es einen Vektor $\vec{v} \notin [\mathcal{B}]$, so wäre auch die Menge $\mathcal{C} = \mathcal{B} \cup \{\vec{v}\}$ linear unabhängig, im Widerspruch zur Maximalität von $\mathcal{B}$. Damit ist der Basisergänzungssatz bewiesen, und das Korollar folgt wie oben.

Um wenigstens anhand eines Beispiels zu sehen, daß auch unendlichdimensionale Vektorräume Basen haben, betrachten wir den Vektorraum V aller Polynome mit reellen Koeffizienten. Da sich ein Polynom P vom Grad d als

$$P = a_0 + a_1x + a_2x^2 + \cdots + a_dx^d$$

schreiben läßt, erzeugt das System $\mathcal{B}$ der x-Potenzen $1, x, x^2, x^3, \ldots$ diesen Vektorraum. Jede Linearkombination des Nullvektors, des Polynoms $P \equiv 0$ also, aus Elementen von $\mathcal{B}$ wäre ein Polynom

$$\lambda_0 + \lambda_1x + \cdots + \lambda_nx^n\,,$$

dessen Koeffizienten zumindest teilweise von Null verschieden sind, während es selbst identisch Null wäre. Da es kein solches Polynom gibt, ist $\mathcal{B}$ linear unabhängig, also eine Basis von V.

Die Schwierigkeiten, die bei unendlichdimensionalen Vektorräumen auftreten können, sieht man, wenn man in diesem Beispiel die Polynome durch Potenzreihen (egal ob formal oder konvergent) ersetzt: Da Potenzreihen *unendliche* Summen sind, während bei Linearkombinationen nur *endliche* Summen erlaubt sind, bilden nun die x-Potenzen kein Erzeugendensystem mehr. Nach dem Basisergänzungssatz, der, auch wenn wir das nicht bewiesen haben, auch für unendlichdimensionale Vektorräume gilt, gibt es eine Menge von Potenzreihen, die zusammen mit der obigen Menge $\mathcal{B}$ eine Basis bilden; explizit angeben konnte diese Menge aber noch niemand, genauso wenig wie eine explizite Basis für einen der Räume $\mathcal{C}^n(\mathbb{R},\mathbb{R})$.

Kehren wir also zurück zum überschaubareren endlichdimensionalen Fall, und beweisen wir dort zunächst die anschaulich fast selbstverständliche Aussage, daß jede Basis eines n-dimensionalen Vektorraums aus n Vektoren besteht. Dazu benötigen wir eine leichte Verschärfung

des Basisergänzungssatzes; er geht zurück auf ERNST STEINITZ, den wir bereits von der Körperdefinition aus §1*b)* kennen:

Austauschsatz von STEINITZ: M sei eine endliche linear unabhängige Teilmenge des endlichdimensionalen Vektorraums V, und $\mathcal{B}$ sei eine Basis von V. Dann gibt es eine Teilmenge $\mathcal{B}'$ von $\mathcal{B}$, so daß $M \cup \mathcal{B}'$ eine Basis von V ist. Diese hat genauso viele Elemente wie $\mathcal{B}$.

Mit anderen Worten: Man kann Vektoren aus $\mathcal{B}$ finden, die sich Stück für Stück gegen die Vektoren aus M *austauschen* lassen.

Der *Beweis* ist dem des Basisergänzungssatzes sehr ähnlich; mit Rücksicht auf die Anzahlaussage führen wir ihn aber durch Induktion nach der Elementanzahl m von M.

Für $m = 0$ ist $M = \emptyset$ und wir setzen einfach $\mathcal{B}' = \mathcal{B}$.

Für $m \geq 1$ entfernen wir einen Vektor $\vec{v}$ aus M und wenden den Satz auf die Menge $M' = M \setminus \{\vec{v}\}$ an. Für diese gilt er nach Induktionsannahme, es gibt also eine Teilmenge $\mathcal{C}'$ von $\mathcal{B}$, so daß $\mathcal{C} = M' \cup \mathcal{C}'$ eine Basis von V ist mit gleicher Elementanzahl wie $\mathcal{B}$. Bezüglich dieser Basis habe $\vec{v}$ die Darstellung

$$\vec{v} = \lambda_1\vec{v}_1 + \cdots + \lambda_{m-1}\vec{v}_{m-1} + \mu_1\vec{c}_1 + \cdots + \mu_r\vec{c}_r\,,$$

wobei $M' = \{\vec{v}_1, \ldots, \vec{v}_{m-1}\}$ und $\mathcal{C}' = \{\vec{c}_1, \ldots, \vec{c}_r\}$ sein soll.

Da $M = M' \cup \{\vec{v}\}$ linear unabhängig ist, muß in dieser Darstellung mindestens ein $\vec{\mu}_i$ von Null verschieden sein. Daher läßt sich der zugehörige Vektor $\vec{c}_i$ als Linearkombination aus den restlichen $\vec{c}_j$, den $\vec{v}_\ell$ und dem Vektor $\vec{v}$ schreiben, d.h. auch die durch den *Austausch* von $\vec{c}_i$ durch $\vec{v}$ entstehende Menge

$$M' \cup \big(\mathcal{C}' \setminus \{\vec{c}_i\}\big) \cup \{\vec{v}\} = M \cup \big(\mathcal{C}' \setminus \{\vec{c}_i\}\big)$$

erzeugt ganz V. Diese Menge ist auch linear unabhängig und somit eine Basis, denn ist

$$\alpha\vec{v} + \sum_{\ell=1}^{m-1} \alpha_\ell\vec{v}_\ell + \sum_{\substack{j=1\\ j\neq i}}^{n} \beta_j\vec{c}_j = \vec{0}\,,$$

so muß zunächst α verschwinden, da $\vec{v}$ sonst als Linearkombination der $\vec{v} \in M'$ und der $\vec{c}_j$ mit $j \neq i$ dargestellt werden könnte, was wir oben durch die Wahl eines i mit $\mu_i \neq 0$ ausgeschlossen haben. Also steht hier nur eine Linearkombination von Elementen einer Basis, so daß alle α_ℓ und β_j verschwinden müssen. Mit

$$\mathcal{B}' = \left(\mathcal{C}' \setminus \{\vec{c}_i\}\right)$$

ist somit die Behauptung des Satzes erfüllt. ■

Aus dem STEINITZschen Austauschsatz folgt

Satz: *a)* Jede Basis $\mathcal{B}$ eines n-dimensionalen Vektorraums V besteht aus n Vektoren.
b) Jede Teilmenge von V mit mehr als n Elementen ist linear abhängig.
c) Keine Teilmenge von V mit weniger als n Elementen ist ein Erzeugendensystem.

Beweis: a) Da V die Dimension n hat, gibt es ein Erzeugendensystem $M = \{\vec{v}_1, \ldots, \vec{v}_n\}$ mit n-Elementen, aber keines mit weniger Elementen. Also ist M ein minimales Erzeugendensystem und somit eine Basis.

Nun sei $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_m\}$ irgendeine andere Basis von V. Nach dem Austauschsatz läßt sich M zu einer Basis von V ergänzen, die genauso viele Elemente hat wie $\mathcal{B}$. Da es keine Basis geben kann, die M echt enthält, muß M genauso viele Elemente enthalten wie $\mathcal{B}$, also n.

b) Jede linear unabhängige Teilmenge läßt sich zu einer Basis ergänzen, und jede Basis besteht aus n Vektoren. Also kann eine linear unabhängige Menge höchstens n Vektoren enthalten.

c) Das ist die Definition der Dimension. ■

Nach diesem Satz läßt sich die Dimension eines Vektorraums einfach dadurch bestimmen, daß man eine Basis findet und deren Elemente zählt. Insbesondere hat $\mathbb{R}^n$ als $\mathbb{R}$-Vektorraum die Dimension n, da die n Einheitsvektoren eine Basis bilden.

Weniger offensichtlich ist, daß $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum unendlichdimensional ist: Dazu betrachten wir die unendliche Menge M aller Logarithmen

$\ln p$ der Primzahlen. Wäre diese Menge linear abhängig, gäbe es eine nichttriviale Linearkombination

$$\lambda_1 \ln p_1 + \cdots + \lambda_r \ln p_r = 0$$

mit $\lambda_i \in \mathbb{Q}$. Multipliziert man diese Gleichung mit dem Hauptnenner der λ_i, so erhält man eine entsprechende Gleichung mit Koeffizienten $\mu_i \in \mathbb{Z}$. Dann ist

$$\mu_1 \ln p_1 + \cdots + \mu_r \ln p_r = \ln(p_1^{\mu_1} \cdot \cdots \cdot p_r^{\mu_r}) = 0$$

gleichbedeutend mit

$$p_1^{\mu_1} \cdot \cdots \cdot p_r^{\mu_r} = 1\,,$$

was wegen der Eindeutigkeit der Primzerlegung in $\mathbb{Z}$ nur gelten kann, wenn alle μ_i und damit auch alle λ_i verschwinden.

Also ist $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum unendlichdimensional, und dies erklärt, warum Computer so große Schwierigkeiten mit reellen Zahlen haben: Exakt rechnen kann ein Computer nur in Teilmengen von $\mathbb{R}$, die endlichdimensionale $\mathbb{Q}$-Vektorräume sind – und selbst da gibt es zumindest theoretisch noch das Problem der potentiell beliebig großen Zähler und Nenner.

i) Dimensionen und lineare Abbildungen

Als nächstes wollen wir uns mit Dimensionen von Untervektorräumen, insbesondere auch Kernen und Bildern beschäftigen. Anschaulich klar und auch recht einfach zu beweisen ist der folgende

Satz: Für einen echten Untervektorraum $U < V$ eines endlichdimensionalen Vektorraums V ist $\dim U < \dim V$.

Beweis: Eine Basis von U ist auch in V linear unabhängig, läßt sich also ergänzen zu einer Basis von V. Da die Dimension eines Vektorraums gleich der Elementanzahl einer beliebigen Basis ist, folgt sofort, daß $\dim U \leq \dim V$ sein muß, und wenn beide gleich sind, ist $U = V$. ■

Hier haben wir ganz wesentlich benutzt, daß V endlichdimensional ist; in einen unendlichdimensionalen Vektorraum gibt es stets Untervektorräume, die ebenfalls unendlichdimensinal sind, im Vektorraum V

aller reeller Polynome in x beispielsweise den Untervektorraum aller Polynome in x^2.

Satz: Für endlichdimensionale Vektorräume V, W und eine lineare Abbildung $\varphi\colon V \to W$ ist $\dim \operatorname{Bild} \varphi = \dim V - \dim \operatorname{Kern} \varphi$.

Beweis: $\vec{b}_1, \dots, \vec{b}_r$ sei eine Basis von Kern φ; falls φ injektiv ist, setzen wir $r = 0$. Nach dem Basisergänzungssatz oder (falls $r = 0$) wegen der Existenz von Basen lassen sich dann $n - r$ Vektoren $\vec{b}_{r+1}, \dots, \vec{b}_n$ finden mit $n = \dim V$, so daß $\vec{b}_1, \dots, \vec{b}_n$ eine Basis von V ist.

Das Bild eines beliebigen Vektors $\vec{v} = \lambda_1 \vec{b}_1 + \cdots + \lambda_n \vec{b}_n$ ist dann

$$\varphi(\vec{v}) = \lambda_1 \varphi(\vec{b}_1) + \cdots + \lambda_n \varphi(\vec{b}_n) = \lambda_{r+1} \varphi(\vec{b}_{r+1}) + \cdots + \lambda_n \varphi(\vec{b}_n),$$

da $\vec{b}_1 \dots, \vec{b}_r$ ja auf den Nullvektor abgebildet werden. Also wird Bild φ von den Vektoren $\varphi(\vec{b}_{r+1}), \dots, \varphi(\vec{b}_n)$ erzeugt.

Diese Vektoren sind auch linear unabhängig in W, denn ist

$$\lambda_{r+1} \varphi(\vec{b}_{r+1}) + \cdots + \lambda_n \varphi(\vec{b}_n) = \varphi(\lambda_{r+1} \vec{b}_{r+1} + \cdots + \lambda_n \vec{b}_n) = \vec{0},$$

so liegt $\lambda_{r+1} \vec{b}_{r+1} + \cdots + \lambda_n \vec{b}_n$ im Kern von φ.

Im Fall einer injektiven Abbildung ist $\lambda_{r+1} \vec{b}_{r+1} + \cdots + \lambda_n \vec{b}_n$ daher gleich dem Nullvektor, und damit müssen alle $\lambda_i = 0$ sein, denn die $\vec{b}_i$ sind als Basisvektoren insbesondere linear unabhängig.

Falls φ nicht injektiv ist, können wir nur sagen, daß der Vektor $\lambda_{r+1} \vec{b}_{r+1} + \cdots + \lambda_n \vec{b}_n$ im Kern von φ liegt; er ist also darstellbar als Linearkombination der Basisvektoren $\vec{b}_1, \dots, \vec{b}_r$ des Kerns:

$$\lambda_{r+1} \vec{b}_{r+1} + \cdots + \lambda_n \vec{b}_n = \lambda_1 \vec{b}_1 + \cdots + \lambda_r \vec{b}_r .$$

Auch daraus folgt wegen der linearen Unabhängigkeit der $\vec{b}_i$, daß alle λ_i Null sein müssen.

Damit ist $\{\varphi(\vec{b}_{r+1}), \dots, \varphi(\vec{b}_n)\}$ eine Basis von Bild φ, d.h.

$$\dim \operatorname{Bild} \varphi = n - r = \dim V - \dim \operatorname{Kern} \varphi,$$

wie behauptet. ■

Wir werden diese Aussage im folgenden als *Dimensionsformel* bezeichnen. Da wir hier mit Dimensionen rechnen, ist klar, daß sie nicht auf unendlichdimensionale Vektorräume verallgemeinert werden kann: Sind etwa sowohl V als auch Kern φ unendlichdimensional, kann Bild φ jede beliebige Dimension haben – einschließlich null und unendlich. Der sogenannte *Homomorphiesatz* macht eine genauere Aussage über Bild φ, die auch für unendlichdimensionale Vektorräume gilt. Da wir die zu seiner Formulierung benötigten Begriffe nur teilweise kennen und für diese Vorlesung auch nicht brauchen, sei auf Einzelheiten verzichtet.

Korollar: Eine lineare Selbstabbildung $\varphi\colon V \to V$ eines endlichdimensionalen Vektorraums V ist genau dann injektiv, wenn sie surjektiv ist.

Beweis: φ ist genau dann injektiv, wenn $\dim \operatorname{Kern} \varphi = 0$ ist und genau dann surjektiv, wenn $\dim \operatorname{Bild} \varphi = \dim V$ ist. Diese beiden Dimensionsaussagen sind nach dem gerade bewiesenen Satz äquivalent. ■

Man beachte, daß es in diesem Korollar sehr wesentlich ist, daß wir von einem *endlichdimensionalen* Vektorraum ausgehen: Für den Vektorraum V aller reeller Polynome ist die Abbildung

$$\varphi\colon V \to V; \quad \sum_{i=0}^{d} a_i x^i \mapsto \sum_{i=0}^{d} a_i x^{2i}$$

linear *(warum?)* und injektiv, aber nicht surjektiv. Umgekehrt ist die Ableitung

$$\psi\colon V \to V; \quad f \mapsto f'$$

linear und surjektiv, aber nicht injektiv.

§3: Vektorräume und endliche Körper

Bislang hatten wir in fast allen Beispielen nur Vektorräume über dem Körper der reellen Zahlen betrachtet; in der Informationsverarbeitung treten aber oftmals auch Probleme auf, für die Vektorräume über endlichen Körpern nützlich sind. Als einfachstes Beispiel kennen wir bereits aus §1*e)* den Körper $\mathbb{F}_2 = \{0, 1\}$ der Bits; erstes Thema dieses Paragraphen sind Vektorräume über $\mathbb{F}_2$.