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

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

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 $\lambda_1, \ldots, \lambda_r$ und $\mu_1, \ldots, \mu_s$, 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

$$P = a_0 + a_1 x + a_2 x^2 + \cdots + a_d x^d$$

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

$$\lambda_0 + \lambda_1 x + \cdots + \lambda_n x^n \,,$$

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 μ_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

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

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 $\dim \text{Bild}\,\varphi = \dim V - \dim \text{Kern}\,\varphi$.

Beweis: $\vec{b}_1, \ldots, \vec{b}_r$ 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 $\vec{b}_{r+1}, \ldots, \vec{b}_n$ finden mit $n = \dim V$, so daß $\vec{b}_1, \ldots, \vec{b}_n$ 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 $\vec{b}_1 \ldots, \vec{b}_r$ ja auf den Nullvektor abgebildet werden. Also wird Bild φ von den Vektoren $\varphi(\vec{b}_{r+1}), \ldots, \varphi(\vec{b}_n)$ 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 $\vec{b}_1, \ldots, \vec{b}_r$ 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 $\{\varphi(\vec{b}_{r+1}), \ldots, \varphi(\vec{b}_n)\}$ eine Basis von Bild φ, d.h.

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

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 $\dim \text{Kern}\,\varphi = 0$ ist und genau dann surjektiv, wenn $\dim \text{Bild}\,\varphi = \dim V$ 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.

§ 2: 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 $\mathbb{F}_2$.

a) Bitfolgen als Vektoren

Mit einem einzigen Bit läßt sich nicht viel Information darstellen und verarbeiten; interessant wird es erst mit Bitfolgen. Natürlich können wir Folgen von N Bits als Elemente des Vektorraums $\mathbb{F}_2^N$ betrachten. Da im Körper $\mathbb{F}_2$ die Summen $0+0$ und $1+1$ beide gleich 0 sind, hat dieser Vektorraum die Eigenschaft

$$\vec{v} + \vec{v} = \vec{0} \quad \text{für alle } \vec{v} \in \mathbb{F}_2^N \ ,$$

jeder Vektor ist also zu sich selbst invers, und genau wie auch in $\mathbb{F}_2$ gibt es keinen Unterschied zwischen plus und minus.

Der Vektorraum $\mathbb{F}_2^N$ hat eine sehr einfache Struktur: Die Vektoraddition ist in jeder Komponente einfach die logische Antivalenz, und bitweise logische Antivalenz für ganze Wörter gehört zu den Grundbefehlen der meisten Prozessoren und auch Programmiersprachen. Bei einer Maschine mit 32 Bit-Prozessor läßt sich also eine Vektoraddition in $\mathbb{F}_2^{32}$ mit einem einzigen Befehl ausführen; in C oder C++ wäre der entsprechende Ausdruck gleich `a^b`.

Noch einfacher ist die Multiplikation mit einem Skalar, denn es gibt nur zwei Skalare: Multiplikation mit Eins ändert nichts, Multiplikation mit Null hat immer die Bitfolge aus lauter Nullen als Ergebnis.

Das Rechnen in $\mathbb{F}_2^N$ ist also sehr einfach und effizient, und es kann schon in dieser ganz trivialen Form auch nützlich sein:

Eine Anwendung ist etwa die Fehlererkennung in der Informationsübertragung: Dazu werden Daten beispielsweise oft zusammen mit einem „Paritätsbit“ übertragen, d.h. jede Folge von sieben Bits wird um ein achtes „Prüfbit“ erweitert, so daß im entstehenden Byte immer eine gerade Anzahl von Einsen vorkommt; es hat also gerade Parität. Vor der Übertragung wird also auf jede Folge von sieben Bit die lineare Abbildung

$$\varphi\colon \begin{cases} \mathbb{F}_2^7 \to \mathbb{F}_2^8 \\ (x_1, \ldots, x_7) \mapsto (x_1, \ldots, x_7,\ x_1 + \cdots + x_7) \end{cases}$$

angewendet. Auch die Überprüfung, ob ein gegebenes Byte tatsächlich gerade Parität hat, läßt sich mit einer linearen Abbildung realisieren:

Die Bytes mit gerader Parität sind offenbar gerade die aus dem Kern der linearen Abbildung

$$\psi\colon \begin{cases} \mathbb{F}_2^8 \to \mathbb{F}_2 \\ (x_1, \ldots, x_8) \mapsto (x_1 + \cdots + x_8) \end{cases}$$

Mit etwas mehr Aufwand kann man Fehler nicht nur erkennen, sondern auch korrigieren: Als Beispiel dafür konstruieren wir eine Abbildung

$$\varphi\colon \mathbb{F}_2^{nm} \to \mathbb{F}_2^{(n+1)(m+1)}$$

wie folgt: Wir schreiben die Elemente von $\mathbb{F}_2^{nm}$ in der Form

$$X = \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1n} \\ x_{21} & x_{22} & \ldots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \ldots & x_{mn} \end{pmatrix}$$

und bilden ein solches Element ab auf

$$\varphi(X) = \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1n} & x_{1,n+1} \\ x_{21} & x_{22} & \ldots & x_{2n} & x_{2,n+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{m1} & x_{m2} & \ldots & x_{mn} & x_{m,n+1} \\ x_{m+1,1} & x_{m+1,2} & \ldots & x_{m+1,n} & x_{m+1,n+1} \end{pmatrix} ,$$

wobei

$$x_{i,n+1} = \sum_{j=1}^{n} x_{ij} \quad \text{und} \quad x_{m+1,j} = \sum_{i=1}^{m} x_{ij}$$

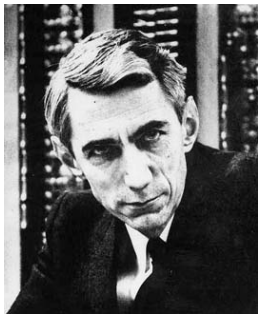
sein soll. Es braucht uns dabei nicht stören, daß $x_{n+1,m+1}$ hier auf zwei verschiedene Weisen definiert ist: Wie man sich leicht überlegt, führen beide Definitionen ausgeschrieben zu

$$x_{n+1,m+1} = \sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij} \ .$$

Hier gibt es also $n+m+1$ Prüfbits; in $\varphi(X)$ sind alle Zeilensummen und alle Spaltensummen Null. Falls nun durch einen Übertragungsfehler das Bit x_{ij} (und sonst keines) verfälscht wurde, ist genau in der i-ten Zeile

und der j-ten Spalte die entsprechende Summe gleich eins, es ist also klar, daß x_{ij} korrigiert werden muß.

Mit entsprechend größerem Aufwand lassen sich auch mehr Fehler korrigieren; tatsächlich können nach zwei Sätzen von SHANNON, wenn man nur genügend lange Codewörter zuläßt, mit beliebig geringem (relativem) Aufwand beliebig hohe (vorgegebene) Fehlerraten korrigiert werden – vorausgesetzt natürlich, diese Raten sind echt kleiner als $1/2$. Bei einer Fehlerrate von $1/2$ kommen nur Zufallsbits ohne jeglichen Informationsgehalt an.



CLAUDE ELWOOD SHANNON (1916–2001) wurde in Petoskey im US-Bundesstaat Michigan geboren; 1936 verließ er die University of Michigan mit sowohl einem Bachelor der Mathematik als auch einem Bachelor der Elektrotechnik, um am M.I.T. weiterzustudieren. Seine 1938 geschriebene Diplomarbeit *A symbolic analysis of relay and switching circuits* bildet die Grundlage der digitalen Informationsverarbeitung auf der Grundlage der hier entwickelten Schaltlogik; seine Dissertation 1940 befaßte sich mit Anwendungen der Algebra auf die MENDELSchen Gesetze. Danach arbeitete er bis 1956 bei den Bell Labs, wo er während des zweiten Weltkriegs insbesondere über die Sicherheit kryptographischer Systeme forschte. Seine *Mathematical theory of cryptography* wurde aus Geheimhaltungsgründen erst 1949 zur Veröffentlichung freigegeben. Seine wohl bekannteste Arbeit ist die 1948 erschienene *Mathematical theory of communication,* in der er die fehlerfreie Übertragung von Nachrichten über einen gestörten Kanal untersuchte. Von 1956 bis zu seiner Emeritierung 1978 lehrte er am M.I.T., das er dadurch zur führenden Universität auf dem Gebiet der Informationstheorie und Kommunikationstechnik machte. Zu seinen zahlreichen Arbeiten zählt auch eine über die mathematische Theorie der Jongliermuster, anhand derer Jongleure eine Reihe neuer Muster gefunden haben; auch konstruierte er mehrere Jonglierroboter.

Beim nächsten Beispiel geht es um die Sicherung von Information gegen *absichtliche* Manipulation und unberechtigtes Mithören:

Während des kalten Kriegs hielten viele (wohl zu Recht) die Gefahr eines Atomkriegs aus Versehen für erheblich größer als die eines absichtlichen Atomkriegs. Um ersteren weniger wahrscheinlich zu machen, einigten sich die beiden Großmächte im Juni 1963 in Genf darauf, das sogenannte *Rote Telephon* einzurichten; es funktioniert seit dem 30. August 1963.

Natürlich handelt es sich dabei nicht wirklich um ein Telephon, denn zu keinem Zeitpunkt des kalten Krieges reichten die Sprachkenntnisse eines amerikanischen Präsidenten oder eines Generalsekretärs der KPdSU auch nur für ein direktes Gespräch über das Wetter.

Tatsächlich war das *Rote Telephon* eine Fernschreibverbindung mit je vier Fernschreibern (geliefert von Siemens Mannheim) an beiden Enden: jeweils zwei mit lateinischem und zwei mit kyrillischem Alphabet. Bislang verbrachten sie ihre meiste Zeit damit, stündliche Testnachrichten zu drucken wie amerikanische Baseball-Ergebnisse oder TURGENJEWs *Aufzeichnungen einer Jägers.*

Aus Sicherheitsgründen wurden zwei Leitungen eingerichtet, eine entlang der Route Washington-London-Kopenhagen-Stockholm-Helsinki-Moskau, die andere via Tanger. Natürlich war es unmöglich, diese Leitungen auf ihrer ganzen Länge zu überwachen, so daß niemand ausschließen konnte, daß irgendwo zwischen Moskau und Washington eine vertrauliche Kommunikation abgehört oder – schlimmer noch – eine gefälschte Nachricht eingespielt wurde.

Zum Schutz davor wurde die gesamte Kommunikation verschlüsselt. Wegen der hohen Sicherheitsanforderungen konnte dazu allerdings keines der üblicherweise in heutiger Office-Software eingebauten Verfahren verwendet werden: Wer noch irgendwelche Illusionen über die Sicherheit gängiger kommerzieller Programme hat, sollte unter

`http://pwcrack.com`

nachlesen, für welche vergleichsweise bescheidenen Beträge spezialisierte Unternehmen dazu bereit sind, „vergessene“ Paßwörter zu rekonstruieren.

Das *Rote Telephon* benutzte stattdessen eine Variante eines alten, absolut sicheren, Verschlüsselungsverfahrens, des sogenannten *one time pads:* Von Zeit zu Zeit tauschten die beiden Seiten per Kurier Magnetbänder mit zufallserzeugten Bitfolgen aus. Jedesmal, wenn eine Nachricht übermittelt werden sollte, übersetzte der Fernschreiber diese in eine Bitfolge, d.h. in einen Vektor $\vec{v}$ aus einem Vektorraum $\mathbb{F}_2^N$. Aus den ersten N bislang noch nicht benutzten Bits auf dem Magnetband wurde dazu ein

weiterer Vektor $\vec{w} \in \mathbb{F}_2^N$ gebildet, und tatsächlich übertragen wurde die Summe $\vec{s} = \vec{v} + \vec{w}$.

Am anderen Ende der Leitung, wo eine Kopie des Magnetbands vorlag, war $\vec{w}$ bekannt, so daß die Nachricht

$$\vec{v} = \vec{v} + \vec{0} = \vec{v} + (\vec{w} + \vec{w}) = (\vec{v} + \vec{w}) + \vec{w} = \vec{s} + \vec{w}$$

rekonstruiert werden konnte.

Ein Lauscher ohne Magnetband konnte nur die Länge N der Nachricht ermitteln, was bei den seitenlangen in Diplomatensprache formulierten Texten, die über diese Leitung liefen, so gut wie keine konkrete Information lieferte. In der Tat können auch schon sehr kurze Nachrichten gleicher Länge völlig verschiedenen Inhalt haben: Im Deutschen etwa besteht der Satz „Herzlichen Glückwunsch zu Ihrem sehr guten Klausurergebnis!“ aus genauso vielen Zeichen wie „Mit 3 von 2000 Punkten haben Sie das schlechteste Ergebnis.“ Entsprechend hat auch jemand, der irgendeinen Vektor $\vec{s}$ in die Leitung einspielt, so gut wie keine Chance, daß nach Addition von $\vec{w}$ daraus verständlicher Text wird, so daß die Manipulation mit an Sicherheit grenzender Wahrscheinlichkeit entdeckt wird.

Diese Art der Kommunikation ist also sehr sicher, aber leider auch sehr aufwendig: Wer einfach ein Buch im Internet bestellen will, hat üblicherweise keine Möglichkeit, vorher über Kurier ein Magnetband oder eine CD-ROM mit dem Versandhaus auszutauschen, bevor er seine Kontendaten dorthin schickt. Für Alltagsanwendungen braucht man daher Verfahren, die einfacher anwendbar sind. Leider sind die wirklich guten darunter mathematisch deutlich anspruchsvoller als der *one time pad;* zwei davon werden wir im Laufe dieses Paragraphen noch kennenlernen.

b) Körper von Primzahlordnung

Der Körper mit zwei Elementen ist nur einen von vielen endlichen Körpern; beispielsweise gibt es zu jeder Primzahl p einen solchen Körper; wir bezeichnen ihn in Analogie zu $\mathbb{F}_2$ mit $\mathbb{F}_p$.

Als Menge ist $\mathbb{F}_p \underset{\text{def}}{=} \{0, 1, \ldots, p-1\}$; Addition und Multiplikation werden definiert durch die Vorschriften

$$a \oplus b \underset{\text{def}}{=} (a + b) \bmod p \quad \text{und} \quad a \odot b \underset{\text{def}}{=} ab \bmod p\,,$$

d.h. für führen zunächst die entsprechenden Operationen für ganze Zahlen aus und betrachten dann den Divisionsrest des Ergebnisses bei Division durch die Primzahl p. Für $p = 2$ führt das auf die inzwischen wohlbekannten Rechenoperationen von $\mathbb{F}_2$.

Da das Kommutativ- und das Assoziativgesetz für Addition und Multiplikation ganzer Zahlen gelten und zwei gleiche Zahlen insbesondere den gleichen Divisionsrest modulo p haben, gelten diese Gesetze auch für $\oplus$ und $\odot$; aus demselben Grund gilt auch das Distributivgesetz, und 0 und 1 sind Neutralelemente bezüglich $\oplus$ und $\odot$.

Das zu a inverse Element bezüglich der Addition ist für $a = 0$ natürlich a selbst, ansonsten $p - a$, denn

$$a \oplus (p - a) = a + (p - a) \bmod p = p \bmod p = 0\,.$$

Multiplikative Inverse lassen sich nicht so leicht finden, aber immerhin läßt sich relativ einfach sehen, daß die existieren: Für $a \in \mathbb{F}_p \setminus \{0\}$ betrachten wir die sämtlichen Elemente $a \cdot x$ mit $x \in \mathbb{F}_p$. Ist $a \cdot x = a \cdot y$, so ist $a \cdot (x - y) = 0$, d.h. $a(x - y)$ ist durch p teilbar. Da p eine Primzahl ist (das verwenden wir hier zum ersten Mal!), muß dann auch mindestens einer der beiden Faktoren durch p teilbar sein. a aber ist eine Zahl zwischen 1 und $p - 1$, also sicherlich nicht durch p teilbar. Somit teilt p die Differenz $x - y$. Da x, y als Elemente von $\mathbb{F}_p$ höchstens gleich $p - 1$ sind, hat ihre Differenz höchstens Betrag $p - 1$, also kann sie nur dann durch p teilbar sein, wenn sie verschwindet. Somit gilt: Für $x, y \in \mathbb{F}_p$ ist $a \cdot x = a \cdot y$ genau dann, wenn $x = y$. Die p Elemente $a \cdot x$ sind somit allesamt verschieden; da es in $\mathbb{F}_p$ nur p Elemente gibt, läßt sich also jedes von diesen in der Form $a \cdot x$ mit einem geeigneten $x \in \mathbb{F}_p$ schreiben. Dies gilt insbesondere für die Eins, und damit ist auch die Existenz multiplikativer Inverser als letztes des Körpereaxiome nachgewiesen.

$\mathbb{F}_p$ ist also in der Tat ein Körper.