Leitfaden: Lineare Algebra

Mit $\mathbb{Q}$ wird die Menge der rationalen Zahlen bezeichnet, also der Brüche $\frac{a}{b}$, wobei a, b ganze Zahlen sind mit $b \neq 0$.

Teil 1. Matrizen.

Eine $(m \times n)$-*Matrix* $A = (a_{ij})_{ij}$ *mit Koeffizienten in* $\mathbb{Q}$ ist eine Abbildung $A\colon \{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\} \to \mathbb{Q}$, die dem Paar (i, j) mit $1 \leq i \leq m$ und $1 \leq j \leq n$ eine Zahl $a_{ij} \in \mathbb{Q}$ zuordnet; man nennt a_{ij} den (i, j)-*Koeffizienten, i* seinen *Zeilenindex, j* seinen *Spaltenindex.* Eine derartige Matrix $A = (a_{ij})_{ij}$ wird üblicherweise in der Form

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

notiert. Ist $1 \leq i \leq m$, so nennt man

$$\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix}$$

die i-*te Zeile* von A; entsprechend ist

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \cdots \\ a_{mj} \end{bmatrix}$$

die j-*te Spalte* von A. Sei $M(m \times n, \mathbb{Q})$ die Menge der $(m \times n)$-Matrizen mit Koeffizienten in $\mathbb{Q}$. (Wir betrachten im Augenblick nur Matrizen mit Koeffizienten in $\mathbb{Q}$, später werden wir mit Matrizen mit Koeffizienten in einem beliebigen Ring arbeiten.)

Addition von Matrizen. Sind $A = (a_{ij})_{ij}, B = (b_{ij})_{ij}$ zwei $(m \times n)$-Matrizen, so definiert man $A + B = (c_{ij})_{ij}$ mit $c_{ij} = a_{ij} + b_{ij}$. Offensichtlich gelten folgende Rechenregeln: Diese Addition ist assoziativ (d.h. $(A + B) + C = A + (B + C)$ für alle $(m \times n)$-Matrizen A, B, C) und kommutativ (d.h. $A + B = B + A$ für alle $(m \times n)$-Matrizen A, B).

Skalar-Multiplikation von Matrizen. Ist $\lambda \in \mathbb{Q}$ und $A = (a_{ij})_{ij}$ eine $(m \times n)$-Matrix, so definiert man $\lambda A = (d_{ij})_{ij}$ mit $d_{ij} = \lambda a_{ij}$. Es gelten folgende Rechenregeln: $(\lambda + \lambda')A = \lambda A + \lambda' A$, $\lambda(A + A') = \lambda A + \lambda A'$, $(\lambda\lambda')A = \lambda(\lambda' A)$ für alle $\lambda, \lambda' \in \mathbb{Q}$ und alle $(m \times n)$-Matrizen A, A'.

Null-Matrix. Die $(m \times n)$-Matrix, deren Koeffizienten alle Null sind, wird die *Null-Matrix* genannt, sie wird mit 0_{mn} oder einfach mit 0 bezeichnet. Es ist $A + 0_{mn} = A = 0_{mn} + A$, für alle $A \in M(m \times n, \mathbb{Q})$.

(1.1) Elementare Zeilen-Umformungen.

Typ I: Multiplikation der i-ten Zeile mit einem Faktor $\lambda \neq 0$.
Typ II: Addition der j-ten Zeile zur i-ten Zeile; dabei sei $i \neq j$.

Als Hintereinanderschaltung elementarer Zeilen-Umformungen erhalten wir weitere Zeilen-Umformungen:

Typ III: Addition des μ-fachen der j-ten Zeile zur i-ten Zeile, dabei sei $i \neq j$.
Typ IV: Vertauschen der i-ten und der j-ten Zeile.

Satz. *Durch Zeilen-Umformungen vom Typ III und IV kann jede Matrix in* **Zeilenstufenform** *gebracht werden.*

Zeilenstufenform bedeutet: Es gibt $1 \leq j_1 < j_2 < \cdots < j_r \leq n$ mit:

$$\begin{aligned} b_{i,j_i} &\neq 0 \quad \text{falls} \quad 1 \leq i \leq r, \\ b_{ij} &= 0 \quad \text{falls} \quad j < j_i \quad \text{und} \quad 1 \leq i \leq r, \\ b_{ij} &= 0 \quad \text{falls} \quad r < i. \end{aligned}$$

Die Paare (i, j_i) heißen *Pivot-Positionen,* die Koeffizienten b_{i,j_i} heißen *Pivot-Koeffizienten.*

(1) Mit Hilfe von Umformungen vom Typ I kann man zusätzlich erreichen, daß die Pivot-Koeffizienten 1 sind.

(2) Mit Hilfe von weiteren Umformungen vom Typ III kann man zusätzlich erreichen, daß gilt

$$b_{t,j_i} = 0 \quad \text{falls} \quad t \neq i \quad \text{und} \quad 1 \leq i \leq r.$$

Eine Matrix in Zeilenstufenform mit den Zusatzbedingungen (1) und (2) ist eine **Schubert-Normalform.** *Zu jeder Matrix A gibt es eine und nur eine Schubert-Normalform B, die durch elementare Zeilen-Umformungen aus A erhalten werden kann.*

Spalten-Umformungen. Man kann entsprechend auch elementare Spalten-Umformungen definieren. Durch elementare Zeilen-Umformungen und elementare Spalten-Umformungen läßt sich jede Matrix in folgende Form bringen:

$$\left[\begin{array}{ccc|c} 1 & & & \\ & 1 & 0 & \\ & \ddots & & 0 \\ & 0 & & \\ & & 1 & \\ \hline & 0 & & 0 \end{array}\right]$$

(1.2) Lineare Gleichungssysteme (I).

Ein lineares Gleichungssystem mit Koeffizienten in $\mathbb{Q}$ ist von folgender Form:

$$\sum_{j=1}^{n} a_{ij} X_j = b_i \quad \text{mit} \quad 1 \le i \le m$$

dabei sind $a_{ij}, b_i \in \mathbb{Q}$, für alle $1 \le i \le m,\ 1 \le j \le n$.

Wir bezeichnen mit $\text{Lös}(A, b)$ die Menge der n-Tupel $(z_1, \ldots, z_n)$ rationaler Zahlen, für die gilt:

$$\sum_{j=1}^{n} a_{ij} z_j = b_i \quad \text{für alle} \quad 1 \le i \le m.$$

Wir bezeichnen mit $[A, b]$ die *erweiterte Koeffizientenmatrix*, es ist

$$[A, b] = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{bmatrix}$$

Lemma. *Die Matrix $[A', b']$ entstehe aus der Matrix $[A, b]$ durch elementare Zeilen-Umformungen. Dann ist* $\text{Lös}(A, b) = \text{Lös}(A', b')$.

Folgerung: Um ein lineares Gleichungssystem zu lösen, genügt es, die erweiterte Koeffizientenmatrix in Zeilenstufenform mit Pivot-Koeffizienten 1 zu bringen und dieses modifizierte Gleichungssystem zu lösen.

Lösung linearer Gleichungssysteme in Zeilenstufenform mit Pivot-Koeffizienten 1. Es sei $[A, b]$ die erweiterte Koeffizientenmatrix. Sie sei in Zeilenstufenform mit Pivot-Koeffizienten 1. Die Pivot-Positionen seien $(1, j_1), \ldots, (r, j_r)$.

Fall 1. $j_r = n + 1$. Dann gibt es **keine Lösung.**

Fall 2. Es ist $j_r \le n$. Wir bezeichnen die X_{j_i} als *gebundene Variablen,* die X_j mit $j \ne j_i$ für alle i (und $1 \le j \le n$) als *freie Variablen.*

Wähle beliebige Zahlen z_j, falls X_j eine freie Variable ist. Ist X_{j_i} eine gebundene Variable, so berechne z_{j_i} mit Hilfe der i-ten Gleichung, rückwärts, das heißt zuerst für $i = r$, dann für $i = r - 1$, und so weiter, schließlich für $i = 2$ und $i = 1$, und zwar wie folgt:

$$z_{j_i} = -\sum_{j_i < j} a_{ij} z_j + b_i;$$

verwendet wird hier nur, daß bei der Berechnung von z_{j_i} schon alle Werte z_j mit $j > j_i$ gegeben sind (die an den freien Positionen durch Vorgabe, die an den gebundenen Positionen durch Berechnung).

Wir sehen: es gibt im Fall 2 immer Lösungen, und zwar **entweder unendlich viele Lösungen** (falls es freie Variablen gibt, falls also $r < n$ gilt) **oder eine einzige Lösung** (falls es keine freie Variablen gibt, falls also $r = n$ gilt).

Dieses Verfahren wird manchmal Gauß-Elimination genannt (GAUSS 1777-1855), es war aber vor 2000 Jahren schon den Chinesen bekannt und wird dort die Fang-Cheng-Methode genannt.

(1.3) Determinanten.

Eine *Permutation* σ der Menge S ist eine bijektive Abbildung $\sigma\colon S \to S$. Ist S eine endliche Menge, so reicht es zu verlangen, daß σ injektiv ist, daß also gilt: ist $i \neq i'$, so ist $\sigma(i) \neq \sigma(i')$. Insbesondere interessieren uns Permutationen der Menge $\{1, 2, \ldots, n\}$, wir notieren eine derartige Permutation zum Beispiel in der Form $\begin{bmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{bmatrix}$. Die Menge aller Permutationen der Menge $\{1, 2, \ldots, n\}$ bezeichnen wir mit S_n. Ist $\sigma \in S_n$, so nennen wir jedes Paar (i, j) mit $1 \leq i < j \leq n$ und $\sigma(j) < \sigma(i)$ eine *Fehlstellung.* Ist die Anzahl der Fehlstellungen gerade, so nennt man σ eine *gerade* Permutation, andernfalls eine *ungerade* Permutation. Man schreibt

$$\operatorname{sign}(\sigma) = \begin{cases} 1 & \text{falls } \sigma \text{ gerade} \\ -1 & \text{falls } \sigma \text{ ungerade} \end{cases}.$$

und nennt $\operatorname{sign}(\sigma)$ das *Signum* der Permutation σ. Zur Berechnung des Signums: Sei $f(\sigma)$ die Anzahl der Fehlstellungen. Dann gilt

$$\operatorname{sign}(\sigma) = (-1)^{f(\sigma)} = \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Leibniz-Formel (LEIBNIZ 1646-1716). Sei $A = (a_{ij})_{ij}$ eine $n \times n$-Matrix. Wir setzen

$$\det A = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma)\, a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

Die einzelnen Summanden sind folgendermaßen gebildet: Man bildet Produkte mit n Faktoren, dabei kommt jeweils ein Faktor aus jeder der n Reihen, aber auch jeweils aus jeder der n Spalten. Die Positionen sind also die möglichen Positionen von n Türmen auf dem $(n \times n)$-Schachbrett, die sich gegenseitig nicht schlagen. Eine derartige Verteilung von n Türmen liefert gerade eine Permutation σ, wenn wir notieren, daß in der i-ten Zeile das $\sigma(i)$-te Feld belegt ist. Jedes dieser Produkte ist mit einem Vorzeichen versehen.

Eine $(n \times n)$-Matrix $A = (a_{ij})_{ij}$ heißt *obere Dreiecksmatrix*, falls $a_{ij} = 0$ für $i > j$ gilt; entsprechend spricht man von einer *unteren Dreiecksmatrix*, falls a_{ij} für $i < j$ gilt; man nennt A eine *Diagonalmatrix* falls $a_{ij} = 0$ für $i \neq j$ gilt. Diagonalmatrizen mit $a_{11} = a_{22} = \cdots = a_{nn}$ heißen *Skalarmatrizen.*

Beispiel: *Ist die $(n \times n)$-Matrix $A = (a_{ij})_{ij}$ eine obere Dreiecksmatrix oder eine untere Dreiecksmatrix, so gilt* $\det A = a_{11} a_{22} \cdots a_{nn}$.

Wichtigste Eigenschaft: *Ist A eine quadratische Matrix mit zwei gleichen Zeilen, so ist* $\det A = 0$. Zum Beweis zeigt man, daß in der Leibniz-Formel jeweils zwei Summanden den gleichen Betrag,aber verschiedenes Vorzeichen haben.

Verhalten der Determinanten bei elementaren Zeilen-Unformungen.

Typ I. Entsteht A' aus A dadurch, daß die i-te Zeile von A mit λ multipliziert wird, so ist $\det A' = \lambda \det A$.

Typ II. Entsteht A' aus A dadurch, daß die j-te Zeile von A zur i-ten Zeile addiert wird, so ist $\det A' = \det A$.

Typ III. Entsteht A' aus A dadurch, daß man das λ-fache der j-ten Zeile von A zur i-ten Zeile addiert wird, so ist $\det A' = \det A$.

Typ IV. Entsteht A' aus A dadurch, daß man die i-te Zeile und die j-te Zeile mit $i \neq j$ miteinander vertauscht, so ist $\det A' = -\det A$.

Folgerung. *Die Matrix A' entstehe aus der Matrix A durch s Zeilen-Umformungen der Art III und t Zeilen-Umformungen der Art IV. Dann ist* $\det A' = (-1)^t \det A$. Entsprechendes gilt für Spalten-Umformungen.

Ist A eine Matrix, so entsteht die Matrix A_{ij} durch Streichen der i-ten Zeile und der j-ten Spalte (bei einer $m \times n$ Matrix können wir also A_{ij} für $1 \leq i \leq m$ und $1 \leq j \leq n$ bilden.

Laplace-Entwicklung. (LAPLACE 1749-1827). *Sei A eine $(n \times n)$-Matrix. Für jedes i gilt*

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}.$$

Beweis. Wir betrachten zusätzlich auch die Matrix $B_{ij} = (b_{ij})_{ij}$, deren i-te Zeile die folgende Form hat: an der Stelle (i,j) steht der Koeffizient 1, alle anderen Koeffizienten der i-ten Zeile sind 0, und die restlichen Koeffizienten sind diejenigen der Matrix A (also $b_{st} = a_{st}$ falls $s \neq i$). Es ist

$$\begin{aligned}
\det B_{ij} &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma)\, b_{1,\sigma(1)} \cdots b_{i,\sigma(i)} \cdots b_{n,\sigma(n)} \\
&= \sum_{\sigma \in S_n,\, \sigma(i)=j} \operatorname{sign}(\sigma)\, b_{1,\sigma(1)} \cdots b_{i,\sigma(i)} \cdots b_{n,\sigma(n)} \\
&= \sum_{\sigma \in S_n,\, \sigma(i)=j} \operatorname{sign}(\sigma)\, a_{1,\sigma(1)} \cdots a_{i-1,\sigma(i-1)} \cdot 1 \cdot a_{i+1,\sigma(i+1)} \cdots a_{n,\sigma(n)},
\end{aligned}$$

andererseits können wir B_{ij} durch Vertauschen von Zeilen und Spalten in eine Matrix der folgenden Form überführen:

$$A'_{ij} = \begin{bmatrix} A_{ij} & * \\ 0 & 1 \end{bmatrix}$$

dazu müssen wir die i-te Zeile mit allen nachfolgenden Zeilen vertauschen, insgesamt sind dies $n-i$ Zeilenvertauschungen, und wir müssen die j-te Spalte mit allen nachfolgenden Spalten vertauschen, insgesamt sind dies $n-j$ Spaltenvertauschungen. Es ist also:

$$\det B = (-1)^{n-i+n-j} \det A'_{ij} = (-1)^{i+j} \det A_{ij}.$$

(Wir haben verwendet, daß $(-1)^{n-i+n-j} = (-1)^{i+j}$ gilt.)

Mit diesen Vorbemerkungen ist der Beweis einfach:

$$\begin{aligned}\det A &= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)\, a_{1,\sigma(1)}\cdots a_{i,\sigma(i)}\cdots a_{n,\sigma(n)}\\ &= \sum_{j=1}^{n} \sum_{\sigma\in S_n,\sigma(i)=j} \operatorname{sign}(\sigma)\, a_{1,\sigma(1)}\cdots a_{i,\sigma(i)}\cdots a_{n,\sigma(n)}\\ &= \sum_{j=1}^{n} a_{ij} \sum_{\sigma\in S_n,\sigma(i)=j} \operatorname{sign}(\sigma)\, a_{1,\sigma(1)}\cdots a_{i-1,\sigma(i-1)}\cdot 1\cdot a_{i+1,\sigma(i+1)}\cdots a_{n,\sigma(n)}\\ &= \sum_{j=1}^{n} a_{ij}\det B_{ij} = \sum_{j=1}^{n} a_{ij}(-1)^{i+j}\det A_{ij}\end{aligned}$$

Dies ist die Laplace-Entwicklung nach Zeilen. Es gibt eine entsprechende Laplace-Entwicklung nach Spalten: *Für jedes j gilt*

$$\det A = \sum_{i=1}^{n}(-1)^{i+j}a_{ij}\det A_{ij}.$$

Wofür braucht man Determinanten? Zum Beispiel, wenn man ein lineares Gleichungssystem mit n Zeilen und n Variablen gegeben hat. Erinnerung: Manchmal gibt es keine Lösung, manchmal gibt es viele Lösungen.

Satz. *Ein lineares Gleichungssystem mit n Zeilen und n Variablen besitzt genau dann eine und nur eine Lösung, wenn die Determinante der Koeffizientenmatrix von Null verschieden ist.*

(1.4) Matrizenmultiplikation.

Definition. Sei $A = (a_{ij})_{ij}$ eine $(m\times n)$-Matrix und $B = (b_{jk})_{jk}$ eine $(n\times r)$-Matrix. Das *Produkt* AB ist definiert als $AB = (c_{ik})_{ik}$ mit $c_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk}$, dies ist eine $(m\times r)$-Matrix.

Assoziativität: Sei $A = (a_{ij})_{ij}$ eine $(m\times n)$-Matrix, $B = (b_{jk})_{jk}$ eine $(n\times r)$-Matrix, $C = (c_{kl})_{kl}$ eine $(r\times s)$-Matrix. Dann gilt $(AB)C = A(BC)$. Beweis: Der (i,l)-Koeffizient von $(AB)C$ wie auch von $A(BC)$ ist $\sum_{j,k} a_{ij}b_{jk}c_{kl}$.

Die Matrix $I_n = (\delta_{ij})_{ij}$ mit $\delta_{ii} = 1$ und $\delta_{ij} = 0$, für alle i,j mit $i\neq j$ nennt man *Einheitsmatrix*, es gilt $AI_n = I_nB$ für alle $(m\times n)$-Matrizen A und alle $(n\times r)$-Matrizen B. (Man nennt diese Koeffizienten δ_{ij} auch die Werte der *Kronecker-Delta-Funktion*). Die Skalarmatrizen sind gerade die skalaren Vielfache der Einheitsmatrix.

Eine $(n\times n)$-Matrix A heißt *invertierbar*, wenn es eine Matrix A' mit $AA' = A'A = I_n$ gibt. Man zeigt leicht, daß es zu A höchstens eine derartige Matrix A' geben kann, man nennt sie (wenn sie existiert) *die zu A inverse Matrix* und schreibt A^{-1} statt A'. *Sind A,B invertierbare $(n\times n)$-Matrizen, so ist auch AB invertierbar, und es gilt $(AB)^{-1} = B^{-1}A^{-1}$.*

Sei $E_{ij} = (a_{rs})_{rs}$ mit $a_{ij} = 1$, und $a_{rs} = 0$ falls $(r,s) \neq (i,j)$ (dies ist eine $(n \times n)$-Matrix mit (i,j)-Koeffizienten 1, alle anderen Koeffizienten sind Null). Mit Hilfe der Matrizen E_{ij} definieren wir die **Elementar-Matrizen:** sie unterscheiden sich von der Einheitsmatrix an höchstens einer Stelle:

$S_i(\lambda) = I + (\lambda - 1)E_{ii}$. Ist A eine $(n \times r)$-Matrix, so entsteht $S_i A$ aus A durch Multiplikation der i-ten Zeile mit λ (ist also $\lambda \neq 0$, so ist dies gerade eine elementare Zeilen-Umformung von Typ I). Für $\lambda \neq 0$ ist die Matrix $S_i(\lambda)$ invertierbar und es gilt $S_i(\lambda)^{-1} = S_i(\frac{1}{\lambda})$.

$Q_{ij} = I + E_{ij}$ für $i \neq j$. Ist A eine $(n \times r)$-Matrix, so entsteht $Q_{ij}A$ aus A durch das Addieren der j-ten Zeile von A zur i-ten Zeile (dies ist eine elementare Zeilen-Umformung vom Typ II).

$Q_{ij}(\mu) = I + \mu E_{ij}$ für $i \neq j$. Ist A eine $(n \times r)$-Matrix, so entsteht $Q_{ij}A$ aus A durch Addition des μ-fachen der j-ten Zeile von A zur i-ten Zeile (dies ist gerade eine Zeilen-Umformung von Typ III). Es ist natürlich $Q_{ij} = Q_{ij}(1)$. Jede Matrix der Form $Q_{ij}(\mu)$ ist invertierbar und es gilt $Q_{ij}(\mu)^{-1} = Q_{ij}(-\mu)$.

$P_{ij} = I - E_{ii} - E_{jj} + E_{ij} + E_{ji}$ Ist A eine $(n \times r)$-Matrix, so entsteht $P_{ij}A$ aus A durch Vertauschung der i-ten und der j-ten Zeile (also Zeilen-Umformung vom Typ IV).

Die Formeln

$$Q_{ij}(\mu) = S_j(\tfrac{1}{\mu})Q_{ij}S_j(\mu) \quad \text{für} \quad \mu \neq 0.$$
$$P_{ij} = Q_{ji}(1)Q_{ij}(-1)Q_{ji}(1)S_j(-1)$$

zeigen, wie die Zeilen-Unformungen vom Typ III und IV als Hintereinanderschaltungen von Zeilen-Umformungen vom Typ I und II geschrieben werden können.

Satz. *Jede quadratische Matrix ist Produkt von Elementar-Matrizen.*

Beweis. Sei A eine $(n \times n)$-Matrix. Verwende elementare Zeilen-Umformungen, um die Matrix in obere Dreickecksform zu bringen: seien also $B_1, \ldots, B_s$ invertierbare Elementarmatrizen, so daß $B_s \cdots B_1 A$ eine obere Dreiecksmatrix ist. Verwende elementare Spalten-Umformungen, um eine Diagonalmatrix zu erhalten: seien also $C_1, \ldots, C_s$ invertierbare Elementarmatrizen, so daß $B_s \cdots B_1 A C_1 \cdots C_t$ eine Diagonalmatrix ist. Eine Diagonalmatrix läßt sich immer als Produkt $S_1(\lambda_1) \cdots S_n(\lambda_n)$ schreiben. Also

$$B_s \cdots B_1 A C_1 \cdots C_t = S_1(\lambda_1) \cdots S_n(\lambda_n).$$

Daraus folgt:

$$A = B_1^{-1} \cdots B_s^{-1} S_1(\lambda_1) \cdots S_n(\lambda_n) C_t^{-1} \cdots C_1^{-1},$$

alle Faktoren sind Elementar-Matrizen.

Determinanten-Produktsatz. *Sind A, B zwei $n \times n$ Matrizen, so ist*

$$\det(AB) = \det A \cdot \det B.$$

Beweis: Schreibe A als Produkt von m Elementarmatrizen. Induktion nach m. Ist $m = 1$, so ist A eine Elementar-Matrix, also ist A eine Matrix der Form $S_i(\lambda)$ oder $Q_{ij}(\mu)$. In beiden Fällen sieht man leicht, daß die Behauptung gilt. Induktions-Schritt: Wir nehmen an, daß die Behauptung richtig ist, falls A sich als Produkt von $m-1$ Elementar-Matrizen schreiben läßt. Sei nun $A = A_1 \cdots A_m$ mit Elementarmatrizen A_i. Dann ist

$$\begin{aligned}\det(AB) &= \det(A_1 A_2 \cdots A_m B)\\ &= \det A_1 \cdot \det(A_1 \cdots A_m B)\\ &= \det A_1 \cdot \det(A_2 \cdots A_m) \cdot \det B\\ &= \det(A_1 A_2 \cdots A_m) \cdot \det B = \det A \cdot \det B.\end{aligned}$$

dabei haben wir beim zweiten und beim vierten Gleichheitszeichen den Induktionsanfang verwandt, beim dritten den Fall $m-1$.

Folgerung. *Eine $(n \times n)$-Matrix A mit Koeffizienten in $\mathbb{Q}$ ist genau dann invertierbar, wenn $\det A \neq 0$ gilt.*

Beweis: Aus $AA^{-1} = I_n$ folgt wegen des Produktsatzes $\det A \cdot \det(A^{-1}) = \det I_n = 1$, also muß $\det A \neq 0$ sein. Ist umgekehrt $\det A \neq 0$, so schreiben wir $A = A_1 \cdots A_t$ mit Elementarmatrizen $A_1, \ldots, A_n$. Der Produktsatz zeigt $\det A_i \neq 0$ für $1 \leq i \leq t$. Alle diese Elementarmatrizen sind invertierbar, also ist auch A invertierbar.

Die komplementäre Matrix zu A. Sei A eine $n \times n$ Matrix. Setze $A^\sharp = (a'_{ij})_{ij}$ mit $a'_{ij} = (-1)^{i+j} \det A_{ji}$ (beachte die Indexvertauschung); man nennt $A^\sharp$ die *komplementäre Matrix* zu A.

Satz (Laplace). *Sei $A \in M(n \times n, \mathbb{Q})$. Es ist $AA^\sharp = A^\sharp A = \det A \cdot I_n$.*

Beweis: Sei $A = (a_{ij})_{ij}$ und $A^\sharp = (a'_{ij})_{ij}$ mit $a'_{ij} = (-1)^{i+j} \det A_{ji}$. Setze $AA^\sharp = (c_{ik})_{ik}$. Also ist

$$c_{ik} = \sum_j a_{ij} a'_{jk} = \sum_j (-1)^{i+j} a_{ij} \det A_{kj}.$$

Die Laplace-Entwicklung nach der i-ten Reihe zeigt, daß wir für $i = k$ gerade $\det A$ erhalten. Sei nun $i \neq k$. Die Matrix B entstehe aus A, indem wir die k-te Zeile von A durch die i-te Zeile von A ersetzen, alle anderen Zeilen aber übernehmen. Da B zwei gleiche Zeilen (nämlich die i-te und die k-te) hat, liefert die Laplace-Entwicklung der Matrix B nach der k-ten Zeile:

$$\sum_j (-1)^{i+j} b_{kj} \det B_{kj} = \det B = 0.$$

Andererseits ist aber $b_{kj} = a_{ij}$ und $A_{kj} = B_{kj}$ für alle j, demnach ist $c_{ik} = 0$. Dies zeigt $AA^\sharp = \det A \cdot I_n$.

Entsprechend sieht man $A^\sharp A = \det A \cdot I_n$ mit Hilfe der Laplace-Entwicklung nach Spalten.

Wie wir gesehen haben, beruht der Beweis auf den Laplace-Entwicklungen, und auf der Tatsache, daß die Determinante einer Matrix mit zwei gleichen Zeilen oder zwei gleichen Spalten Null ist. Natürlich kann man die Laplace-Entwicklungen aus der Formel wieder ableiten.

(1.5) Das charakteristische Polynom.

Der Polynom-Ring $\mathbb{Q}[T]$. Ein Polynom mit Koeffizienten in $\mathbb{Q}$ ist von der Form $f(T) = \sum_{i=0}^{n} c_i T^i$ mit Koeffizienten $c_i \in \mathbb{Q}$. (Dabei ist T eine "Variable"; ein Polynom ist durch seine Koeffizienten gegeben. Zur Definition von Polynomringen siehe auch Abschnitt 2.2.) Ist $f(T) = \sum_{i=0}^{n} c_i T^i$ mit Koeffizienten $c_i \in \mathbb{Q}$ und ist $c_n \neq 0$, so nennt man n den *Grad von $f(T)$;* dem Null-Polynom (= alle Koeffizienten sind Null) ist auf diese Weise kein Grad zugeordnet (manchmal gibt man ihm den Grad -1 oder $-\infty$).

Matrizen mit Koeffizienten in $\mathbb{Q}[T]$. Wir betrachten nun auch die Menge $M(m \times n, \mathbb{Q}[T])$ aller $(m \times n)$-Matrizen $A = (a_{ij})_{ij}$ mit $a_{ij} \in \mathbb{Q}[T]$. Zwei Matrizen in $M(m \times n, \mathbb{Q}[T])$ können addiert werden (komponentenweise), ist A eine Matrix in $M(m \times n, \mathbb{Q}[T])$ und $f \in \mathbb{Q}[T]$, so entsteht $f \cdot A$ aus A, indem ich jeden Koeffizienten von A mit f multipliziere (man nennt dies die Skalar-Multiplikation, sie ist also ebenfalls komponentenweise definiert). Insbesondere kann jede Matrix in $M(m \times n, \mathbb{Q}[T])$ in der Form $A = \sum_{i=0}^{n} T^i A_i$ mit Matrizen $A_i \in M(m \times n, \mathbb{Q})$ geschrieben werden (dabei bedeutet $T^i A_i$ die Skalarmultiplikation der Matrix A_i mit T^i).

Sei A eine $(n \times n)$-Matrix mit Koeffizienten in $\mathbb{Q}$. Wir betrachten die Matrix

$$A - T \cdot I_n = \begin{bmatrix} a_{11} - T & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - T & \cdots & a_{2n} \\ \cdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} - T \end{bmatrix},$$

dies ist eine Matrix mit Koeffizienten im Polynomring $\mathbb{Q}[T]$. Setze

$$\chi_A(T) = \det(A - T \cdot I_n),$$

dies ist ein Polynom vom Grad n mit Koeffizieten in $\mathbb{Q}$, man nennt es das *charakteristische Polynom* der Matrix A.

Das **Einsetzen von quadratischen Matrizen in Polynome:** Sei $A \in M(n \times n, \mathbb{Q})$. Sei $f(T) \in \mathbb{Q}[T]$, etwa $f(T) = \sum_{i=0}^{n} c_i T^i$ mit $c_i \in \mathbb{Q}$. Wir können die Potenzen A^i der Matrix A bilden, wir können diese Potenzen A^i mit c_i multiplizieren (Skalar-Multiplikation), wir können die entstandenen Matrizen $c_i A^i$ addieren: wir erhalten $\sum_{i=0}^{n} c_i A^i$. Dieser Vorgang entspricht dem üblichen Einsetzen von Zahlen in Polynomen, hier haben wir "die Matrix A in das Polynom $f(T)$ eingesetzt", man schreibt demnach $f(A) = \sum_{i=0}^{n} c_i A^i$.

Satz von Cayley-Hamilton (CAYLEY 1821-1895, HAMILTON 1805-1865). *Sei $A \in M(n \times n, \mathbb{Q})$. Es ist $\chi_A(A) = 0_{nn}$.*

Was besagt dieser Satz? Zur Matrix A ist das charakteristische Polynom $\chi_A(T)$ definiert; in dieses Polynom können wir die Matrix A einsetzen, also $\chi_A(A)$ bilden.

Was wir erhalten, ist die **Null-Matrix**. (Um dies zu betonen, haben wir 0_{nn} und nicht etwa nur 0 geschrieben.)

Beweis des Satzes: Sei $B = A - T \cdot I_n$. Es ist $\det B$ das charakteristische Polynom von $\chi_A(T)$, dies sei das Polynom $\det B = \chi_A(T) = \sum_{i=0}^{n} c_i T^i$.

Wir bilden wie üblich die zu B komplementäre Matrix $B^\sharp$. Beachte: die Koeffizienten von $B^\sharp$ sind Polynome mit Koeffizienten in $\mathbb{Q}$, und ihr Grad ist höchstens $n-1$. Also kann man $B^\sharp$ in der Form

$$(*) \qquad B^\sharp = \sum_{i=0}^{n-1} B_i T^i \quad \text{mit} \quad B_i \in M(n \times n, \mathbb{Q})$$

schreiben. Der Satz von Laplace besagt:

$$BB^\sharp = \det(B) \cdot I_n = \chi_A \cdot I_n = \sum_{i=0}^{n} c_i T^i I_n.$$

Das Produkt $BB^\sharp$ können wir umformen (einsetzen, distributiv rechnen):

$$\begin{aligned} BB^\sharp &= \big(A - T \cdot I_n\big)\Big(\sum_{i=0}^{n-1} T^i B_i\Big) = \sum_{i=0}^{n-1} T^i A B_i - \sum_{i=0}^{n-1} T^{i+1} B_i \\ (**) \qquad &= AB_0 + \sum_{i=1}^{n-1} T^i \big(AB_i - B_{i-1}\big) - T^n B_{n-1}. \end{aligned}$$

Die Matrix $BB^\sharp$ ist in $(*)$ und $(**)$ auf zwei verschiedene Weisen nach Potenzen von T entwickelt worden. Koeffizienten-Vergleich liefert:

$$\begin{aligned} c_0 I_n &= AB_0 \\ c_1 I_n &= AB_1 - B_0 \\ c_2 I_n &= AB_2 - B_1 \\ &\dots \\ c_{n-1} I_n &= AB_{n-1} - B_{n-2} \\ c_n I_n &= - B_{n-1} \end{aligned}$$

Wir multiplizieren die Gleichung $c_i I_n = \dots$ von links mit A^i und erhalten

$$\begin{aligned} c_0 A^0 &= AB_0 \\ c_1 A^1 &= A^2 B_1 - AB_0 \\ c_2 A^2 &= A^3 B_2 - A^2 B_1 \\ &\dots \\ c_{n-1} A^{n-1} &= A^n B_{n-1} - A^{n-1} B_{n-2} \\ c_n A^n &= - A^n B_{n-1} \end{aligned}$$

Wenn wir nun die Gleichungen addieren, erhalten wir links $\chi_A(A)$, und rechts die Null-Matrix.

Teil 2. Grundbegriffe der Algebra.

(2.0) Mengen.

Menge, Element, $\in$*:* Dies sind Grundbegriffe, die wir nicht weiter hinterfragen. Wichtiges Beispiel: Die Menge $\mathbb{N}_0$ der natürlichen Zahlen $0, 1, 2, \ldots$. Mengen können gegeben werden durch eine Aufzählung ihrer Elemente: zum Beispiel kann die Menge S der natürlichen Zahlen, die kleiner oder gleich 4 sind, durch $S = \{0, 1, 2, 3, 4\}$ oder durch $\{1, 1, 4, 0, 1, 3, 2\}$ beschrieben werden (bei dieser Beschreibung mit Hilfe der der Mengenklammern $\{\,\}$ kommt es nicht auf die Reihenfolge an, auch prüft man nicht, ob Elemente mehrfach notiert sind). Ist S eine Menge, die nur endlich viele Elemente enthält, so schreibt man $|S|$ für die Anzahl der Elemente.

Sind U, M Mengen, so nennt man U eine *Teilmenge* von M, falls jedes Element von U auch Element von M ist. Untermengen beschreibt man oft durch die Angabe der definierenden Eigenschaften, die Menge $S = \{0, 1, 2, 3, 4\}$ kann zum Beispiel durch $S = \{z \in \mathbb{N}_0 \mid z \leq 4\}$ beschrieben werden. Die leere Menge $\emptyset$ (sie enthält kein Element) ist Teilmenge jeder Menge.

Sind M_1, M_2 Untermengen einer Menge M, so nennt man

$$M_1 \cap M_2 = \{x \in M \mid x \in M_1 \quad \text{und} \quad x \in M_2\}$$

den *Durchschnitt* der beiden Mengen, und

$$M_1 \cup M_2 = \{x \in M \mid x \in M_1 \quad \text{oder} \quad x \in M_2\}$$

die *Vereinigung* der beiden Mengen M_1, M_2.

Seien A, B Mengen. Das *Produkt* $A \times B$ der Mengen A, B ist die Menge aller Paare (a, b) mit $a \in A$, $b \in B$. Eine *Abbildung* $f\colon A \to B$ ist eine Teilmenge $f \subseteq A \times B$ mit folgenden Eigenschaften:

(A1) Zu jedem $a \in A$ gibt es ein $b \in B$ mit $(a, b) \in f$.

(A2) Gehören die Paare (a, b_1) und (a, b_2) zu f, mit $a \in A$ und $b_1, b_2 \in B$, so ist $b_1 = b_2$.

Sei $f\colon A \to B$ eine Abbildung. Meist schreibt man statt $(a, b) \in f$ lieber $f(a) = b$ oder auch $a \mapsto b$. Man sagt, daß f *injektiv* ist, falls gilt: Sind Elemente $a_1, a_2 \in A$ gegeben mit $f(a_1) = f(a_2)$, so ist $a_1 = a_2$. Man sagt, daß f *surjektiv* ist, falls es zu jedem $b \in B$ ein $a \in A$ gibt mit $f(a) = b$. Ist $f\colon A \to B$ sowohl injektiv als auch surjektiv, so sagt man, daß f *bijektiv* ist.

Unsere Definition interpretiert eine Abbildung f einfach als die Menge der Paare $(a, f(a))$ mit $a \in A$. Manchmal unterscheidet man auch zwischen der Abbildung f und der Menge $\{(a, f(a)) \mid a \in A\}$, die man den *Graph* der Abbildung f nennt; aber dann muß man irgenwie anders erklären, was eine Abbildung sein soll: man spricht von einer Zuordnung oder so, aber wie definiert man, was eine Zuordnung ist? Die Zuordnung $a \mapsto f(a)$ sollte man keinesfalls als einen dynamischen Vorgang auffassen, sondern statisch: durch f ist mit $a \in A$ das Element $f(a) \in B$ verbunden; gegeben sind die Paare $(a, f(a))$ in $A \times B$, also der Graph.

(2.1) Halbgruppen und Gruppen.

Sei S eine Menge. Eine *Verknüpfung* auf S ist eine Abbildung $\mu\colon S \times S \to S$, statt $\mu(s_1, s_2)$ schreibt man manchmal $s_1 + s_2$ oder $s_1 s_2$ oder ... .

Eine *Halbgruppe* $H = (H, *)$ ist eine Menge H mit einer Verknüpfung (die $(h_1, h_2) \mapsto h_1 * h_2$ geschrieben wird), mit folgenden Eigenschaften:

(H1) Für alle $h_1, h_2, h_3 \in H$ gilt $(h_1 * h_2) * h_3 = h_1 * (h_2 * h_3)$ (Assoziativität).

(H2) Es gibt ein Element $e \in H$ mit $e * h = h = h * e$ für alle $h \in H$. (Einselement).

Eine Halbgruppe H hat nur ein Einselement. (Beweis: Sind Elemente $e, e' \in H$ gegeben mit $e * h = h = h * e$ und $e' * h = h = h * e'$ für alle $h \in H$, so ist $e = e * e' = e'$.) Dieser Beweis zeigt sogar: Das Einselement einer Halbgruppe ist die einzige "Rechts-Eins" und die einzige "Links-Eins"; dabei ist eine "Rechts-Eins" ein Element r mit $h * r = h$ für alle $h \in H$, eine "Links-Eins" .... Man schreibt manchmal 1_H für das Einselement der Halbgruppe H. Ist H eine Halbgruppe, so heißt $h \in H$ *idempotent*, falls $h * h = h$ gilt. Das Einselement einer Halbgruppe ist idempotent, im allgemeinen wird es aber in einer Halbgruppe weitere idempotente Elemente geben.

Eine Halbgruppe H heißt *kommutativ* oder auch *abelsch* falls gilt: $h_1 * h_2 = h_2 * h_1$ für alle $h_1, h_2 \in H$. In abelschen Halbgruppen bezeichnet man oft die Verknüpfung mit dem Symbol $+$ und man spricht dann statt vom Einselement der Halbgruppe von der Null der Halbgruppe (Rechenregel: $0 + h = h = h + 0$, für alle $h \in H$).

Ist $H = (H, \cdot)$ eine Halbgruppe und U eine Untermenge von H, so sagt man, daß U *unter der Multiplikation abgeschlossen* ist, falls gilt: Sind $u_1, u_2 \in U$, so ist auch $u_1 u_2 \in U$. Eine Untermenge, die unter der Multiplikation abgeschlossen ist und die das Einselement enthält, ist selbst wieder eine Halbgruppe.

Einzelbeispiele von Halbgruppen sollen hier nicht weiter thematisiert werden. Man sehe sich die Beispiele von Gruppen an (jede Gruppe ist eine Halbgruppe), auch sehe man sich die Beispiele von Ringen und von Körpern an (ist R ein Ring, so sind $(R, +)$ und $(R, \cdot)$ Halbgruppen; ist R sogar ein Körper, so ist zusätzlich auch $(R^*, \cdot)$ eine Halbgruppe). Insbesondere ist die Menge $M(n \times n, R)$ für jeden Ring R bezüglich der Multiplikation eine Halbgruppe (natürlich auch bezüglich der Addition). Und $(\mathbb{N}_0, +)$ und $(\mathbb{N}_0, \cdot)$ sind ebenfalls Halbgruppen.

Ein Element h einer Halbgruppe $(H, \cdot)$ heißt *invertierbar*, wenn es ein $h' \in H$ mit $hh' = h'h = 1_H$ gibt. Existiert ein derartiges Element h', so ist es eindeutig bestimmt und wird das *zu h inverse* Element genannt, und man schreibt h^{-1} statt h' (die Eindeutigkeit sieht man so: ist auch $hh'' = h''h = 1_H$, so ist $h' = h' \cdot 1_H = h'(hh'') = (h'h)h'' = 1_H \cdot h'' = h''$). Sind die Elemente $h_1, h_2 \in H$ invertierbar, so ist auch $h_1 h_2$ invertierbar und es gilt $(h_1 h_2)^{-1} = h_2^{-1} h_1^{-1}$ (beachte die Reihenfolge).

Gruppen. Eine *Gruppe* $G = (G, *)$ ist eine Halbgruppe, in der zusätzlich gilt:

(G) Zu jedem Element $g \in G$ gibt es ein $g' \in G$ mit $gg' = 1_G$.

Ist G eine Gruppe, und gilt $gg' = 1_G$, so gilt auch $g'g = 1_G$, es ist also $g' = g^{-1}$; jedes Element in G ist also invertierbar. (Beweis: Sei $gg' = 1_G$. Zu g'

gibt es ebenfalls ein Element $g'' \in G$ mit $g'g'' = 1_G$. Dann ist aber $g = g \cdot 1_G = g(g'g'') = (gg')g'' = 1_G \cdot g'' = g''$. Also gilt $g'g = g'g'' = 1_G$.)

In einer Gruppe G ist das Einselement e das einzige idempotente Element. (In einer Halbgruppe kann es viele idempotente Elemente geben, wie die Halbgruppe $(M(2 \times 2), \mathbb{Q}), \cdot)$ zeigt.)

Gruppen-Homomorphismen. Sind $G = (G, *)$ und $H = (H, \circ)$ Gruppen, und $f\colon G \to H$ eine Abbildung, so nennt man f einen Gruppen-Homomorphismus, falls für alle $g_1, g_2 \in H$ gilt

$$f(g_1 * g_2) = f(g_1) \circ f(g_2).$$

Einen Homomorphismus, der bijektiv ist, nennt man einen *Isomorphismus.*

Ist $f\colon G \to H$ ein Gruppen-Homomorphismus, so ist $f(1_G) = 1_H$ und $f(g^{-1}) = f(g)^{-1}$, für jedes $g \in G$. (Beweis: Es ist $f(1_G)^2 = f(1_G^2) = f(1_G)$. Und $f(g^{-1})f(g) = f(g^{-1}g) = f(1_G) = 1_H$.)

Lemma. *Sei H eine endliche Halbgruppe mit folgender Eigenschaft: Sind g, h_1, h_2 Elemente von H mit $gh_1 = gh_2$, so ist $h_1 = h_2$ (Linkskürzungs-Eigenschaft). Dann ist H eine Gruppe.* Beweis: Sei $g \in H$, wir suchen ein Element g' mit $gg' = 1$. Die Halbgruppe H habe n Elemente und $h_1, \ldots, h_n$ sei die Liste aller Elemente von H. Da die Elemente $gh_1, \ldots, gh_n$ paarweise verschieden sind, sind dies wieder alle Elemente, insbesondere ist das Einselement 1_H eines dieser Elemente. Dies zeigt, daß es ein Element h_i mit $gh_i = 1$ gibt.

Ist $f\colon G \to H$ ein Gruppen-Homomorphismus, so nennt man $\mathrm{Ker}(f) = \{g \in G \mid f(g) = 1_H\}$ den *Kern* von f. *Dies ist eine Untergruppe von G.*

BEISPIELKLASSEN VON GRUPPEN UND GRUPPEN-HOMOMORPHISMEN:

Zahlbereiche, die Gruppen sind: $(\mathbb{Z}, +), (\mathbb{Q}, +), \mathbb{Q}^* = (\mathbb{Q} \setminus \{0\}, \cdot)$, usw.

Permutationsgruppen. Mit S_n bezeichnen wir die Gruppe aller Permutationen der Menge $\{1, 2, \ldots, n\}$ mit der Hintereinanderschaltung von Permutationen als Verknüpfung. Sei A_n die Untergruppe von S_n aller geraden Permutationen. Die Abbildung $\mathrm{sign}\colon S_n \to (\{1, -1\}, \cdot)$ ist ein Gruppen-Homomorphismus, A_n ist sein Kern.

Geometrisch motivierte Gruppen. Sei $n \geq 3$. Sei E_n ein regelmäßiges n-Eck in der Ebene. Sei D_n die Menge aller Symmetrien der Ebene, die das n-Eck auf sich abbilden (also Drehungen und Spiegelungen). Insgesamt gibt es n Drehungen (nämlich mit den Drehwinkeln $\frac{i}{n}360$ mit $0 \leq i < n$) und n Spiegelungen. Es gilt also $|D_n| = 2n$. Ist $n \geq 3$, so ist D_n nicht abelsch.

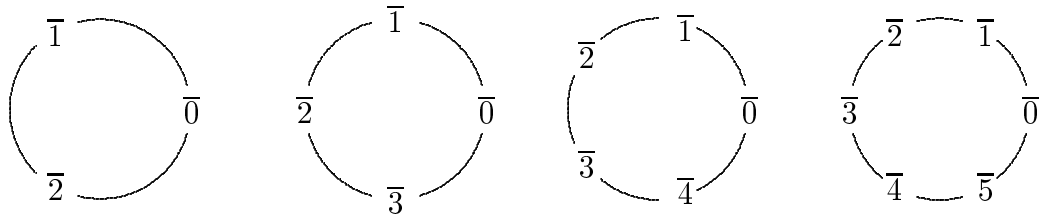
Matrizengruppen. Sei $GL(n, \mathbb{Q})$ die Menge aller invertierbaren $(n \times n)$-Matrizen mit Koeffizienten in $\mathbb{Q}$. Dies ist eine Gruppe $GL(n, \mathbb{Q}) = (GL(n, \mathbb{Q}), \cdot)$, die *volle lineare Gruppe.* Die Abbildung $\det\colon GL(n, \mathbb{Q}) \to \mathbb{Q}^*$ ist ein Gruppen-Homomorphismus, der Kern von det wird mit $SL(n, \mathbb{Q})$ bezeichnet, man nennt dies die *spezielle lineare Gruppe.*

Weitere wichtige Untergruppen der $GL(n,\mathbb{Q})$ sind die Menge $B(n,\mathbb{Q})$ aller invertierbaren oberen Dreiecksmatrizen und die Menge $U(n,\mathbb{Q})$ der unipotenten oberen Dreiecksmatrizen (eine obere Dreiecksmatrix heißt *unipotent*, wenn alle Koeffizienten auf der Hauptdiagonalen gleich 1 sind).

Wichtige Gruppen-Homomorphismen wurdem im Abschnitt 2.4 eingeführt:

$$S_i(-)\colon \mathbb{Q}^* \to \mathrm{GL}(n,\mathbb{Z}) \quad \text{und} \quad Q_{ij}(-)\colon (\mathbb{Q},+) \to \mathrm{GL}(n,\mathbb{Z}).$$

Die zyklischen Gruppen. Sei $n \geq 1$. Sei $n\mathbb{Z}$ die Menge der Vielfachen von n, also die Menge der ganzen Zahlen, die durch n ohne Rest teilbar sind. Für jede ganze Zahl a setze $\overline{a} = a + n\mathbb{Z}$. Beachte: *Es gilt* $\overline{a_1} = \overline{a_2}$ *genau dann, wenn* $a_1 - a_2$ *durch* n *teilbar ist.* Ist $0 \leq a < n$, so ist $\overline{a}$ die Menge der ganzen Zahlen, die bei Division durch n den Rest a liefern (man nennt dies eine Restklasse modulo n). Man schreibt $\mathbb{Z}/n\mathbb{Z}$ für die Menge der Restklassen modulo n, und man definiert auf dieser Menge eine Addition vermöge $\overline{a_1} + \overline{a_2} = \overline{a_1 + a_2}$ (Zu zeigen: dies ist wohl-definiert). *Mit dieser Addition ist* $\mathbb{Z}/n\mathbb{Z}$ *eine Gruppe,* die zyklische Gruppe der Ordnung n. Für $n \geq 3$ ist dies gerade die Drehgruppe des regulären n-Ecks. Man nennt diese Gruppen $(\mathbb{Z}/n\mathbb{Z},+)$ die *endlichen zyklischen Gruppen.* Zusätzlich nennt man $(\mathbb{Z},+)$ die *unendliche zyklische Gruppe.* Um zu verstehen, warum die endlichen zyklischen Gruppen "zyklisch" heißen, empfiehlt es sich, die Elemente im Kreis anzuorden und die Abbildung $+\overline{1}$ zu betrachten: Hier die Bilder für $n = 3, 4, 5, 6$.



Das entsprechende Bild für die unendliche zyklische Gruppe wäre die (ganzzahlige) Zahlengerade, also ein "Zykel mit unendlichem Radius":

$$\cdots \quad — \quad -1 — 0 — 1 — 2 — \quad \cdots$$

Die Abbildung $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ mit $z \mapsto \overline{z}$ ist ein Gruppen-Homomorphismus, sein Kern ist $n\mathbb{Z}$.

Ist G eine Gruppe und $g \in G$, so erhält man durch $\phi_g\colon \mathbb{Z} \to G$ mit $\phi_g(z) = g^z$ einen Gruppen-Homomorphismus. Ist ϕ_g injektiv, so erhält man einen Isomorphismus $\mathbb{Z} \to \{g^z \mid z \in \mathbb{Z}\}$ und $\{g^z \mid z \in \mathbb{Z}\}$ ist die von g erzeugte Untergruppe von G. Andernfalls wähle die minimale natürliche Zahl n mit $\phi_g(n) = 1_G$. Dann erhalten wir einen Isomorphismus $\mathbb{Z}/n\mathbb{Z} \to \{g^z \mid z \in \mathbb{Z}\}$.

(2.2). Ringe.

Definition: Ein *Ring* $R = (R, +, \cdot)$ ist eine Menge R mit zwei Verknüpfungen $+$ und $\cdot$, so daß die folgenden Eigenschaften erfüllt sind:

(R1) $(R, +)$ ist eine abelsche Gruppe.

(R2) $(R, \cdot)$ ist eine Halbgruppe.

(R3) Sind r, r_1, r_2 Elemente von R, so gilt $r(r_1 + r_2) = rr_1 + rr_2$ und $(r_1 + r_2)r = r_1 r + r_2 r$.

Das Einselement von $(R, +)$ bezeichnet man mit 0_R oder einfach mit 0 und nennt es die *Null* des Rings. Das Einselement von $(R, \cdot)$ bezeichnet man mit 1_R oder einfach mit 1 und nennt es die *Eins* von R. Ein Element $r \in R$ heißt *invertierbar,* wenn es als Element der Halbgruppe $(R, \cdot)$ invertierbar ist, wenn es also ein $r' \in R$ mit $rr' = 1_R = r'r$ gibt, und man schreibt dann r^{-1} statt r'. Ist $(R, \cdot)$ abelsch, so nennt man R einen *kommutativen* Ring.

Einfach zu zeigen ist: *Ist R ein Ring, und $r \in R$, so ist $0 \cdot r = 0 = r \cdot 0$.* Beweis: Es ist $0 \cdot r = (0+0) \cdot r = 0 \cdot r + 0 \cdot r$, also ist $0 \cdot r$ ein idempotentes Element der Gruppe $(R, +)$. Das einzige idempotente Element einer Gruppe ist aber ihr Einselement.

Ist R ein Ring und gibt es eine natürliche Zahl $n \geq 1$ mit $\underbrace{1 + 1 + \cdots + 1}_{n} = 0$,

so nennt man die kleinste derartige Zahl n die *Charakteristik* des Rings R und man schreibt $\operatorname{char} R = n$. Gibt es keine solche Zahl n, so schreibt man $\operatorname{char} R = 0$.

Sind $R = (R, +, \cdot)$ und $S = (S, +, \cdot)$ Ringe, so ist ein *Ring-Homomorphismus* $f \colon R \to S$ eine Abbildung mit folgenden Eigenschaften:

(1) $f \colon (R, +) \to (S, +)$ ist ein Gruppen-Homomorphismus.

(2) $f(r_1 \cdot r_2) = f(r_1) \cdot f(r_2)$ für alle $r_1, r_2 \in R$.

(3) $f(1_R) = 1_S$.

BEISPIELKLASSEN VON RINGEN UND RING-HOMOMORPHISMEN.

A) Die **Zahlbereiche** $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ sind kommutative Ringe

B) Sei $n \geq 1$. Auf der abelschen Gruppe $\mathbb{Z}/n\mathbb{Z}$ definiert man eine Multiplikation durch $\overline{a_1} \cdot \overline{a_2} = \overline{a_1 a_2}$ (wieder ist zu zeigen, daß dies wohl-definiert ist). Auf diese Weise wird $\mathbb{Z}/n\mathbb{Z}$ zu einem kommutativen Ring. Beachte: Im Ring $\mathbb{Z}/n\mathbb{Z}$ gilt:

$$\underbrace{1 + 1 + \cdots + 1}_{n} = 0,$$

$\mathbb{Z}/n\mathbb{Z}$ ist ein Ring der Characteristik n. Die Abbildung $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ mit $z \mapsto \overline{z}$ ist ein Ring-Homomorphismus.

Sei nun ein Ring R gegeben. Es gibt eine Vielzahl von Möglichkeiten, um mit Hilfe von R neue Ringe zu konstruieren.

C) Der Matrizenring $M(n \times n, R)$**.** Sei $n \geq 1$ eine natürliche Zahl. Die Menge der $(n \times n)$-Matrizen mit Koeffizienten in R ist bezüglich der Addition und der Multiplikation von Matrizen wieder ein Ring. Ist $n \geq 2$ und hat R mindestens zwei Elemente, so ist $M(n \times n, R)$ nicht kommutativ.

D) Der Funktionenring $\mathrm{Abb}(S,R)$**.** Sei S eine Menge, sei $\mathrm{Abb}(S,R)$ die Menge der Abbildungen $S \to R$. Definiere auf $\mathrm{Abb}(S,R)$ Addition und Multiplikation komponentenweise (d.h.: sind $f,g\colon S \to R$ Abbildungen, so definiere $f+g$ und fg durch

$$(f+g)(s) = f(s)+g(s), \quad \text{und} \quad (fg)(s) = f(s)g(s) \quad \text{für} \quad s \in S.$$

Mit diesen Verknüpfungen ist $\mathrm{Abb}(S,R)$ ein Ring. Ist R kommutativ, so ist der Ring $\mathrm{Abb}(S,R)$ kommutativ.

Ist $s \in S$, so definiert man $e_s\colon \mathrm{Abb}(S,R) \to R$ durch $e_s(f) = f(s)$ (man nennt dies die Auswertung von f an der Stelle s). Die Abbildung $e_s\colon \mathrm{Abb}(S,R) \to R$ ist ein Ring-Homomorphismus.

E) Der Polynomring $R[T]$**.** Sei $R[T]$ die Menge der Folgen $(a_0,a_1,\dots)$ von Elementen $a_i \in R$ für die $a_i = 0$ für $i \gg 0$ gilt (es gibt also ein $n \in \mathbb{N}_0$ mit $a_i = 0$ falls $i > n$). Wir betrachten auf $R[T]$ die komponentenweise Addition (also $(a_0,a_1,\dots)+(b_0,b_1,\dots) = (a_0+b_0,a_1+b_1,\dots)$. Die Multiplikation ist folgendermaßen definiert:

$$(a_0,a_1,\dots)\cdot(b_0,b_1,\dots) = (c_0,c_1,\dots)$$
$$\text{mit} \quad c_k = a_0b_k + a_1b_{k-1} + \cdots + a_kb_0 = \sum_{i+j=k} a_ib_j.$$

Elemente in $R[T]$ werden üblicherweise in folgenderweise notiert: Statt $(a_0,a_1,\dots)$ schreibt man $\sum_i a_iT^i$ (dabei ist also $1 = T^0 = (1,0,\dots), T = T^1 = (0,1,0,\dots)$ und so weiter).

Die Abbildung $e\colon R[T] \to \mathrm{Abb}(R,R)$ mit $e(f)(r) = f(r)$ für $f = f(T) \in R[T]$ und $r \in R$ ist ein Ring-Homomorphismus.

Matrizen mit Koeffizienten in einem kommutativen Ring.

Ist R ein kommutativer Ring, so kann man durch die Leibniz-Formel wie im Spezialfall $R = \mathbb{Q}$ für jedes $A \in M(n\times n,R)$ die Determinante $\det A$ definieren und es gelten die bekannten Regeln. Insbesondere ist zu jeder $(n\times n)$-Matrix A die komplementäre Matrix $A^\sharp$ definiert. Zu jeder Matrix $A \in M(n\times n,R)$ kann man wie oben das charakteristische Polynom $\chi_A(T) \in R[T]$ definieren. Es gelten die folgenden Sätze:

Satz (Laplace). *Sei R ein kommutativer Ring. Sei $A \in M(n\times n,R)$. Dann ist $AA^\sharp = A^\sharp A = \det A \cdot I$.*

Satz (Cayley-Hamilton). *Sei R ein kommuativer Ring. Sei $A \in M(n\times n,R)$. Dann ist $\chi_A(A) = 0$.*

Wir haben diese beiden Sätze nur im Spezialfall $R = \mathbb{Q}$ bewiesen. Man mache sich aber klar, daß wir beim Beweis dieser Sätze nur verwendet haben, daß $\mathbb{Q}$ ein kommutativer Ring ist! Das Invertierbarkeitskriterium ist folgendermaßen zu formulieren:

Satz. *Sei R ein kommutativer Ring. Genau dann ist $A \in M(n\times n,R)$ invertierbar, wenn $\det A \in R$ invertierbar ist.*

(2.3) Körper.

Ein *Körper* $K = (K, +, \cdot)$ ist ein Ring, in dem die Menge $K^* = K \setminus \{0\}$ unter der Multiplikation abgeschlossen ist und $(K^*, \cdot)$ eine abelsche Gruppe bildet.

BEISPIELKLASSEN.

A) Die **Zahlbereiche** $\mathbb{Q}$, $\mathbb{R}$ sind Körper.

B) Der Körper $\mathbb{C}$ der komplexen Zahlen. Sei $\mathbb{C}$ die Menge aller (2×2)-Matrizen der Form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ mit $a, b \in \mathbb{R}$. Dies ist ein kommutativer Ring. Ist $(a, b) \neq (0, 0)$, so ist die Matrix $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ invertierbar, ihr Inverses ist $\frac{1}{a^2+b^2} \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$. Bezeichnen wir mit $1_{\mathbb{C}}$ die Einheitsmatrix $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ und mit i die Matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, so können wir schreiben

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = a \cdot 1_{\mathbb{C}} + b \cdot i$$

oder einfacher $a + bi$. Beachte, daß $i^2 = -1_{\mathbb{C}}$ gilt, i ist also so etwas wie $\sqrt{-1}$.

Den folgenden Satz nennt man den **Fundamentalsatz der Algebra:** *Jedes nicht-konstante Polynom mit Koeffizienten in $\mathbb{C}$ läßt sich als Produkt von Polynomen vom Grad 1 schreiben.* Diesen Satz werden wir später ganz entscheidend brauchen, dann soll auch etwas zum Beweis des Satzes gesagt werden.

C) Satz. *Sei $n \geq 1$. Genau dann ist $\mathbb{Z}/n\mathbb{Z}$ ein Körper, wenn n eine Primzahl ist.*

Beweis: Ist $n = 1$, so ist $(\mathbb{Z}/n\mathbb{Z})^*$ die leere Menge, also sicher keine Gruppe. Ist $n > 1$ und $n = n_1 n_2$ mit ganzen Zahlen $1 < n_1 < n$ und $1 < n_2 < n$, und betrachten wir die Restklassen $\overline{n_1}, \overline{n_2}$, so ist $\overline{n_1} \neq 0$ und $\overline{n_2} \neq 0$, aber $\overline{n_1} \cdot \overline{n_2} = 0$, demnach ist $(\mathbb{Z}/n\mathbb{Z})^*$ unter der Multiplikation nicht abgeschlossen, also ist $\mathbb{Z}/n\mathbb{Z}$ kein Körper.

Sei nun p eine Primzahl. Aus dem Satz über die eindeutige Primfaktorzerlegung für $\mathbb{Z}$ (der hier vorausgesetzt wird) folgt, daß $(\mathbb{Z}/n\mathbb{Z})^*$ unter der Multiplikation abgeschlossen ist (teilt p ein Produkt $n_1 n_2$ zweier ganzer Zahlen, so teilt p mindestens eine der beiden Zahlen n_1, n_2). Daraus folgt: Sind n, m_1, m_2 ganze Zahlen, die nicht durch p teilbar sind, und ist $\overline{nm_1} = \overline{nm_2}$, so ist $\overline{m_1} = \overline{m_2}$ (denn aus $\overline{nm_1} = \overline{nm_2}$ folgt $\overline{n}(\overline{m_1} - \overline{m_2}) = 0$, also $\overline{m_1} - \overline{m_2} = 0$.) Eine endliche Halbgruppe, in der die Linkskürzungseigenschaft gilt, ist aber eine Gruppe.

Teil 3. Vektorräume und lineare Abbildungen.

Sei K ein Körper. Ein K-Vektorraum $V = (V, +, *)$ ist eine Menge V mit einer Verknüpfung $+$ und einer Abbildung $*\colon K \times V \to V$, so daß $(V, +)$ eine abelsche Gruppe ist und für alle $v, v_1, v_2 \in V$ und alle $\lambda, \lambda_1, \lambda_2 \in K$ folgende Bedingungen erfüllt sind:

(V1) $\lambda * (v_1 + v_2) = \lambda * v_1 + \lambda * v_2$.
(V2) $(\lambda_1 + \lambda_2) * v = \lambda_1 * v + \lambda_2 * v$.
(V3) $(\lambda_1 \lambda_2) * v = \lambda_1 * (\lambda_2 * v)$.
(V3) $1_K * v = v$.

Die Elemente von V heißen *Vektoren,* die von K *Skalare.*

Ist $V = (V, +, *)$ ein K-Vektorraum, so nennt man $*$ die *Skalar-Multiplikation*, statt $\lambda * v$ schreibt man meist einfach λv. Die Null von $(V, +)$ wird mit 0_V oder einfach mit 0 bezeichnet (manchmal auch mit $\mathbf{0}$ oder $\vec{0}$).

Einfache Folgerungen:

(1) Es ist $\lambda * 0_V = 0_V$ für alle $\lambda \in K$.
(2) Es ist $0_K * v = 0_V$ für alle $v \in V$.
(3) Ist $\lambda \in K$ und $v \in V$ mit $\lambda v = 0$, so ist $\lambda = 0$ oder $v = 0$.
(4) Es ist $(-\lambda) * v = -\lambda * v$ für alle $\lambda \in K, v \in V$.

Standard-Beispiel: Der K^n. Wir bezeichnen mit K^n die Menge der n-Tupel $(a_1, \ldots, a_n)$ mit $a_1, \ldots, a_n \in K$. *Bezüglich punktweiser Addition und Skalar-Multiplikation ist K^n ein Vektorraum.* (Punktweise Addition bedeutet:

$$(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n);$$

punktweise Skalar-Multiplikation bedeutet

$$\lambda * (a_1, \ldots, a_n) = (\lambda a_1, \ldots, \lambda a_n);$$

für alle $a_1, \ldots, a_n, b_1, \ldots, b_n \in V, \lambda \in K$.)

Warum nennt man K^n den Standard-Raum? Es wird sich zeigen, daß jeder "endlich-dimensionale" K-Vektorraum zu einem Standard-Raum isomorph ist!

Beispiel 2. Die Menge $M(m \times n, K)$ ist (bezüglich der Addition von Matrizen und der Skalarmultiplikation) ein K-Vektorraum.

Beispiel 3. Der Polynomring $K[T]$ ist (bezüglich Addition und Skalar-Multiplikation) ein K-Vektorraum.

Beispiel 4. Sei S eine Menge. Die Menge $\mathrm{Abb}(S, K)$ ist (bezüglich punktweiser Addition und punktweiser Skalarmultiplikation) ein Vektorraum.

Linearkombination. Sei K ein Körper, sei V ein K-Vektorraum. Seien $v_1, \ldots, v_n \in V$. Eine *Linearkombination* der Vektoren $v_1, \ldots, v_n$ ist ein Element der Form

$$\sum_{i=1}^{n} \lambda_i v_i \quad \text{mit} \quad \lambda_i \in K.$$

Sei $\operatorname{span}(v_1, \ldots, v_n)$ die Menge der Linearkombinationen der Vektoren $v_1, \ldots, v_n$ (falls $n = 0$ sei dies $\{0\}$). Es gilt: $\operatorname{span}(v_1, \ldots, v_n)$ *ist ein Unterraum von* V. (Genauer: $\operatorname{span}(v_1, \ldots, v_n)$ ist ein Unterraum von V, der die Vektoren $v_1, \ldots, v_n$ enthält und dieser Unterraum ist in jedem Unterraum von V, der alle Vektoren v_i enthält, enthalten.)

Erzeugendensystem. Die Folge $(v_1, \ldots, v_n)$ heißt *(endliches) Erzeugendensystem von* V, falls gilt: $\operatorname{span}(v_1, \ldots, v_n) = V$. Der Vektorraum V heißt *endlich erzeugt*, falls er ein endliches Erzeugendensystem besitzt.

Lineare Unabhängigkeit. Die Folge $(v_1, \ldots, v_n)$ heißt *linear unabhängig*, falls gilt: Ist $\sum_{i=1}^{n} \lambda_i v_i = 0$, so ist $\lambda_1 = \cdots = \lambda_n = 0$. Ist die Folge nicht linear unabhängig, so nennt man sie *linear abhängig.* Wichtig: *Ist die Folge* $(v_1, \ldots, v_n)$ *linear unabhängig, und ist* $w \in \operatorname{span}(v_1, \ldots, v_n)$, *etwa* $w = \sum \lambda_i v_i$, *so sind die Skalare* λ_i *eindeutig bestimmt.* (Beweis: Ist $\sum \lambda_i v_i = \sum \lambda_i' v_i$, so ist $\sum (\lambda_i - \lambda_i') v_i = 0$, also $\lambda_i - \lambda_i' = 0$ für alle i.)

Basis. Eine linear unabhängige Folge $(v_1, \ldots, v_n)$, die V erzeugt, nennt man eine *(endliche) Basis.*

Beispiel: Die kanonische Basis des K^n. Die Elemente $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0)$, $\ldots$, $e_n = (0, \ldots, 0, 1)$ bilden eine Basis des Standard-Raums K^n, man nennt sie die *kanonische Basis* des K^n.

(1) *Jedes minimale Erzeugendensystem ist eine Basis.* Beweis: Gilt $\sum \lambda_i v_i = 0$ und ist $\lambda_s \neq 0$, so ist

$$v_s = -\sum_{i \neq s} \frac{\lambda_i}{\lambda_s} v_i,$$

also ist $v_s \in \operatorname{span}(v_1, \ldots, v_{s-1}, v_{s+1} \ldots, v_n)$. Also ist auch $(v_1, \ldots, v_{s-1}, v_{s+1} \ldots, v_n)$ ein Erzeugendensystem.

(1′) *Jeder endlich erzeugte Vektorraum hat eine Basis.* Einfach weglassen, weglassen, usw.

(2) *Jede maximale Folge linear unabhängiger Elemente ist eine Basis.* Beweis: Sei $(v_1, \ldots, v_n)$ linear unabhängig, und sei $(v_1, \ldots, v_n, v)$ linear abhängig für alle $v \in V$. Zu zeigen ist: Jedes $v \in V$ gehört zu $\operatorname{span}(v_1, \ldots, v_n)$. Sei also $v \in V$. Es ist $\sum_i \lambda_i v_i + \lambda v = 0$ für gewisse λ, λ_i, nicht alle 0. Wäre $\lambda = 0$, so wäre $(v_1, \ldots, v_n)$ linear abhängig, Unsinn. Also $\lambda \neq 0$, also

$$v = -\sum_i \frac{\lambda_i}{\lambda} v_i.$$

(3) **Austauschlemma:** *Sei* $(v_1, \ldots, v_n)$ *eine Basis, sei* $w = \sum \lambda_i v_i$ *eine Linearkombination mit* $v_t \neq 0$. *Dann ist* $(v_1, \ldots, v_{t-1}, v_{t+1}, \ldots, v_n, w)$ *eine Basis.*

Beweis: Lineare Unabhängigkeit: Sind Skalare μ, μ_i mit $\sum_{i \neq t} \mu_i v_i + \mu w = 0$ gegeben, so ersetzen wir w durch $\sum \lambda_i v_i$, also $\sum_{i \neq t} (\mu_i + \mu \lambda_i) v_i + \mu \lambda_t = 0$. Verwende nun, daß die Folge $(v_1, \ldots, v_n)$ linear unabhängig ist. Wir sehen: $\mu \lambda_t = 0$, also $\mu = 0$ (denn $\lambda_t \neq 0$). Für $i \neq t$ haben wir $\mu_i + \mu \lambda_i = 0$, also $\mu_i = 0$. Erzeugendensystem: Ist $u \in V$ gegeben, so können wir u als Linearkombination $u = \sum \mu_i v_i$ schreiben. Ersetze hier v_t durch $\frac{1}{\lambda_t} w - \sum_{i \neq t} \frac{\lambda_i}{\lambda_t} v_i$. Umrechnen zeigt, daß sich u als Linearkombination der Vektoren w und v_i mit $i \neq t$ schreiben läßt.

(3′) Zusatz: *Sei* $(v_1, \ldots, v_n)$ *eine Basis von* V, *sei* $w \in V$. *Ist* $(v_1, \ldots, v_{s-1}, w)$ *eine linear unabhängige Folge, und schreiben wir* $w = \sum \lambda_i v_i$, *so gibt es immer ein* $t \geq s$ *mit* $\lambda_t \neq 0$. *Also: Wir können eines der* v_t *(mit* $t \geq s$*) durch* w *ersetzen.*

(4) **Austauschsatz:** *Ist* $(v_1, \ldots, v_n)$ *eine Basis und ist die Folge* $(w_1, \ldots, w_s)$ *linear unabhängig, so ist* $s \leq n$ *und es gibt Indizes* $i_1, \ldots, i_{n-s}$, *so daß die Folge* $(w_1, \ldots, w_s, v_{i_1}, \ldots, v_{i_{n-s}})$ *eine Basis bildet.*

Beweis: Sei zuerst $s \leq n$. Induktion nach s. Ist $s = 0$, so ist nichts zu zeigen. Sei schon gezeigt: $(w_1, \ldots, w_{s-1}, v_{i_1}, \ldots, v_{i_{n-s+1}})$ ist Basis. Verwende den Austauschsatz und seinen Zusatz für diese Basis und den Vektor $w = w_s$. Wir sehen, daß wir einen der Vektoren v_{i_r} durch w_s ersetzen können. Also erhalten wir eine Basis, die mit $w_1, \ldots, w_s$ beginnt und durch $n-s$ Vektoren der Form v_i fortgesetzt wird.

Angenommen es ist $s > n$. Im Schritt n wird gezeigt, daß $w_1, \ldots, w_n$ eine Basis ist. Daraus folgt, daß die Vektoren $(w_1, \ldots, w_{n+1})$ linear abhängig sind. Dies aber ist ein Widerspruch zur Annahme, daß die Vektoren $(w_1, \ldots, w_s)$ linear unabhängig sind. Also kann der Fall $s > n$ gar nicht auftreten.

Folgerung 1. *Ist* V *endlich erzeugt, so haben je zwei Basen gleichviel Elemente.*

Folgerung 2. *Ist* V *endlich erzeugt, so ist jeder Unteraum* U *endlich erzeugt. Eine Basis von* U *hat höchstens so viel Elemente wie eine Basis von* V. *Hat eine Basis von* U *genau so viel Elemente wie eine Basis von* V, *so ist* $U = V$.

Folgerung 3. (Basis-Ergänzungs-Satz.) *Sei* V *endlich erzeugter Vektorraum. Ist die Folge* $(v_1, \ldots, v_m)$ *linear unabhängig, so gibt es immer Vektoren* $v_{m+1}, \ldots, v_n$, *so daß* $(v_1, \ldots, v_n)$ *eine Basis ist.*

Dimension. Ist V ein endlich erzeugter Vektorraum mit einer Basis $v_1, \ldots, v_n$, so nennt man n die *Dimension* von V, und man schreibt $\dim V = n$. Man sagt auch, daß V n-dimensional ist (und daß V endlich-dimensional ist). Endlich-dimensionale Vektorräume sind endlich erzeugt; wegen (1') gilt umgekehrt auch, daß endlich erzeugte Vektorräume endlich-dimensional sind.

Umformulierung von Folgerung 2: *Ist* V *ein* n*-dimensionaler Vektorraum, und ist* U *ein Unterraum, so ist* $\dim U \leq \dim V$. *Zweitens gilt: Aus* $\dim U = \dim V$ *folgt* $U = V$.

Der Durchschnitt $U_1 \cap U_2$ zweier Unterräume U_1, U_2 eines Vektorraums V ist ein Unterraum von V.

Warnung: Im allgemeinen ist $U_1 \cup U_2$ **kein** Unterraum. Sind U_1, U_2 Unterräume eines Vektorraums V, so setzt man

$$U_1 + U_2 = \{u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2\}$$

Man zeigt leicht: Dies ist ein Unterraum.

Satz. *Seien* U_1, U_2 *endlich-dimensionale Unterräume eines Vektorraums. Dann gilt:*

$$\dim U_1 + \dim U_2 = \dim(U_1 \cap U_2) + \dim(U_1 + U_2)$$

.

Beweisidee: Wähle Basis von $U_1 \cap U_2$, ergänze sie zu einer Basis von U_1 und auch zu einer Basis von U_2. Zeige: Die Gesamtmenge der so erhaltenen Vektoren ist eine Basis von $U_1 + U_2$.

Lineare Abbildungen.

Seien V, W K-Vektorräume. Eine Abbildung $f\colon V \to W$ heißt *linear* (oder genauer: *K-linear,* oft sagt man auch *Homomorphismus von K-Vektorräumen* oder auch *lineare Transformation*) falls gilt

(1) $f(v_1 + v_2) = f(v_1) + f(v_2)$ für alle $v_1, v_2 \in V$.
(2) $f(\lambda v) = \lambda f(v)$ für $v \in V$, $\lambda \in K$.

Eine lineare Abbildung $f\colon V \to W$ heißt *Monomorphismus*, falls f injektiv ist; *Epimorphismus*, falls f surjektiv ist; *Isomorphismus*, falls f sowohl injektiv als auch surjektiv ist. Ist $V = W$, so nennt man f einen *Endomorphismus*; ein *Automorphismus* ist ein Endomorphismus, der gleichzeitig ein Isomorphismus ist.

Leicht zu verifizieren: Sei $f\colon V \to W$ K-linear.

(a) *Es ist $f(0) = 0$ und $f(-v) = -f(v)$.*
(b) *Ist $(v_1, \ldots, v_n)$ eine Familie von Vektoren in V, die linear abhängig ist, so ist $(f(v_1), \ldots, f(v_n))$ linear abhängig in W.*
(c) *Ist f ein Isomorphismus, so ist auch $f^{-1}\colon W \to V$ K-linear (also ebenfalls ein Isomorphismus.*
(d) *Ist V' ein Unterraum von V, so ist $f(V')$ ein Unterraum von W. Ist W' ein Unterraum von W, so ist $f^{-1}(W')$ ein Unterraum von V.*
(e) *Ist $g\colon U \to V$ auch K-linear, so ist $f \circ g\colon U \to W$ ebenfalls K-linear.*

Ist $f\colon V \to W$ eine K-lineare Abbildung, so setzt man

$$\operatorname{Im} f = f(V) \quad \text{und} \quad \operatorname{Ker} f = f^{-1}(0);$$

man nennt $\operatorname{Im} f$ das *Bild* von f, und $\operatorname{Ker} f$ den *Kern* von f. (Beachte, daß jede lineare Abbildung f insbesondere ein Gruppen-Homomorphismus ist und $\operatorname{Ker} f$ ist der Kern, wie er für Gruppen-Homomorphismen definiert wurde; unsere Zusatzvoraussetzung an f, daß f auch mit der Skalar-Multiplikation vertauscht, impliziert, daß $\operatorname{Ker} f$ nicht nur eine Untergruppe, sondern ein Unterraum von V ist.)

Ist $f\colon V \to W$ eine lineare Abbildung, so nennt man die Dimension $\dim \operatorname{Im} f$ den *Rang* von f.

Satz. *Seien V, W Vektorräume, sei $f\colon V \to W$ eine linear Abbildung. Ist V endlich-dimensional, so gilt*

$$\dim \operatorname{Im} f + \dim \operatorname{Ker} f = \dim V.$$

Ist $\dim V = n$ und $\dim \operatorname{Im} f = r$, so gibt es eine Basis $(v_1, \ldots, v_n)$ von V, mit folgenden Eigenschaften: Die Folge $(f(v_1), \ldots, f(v_r))$ ist eine Basis von $\operatorname{Im} f$ und die Folge $(v_{r+1}, \ldots, v_n)$ ist eine Basis von $\operatorname{Ker} f$.

Beweisidee: Wähle eine Basis $(u_1, \ldots, u_s)$ des Kerns von f und ergänze diese Folge durch Vektoren $u_{s+1}, \ldots, u_n$ zu einer Basis von V.
Zeige: Die Folge $(f(u_{s+1}, \ldots, f(u_n))$ ist eine Basis von $\operatorname{Im} f$. Daraus folgt nun $r = n-s$, also die Behauptung des Satzes. Der Zusatz folgt einfach durch Umbenennung der Vektoren: Setze $(v_1, \ldots, v_r) = (u_{s+1}, \ldots, u_n)$ und $(v_{r+1}, \ldots, v_n) = (u_1, \ldots, u_s)$.

Die Entsprechung: Matrizen — Lineare Abbildungen.

Satz. *Jede Matrix $C \in M(m \times n, K)$ liefert (durch Linksmultiplikation) eine lineare Abbildung $f_C \colon K^n \to K^m$. Umgekehrt ist jede lineare Abbildung $K^n \to K^m$ von dieser Form.* Statt f_C schreibt man manchmal $C\cdot$, deutet also auf diese Weise an, daß man die Linksmultiplikation mit C meint. Man nennt C die *darstellende Matrix* zur linearen Abbildung f_C.

Merkregel. *Die j-te* **Spalte** *der darstellenden Matrix C zu einer linearen Abbildung $f \colon K^n \to K^m$ ist gerade $f(e_j)$, also das Bild des j-ten kanonischen Basisvektors e_j von K^n unter f.*

Zusatz: *Ist $C \in M(m \times n, K)$ und $D \in M(n \times r, K)$, so ist $f_{CD} = f_C \circ f_D$.*

Zum Beweis (Satz, Merkregel, Zusatz): Zu zeigen ist $f_C(\lambda v) = \lambda f_C(v)$ und $f_C(v_1 + v_2) = f_C(v_1) + f_C(v_2)$ für $v, v_1, v_2 \in K^n$ und $\lambda \in K$. Beides folgt aus Rechenregeln für die Matrizen-Multiplikation. Die Merkregel natürlich ebenfalls. Der Zusatz folgt aus der Assoziativität der Matrizen-Multiplikation.

Entsprechungen:

Assoziativität der Matrizen-Multiplikation	Assoziativität der Hintereinanderschaltung von Abbildungen
Sei C eine $(m \times n)$-Matrix	und f_C die zugehörige lineare Abbildung
Die Spalten von C sind linear unabh.	f_C ist ein Monomorphismus.
Die Spalten von C erzeugen K^m	f_C ist ein Epimorphismus.
Die Spalten von C bilden eine Basis (d.h. C ist invertierbar)	f_C ist ein Isomorphismus.

Höherer Standpunkt: Wir betrachten nicht nur einzelne lineare Abbildungen $f \colon V \to W$, sondern **alle** derartigen Abbildungen. Die Menge dieser Abbildungen werde mit $\operatorname{Hom}_K(V, W)$ bezeichnet, *dies ist ein Unterraum von* $\operatorname{Abb}(V, W)$. *Die Zuordnung $C \mapsto f_C$ ist ein Vektorraum-Isomorphismus*

$$M(m \times n, K) \to \operatorname{Hom}_K(K^n, K^m).$$

Zum Beweis: Um zu zeigen, daß $\operatorname{Hom}_K(V, W)$ ein Unterraum von $\operatorname{Abb}(V, W)$ ist, ist zu zeigen: Sind $f, g \in \operatorname{Hom}_K(V, W)$ und ist $\lambda \in K$, so sind auch die Abbildungen $f+g$ und λf K-linear. Der zweite Teil ("umgekehrt ... ") des vorangehenden Satzes besagt, daß diese Zuordnung $C \mapsto f_C$ surjektiv ist. Sie ist injektiv, da die Matrix C durch f_C eindeutig bestimmt ist (dies besagt gerade die Merkregel). Um zu sehen, daß die Zuordnung K-linear ist, muß gezeigt werden, daß für $C, C' \in M(m \times n, K)$ und $\lambda \in K$ gilt:

$$f_{C+C'} = f_C + f_{C'} \quad \text{und} \quad f_{\lambda C} = \lambda f_C.$$

Was bedeutet die Wahl einer Basis?

Satz. *Ist $\mathcal{A} = (v_1, \ldots, v_n)$ eine Basis des K-Vektorraums V, so ist*

$$\Phi_{\mathcal{A}} \colon K^n \to V \quad \textit{mit} \quad \Phi_{\mathcal{A}}(\lambda_1, \ldots, \lambda_n) = \sum \lambda_i v_i$$

(für $\lambda_1, \ldots, \lambda_n \in K$) ein Isomorphismus von Vektorräumen. Um $\Phi_{\mathcal{A}}$ zu definieren, reicht es, $\Phi_{\mathcal{A}}$ auf einer Basis zu definieren: Wir nehmen die kanonische Basis $e_1, \ldots, e_n$ von K^n und definieren

$$\Phi_{\mathcal{A}}(e_i) = v_i.$$

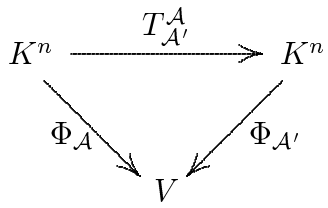
Beweis, daß $\Phi_{\mathcal{A}}$ ein Isomorphismus ist: Die Abbildung $\Phi_{\mathcal{A}}$ ist injektiv, weil $\mathcal{A}$ eine linear unabhängige Folge ist; sie ist surjektiv, weil $\mathcal{A}$ ein Erzeugendensystem ist.

Basiswechsel: *Seien $\mathcal{A}$, $\mathcal{A}'$ zwei Basen eines n-dimensionalen Vektorraums V, dann ist*

$$T^{\mathcal{A}}_{\mathcal{A}'} = \Phi^{-1}_{\mathcal{A}'} \circ \Phi_{\mathcal{A}} \colon K^n \to K^n$$

eine lineare Abbildung $K^n \to K^n$, also durch die Linksmultiplikation mit einer $(n \times n)$-Matrix A gegeben; da $T^{\mathcal{A}}_{\mathcal{A}'}$ ein Isomorphismus ist, muß A invertierbar sein, also eine Matrix in $\mathrm{GL}(n, K)$. Man sagt, daß $\Phi^{-1}_{\mathcal{A}'} \circ \Phi_{\mathcal{A}}$ (oder auch die zugehörige Matrix A) den Basiswechsel beschreibt.

Man hat das folgende (offensichtlich kommutative) Diagramm:



Ist $\mathcal{A} = (v_1, \ldots, v_n)$, $\mathcal{A}' = (v'_1, \ldots, v'_n)$, und schreiben wir einen Vektor $v \in V$ in den beiden Basen in der Form

$$v = \sum \lambda_i v_i = \sum \lambda'_i v'_i,$$

so gilt

$$\begin{bmatrix} \lambda'_1 \\ \vdots \\ \lambda'_n \end{bmatrix} = T^{\mathcal{A}}_{\mathcal{A}'} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}.$$

Beweis: Es ist

$$\Phi_{\mathcal{A}} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = \sum \lambda_i v_i = \sum \lambda'_i v'_i = \Phi_{\mathcal{A}'} \begin{bmatrix} \lambda'_1 \\ \vdots \\ \lambda'_n \end{bmatrix}$$

also

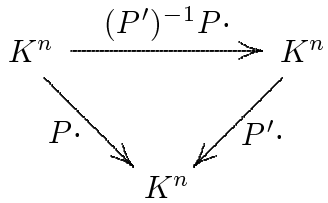
$$T^{\mathcal{A}}_{\mathcal{A}'} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = \Phi^{-1}_{\mathcal{A}'} \circ \Phi_{\mathcal{A}} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = \begin{bmatrix} \lambda'_1 \\ \vdots \\ \lambda'_n \end{bmatrix}$$

Merkregel. *Die j-te* **Spalte** *der darstellenden Matrix $(a_{ij})_{ij}$ zu $T^{\mathcal{A}}_{\mathcal{A}'}$ erhält man dadurch, daß man v_j als Linearkombination in der Basis $\mathcal{A}'$ schreibt: $v_j = \sum_i a_{ij} v'_i$.*

Die linearen Abbildungen $T^{\mathcal{A}}_{\mathcal{A}'}$ und $T^{\mathcal{A}'}_{\mathcal{A}}$ beschreiben den Übergang zwischen zwei verschiedenen Basen $\mathcal{A}$ und $\mathcal{A}'$; natürlich gilt:

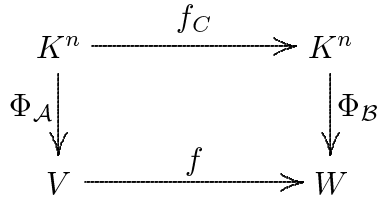
$$T^{\mathcal{A}'}_{\mathcal{A}} = \left(T^{\mathcal{A}}_{\mathcal{A}'}\right)^{-1}.$$

Insbesondere braucht man derartige Basiswechsel im Fall $V = K^n$. Eine Basiswahl $\mathcal{A} = (v_1, \ldots, v_n)$ liefert zu $\Phi_{\mathcal{A}} \colon K^n \to K^n$ eine darstellende Matrix P, deren j-te Spalte gerade das n-Tupel v_j ist (es ist also $\Phi_{\mathcal{A}} = P\cdot$). Ist eine zweite Basis $\mathcal{A}' = (v'_1, \ldots, v'_n)$ gegeben und ist P' die darstellende Matrix zu $\Phi_{\mathcal{A}'}$, also die Matrix, deren j-te Spalte das n-Tupel v'_j ist, so erhalten wir als Basiswechsel-Matrix die Matrix $T^{\mathcal{A}}_{\mathcal{A}'} = (P')^{-1}P$ und das offensichtliche kommutative Diagramm



Wir sehen also: **Jeder endlich erzeugte K-Vektorraum V ist zu einem Vektorraum der Form K^n isomorph.** Durch eine Basiswahl $\mathcal{A}$ von V erhält man einen derartigen Isomorphismus.

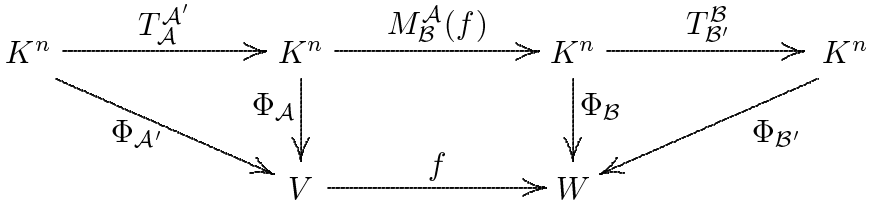
Sind V und W K-Vektorräume mit Basen $\mathcal{A} = (v_1, \ldots, v_n)$, $\mathcal{B} = (w_1, \ldots, w_n)$, und ist $f \colon V \to W$ eine lineare Abbildung, so können wir die lineare Abbildung $\Phi^{-1}_{\mathcal{B}} \circ f \circ \Phi_{\mathcal{A}}$ in der Form f_C mit $C \in M(m \times n, K)$ schreiben. Man notiert dies auch durch das folgende "kommutative Diagramm":



Wir erhalten also durch die Wahl der Basen $\mathcal{A}, \mathcal{B}$ einen Vektorraum-Isomorphismus

$$M^{\mathcal{A}}_{\mathcal{B}} \colon \operatorname{Hom}_K(V, W) \to M(m \times n, K), \quad \text{mit} \quad M^{\mathcal{A}}_{\mathcal{B}}(f) = \Phi^{-1}_{\mathcal{B}} \circ f \circ \Phi_{\mathcal{A}} = f_C.$$

Wie ändert sich die f zugeordnete Matrix $M_{\mathcal{B}}^{\mathcal{A}}(f)$, wenn wir in V und in W Basiswechsel vornehmen? Seien also in V eine zweite Basis $\mathcal{A}'$, in W eine zweite Basis $\mathcal{B}'$ gegeben. Wir setzen die entsprechenden kommutativen Diagramme zusammen:



Wir sehen demnach: Die Matrizen-Darstellung von f bezüglich der Basen $\mathcal{A}'$ und $\mathcal{B}'$ ist durch

$$M_{\mathcal{B}'}^{\mathcal{A}'}(f) = T_{\mathcal{B}'}^{\mathcal{B}} \circ M_{\mathcal{B}}^{\mathcal{A}}(f) \circ T_{\mathcal{A}}^{\mathcal{A}'}$$

gegeben. Ist also f bezüglich der Basen $\mathcal{A}$, $\mathcal{B}$ durch die Matrix C gegeben, so ist f bezüglich der Basen $\mathcal{A}'$ und $\mathcal{B}'$ durch eine Matrix der Form

$$BCA \quad \text{mit invertierbaren Matrizen} \quad A, B$$

gegeben (es ist $f_A = T_{\mathcal{A}}^{\mathcal{A}'}$, $f_C = M_{\mathcal{B}}^{\mathcal{A}}(f)$, und $f_B = T_{\mathcal{B}'}^{\mathcal{B}}$).

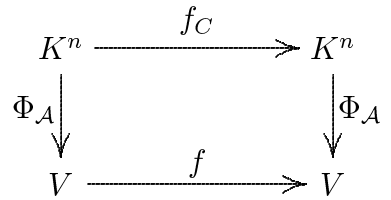
Hier ist nun wie oben herauszuarbeiten, wie sich Aussagen über Matrizen und über lineare Abbildungen entsprechen.

Insbesondere: Sind V, W endlich-dimensionale Vektorräume und ist $f\colon V \to W$ linear, so wissen wir, daß es eine Basis $\mathcal{A} = (v_1, \ldots, v_n)$ von V gibt, so daß die Folge $(v_{r+1}, \ldots, v_n)$ eine Basis von $\operatorname{Ker} f$ und $(f(v_1), \ldots, f(v_r))$ eine Basis von $\operatorname{Im} f$ ist. Setzen wir diese Basis von $\operatorname{Im} f$ zu einer Basis $\mathcal{B} = (f(v_1), \ldots, f(v_r), w_{r+1}, \ldots, w_m)$ von W fort, so erhält man die folgende Matrizendarstellung:

$$M_{\mathcal{B}}^{\mathcal{A}}(f) = \left[\begin{array}{ccc:c} 1 & & & \\ & 1 & 0 & \\ & \ddots & & 0 \\ & 0 & 1 & \\ \hdashline & 0 & & 0 \end{array}\right]$$

Dies ist also ein neuer Beweis, daß jede Matrix durch Multiplikation von rechts und links mit invertierbaren Matrizen in diese Form gebracht werden kann. Und umgekehrt hätten wir auf unsere Kenntnisse über Matrizen zurückgreifen können, um die Formel $\dim \operatorname{Im} f + \dim \operatorname{Ker} f = \dim V$ zu beweisen!

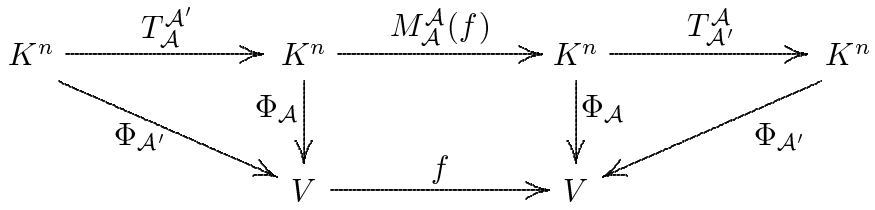
Sei nun $V = W$ und sei $f\colon V \to V$ ein Endomorphismus. Wählen wir zwei verschiedene Basen $\mathcal{A}$ und $\mathcal{B}$ in V, so ändert sich nichts an den obigen Überlegungen. Allerdings ist dies unredlich: Nichts spricht dagegen, V mit K^n zu identifizieren, aber dies besagt, daß man **eine** Basis $\mathcal{A}$ in V wählen darf; wollen wir f eine Matrix zuordnen, so sollte man jeweils nur mit einer Basis arbeiten, also:



Es ist also

$$f_C = M_{\mathcal{A}}^{\mathcal{A}}(f).$$

Ein Basiswechsel von der Basis $\mathcal{A}$ zur Basis $\mathcal{A}'$ liefert folgendes Diagramm:



also

$$M_{\mathcal{A}'}^{\mathcal{A}'}(f) = T_{\mathcal{A}'}^{\mathcal{A}} \circ M_{\mathcal{A}}^{\mathcal{A}}(f) \circ T_{\mathcal{A}}^{\mathcal{A}'}.$$

Demnach ändert sich die Matrix C mit $f_C = M_{\mathcal{A}}^{\mathcal{A}}(f)$ durch einen Basiswechsel (von der Basis $\mathcal{A}$ zur Basis $\mathcal{A}'$) wie folgt: man erhält eine Matrix der Form

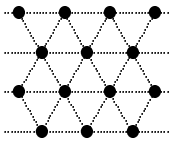
$$ACA^{-1} \quad \text{mit einer invertierbaren Matrizen} \quad A$$

(dabei ist $f_A = T_{\mathcal{A}'}^{\mathcal{A}}$ und $f_{A^{-1}} = T_{\mathcal{A}}^{\mathcal{A}'}$.)

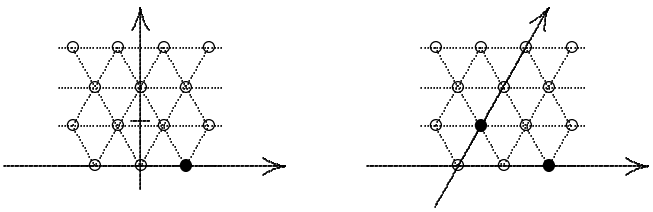
Beispiele: Vektorräume und Basen

Warum betrachtet man neben dem K^n auch "andere" endlich-dimensionale Vektorräume?

(1) Es ist oft wenig hilfreich, mit der kanonischen Basis des K^n zu arbeiten. Beispiel: Sei in der rellen Ebene $\mathbb{R}^2$ folgendes Punkteraster gegeben (die kleinen Dreiecke seien regulär):



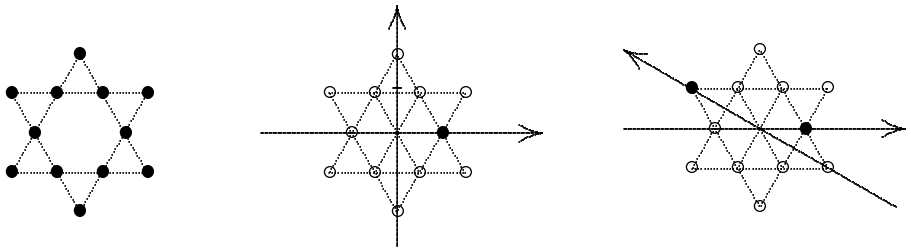
also etwa alle Vektoren in $\mathbb{R}^2$ mit den Koordinaten $(x, 2cy)$ und $(x + \frac{1}{2}, c(2y + \frac{1}{2}))$ mit $x, y \in \mathbb{Z}$ und $c = \frac{1}{2}\sqrt{3}$, wenn wir ein Koordinatensystem wie links gezeigt verwenden.



Viel einfacher ist die Beschreibung der gegebenen Punkte allerdings, wenn wir (wie rechts) als Basis die beiden Vektoren $(1, 0)$ und $(\frac{1}{2}, c)$ verwenden: bezüglich der neuen Basis haben alle Punkte ganzzahlige Kooffizienten.

Wir wollen innerhalb dieses Rasters einen "Davidstern" betrachten, er spielt in der Mathematik als "Wurzelsystem" G_2 eine wichtige Rolle. Es handelt sich dabei um das folgende linke Bild. Zum Zeichnen ist die kanonische Basis sehr gut

geeignet (mittleres Bild), zum Rechnen dagegen ist es schöner mit der Basis rechts zu arbeiten:



Betrachtet werden also die Punkte der Ebene mit den Koordinaten

$$\begin{array}{cccc} & & (0,a) & \\ (-\frac{3}{2},\frac{a}{2}) & (-\frac{1}{2},\frac{a}{2}) & (\frac{1}{2},\frac{a}{2}) & (\frac{3}{2},\frac{a}{2}) \\ & & (-1,0)\ \ (1,0) & \\ (-\frac{3}{2},-\frac{a}{2}) & (-\frac{1}{2},-\frac{a}{2}) & (\frac{1}{2},-\frac{a}{2}) & (\frac{3}{2},-\frac{a}{2}) \\ & & (0,-a) & \end{array}$$

mit $a = \sqrt{3}$ (mittleres Bild; einer der Basisvektoren gehört zum Wurzelsystem, er ist als schwarzer Punkt gezeichnet; der zweite Basisvektor gehört nicht dazu). Wählt man als Basis die Vektoren $(1,0)$ und $(-\frac{3}{2},\frac{a}{2})$ (sie sind rechts als schwarze Punkte gezeichnet) so haben die genannten Vektoren die neuen Koordinaten

$$\begin{array}{cccc} & & (3,2) & \\ (0,1) & (1,1) & (2,1) & (3,1) \\ & -(1,0) & (1,0) & \\ -(3,1) & -(2,1) & -(1,1) & -(0,1) \\ & & -(3,2) & \end{array}$$

(alle Koordinaten sind nun ganzzahlig; die Koordinaten jedes Punkts sind jetzt entweder beide nicht-negativ oder beide nicht-positiv).

(2) Vektorräume sind häufig als Unterräume eines K^n definiert, und derartige Unterräume haben meist keine irgendwie ausgezeichnete Basis. Zum Beispiel: der Unterraum $\{(x_1,\ldots,x_n) \in \mathbb{Q}^n \mid \sum x_i = 0\}$.

Lineare Gleichungssysteme (II).

Sei K ein Körper. Ein lineares Gleichungssystem mit Koeffizienten in K ist von folgender Form:

$$\sum_{j=1}^{n} a_{ij} X_j = b_i \quad \text{mit} \quad 1 \leq i \leq m$$

dabei sind $a_{ij}, b_i \in K$, für alle $1 \leq i \leq m$, $1 \leq j \leq n$. Ist $b_i = 0$ für alle i, so spricht man von einem *homogenen* linearen Gleichungssystem. Wie im Fall $K = \mathbb{Q}$ bezeichnen wir mit A die Koeffizientenmatrix, mit $[A, b]$ die erweiterte Koeffizientenmatrix, und wir setzen

$$\text{Lös}(A, b) = \{\, (z_1, \ldots, z_n) \in K^n \mid \sum_{j=1}^{n} a_{ij} z_j = b_i \quad \text{für alle} \quad 1 \leq i \leq m \,\}.$$

Die Matrix A liefert eine Abbildung

$$f_A \colon K^n \to K^m$$

und es gilt:

(1) $\text{Lös}(A, 0) = \text{Ker}\, f_A$, also: *Die Lösungsmenge eines homogenen Gleichungssystems ist ein Unterraum.*

(2) *Ist der Rang der Matrix A gleich r, so hat* $\text{Lös}(A, 0)$ *die Dimension $n - r$.*

(3) *Genau dann ist* $\text{Lös}(A, b)$ *nicht leer, wenn gilt:* $b \in \text{Im}\, f_A$ (also genau dann, wenn b eine Linearkombination der Spalten von A ist).

(4) *Ist* $y \in \text{Lös}(A, b)$, *so ist* $\text{Lös}(A, b) = y + \text{Lös}(A, 0)$ (also: kennt man **eine** spezielle Lösung y in $\text{Lös}(A, b)$, so erhält man **alle** Lösungen, indem man zu y alle Lösungen des zugehörigen homogenen Gleichungssytems addiert).

Satz. *Jeder Unterraum von K^n ist Lösungsraum eines homogenen linearen Gleichungssystems.*

Beweis: Sei $U \subseteq K^n$ ein Unterraum, sei $u_1, \ldots, u_s$ eine Basis von U. Dabei sei $u_i = (u_{i1}, \ldots, u_{in})$ mit Koeffizienten $u_{ij} \in K$ und $1 \leq i \leq s$. Betrachte das homogene Gleichungssystem mit Koeffizientenmatrix $C = (u_{ij})_{ij}$. Der Rang der Matrix f_C ist s, also hat $\text{Lös}(C, 0)$ die Dimension $r = n - s$. Sei $a_1, \ldots, a_r$ eine Basis von $\text{Lös}(C, 0)$, mit Vektoren $a_i = (a_{i1}, \ldots, a_{in}) \in K^n$, dabei ist $1 \leq i \leq r$. Sei $A = (a_{ij})_{ij}$. Dann gilt: $\text{Lös}(A, 0) = U$ (offensichtlich sind die Vektoren u_i in $\text{Lös}(A, 0)$; andererseits hat $\text{Lös}(A, 0)$ die Dimension $n - r = s$.)

Komplementärbasen. Sei V ein Vektorraum, U ein Unterraum. Eine Folge $(v_1, \dots, v_t)$ von Vektoren aus V heißt *linear unabhängig modulo* U, falls folgendes gilt: sind λ_i Skalare in K und ist $\sum_{i=1}^{t} \lambda_i v_i \in U$, so sind alle $\lambda_i = 0$. Ist die Folge $(v_1, \dots, v_t)$ linear unabhängig modulo U, so ist diese Folge natürlich linear unabhängig. Die Folge $(v_1, \dots, v_t)$ von Vektoren aus V heißt *Komplementärbasis* von U in V, falls erstens die Folge $(v_1, \dots, v_t)$ linear unabhängig modulo U ist, und zweitens U zusammen mit den Vektoren $v_1, \dots, v_t$ den Vektorraum V erzeugt.

(1) *Sei U ein Unterraum von V und $(u_1, \dots, u_m)$ eine Basis von U. Seien $v_1, \dots, v_t$ Vektoren in V. Genau dann ist $(v_1, \dots, v_t)$ eine Komplementärbasis von U in V, wenn $(u_1, \dots, u_m, v_1, \dots, v_t)$ eine Basis von V ist.*

(2) *Ist $0 = V_0 \subseteq V_1 \subseteq V_2 \subseteq \cdots \subseteq V_{m-1} \subseteq V_m = V$ eine Kette von Unterräumen, und sind $(v_{i,1}, \dots, v_{i,n_i})$ Folgen von Vektoren in V_i, so gilt: Genau dann ist $(v_{i,1}, \dots, v_{i,n_i})$ eine Komplementärbasis von V_{i-1} in V_i, für $1 \leq i \leq m$, wenn die Folge*

$$(v_{1,1}, \dots, v_{1,n_1}, v_{2,1}, \dots, v_{2,n_2}, \dots, v_{m,1}, \dots, v_{m,n_m})$$

eine Basis von V ist.

(3) *Sei V ein endlich-dimensionaler Vektorraum. Sei U ein Unterraum von V. Eine Folge von Vektoren in V, die linear unabhängig modulo U sind, läßt sich zu einer Komplementärbasis von U in V ergänzen.*

Teil 4. Endomorphismen

Seien V, W Vektorräume. Sei $f\colon V \to W$ eine lineare Abbildung.
(a) *Genau dann ist f injektiv, wenn* $\operatorname{Ker} f = 0$ *gilt.*
(b) *Genau dann ist f surjektiv, wenn* $\operatorname{Im} f = W$ *gilt.*
(c) *Sind V, W endlich-dimensional und ist* $\dim V = \dim W$*, so ist f genau dann injektiv, wenn f surjektiv ist (und dann also ein Isomorphismus).*

Zum Beweis: Die Äquivalenz in (a) gilt ganz allgemein für Gruppen (wobei man bei einer nicht-abelschen Gruppe $\operatorname{Ker} f = 1$ schreiben würde): Ist nämlich $\operatorname{Ker} f = 0$ und ist $f(v) = f(v')$, so ist $f(v - v') = f(v) - f(v') = 0$, also $v - v' = 0$, also $v = v'$. Die umgekehrte Implikation ist natürlich trivial. Zu (b) ist nichts zu zeigen. (c) ist eine unmittelbare Folgerung der Dimensiongleichung

$$\dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim V$$

(und der trivialen Implikationen $\dim U = 0 \implies U = 0$ und $\dim U = \dim V \implies U = V$, für jeden Unterraum U von V).

Zur Erinnerung: Sei V ein n-dimensionaler Vektorraum und $f\colon V \to V$ ein Endomorphismus. Ist $\mathcal{A} = (v_1, \dots, v_n)$ eine Basis von V, so nennt man $M_{\mathcal{A}}^{\mathcal{A}}(f)$ die Matrizendarstellung von f bezüglich der Basis $\mathcal{A}$. Man erhält $M_{\mathcal{A}}^{\mathcal{A}}(f) = C = (c_{ij})_{ij}$ wie folgt: schreibe $f(v_j)$ als Linearkombination $f(v_j) = \sum_i c_{ij} v_i$ der Vektoren $v_1, \dots, v_n$, die Koeffizienten liefern die j-te Spalte der Matrix. *Wählen wir eine*

andere Basis von V, so erhalten wir eine Matrix der Form $C' = PCP^{-1}$, dabei ist P eine invertierbare $(n \times n)$-Matrix.

Definition: Zwei $(n \times n)$-Matrizen C, C' mit Koeffizienten im Körper K heißen *ähnlich,* wenn es eine invertierbare $(n \times n)$-Matrix P mit Koeffizienten in K gibt, so daß $C' = PCP^{-1}$ gibt.

Sei A eine $(n \times n)$-Matrix. Die folgenden Aussagen sind äquivalent:

(i) *A ist invertierbar.*
(ii) *Die Spalten von A sind linear unabhängig in K^n.*
(ii*) *Die Zeilen von A sind linear unabhängig in K^n.*
(iii) $\det A \neq 0$.
(iv) *Die lineare Abbildung $f_A \colon K^n \to K^n$ ist ein Isomorphismus.*

Zum Beweis: (i) $\Longrightarrow$ (ii): Sind die Spalten linear abhängig, so gibt es einen von Null verschiedenen Vektor v mit $Av = 0$. Wäre A invertierbar, so würde gelten $v = A^{-1}Av = 0$, Widerspruch.

(ii) $\Longrightarrow$ (i): Die Spalten von A seien linear unabhängig, sie bilden eine Basis von K^n. Also kann ich die kanonischen Basisvektoren e_i als Linearkombination der Spalten ausdrücken. Es gibt also einen Vektor v_i mit $Av_i = e_i$ (denn Linearkombinationen der Spalten von A zu bilden, heißt gerade, mit einem Spaltenvektor zu multiplizieren). Die Matrix A', deren Spalten $v_1, \ldots, v_n$ sind, liefert demnach $AA' = I_n$.

Entsprechend sieht man, daß (i) und (ii*) äquivalent sind. Die Äquivalenz von (i), (iii) und (iv) wurde schon früher bewiesen.

(4.1) Eigenwerte, Eigenvektoren.

Sei V ein K-Vektorraum, sei $f \colon V \to V$ ein Endomorphismus. Sei $\lambda \in K$. Ein Vektor $v \neq 0$ in V heißt *Eigenvektor zu f mit Eigenwert λ* falls gilt

$$f(v) = \lambda \cdot v.$$

Also: v wird unter f auf ein Vielfaches von sich abgebildet. Nach Voraussetzung sind Eigenvektoren immer von Null verschieden (dagegen darf $\lambda = 0$ sein; Eigenvektoren mit Eigenwert 0 sind gerade alle von 0 verschiedene Vektoren im Kern von f).

Satz. *Eigenvektoren zu paarweise verschiedenen Eigenwerten sind linear unabhängig.*

Beweis: Sei v_i ein Eigenvektor mit Eigenwert λ_i, mit $1 \leq i \leq t$. Seien die Elemente $\lambda_i \in K$ paarweise verschieden. Sei $\sum_{i=1}^{t} \mu_i v_i = 0$. Zu zeigen: $\mu_i = 0$ für alle i. Wir führen Induktion nach t. Induktionsanfang: Ist $t = 1$, so ist $\mu_1 = 0$.

Induktionsschritt. Sei $t \geq 2$. Sei also $0 = \sum_{i=1}^{t} \mu_i v_i$ gegeben, dabei können wir voraussetzen $\mu_i \neq 0$ für alle i. Wende auf diese Gleichung den Endomorphismus f an, wir erhalten

$$0 = f(0) = f(\sum_{i=1}^{t} \mu_i v_i) = \sum_{i=1}^{t} \mu_i f(v_i) = \sum_{i=1}^{t} \mu_i \lambda_i v_i.$$

Andererseits ist mit $0 = \sum_{i=1}^{t} \mu_i v_i$ auch $0 = \sum_{i=1}^{t} \mu_i \lambda_t v_i$, also

$$0 = \sum_{i=1}^{t} \mu_i (\lambda_i - \lambda_t) v_i$$

Dies ist aber eine Linearkombination der $t-1$ Eigenvektoren $v_1, \ldots, v_{t-1}$, denn der Koeffizient von v_t ist Null. Also folgt nach Induktion

$$\mu_i(\lambda_i - \lambda_t) = 0 \quad \text{für} \quad 1 \leq i < t.$$

Für diese i gilt aber $\lambda_i \neq \lambda_t$, also folgt $\mu_i = 0$ für $1 \leq i < t$. Widerspruch zur Annahme, daß $\mu_i \neq 0$ für alle i gilt (hier verwenden wir $t \geq 2$).

Korollar. *Sei V ein n-dimensionaler Vektorraum. Jeder Endomorphismus $f\colon V \to V$ besitzt höchstens n Eigenwerte.*

Eigenraum. Sei $f\colon V \to V$ Endomorphismus eines K-Vektorraums. Ist $\lambda \in K$, so sei $\mathrm{Eig}(f, \lambda)$ die Menge der Vektoren $v \in V$ mit $f(v) = \lambda v$. Man nennt dies den *Eigenraum zu f und λ*. Die von Null verschiedenen Elemente in $\mathrm{Eig}(f, \lambda)$ sind gerade die Eigenvektoren zu f mit Eigenwert λ. *Die Menge* $\mathrm{Eig}(f, \lambda)$ *ist ein Unterraum von V* (einerseits gehört die Null dazu, andererseits ist diese Menge abgeschlossen unter Addition und Skalarmultiplikation).

Satz. *Sei V ein n-dimensionaler Vektorraum. Sei $f\colon V \to V$ ein Endomorphismus. Es ist*

$$\sum_{\lambda \in K} \dim \mathrm{Eig}(f, \lambda) \leq n.$$

Beweis: Wähle für jedes $\lambda \in K$ eine Basis $\mathcal{A}(\lambda) = (v_1^{(\lambda)}, \ldots, v_{n(\lambda)}^{(\lambda)})$ von $\mathrm{Eig}(f, \lambda)$. Behauptung: Die Vektoren $v_i^{(\lambda)}$ mit $1 \leq i \leq n(\lambda)$ und $\lambda \in K$ sind linear unabhängig. Sei nämlich eine Linearkombination $\sum_{\lambda,i} \mu_{\lambda,i} v_i^{(\lambda)} = 0$ gegeben (dies soll eine **endliche** Summe sein, es seien also nur endlich viele Koeffizienten $\mu_{\lambda,i} \neq 0$). Setze $v_\lambda = \sum_i \mu_{\lambda,i} v_i^{(\lambda)}$, für jedes λ. Dies ist entweder der Nullvektor oder ein Eigenvektor von f mit Eigenwert λ. Aber es ist $\sum_\lambda v_\lambda = 0$, dies ist nur möglich wenn $v_\lambda = 0$ ist für jedes λ (denn Eigenvektoren zu paarweise verschiedenen Eigenwerten sind linear unabhängig). Aus $\sum_i \mu_{\lambda,i} v_i^{(\lambda)} = v_\lambda = 0$ folgt nun $\mu_{\lambda,i} = 0$ für alle i.

(4.2) Die Eigenwerte einer Matrix A und das charakteristische Polynom χ_A von A.

Ist $A \in M(n \times n, K)$, so haben wir für die durch A gegebene lineare Abbildung $f_A\colon K^n \to K^n$ geschrieben; dabei verstehen wir unter K^n den Raum $K^n = M(n \times 1, K)$ aller Spaltenvektoren der Länge n und f_A ist die Matrizenmultiplikation $M(n \times n, K) \times M(n \times 1, K) \to M(n \times 1, K)$. Genauer müßten wir aber $f_{A,K}$ statt f_A schreiben, denn man kann der Matrix A nicht ansehen, mit welchem K wir arbeiten! Dabei ist nicht so sehr daran gedacht, daß wir für jeden Körper K das Nullelement mit 0, das Einselement mit 1 bezeichnen, also über

jedem Körper K eine Matrix zum Beispiel der Form $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ haben; das Nullelement wie das Einselement von $\mathbb{Q}$ und von $\mathbb{Z}/p\mathbb{Z}$ werden zwar gleich bezeichnet, sind jedoch völlig verschiedene Elemente. Wegen $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ sind aber das Nullelement und das Einselement von $\mathbb{Q}$ auch Elemente von $\mathbb{R}$ wie von $\mathbb{C}$, die Matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \in M(2\times 2, \mathbb{Q})$ gehört demnach auch zu $M(2\times 2, \mathbb{R})$ und zu $M(2\times 2, \mathbb{C})$.

Ist λ ein Eigenwert von $f_{A,K}$, so nennt man λ einen *Eigenwert der Matrix* $A \in M(n \times n, K)$. Hier ist die Fixierung des Körpers K ganz wichtig, wie wir gleich sehen werden: **Beispiel.** Betrachte die Matrix $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in M(2 \times 2, \mathbb{Q})$. Sie besitzt **keinen** Eigenwert, auch nicht, wenn wir sie als Matrix in $M(2 \times 2, \mathbb{R})$ auffassen. Betrachten wir sie dagegen als Matrix in $M(2 \times 2, \mathbb{C})$, so besitzt sie die Eigenwerte $i, -i$.

Sei $A \in M(n \times n, K)$. Genau dann ist $\lambda \in K$ Eigenwert von A, wenn $A - \lambda I$ nicht invertierbar ist, wenn also $\det(A - \lambda I) = 0$ ist. Man setzt $\mathrm{Eig}(A, \lambda) = \mathrm{Eig}(f_{A,K}, \lambda)$. *Es ist* $\mathrm{Eig}(A, \lambda)$ *der Kern von* $f_{A-\lambda 1}$. Beweis: Sei $\lambda \in K$. Es ist $v \in \mathrm{Eig}(A, \lambda)$ genau dann, wenn $Av = \lambda v$ gilt, also genau dann, wenn $(A-\lambda I)v = 0$ gilt, also genau dann, wenn v zu $\mathrm{Ker}(f_{A-\lambda I})$ gehört.

Nullstellen von Polynomen. Sei $f \in K[T]$ ein von Null verschiedenes Polynom. Sei $\lambda \in K$. Man nennt λ eine *Nullstelle* von f, falls $f(\lambda) = 0$ gilt. Ist $f = (T-\lambda)g$ für ein $g \in K[T]$, so ist λ Nullstelle von f (es gilt auch die Umkehrung). Ist $f = (T-\lambda)^m h$ für ein Polynom $h \in K[T]$ und $m \in \mathbb{N}_0$, so sagt man, daß λ eine Nullstelle der *Vielfachheit mindestens* m ist. Die Bestimmung von Nullstellen eines Polynoms ist ein wichtiges Thema der Mathematik.

Satz. *Die Nullstellen des charakteristischen Polynoms χ_A einer Matrix A sind gerade die Eigenwerte von f_A.*

Beweis: Sei $\lambda \in K$. Es ist $\chi_A(\lambda) = 0$ genau dann, wenn $\det(A - \lambda I) = 0$ gilt, genau dann, wenn $A-\lambda I$ nicht invertierbar ist, also genau dann, wenn $\mathrm{Ker}(f_{A-\lambda I}) \neq 0$ gilt.

Satz. *Sei $A \in M(n\times n, K)$ und $\lambda \in K$. Die Vielfachheit von λ ist mindestens* $\dim \mathrm{Eig}(A, \lambda)$.

Beweis: Sei $(v_1, \ldots, v_t)$ eine Basis von $\mathrm{Eig}(A, \lambda)$. Vervollständige diese Folge zu einer Basis $(v_1, \ldots, v_n)$ von K^n. Sei P die Matrix, deren Spalten die Vektoren $v_1, \ldots, v_n$ sind. Dann ist $AP = PB$, wobei B die folgende Form hat: Für $1 \leq j \leq t$ ist die j-te Spalte von B gerade λe_j, denn v_j ist Eigenvektor mit Eigenwert λ, die restlichen Spalten dagegen sind unübersichtlich (genauer: Ist $j > t$ und $Av_j = \sum_i c_{ij} v_i$, so ist die j-te Spalte gerade $\begin{bmatrix} c_{1j} \\ \vdots \\ c_{nj} \end{bmatrix}$). Da P invertierbar ist, folgt $P^{-1}AP = B$. Da die Matrix B die folgende Form hat:

$$\begin{bmatrix} \underbrace{\begin{matrix} \lambda & & \\ & \ddots & \\ & & \lambda \end{matrix}}_{t} & * \\ 0 & B' \end{bmatrix}$$

ist $\chi_A = (\lambda - T)^t \chi_{B'}$. Also ist die Vielfachheit von λ mindestens $\dim \mathrm{Eig}(A, \lambda)$.

Wichtig. Die Vielfachheit von λ als Nullstelle von χ_A kann echt größer als $\dim \mathrm{Eig}(A, \lambda)$ sein. Beispiel: $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Die Nullstelle $\lambda = 0$ hat die Vielfachheit 2 in $\chi_A = T^2$, aber $\dim \mathrm{Eig}(A, 0) = 1$.

Lemma. *Ähnliche Matrizen haben das gleiche charakteristische Polynom.*

Beweis: Seien A, A' ähnliche $(n \times n)$-Matrizen. Es gibt also eine invertierbare Matrix P mit $A' = PAP^{-1}$. Wegen $T \cdot I_n = P(T \cdot I_n)P^{-1}$ gilt $A' - T \cdot I_n = PAP^{-1} - P(T \cdot I_n)P^{-1} = P(A - T \cdot I_n)P^{-1}$. Der Determinanten-Multiplikationssatz liefert:

$$\begin{aligned} \chi_{A'} &= \det(A' - T \cdot I_n) = \det P(A - T \cdot I_n)P^{-1} \\ &= \det P \cdot \det(A - T \cdot I_n) \cdot \det P^{-1} = \det P \cdot \chi_A \cdot (\det P)^{-1} = \chi_A. \end{aligned}$$

Ist $f \colon V \to V$ ein Endomorphismus eines endlich-dimensionalen Vektorraums und $\mathcal{A}$ eine Basis von V, so hängt das charakteristische Polynom von $M_{\mathcal{A}}^{\mathcal{A}}(f)$ nur von f nicht aber von der ausgewählten Basis $\mathcal{A}$ ab, man nennt es deshalb auch das *charakteristische Polynom von* f.

Wie findet man Eigenwerte und Eigenvektoren? Sei $f \colon V \to V$ Endomorphismus eines endlich-dimensionalen Vektorraums.
(0) Wähle eine Basis $\mathcal{A}$ und berechne die Matrizendarstellung $A = M_{\mathcal{A}}^{\mathcal{A}}(f)$.
(1) Berechne das charakteristische Polynom χ_A der Matrix A.
(2) Berechne die Nullstellen von χ_A.
(3) Ist λ eine Nullstelle, berechne eine Basis von $\mathrm{Eig}(A, \lambda)$.

Beispiele. **(a)** Sei V der Vektorraum der Polynome vom Rang höchstens 4 mit Koeffizienten in $\mathbb{Q}$ und f das Differenzieren.
(0) Wähle die Basis $\mathcal{A} = (1, T, \ldots, T^4)$. Es ist

$$A = M_{\mathcal{A}}^{\mathcal{A}}(f) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

(1) Das charakteristische Polynom ist $\chi_A(T) = T^5$.
(2) Die einzige Nullstelle von T^5 ist $\lambda = 0$.
(3) Der Eigenraum $\mathrm{Eig}(A, 0)$ besteht aus den konstanten Polynomen, ist also eindimensional.

(b) Sei $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \in M(2 \times 2, \mathbb{R})$ (man nennt sie die "Fibonacci-Matrix").
(1) Das charakteristische Polynom ist $\chi_A(T) = T^2 - T - 1$.
(2) In $\mathbb{R}$ besitzt $\chi_A(T)$ die Nullstellen $\lambda_1 = \frac{1}{2}(1 + \sqrt{5})$ und $\lambda_2 = \frac{1}{2}(1 - \sqrt{5})$.

(3) Der Unterraum $E(A, \lambda_i)$ wird von $v_i = \begin{bmatrix} 1 \\ \lambda_i \end{bmatrix}$ erzeugt.

Die Matrix A ist, wie wir sagen werden, "diagonalisierbar".

(4.3) Diagonalisierbare Matrizen.

Eine $n \times n$-Matrix A heißt *diagonalisierbar,* falls A zu einer Diagonalmatrix ähnlich ist (falls es also eine invertierbare Matrix P gibt, so daß $P^{-1}AP$ eine Diagonalmatrix ist).

Satz. *Sei A eine $(n \times n)$-Matrix. Genau dann ist A diagonalisierbar, wenn es in K^n eine Basis von Eigenvektoren zu $f_A \colon K^n \to K^n$ gibt.*

Beweis (und gleichzeitig Regel zur Bestimmung von P): Sei $(v_1, \ldots, v_n)$ eine Basis aus Eigenvektoren, sei λ_i der Eigenwert zu v_i. Bilde die $(n \times n)$-Matrix P, deren j-te Spalte gerade v_j ist. Wegen $Av_j = \lambda_j v_j$ entsteht also AP aus P, indem die j-te Spalte von P mit dem Skalar λ_i multipliziert wird. Ist demnach D die Diagonalmatrix mit Diagonalkoeffizienten $\lambda_1, \ldots, \lambda_n$, so ist $AP = PD$. Nun ist P invertierbar, also $P^{-1}AP = D$.

Umgekehrt gilt: Ist $P^{-1}AP = D = (d_{ij})_{ij}$ eine Diagonalmatrix, so ist die i-te Spalte von A ein Eigenvektor für A mit Eigenwert d_{ii}. (Denn $AP = PD$ und diese Matrix entsteht aus P, in dem die i-te Spalte von P mit d_{ii} multipliziert wird. Ist also v_i die i-te Spalte von P, so ist $f_A(v_i) = d_{ii}v_i$.)

Zweite Formulierung. *Sei V ein n-dimensionaler Vektorraum, sei $f \colon V \to V$ ein Endomorphismus. Genau dann gibt es eine Basis aus Eigenvektoren zu f, wenn f durch eine Diagonalmatrix darstellbar ist.*

Zur Erinnerung 1: Eine $(n \times n)$-Matrix besitzt höchstens n Eigenwerte. *Besitzt sie n Eigenwerte, so ist sie diagonalisierbar.* Beweis: Sind $(v_1, \ldots, v_t)$ Eigenvektoren zu A mit paarweise verschiedenen Eigenwerten, so ist diese Folge linear unabhängig, also $t \leq n$. Ist $t = n$, so ist dies eine Basis aus Eigenvektoren.

Zur Erinnerung 2: Sei A eine $(n \times n)$-Matrix. Es ist $\sum_{\lambda \in K} \dim \mathrm{Eig}(A, \lambda) \leq n$. *Gleichheit gilt genau dann, wenn A diagonalisierbar ist.* Beweis: Sei $n(\lambda) = \dim \mathrm{Eig}(A, \lambda)$. Ist $\sum_{\lambda \in K} n(\lambda) = n$, so haben wir eine Basis aus Eigenvektoren konstruiert. Umgekehrt: Ist A eine Diagonalmatrix, so gibt es eine Basis $(v_1, \ldots, v_n)$ aus Eigenvektoren. Ist $m(\lambda)$ die Anzahl der Vektoren v_i mit Eigenvektor λ, so ist also $\sum_\lambda m(\lambda) = n$. Andererseits ist $m(\lambda) \leq n(\lambda)$. Insgesamt sehen wir:

$$n = \sum_\lambda m(\lambda) \leq \sum_\lambda n(\lambda) \leq n,$$

und demnach gilt überall das Gleichheitszeichen (und es ist auch $m(\lambda) = n(\lambda)$ für alle λ).

Zur Erinnerung 3: Ist $p(\lambda)$ die Vielfachheit der Nullstelle λ im Polynom χ_A, so ist $\dim \mathrm{Eig}(A, \lambda) \leq p(\lambda)$. Ist $p(\lambda) \neq 0$, so ist

$$1 \leq \dim \mathrm{Eig}(A, \lambda) \leq p(\lambda).$$

Läßt sich χ_A als Produkt von Linearfaktoren schreiben, so ist A genau dann diagonalisierbar, wenn $\dim \mathrm{Eig}(A, \lambda) = p(\lambda)$ für alle $\lambda \in K$ gilt.

Noch einmal das Beispiel $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Nur $\lambda = 0$ ist Eigenwert, und $\mathrm{Eig}(A, 0)$ ist eindimensional. Also ist $\sum \dim \mathrm{Eig}(A, \lambda) = 1$, aber $n = 2$. Es ist $\chi_A = T^2$, einzige Nullstelle ist $\lambda = 0$ und $p(0) = 2$. Die Matrix A ist nicht diagonalisierbar.

Beispiel: Fibonacci-Zahlen. Dies ist die Folge der Zahlen

$$a_0 = 0,\ a_1 = 1,\ a_2 = 1,\ a_3 = 2,\ a_4 = 3,\ a_5 = 5,\ a_6 = 8, \ldots,$$

dabei ist $a_0 = 0, a_1 = 1$, und man erhält die weiteren Zahlen durch $a_{n+1} = a_{n-1} + a_n$, für $n \geq 1$.

Die Fibonacci-Matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ liefert diese Folge vermöge

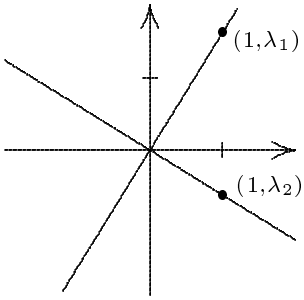
$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} a_{n-1} \\ a_n \end{bmatrix} = \begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix}, \quad \text{also} \quad \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix}.$$

Fassen wir A als Matrix in $M(2 \times 2, \mathbb{R})$ auf, so besitzt A zwei Eigenwerte

$$\lambda_1 = \frac{1}{2}(1 + \sqrt{5}) \approx 1,6180, \quad \lambda_2 = \frac{1}{2}(1 - \sqrt{5}) \approx -0,6180.$$

Ein Eigenvektor zu λ_i ist $v_i = \begin{bmatrix} 1 \\ \lambda_i \end{bmatrix}$, denn $\lambda_i^2 = 1 + \lambda_i$.

Eigenvektoren:



Da wir eine Basis aus Eigenvektoren kennen, wissen wir, daß A diagonalisierbar ist:

$$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = P^{-1}AP = \frac{1}{\lambda_2 - \lambda_1} \begin{bmatrix} \lambda_2 & -1 \\ -\lambda_1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{bmatrix}.$$

Das Diagonalisieren erlaubt es, hohe Potenzen von A zu berechnen:

$$A^{1000} = P \cdot \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}^{1000} \cdot P^{-1} = P \begin{bmatrix} \lambda_1^{1000} & 0 \\ 0 & \lambda_2^{1000} \end{bmatrix} P^{-1},$$

und demnach auch zum Beispiel a_{1000} (denn dies ist gerade der Koeffizient von A^{1000} an den Positionen $(2,1)$ und $(1,2)$).

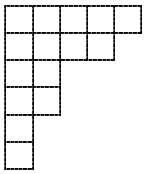
Die Matrix A^2 hat die gleichen Eigenvektoren wie A, ihre Eigenwerte sind beide positiv, nämlich

$$\lambda_1^2 \approx 2,618 \quad \text{und} \quad \lambda_2^2 \approx 0,382\,.$$

Schreiben wir einen beliebigen Vektor $v \in \mathbb{R}^2$ als Linearkombination der Eigenvektoren v_1, v_2, etwa $v = c_1v_1 + c_2v_2$, so läßt sich die Folge der Bilder $v, f^2(v), f^4(v), \ldots$ sehr schön beschreiben: in der v_1-Richtung erfolgt jeweils eine Streckung (mit dem Faktor λ_1^2), in der v_2-Richtung eine Schrumpfung (mit dem Faktor λ_2^2).

(4.4) Nilpotente Matrizen.

Sei n eine natürliche Zahl. Eine *Partition* von n ist eine Folge $\lambda = (\lambda_1, \ldots, \lambda_t)$ von natürlichen Zahlen λ_i, so daß gilt $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_t \geq 1$ und $\sum_i \lambda_i = n$. Jeder Partition ordnet man ein sogenanntes *Young-Diagramm* zu: man betrachtet ein Kästchenmuster mit t Kästchenreihen, linksbündig untereinander gesetzt, wobei die t-te Reihe aus λ_i Kästchen besteht. Analog zur Indizierung der Positionen in einer Matrix kann man diese Kästchen durch die Pare (i, j) mit $1 \leq j \leq \lambda_i$ und $1 \leq i \leq t$ indizieren. Beispiel: Das Young-Digramm zur Partition $\lambda = (5, 4, 2, 2, 1, 1)$ hat die Form



Wir wollen die Kästchen durchnummerieren, und zwar zeilenweise, von links nach rechts, und von oben nach unten:

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)
(2,1)	(2,2)	(2,3)	(2,4)	
(3,1)	(3,2)			
(4,1)	(4,2)			
(5,1)				
(6,1)				

1	2	3	4	5
6	7	8	9	
10	11			
12	13			
14				
15				

jeder Position (i, j) haben wir auf diese Weise eine Zahl $\nu(i, j)$ zugeordnet, ν ist eine Bijektion zwischen der Menge $\{(i, j) \mid 1 \leq j \leq \lambda_i,\ 1 \leq i \leq t\}$ und $\{1, 2, \ldots, n\}$. Natürlich können wir ν durch eine Formel festlegen: es ist $\nu(i, j) = j + \sum_{r<i} \lambda_r$.

Ist $\lambda = (\lambda_1, \ldots, \lambda_t)$ eine Partition von n, so bildet man die duale Partition $\lambda' = (\lambda'_1, \ldots, \lambda'_{\lambda_1})$ wie folgt: Es sei $\lambda'_j = |\{\lambda_i \mid \lambda_i \geq j\}|$, dies ist im Young-Diagramm gerade die Anzahl der Kästchen in der j-ten Spalte. Insbesondere ist λ' ebenfalls eine Partition von n. Die zu $\lambda = (5, 4, 2, 2, 1, 1)$ duale Partition ist $\lambda' = (6, 4, 2, 2, 1)$.

Jeder Partition λ von n ordnet man einen Endomorphismus $f_\lambda \colon K^n \to K^n$ wie folgt zu. Man benennt die kanonischen Basisvektoren $e_1, \ldots, e_n$ um, und zwar setzt man $e_{ij} = e_{ij}^{(\lambda)} = e_{\nu(ij)}$. Setze

$$f_\lambda(e_{ij}) = e_{i,j-1} \text{ falls } j > 1, \text{ und } f_\lambda(e_{i1}) = 0 \quad \text{für alle} \quad i.$$

Wir ordnen der Partition λ auch eine $(n\times n)$-Matrix $J(\lambda)$ zu (die *Jordan-Matrix* zur Partition λ mit Eigenwert 0), hier als typisches Beispiel der Fall $\lambda = (5,4,2,2,1,1)$:

$$J((5,4,2,2,1,1)) = \begin{bmatrix} 0 & 1 & & & & & & & & & & \\ & 0 & 1 & & & & & & & & & \\ & & 0 & 1 & & & & & & & & \\ & & & 0 & 1 & & & & & & & \\ & & & & 0 & & & & & & & \\ & & & & & 0 & 1 & & & & & \\ & & & & & & 0 & 1 & & & & \\ & & & & & & & 0 & 1 & & & \\ & & & & & & & & 0 & & & \\ & & & & & & & & & 0 & & \\ & & & & & & & & & & 0 & \end{bmatrix}$$

wobei alle weiteren Einträge Nullen sind. Die allgemeine Regel lautet: es ist $J(\lambda) = (a_{ij})_{ij}$ mit $a_{r,r+1} = 1$ für alle r, die nicht von der Form $\sum_{i\le s}\lambda_i$ sind, und $a_{ij} = 0$ sonst. Entlang der Diagonale sind also entsprechende $(\lambda_i \times \lambda_i)$-Matrizen aufgereiht.

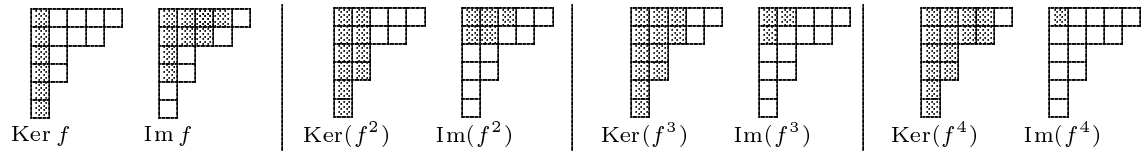
Natürlich ist $J(\lambda)$ die Matrizendarstellung von f_λ bezüglich der kanonischen Basis des K^n.

Satz. *Sei λ eine Partition. Es ist*

$$\begin{aligned}\operatorname{Ker}(f_\lambda^s) &= \operatorname{span}\{e_{ij} \mid 1 \le j \le s,\ 1 \le i \le \lambda'_j\},\\ \operatorname{Im}(f_\lambda^s) &= \operatorname{span}\{e_{ij} \mid 1 \le i \le t,\ 1 \le j \le \lambda_i - s\}.\end{aligned}$$

und

$$\dim\operatorname{Ker}(f_\lambda^s) = \sum_{j=1}^{s}\lambda'_j \quad \text{und} \quad \dim\operatorname{Im}(f_\lambda^s) = \sum_{j\ge s+1}\lambda'_j.$$



Beweis: Sei $f = f_\lambda$. Es ist

$$f^s(e_{ij}) = e_{i,j-s} \quad \text{falls} \quad j > s, \quad \text{und} \quad f^s(e_{ij}) = 0 \quad \text{falls} \quad j \le s.$$

Trivialerweise ist

$$\operatorname{Ker}(f^s) \supseteq \operatorname{span}\{e_{ij} \mid 1 \le j \le s,\ 1 \le i \le \lambda'_j\},$$

also ist $\dim\operatorname{Ker}(f^s) \ge \sum_{j=1}^{s}\lambda'_j$.

Ebenfalls sieht man unmittelbar, daß für $1 \le j \le \lambda_i - s$ jedes Element $e_{ij} = f^s(e_{i,j-s})$ zu $\operatorname{Im}(f^s)$ gehört, also gilt

$$\operatorname{Im}(f^s) \supseteq \operatorname{span}\{e_{ij} \mid 1 \le i \le t,\ 1 \le j \le \lambda_i - s\},$$

und f^s liefert eine Bijektion zwischen den Elementen e_{ij} mit $j > s$ und den Elementen e_{ij} mit $j \le \lambda_i - s$. Die Anzahl der Elemente e_{ij} mit $j > s$ ist aber $\sum_{j>s}\lambda'_j$.

Insgesamt sehen wir:

$$\dim \mathrm{Ker}(f^s) \geq \sum_{j=1}^{s} \lambda'_j \quad \text{und} \quad \dim \mathrm{Im}(f^s) \geq \sum_{j>s} \lambda'_j.$$

Die Summe der beiden rechten Seiten ist n, weil λ' eine Partition von n ist. Die Summe der beiden linken Seiten ist ebenfalls n, nach dem Dimensionssatz für lineare Abbildungen. Also gilt jeweils das Gleichheitszeichen.

Folgerung 1.

$$\dim \mathrm{Ker}(f_\lambda^s) - \dim \mathrm{Ker}(f_\lambda^{s-1}) = \lambda'_s.$$

Folgerung 2. *Seien λ, μ Partitionen von n. Ist $\lambda \neq \mu$, so sind die Matrizen $J(\lambda)$ und $J(\mu)$ nicht ähnlich.*

Beweis: Angenommen, die Matrizen $J(\lambda)$ und $J(\mu)$ sind ähnlich. Dann sind auch die Matrizen $J(\lambda)^s$ und $J(\mu)^s$ für jedes s ähnlich. Sind aber Matrizen A, B ähnlich, so haben die Kerne der linearen Abbildungen f_A und f_B die gleiche Dimension. Es ist $f_{J(\lambda)^s} = f_\lambda^s$ und $f_{J(\mu)^s} = f_\mu^s$. Wir verwenden nun Folgerung 1:

$$\lambda'_s = \dim \mathrm{Ker}(f_\lambda^s) - \dim \mathrm{Ker}(f_\lambda^{s-1}) = \dim \mathrm{Ker}(f_\mu^s) - \dim \mathrm{Ker}(f_\mu^{s-1}) = \mu'_s.$$

Da $\lambda'_s = \mu'_s$ für alle s gilt, ist $\lambda' = \mu'$, also $\lambda = \mu$.

Satz. *Zu jeder nilpotenten $(n \times n)$-Matrix A gibt es eine Partition λ von n, so daß A und $J(\lambda)$ ähnlich sind.* **Zusatz:** Dabei kann λ folgendermaßen bestimmt werden: Die zu λ duale Partition λ' ist durch

$$\lambda'_j = \dim \mathrm{Ker}(f^j) - \dim \mathrm{Ker}(f^{j-1})$$

für alle $j \geq 1$ gegeben (durch λ' ist $\lambda = \lambda''$ eindeutig bestimmt).

Für den Beweis verwenden wir folgendes Lemma:

Lemma. *Sei f nilpotenter Endomorphismus von V, sei $j \geq 2$. Sei $(v_1, \ldots, v_s)$ eine Folge von Elementen in $\mathrm{Ker}(f^j)$, die modulo $\mathrm{Ker}(f^{j-1})$ linear unabhängig ist. Dann ist $(f(v_1), \ldots, f(v_s))$ eine Folge von Elementen in $\mathrm{Ker}(f^{j-1})$, die modulo $\mathrm{Ker}(f^{j-2})$ linear unabhängig ist.*

Beweis: Natürlich gehört jedes Element $f(v_i)$ zu $\mathrm{Ker}(f^{j-1})$, denn $f^{j-1}f(v_i) = f^j(v_i) = 0$. Seien nun Elemente $\lambda_i \in K$ gegeben, so daß $\sum_i \lambda_i f(v_i)$ zu $\mathrm{Ker}(f^{j-2})$ gehört. Dann gehört $\sum_i \lambda_i v_i$ zu $\mathrm{Ker}(f^{j-1})$, denn $f^{j-1}\big(\sum_i \lambda_i v_i\big) = f^{j-2}f\big(\sum_i \lambda_i v_i\big) = f^{j-2}\big(\sum_i \lambda_i f(v_i)\big) = 0$. Da die Folge $(v_1, \ldots, v_s)$ modulo $\mathrm{Ker}\, f^{j-1}$ linear unabhängig ist, folgt $\lambda_i = 0$ für $1 \leq i \leq s$.

Beweis des Satzes und des Zusatzes: Setzen wir

$$\lambda'_j = \dim \mathrm{Ker}(f^j) - \dim \mathrm{Ker}(f^{j-1}),$$

so folgt aus dem Lemma sofort, daß λ_j eine Partition ist: denn ist $(v_1, \ldots, v_s)$ eine Komplementärbasis von $\mathrm{Ker}(f^{j-1})$ in $\mathrm{Ker}(f^j)$, so ist $s = \lambda'_j$. Das Lemma besagt,

daß $(f(v_1), \ldots, f(v_s))$ zu einer Komplementärbasis von $\mathrm{Ker}(f^{j-2})$ in $\mathrm{Ker}(f^{j-1})$ fortgesetzt weren kann, also ist $t \leq \lambda'_{j-1}$.

Sei $f^r = 0$. Betrachte die Kette von Unterräumen:

$$0 = \mathrm{Ker}(f^0) \subseteq \mathrm{Ker}(f) \subseteq \mathrm{Ker}(f^2) \subseteq \cdots \subseteq \mathrm{Ker}(f^{r-1}) \subseteq \mathrm{Ker}(f^r) = V.$$

Wir konstruieren induktiv Komplementärbasen $(v_{1,j}, \ldots, v_{\lambda'_j,j})$ von $\mathrm{Ker}(f^{j-1})$ in $\mathrm{Ker}(f^j)$, und zwar in absteigender Folge, wir beginnen also mit $j = r$, dann kommt $j = r-1$, und so weiter, bis schließlich $j = 1$. Induktionsanfang: Wähle eine beliebige Komplementärbasis $(v_{1,r}, \ldots, v_{\lambda'_r,r})$ von $\mathrm{Ker}(f^{r-1})$ in $\mathrm{Ker}(f^r) = V$.

Induktionsschritt: sei schon $(v_{1,j}, \ldots, v_{\lambda'_j,j})$ konstruiert, dies sei also eine Komplementärbasis von $\mathrm{Ker}(f^{j-1})$ in $\mathrm{Ker}(f^j)$, für ein $1 \leq j \leq r$. Ist $2 \leq j$, so wende f an, wir erhalten eine Folge $(f(v_{1,j}), \ldots, f(v_{\lambda'_j,j}))$, die nach dem Lemma in $\mathrm{Ker}(f^{j-1})$ liegt und modulo $\mathrm{Ker}(f^{j-2})$ linear unabhängig ist. Wir setzen

$$(*) \qquad v_{i,j-1} = f(v_{i,j}) \quad \text{für} \quad 1 \leq i \leq \lambda'_j.$$

Wir können diese Folge $(v_{1,j-1}, \ldots, v_{\lambda'_j,j-1})$ zu einer Komplementärbasis

$$(v_{1,j-1}, \ldots, v_{\lambda'_j,j-1}, v_{\lambda'_j+1,j-1}, \ldots, v_{\lambda'_{j-1},j-1})$$

von $\mathrm{Ker}(f^{j-2})$ in $\mathrm{Ker}(f^{j-1})$ fortsetzen.

Die Elemente $v_{i,j}$ mit $1 \leq i \leq \lambda'_j$ und $1 \leq j \leq r$ bilden eine Basis von V und $(*)$ zeigt, daß die Wirkung von f auf dieser Basis genau der Wirkung von $J(\lambda)$ auf den Basiselementen $e_{ij}^{(\lambda)}$ entspricht.

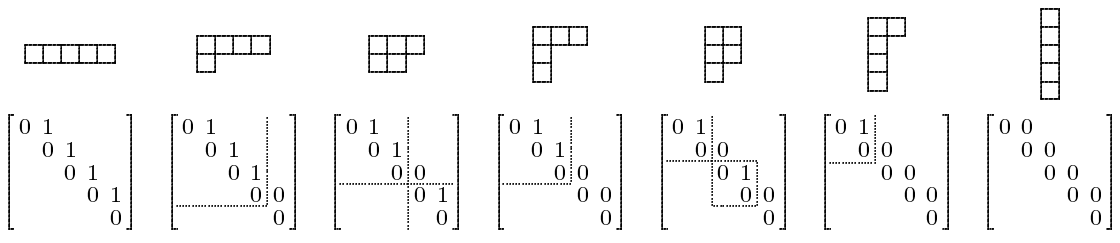
Folgerung. *Sei $A \in M(n \times n, K)$. Die folgenden Aussagen sind äquivalent:*

(i) *A ist nilpotent.*
(ii) *A ist ähnlich zu einer Matrix der Form $J(\lambda)$, mit λ Partition von n.*
(iii) $\chi_A = (-1)^n T^n$.
(iv) $A^n = 0$.

Beweis: (i) $\Longrightarrow$ (ii): Dies wurde gerade bewiesen. (ii) $\Longrightarrow$ (iii): Ist A ähnlich zu $J(\lambda)$, so ist $\chi_A = \chi_{J(\lambda)} = (-1)^n T^n$. (iii) $\Longrightarrow$ (iv): Dies folgt aus dem Satz von Cayley-Hamilton.

Hier alle Partitionen λ von n= 5 und die zugehörigen Young-Diagramme und die Jordanmatrizen $J(\lambda)$:

(5) $(4,1)$ $(3,2)$ $(3,1,1)$ $(2,2,1)$ $(2,1,1,1)$ $(1,1,1,1,1)$



Beispiel. Sei

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & -2 & 1 & -1 \\ 0 & -1 & 0 & 0 \\ -2 & 4 & -2 & 2 \end{bmatrix}.$$

Sei $f = f_A \colon V \to V$, mit $V = K^4$. Es ist

$$V_1 = \operatorname{Ker}(f) = \operatorname{span}\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}\}, \quad V_2 = \operatorname{Ker}(f^2) = \operatorname{span}\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}\},$$

und $\operatorname{Ker} f^3 = V$. Insbesondere ist A nilpotent. Wir sehen also:

$$\begin{aligned} \lambda'_1 &= \dim \operatorname{Ker}(f) = 2, \\ \lambda'_2 &= \dim \operatorname{Ker}(f^2) - \dim \operatorname{Ker}(f) = 3 - 2 = 1, \\ \lambda'_3 &= \dim \operatorname{Ker}(f^3) - \dim \operatorname{Ker}(f^2) = 4 - 3 = 1. \end{aligned}$$

Also

$$\lambda' = (2, 1, 1), \quad \text{daher} \quad \lambda = (3, 1).$$

Wähle $v_{13} \in V \setminus V_2$, zum Beispiel $v_{13} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$. Wir berechnen

$$v_{12} = Av_{13} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -2 \end{bmatrix} \quad \text{und} \quad v_{11} = Av_{12} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}.$$

Wir ergänzen v_{11} durch einen Vektor $v_{21} \in V_1$ zu einer Basis von V_1, zum Beispiel wählen wir $v_{21} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$. Dann haben wir also Vektoren v_{ij} konstruiert, die in das Young-Diagramm zur Partition $\lambda = (3, 1)$ passen:

v_{11}	v_{12}	v_{13}
v_{21}		

Die Matrix P habe als Spalten die Vektoren $v_{11}, v_{12}, v_{13}, v_{21}$. Dann gilt

$$P^{-1}AP = J((3, 1)).$$

Nach Konstruktion muß dies richtig sein (wenn wir uns nicht verrechnet haben). Man kann es natürlich auch nachträglich verifizieren; dafür muß man noch P^{-1} berechnen. Es ist

$$\begin{bmatrix} 0 & 2 & -1 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & -2 & 1 & -1 \\ 0 & 2 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & -2 & 1 & -1 \\ 0 & -1 & 0 & 0 \\ -2 & 4 & -2 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & -2 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$