Inhaltsverzeichnis

1 Analytische Geometrie und Vektoralgebra

1.1 Lineare Gleichungssysteme

Eine lineare Gleichung für eine Unbekannte ist von der Form ax = b mit gegebenen Zahlen a und b.

Beispiel: 3x = 6

genau eine Lösung $\underline{x = 2}$

0x = 6

keine Lösung

0x = 0

unendlich viele Lösungen

→ es gibt nur diese drei Fälle:

a ≠ 0 ⇒ genau eine Lösung, x = $\frac{b}{a}$
a = 0, b ≠ 0 ⇒ keine Lösung
a = 0, b = 0 ⇒ unendlich viele Lösungen

Betrachtet werden nun zwei lineare Gleichungen mit zwei Unbekannten:

$$ax + by = f$$

$$cx + dy = g$$

Beispiel 1 :

$$5x + 2y = 13 \ (1)$$

$$2x + 3y = 14 \ (2)$$

Sei x, y eine Lösung, dann gilt

$$3(1) : 15x + 6y = 39$$

$$2(2) : 4x + 6y = 28$$

⇒ (3(1) − 2(2)) $\qquad 11x = 11 \qquad \Rightarrow \underline{x = 1}$

(1) mit x = 1 ⇒ 5 + 2y = 13 ⇒ $\underline{y = 4}$

Probe zeigt, dass 1,4 wirklich die Lösung ist. Das Gleichungssystem hat also genau eine Lösung.

Beispiel 2 :

$$5x + 2y = 13$$

$$10x + 4y = 14$$

Dieses Gleichungssystem hat keine Lösung, denn
5x + 2y = 13 ⇒ 10x + 4y = 2(5x + 2y) = 2 ∗ 13 = 26 ≠ 14

Beispiel 3 :

$$5x + 2y = 13$$

$$10x + 4y = 26$$

Dieses Gleichungssystem hat unendlich viele Lösungen:

y = t (t ist beliebige Zahl)

$$\underline{x = \frac{13 - 2t}{5}}$$

Ein lineares Gleichungssystem der Form

$$ax + by = f \ (3)$$

$$cx + dy = g \ (4)$$

hat stets entweder keine, genau eine oder unendlich viele Lösungen.

Im folgenden führen wir 3 Beweise dafür. Jeder beruht auf einer anderen Idee und jede dieser Ideen werden wir später ausbauen.

1 Algebraische Herangehensweise

Es wird angenommen, dass x,y eine Lösung ist. Dann gilt:

$d * (3) - b * (4)$:

$$adx + bdy = df$$

$$bcx + bdy = bg$$

$$adx - bcx = df - bg$$

$$(ad - bc)x = df - bg \ (5)$$

$-c * (3) + a(4)$:

$$-acx - bcy = -cf$$

$$acx + ady = ag$$

$$(ad - bc)y = ag - cf \ (6)$$

Fallunterscheidung:

Fall 1: $ad - bc \neq 0 \Rightarrow$ genau eine Lösung:

$$x = \frac{df - bg}{ad - bc}, y = \frac{ag - cf}{ad - bc}$$

Die Probe zeigt, dass dies tatsächlich eine Lösung ist. Zum Beispiel ist

$$ax + by = a\frac{df - bg}{ad - bc} + b\frac{ag - cf}{ad - bc}$$

$$= \frac{adf - abg + abg - bcf}{ad - bc}$$

$$= \frac{f(ad - bc)}{ad - bc} = f$$

Haben also genau eine Lösung.

Fall 2: $ad - bc = 0$

(5) und (6) lauten dann:

$$0x = df - bg$$
$$0y = ag - cf$$

Fall 2.1.: $df - bg \neq 0$ oder $ag - cf \neq 0$

$\Rightarrow$ keine Lösung

Fall 2.2.: $df - bg = 0$ und $ag - cf = 0$

(5) und (6) lauten dann:

$$0x = 0$$
$$0y = 0$$

haben also nichts erreicht.

Fall 2.2.1.: $a^2 + b^2 + c^2 + d^2 > 0$ (wenigstens eine der Zahlen ist von 0 verschieden)

Sei $a \neq 0$, t wird beliebig gewählt und y = t gesetzt,

$$\Rightarrow x = \frac{f - bt}{a}$$

Dann ist

$$ax + by = a\frac{f - bt}{a} + bt = f - bt + bt = f$$
$$cx + dy = c\frac{f - bt}{a} + dt = \frac{cf - cbt + adt}{a}$$
$$= \frac{cf + t(ad - bc)}{a} = \frac{cf}{a} = \frac{ag}{a} = g$$

Weil t beliebig gewählt wird, erhalten wir unendlich viele Lösungen. Dies gilt analog für $b \neq 0$, $c \neq 0$, $d \neq 0$.

Fall 2.2.2.: $a = b = c = d = 0$

(3) und (4) lauten dann:

$$0x + 0y = f$$
$$0x + 0y = g$$

Für $f^2 + g^2 > 0$ hat dieses System keine Lösung, für $f = g = 0$ unendlich viele Lösungen.
#

Dieser Beweis liefert eine Methode, die Lösung auch wirklich zu bestimmen. Der Beweis motiviert folgendes:

Satz 1 *Für Zahlen α, β, γ, δ ist die Determinante* $\begin{vmatrix} \alpha & \beta \\ \gamma & \delta \end{vmatrix}$ *definiert als* $\begin{vmatrix} \alpha & \beta \\ \gamma & \delta \end{vmatrix} = \alpha\delta - \beta\gamma$.

Beweis 1 liefert das folgende präzisere Resultat:

Satz 2 (Cramorsche Regel) *(a) Ist* $\begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$, *so hat das System (3), (4) genau eine Lösung, nämlich*

$$x = \frac{\begin{vmatrix} f & b \\ g & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}, y = \frac{\begin{vmatrix} a & f \\ c & g \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}$$

(b) Ist $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = 0$, *so hat das System (3), (4) entweder keine oder unendlich viele Lösungen.*

2 Operatortheoretische Herangehensweise

Betrachten wir neben dem System

$$ax + by = f \ (3)$$

$$cx + dy = g \ (4)$$

noch das System

$$ax + by = 0 \ (7)$$

$$cx + dy = 0 \ (8)$$

homogenes Gleichungssystem

System (7), (8) hat immer die triviale Lösung x = 0, y = 0.

Fall 1: Das System (7), (8) hat nur die triviale Lösung.

Dann hat das System (3), (4) entweder keine oder genau eine Lösung. In der Tat, nehmen wir an (3), (4) hat zwei verschiedene Lösungen x_1,y_1 und x_2,y_2. Dann gilt:

$$ax_1 + by_1 = f \quad ax_2 + by_2 = f$$
$$cx_1 + dy_1 = g \quad cx_2 + dy_2 = g$$

$$\Rightarrow a(x_1 - x_2) + b(y_1 - y_2) = 0$$
$$c(x_1 - x_2) + d(y_1 - y_2) = 0$$

$\Rightarrow x_1 - x_2 = 0, y_1 = y_2 = 0$, da (7), (8) nur die triviale Lösung hat.

$\Rightarrow x_1 = x_2, y_1 = y_2$ ↙

Fall 2: System (7), (8) hat eine nichttriviale Lösung x_0, y_0. Behauptet wird, dass dann das System (3), (4) keine oder unendlich viele Lösungen hat.

Diese Behauptung ist äquivalent zu folgender Behauptung:

Wenn (3), (4) eine Lösung hat, so hat es unendlich viele Lösungen.

Sei also x_1, y_1 eine Lösung von (3), (4). Dann bilden $x = x_1 + tx_0$, $y = y_1 + ty_0$ für beliebiges t eine Lösung von (3), (4):

$$\begin{aligned} ax + by &= a(x_1 + tx_0) + b(y_1 + ty_0) \\ &= \underbrace{ax_1 + by_1}_{f} + t\underbrace{ax_0 + by_0}_{0} = f \end{aligned}$$

$$\begin{aligned} cx + dy &= c(x_1 + tx_0) + d(y_1 + ty_0) \\ &= \underbrace{cx_1 + dy_1}_{g} + t\underbrace{cx_0 + dy_0}_{0} = g \end{aligned}$$

Da $x_0 \neq 0$ oder $y_0 \neq 0$ ist, erhalten wir wirklich unendlich viele Lösungen für das System (3), (4). # 17.10.05

Beweis 2 lässt sich ohne Mühe auf Gleichungssysteme von m Gleichungen mit n Unbekannten übertragen, d.h. es gilt folgendes:

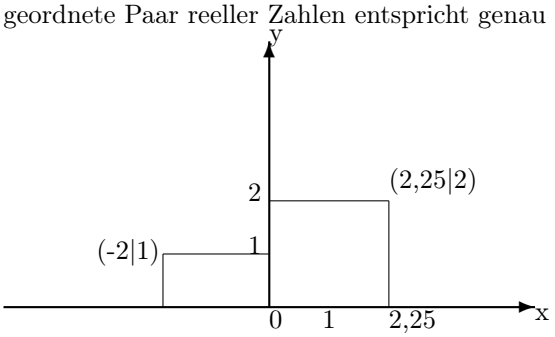
Satz 3 *Für beliebiges $m \geqq 1$, $n \geqq 1$ (m,n $\in$ N) besitzt das Gleichungssystem*

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1n}x_n &= f_1 \\ a_{21}x_1 + \ldots + a_{2n}x_n &= f_2 \\ &\vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n &= f_m \end{aligned}$$

entweder keine, oder genau eine oder unendlich viele Lösungen.

3 Geometrische Herangehensweise

Man wählt in der Ebene ein rechtwinkliges Koordinatensystem. Dann wird jeder Punkt der Ebene durch genau ein geordnetes Paar reeller Zahlen gegeben. Umgekehrt: Jedes geordnete Paar reeller Zahlen entspricht genau einem Punkt der Ebene.

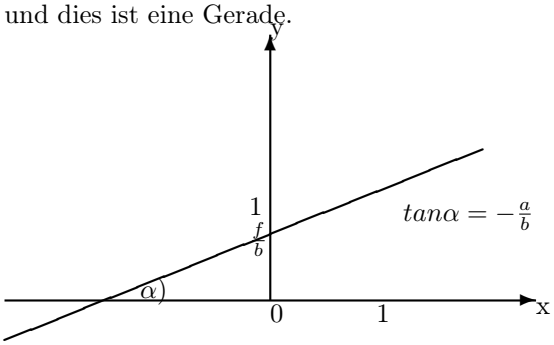


Die Menge aller geordneten Paare (x;y), die eine Gleichung der Form ax + by = f erfüllen, ist entweder die leere Menge ∅, eine Gerade oder die gesamte Ebene.

In der Tat, sie zunächst b $\neq 0$.

$$\Rightarrow y = -\frac{a}{b}x + \frac{f}{b}$$

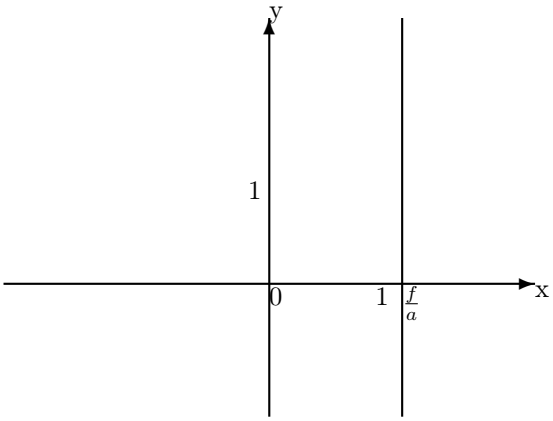
und dies ist eine Gerade.



Sei also b = 0, d.h. die Gleichung lautet

$$ax = f$$

$$a \neq 0 \qquad x = \frac{f}{a}$$

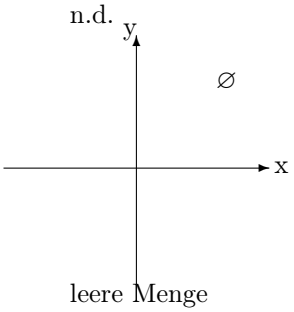


Gerade (keine Funktion)

$a = 0, f \neq 0$

$0 = f$

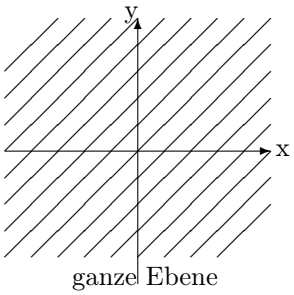
n.d.



leere Menge

$a = 0, f = 0$

$0 = 0$



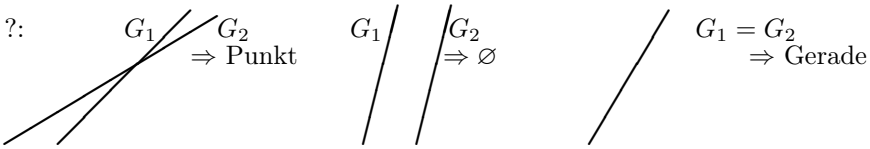
ganze Ebene

Bezeichnen mit G_1 bzw. G_2 die Menge aller Punkte der Ebene, deren Koordinaten, d.h. die Zahlen in der Darstellung als geordnetes Paar, die Gleichung

$$ax + by = f \quad bzw. \quad cx + dy = g$$

erfüllen. Die Menge der Lösungen des Gleichungssystems (3), (4) wird dann in der Ebene durch $G_1 \cap G_2$ dargestellt.

G_2 G_1	Gerade	$\varnothing$	Ebene
Gerade	?	$\varnothing$	Gerade
$\varnothing$	$\varnothing$	$\varnothing$	$\varnothing$
Ebene	Gerade	$\varnothing$	Ebene

?: 

In jedem Fall besteht $G_1 \cap G_2$ aus keinem, genau einem oder unendlich vielen Punkten. #

Der vorherige beweis hat zugleich folgendes ergeben:

Satz 4 *Die Gleichung*

$$ax + by + c = 0$$

beschreibt eine Gerade in der Ebene für $a^2 + b^2 > 0$, die leere Menge für $a = b = 0$ und $c \neq 0$ und die ganz Ebene, wenn $a = b = c = 0$ ist.

1.2 Analytische Geometrie in der Ebene

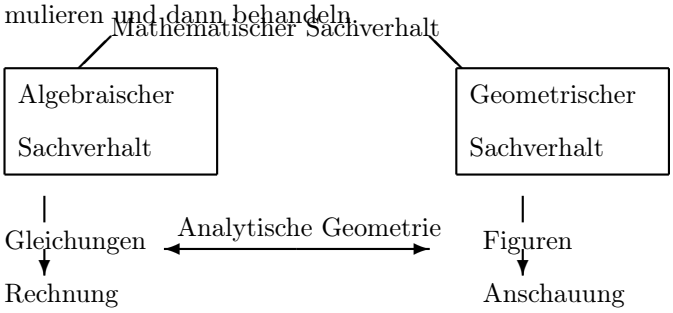
Die im Beweis 3 aus 1.1. benutzte Identifizierung von geordneten Paaren reeller Zahlen und Punkten der Ebene bildet die Grundlage für die Analytische Geometrie.

geordnetes Paar — Punkte

(Koordinaten) $\longrightarrow$

Rechtwinkliges Koordinatensystem

Man kann so mathematische Sachverhalte sowohl algebraisch als auch geometrisch formulieren und dann behandeln.

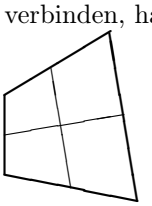


Analytische Geometrie ist die Kunst, die geeignete Darstellung (algebraisch oder geometrisch) eines mathematischen Sachverhalts zu finden und damit zu arbeiten.

Beweis 3 zeigt den Vorteil des Wechsels von algebraischer Sprache zu geometrischer Sprache. Das folgende Beispiel zeigt, dass auch der umgekehrte Wechsel von Vorteil sein kann:

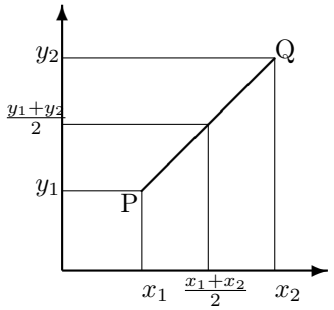
Beispiel 1:

Die Strecken, die die Mittelpunkte gegenüberliegender Seiten eines Vierecks miteinander verbinden, halbieren einander.

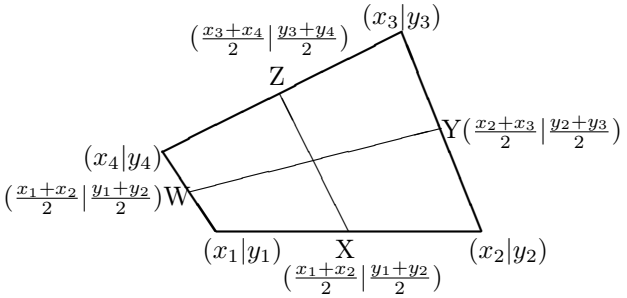


Man übersetzt das Problem in algebraische Sprache:

Vorüberlegung: Wenn 2 Punkte P = (x_1, y_1) und Q = (x_2, y_2) gegeben sind, so hat der Mittelpunkt der Strecke PQ die Koordinaten $(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2})$.



nach dem Strahlensatz



Um zu zeigen, dass sich XZ und YW halbieren, genügt es zu zeigen, dass die Mittelpunkte von XZ und YW übereinstimmen:

Mittelpunkt von XZ:

$$\left(\frac{\frac{x_1+x_2}{2}+\frac{x_3+x_4}{2}}{2}, \frac{\frac{y_1+y_2}{2}+\frac{y_3+y_4}{2}}{2}\right) = \left(\frac{x_1+x_2+x_3+x_4}{4}, \frac{y_1+y_2+y_3+y_4}{4}\right)$$

Mittlpunkt von XW:

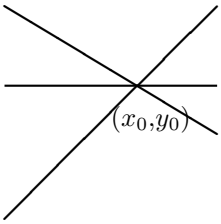
$$\left(\frac{\frac{x_2+x_3}{2}+\frac{x_1+x_4}{2}}{2}, \frac{\frac{y_2+y_3}{2}+\frac{y_1+y_4}{2}}{2}\right) = \left(\frac{x_1+x_2+x_3+x_4}{4}, \frac{y_1+y_2+y_3+y_4}{4}\right)$$

Dabei erhält man tatsächlich beide Male dasselbe.

Zur Lösung von Problemen mittels Analytischer Geometrie ist die Bewältigung einiger Grundaufgaben erforderlich. Betrachten wir einige dieser Grundaufgaben in den folgenden Beispielen.

Beispiel 2

Man bestimme die Gleichungen aller Geraden, die durch einen Punkt (x_0, y_0) gehen.



Geradengleichung ist von der Form:

$$ax + by + c = 0$$

Punkt (x_0,y_0) liegt auf einer solchen Geraden genau dann, wenn gilt:

$$ax_0 + by_0 + c = 0$$

Man erhält somit

$$c = -ax_0 - by_0,$$

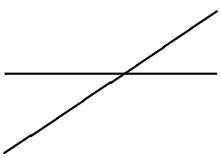
d.h. die gesuchten Gleichungen sind

$$a(x - x_0) + b(y - y_0) = 0$$

für beliebige a und b, $a^2 + b^2 > 0$.

Beispiel 3

Man bestimme die Koordinaten des Schnittpunktes zweier Geraden.



$ax + by + f = 0$

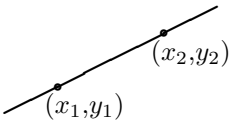
$cx + dy + g = 0$

Dieses Problem wurde im Abschnitt 1.1. behandelt. (Satz 2 → Cramorsche Regel.)

Gleichungssystem mit 2 Gleichungen und 2 Unbekannten

Beispiel 4

Man bestimme die Gleichung der Geraden durch zwei verschiedene Punkte (x_1,y_1) und (x_2,y_2).



Nach Beispiel 2 ist die Gleichung von der Form

$$a(x - x_1) + b(y - y_1) = 0 \ [\ mit\ (x_1, y_1)]$$

Punkt (x_2,y_2) liegt genau dann auf der Geraden, wenn gilt:

$$a(x_2 - x_1) + b(y_2 - y_1) = 0.$$

Sei z.B. $x_1 \neq x_2$ erhält man

$$a = -\frac{b(y_2 - y_1)}{x_2 - x_1}$$

und somit ist die gesuchte Gleichung

$$\frac{-b(y_2 - y_1)}{x_2 - x_1}(x - x_1) + b(y - y_1) = 0$$

$$\boxed{b[(y_2 - y_1)(x - x_1) - (x_2 - x_1)(y - y_1)] = 0, b \neq 0}$$

Man setzt häufig b = 1.

Beispiel: Gerade durch (0,1) und (3,5):

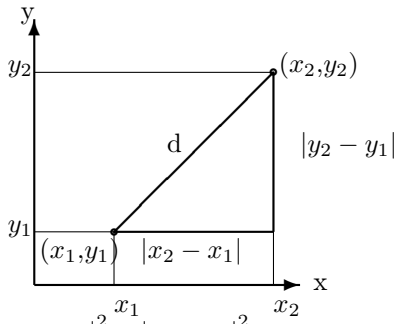
$$b * [(5 - 1)(x - 0) - (3 - 0)(y - 1)] = 0$$

$$b * [4x - 3y + 3] = 0$$

$$\underline{4x - 3y + 3 = 0}$$

Beispiel 5

Man bestimme den Abstand zweier Punkt (x_1,y_1) und (x_2,y_2).



Pythagoras: $d^2 = |x_2 - x_1|^2 + |y_2 - y_1|^2$

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

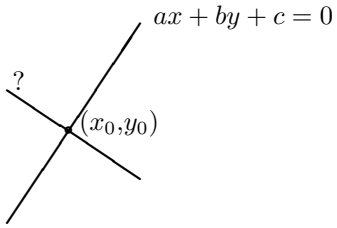
Abstand von (0,1) zu (3,5):

$$d = \sqrt{(3 - 0)^2 + (5 - 1)^2} = \sqrt{3^2 + 4^2} = \sqrt{25} = \underline{5}$$

21.10.05

Beispiel 6

Man bestimme die Gleichung der Geraden, die auf einer gegebenen Geraden zu einem gegebenen Punkte senkrecht steht.



Die gegebene Gerade hat die Gleichung

$$a(x - x_0) + b(y - y_0) = 0.$$

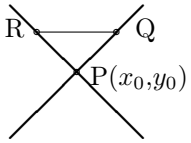
[Beispiel 2]

Wir behaupten, dass die gesuchte Gerade die Gleichung

$$b(x - x_0) - a(y - y_0) = 0$$

hat.

Beweis:



Zum Beweis wählen wir Q = $(x_0 + b|y_0 - a)$ auf der gegebenen Geraden und R = $(x_0 + a|y_0 + b)$ auf der gesuchten Geraden.

Wir zeigen, dass $|RQ|^2 = |PR|^2 + |PQ|^2$ gilt.

Daraus folgt, dass das Dreieck PQR rechtwinklig ist.

[Umkehrung des Satz des Pythagoras (nach dem Kosinussatz)]

Nach Beispiel 5 ist:

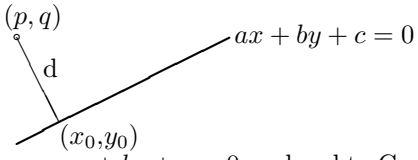
$$|RQ|^2 = (b-a)^2 + (-a-b)^2 = b^2 - 2ab + a^2 + a^2 + 2ab + b^2 = \underline{2b^2 + 2a^2}$$

$$|PR|^2 + |PQ|^2 = (-a)^2 + (-b)^2 + b^2 + (-a)^2 = \underline{2a^2 + 2b^2}$$

#

Beispiel 7

Man bestimme den Abstand eines Punktes (p, q) von einer Geraden $ax + by + c = 0$.



Eine zu $ax + by + c = 0$ senkrechte Gerade hat die Gleichung

$$bx - ay + d = 0$$

[Beispiel 6]

$b(x - x_0) - a(y - y_0) = 0$
$bx - ay - \underbrace{bx_0 + ay_0}_{irgendeine Konstante\ d}$

Der Punkt (p,q) liegt auf dieser senkrechten Geraden genau dann, wenn gilt:

$$bp - aq + d = 0.$$

Die senkrechte Gerade durch (p,q) hat also die Gleichung

$$bx - ay + aq - bp = 0.$$

Die Koordinaten (x_0,y_0) des Fußpunktes des Lotes von (p,q) auf die gegebene Gerade sind somit die Lösung des Gleichungssystems

$$ax_0 + by_0 = -c$$
$$bx_0 + ay_0 = bp - aq$$

Cramorsche Regel:

$$\begin{vmatrix} a & b \\ b & -a \end{vmatrix} = -a^2 - b^2,$$

$-a^2 - b^2 \neq 0$, sonst wäre es keine Gerade

$$x_0 = \frac{\begin{vmatrix} -c & b \\ bp - aq & -a \end{vmatrix}}{-a^2 - b^2} = \frac{ac - b^2p + abq}{-(a^2 + b^2)}$$

$$= \frac{b^2p - ac - abq}{a^2 + b^2} = \frac{a^2p + b^2q - a^2p - ac - abq}{a^2 + b^2}$$

$$= \underline{p - \frac{a(ap + bq + c)}{a^2 + b^2}}$$

$$y_0 = \frac{\begin{vmatrix} a & -c \\ b & bp - aq \end{vmatrix}}{-a^2 - b^2} = \frac{abp - a^2q + bc}{-a^2 - b^2} = \frac{a^2q - abp - bc}{a^2 + b^2}$$
$$= \frac{a^2q + b^2q - b^2q - abp - bc}{a^2 + b^2} = \underline{q - \frac{b(ap + bq + c)}{a^2 + b^2}}$$

Nach Beispiel 5 ist also

$$d^2 = (x_0 - p)^2 + (y_0 - q)^2$$
$$d^2 = \left(\frac{a(ap + bq + c)}{a^2 + b^2}\right)^2 + \left(\frac{b(ap + bq + c)}{a^2 + b^2}\right)^2$$
$$d^2 = (ap + bq + c)^2 \left(\frac{a^2}{(a^2 + b^2)^2} + \frac{b^2}{(a^2 + b^2)^2}\right)$$
$$d^2 = \frac{(ap + bq + c)^2}{a^2 + b^2}$$

$$\boxed{d = \frac{|ap+bq+c|}{\sqrt{a^2+b^2}}}$$

Hessesche Normalform

Zum Beispiel: P = (1,2) 3x + 4y - 1 = 0

„Hessesche Normalform“ der Geraden:

$$\frac{3x+4y-1}{\sqrt{3^2+4^2}}=0$$

$$\frac{3}{5}x+\frac{4}{5}y-\frac{1}{5}=0$$

$$d=\left|\frac{3}{5}*1+\frac{4}{5}*2-\frac{1}{5}\right|=\left|\frac{3}{5}+\frac{8}{5}-\frac{1}{5}\right|=\underline{2}$$

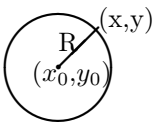
Geometrische Örter

Analytische Geometrie erlaubt es, sogenannte „geometrische Örter“, d.h. Mengen von Punkten mit gegebenen Eigenschaftn zu behandeln. Hier sind grundlegende Beispiele.

Kreis:

= Menge aller Punkte, die von einem gewissen Mittelpunkt den gegebenen Abstand haben

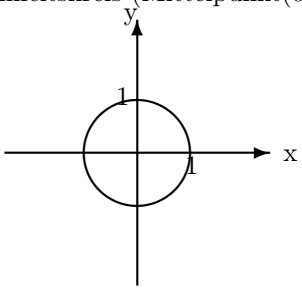
Wir wählen einen Punkt (x_0, y_0) in der Ebene und eine Zahl R > 0. Die Mengen aller Punkte der Ebene, die von (x_0, y_0) den Abstand R haben, ist ein Kreis.



Kreisgleichung: nach dem Abstand zweier Punkte

$$(x-x_0)^2+(y-y_0)^2=R^2$$

Für den Einheitskreis (Mittelpunkt(0,0) und Radius 1) erhalten wir

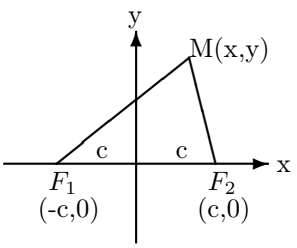


Ellipse:

Wir wählen in der Ebene zwei Punkte F_1 und F_2. Wir wählen ferner eine Zahl a > 0. Die durch diese Vorgaben definierte Ellipse ist die Menge aller Punkte M in der Ebene, für die gilt:

$$|MF_1|+|MF_2|=2a$$

Wir wählen ein Koordinatensystem wie in der Abbildung:



Die Ellipsengleichung lautet dann

$$\overbrace{\sqrt{(x+c)^2+y^2}}^{|MF_1|}+\overbrace{\sqrt{(x-c)^2+y^2}}^{|MF_2|}=2a$$

Wir nehmen an, dass a $>$ c ist. (sonst entsteht keine Ellipse)
Wir formen die obige Gleichung um:

$$\sqrt{(x+c)^2+y^2}=2a-\sqrt{(x-c)^2+y^2}$$
$$(x+c)^2+y^2=(2a-\sqrt{(x-c)^2+y^2})^2 \qquad ist\ äquivalent\ für\ 2a \geqq \sqrt{(x-c)^2+y^2}$$
$$x^2+2cx+c^2+y^2=4a^2-4a\sqrt{(x-c)^2+y^2}+(x-c)^2+y^2$$
$$x^2+2cx+c^2+y^2=4a^2-4a\sqrt{(x-c)^2+y^2}+x^2-2cx+c^2+y^2$$
$$4a\sqrt{(x-c)^2+y^2}=4a^2-4cx$$
$$a\sqrt{(x-c)^2+y^2}=a^2-cx \qquad ist\ äquivalent\ für\ a^2-cx \geqq 0$$
$$a^2((x-c)^2+y^2)=(a^2-cx)^2$$
$$a^2(x^2-2cx+c^2+y^2)=a^4-2a^2cx+c^2x^2$$
$$a^2x^2-2a^2cx+a^2c^2+a^2y^2=a^4-2a^2cx+c^2x^2$$
$$(a^2-c^2)x^2+a^2y^2=a^2(a^2-c^2)$$

Wir setzen $b=\sqrt{a^2-c^2}$ [da a $>$ c wird a^2-c^2 immer positiv] :

$$b^2x^2+a^2y^2=a^2b^2$$

$\boxed{\frac{x^2}{a^2}+\frac{y^2}{b^2}=1}$ *Ellipsengleichung*

Sei nun die Ellipsengleichung gültig, dann ist:

$$x<0 \Rightarrow a^2-cx \geqq 0$$
$$x>0 \Rightarrow x \leqq a$$
$$\Rightarrow cx \leqq ca < a, a < a^2$$
$$\Rightarrow a^2-cx>0$$

Damit ist gezeigt, dass sich das 2. Quadrieren umkehren lässt.
Wir haben desweiteren:

$$-a \leqq x \leqq a [weil\ \frac{x^2}{a^2} \leqq 1]$$
$$\Rightarrow -a-c \leqq x-c \leqq a-c \leqq a+c$$
$$\Rightarrow (x-c)^2 \leqq (a+c)^2$$
$$nach\ Ellipsengleichung y^2 \leqq b^2=a^2-c^2,$$
$$\Rightarrow (x-c)^2+y^2 \leqq (a+c)^2+a^2-c^2$$
$$(x-c)^2+y^2 \leqq a^2+2ac+c^2+a^2-c^2$$
$$(x-c)^2+y^2 \leqq 2y^2+2ac < 2a^+2a*a=4a^2$$

$$\Rightarrow \sqrt{(x-c)^2+y^2} < 2a.$$

Damit ist auch das 1. Quadrieren umkehrbar.

Die eingerahmte Gleichung ist also die Ellipsengleichung.

F_1,F_2... Brennpunkte

a..... große Halbachse

b..... kleine Halbachse

Für c = 0 erhält man einen Kreis. Für c = a erhält man die Verbindung zwischen F_1 und F_2.

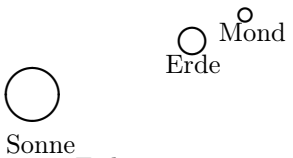
Exzentrizität der Ellipse

$$\epsilon = \frac{c}{a} = \frac{\sqrt{a^2-b^2}}{a} = \sqrt{\frac{a^2-b^2}{a^2}} = \sqrt{1-\frac{b^2}{a^2}}$$

$$\boxed{\epsilon = \sqrt{1-\tfrac{b^2}{a^2}}}$$

→ gibt Abweichung von der Kreisgestalt an.

Beispiel:



Planetenbahnen sind Ellipsen

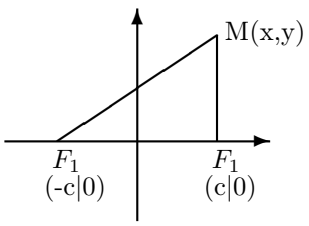
Sonne - Erde : ϵ = 0,017

Erde - Mond : ϵ = 0,056

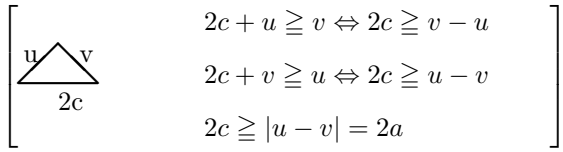
Sonne - Pluto: ϵ = 0.25

Hyperbel: 24.10.05

Gegeben seien in der Ebene zwei Punkte F_1 und F_2 mit dem Abstand 2c. Die Menge aller Punkte M der Ebene mit $||MF_1| - |MF_2|| = 2a$ heißt Hyperbel. Wegen Dreiecksungleichung können wir $0 \leqq a \leqq c$ voraussetzen (ansonsten entsteht die leere Menge).

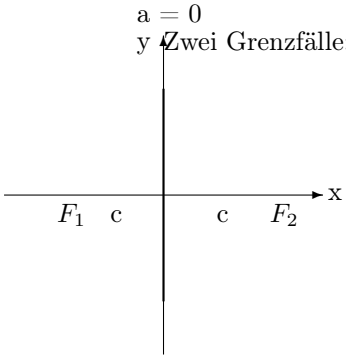


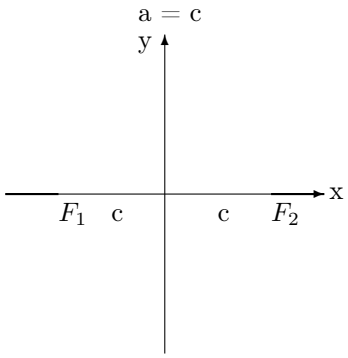
Dreiecksungleichung



$$2c + u \geqq v \Leftrightarrow 2c \geqq v - u$$
$$2c + v \geqq u \Leftrightarrow 2c \geqq u - v$$
$$2c \geqq |u - v| = 2a$$

Zwei Grenzfälle:





Sei also $0 < a < c$, dann setzen wir

$$b = \sqrt{c^2 - a^2}$$

Wie im Falle der Ellipse kann man zeigen, dass bei obiger Wahl des Koordinatensystems die Gleichung der Hyperbel

$\boxed{\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1}$ *Hyperbelgleichung*

lautet.

Die Parameter a, b, c sind wie in der Abbildung. Die Geraden $y = -\frac{b}{a}x$ sind sogenannte Asymptoten der Hyperbel:

Gleichung der Hyperbel ergibt für $x > 0$, $y > 0$:

$$y = b\sqrt{\frac{x^2}{a^2} - 1} = b\sqrt{\frac{x^2}{a^2}\left(1 - \frac{a^2}{x^2}\right)}$$
$$= \frac{b}{a}x\sqrt{1 - \frac{a^2}{x^2}}$$

und in der Analysis wird gezeigt, dass

$$\delta(x) = \frac{b}{a}x - \frac{b}{a}x\sqrt{1 - \frac{a^2}{x^2}} \to 0 \; für \; x \to \infty$$

Unterschiedliche Koordinatensysteme ergeben natürlich unterschiedliche Gleichungen. Sei z.B. $c = 2$, $a = \sqrt{2}$. Wir haben dann

$$b = \sqrt{c^2 - a^2} = \sqrt{4 - 2} = \sqrt{2}.$$

Parabel:

Gegeben sie in der Ebene eine Gerade D und ein Punkt F außerhalb von D. Die Menge aller Punkte M der Ebene, für die gilt

$$|MF| = dist(M, D)$$

heißt Parabel.

Gleichung:

$$\overbrace{\sqrt{x^2 + (y-c)^2}}^{|MF|} = \overbrace{y + c}^{dist(M,D)}$$
$$x^2 + (y-c)^2 = (y+c)^2$$
$$x^2 + y^2 - 2yc + c^2 = y^2 + 2yc + c^2$$

$\boxed{y = \frac{1}{4c}x^2}$

Parabolspiegel: (entsteht bei Rotation einer Parabel um die y-Achse)

Diese haben die Eigenschaft, dass parallel einfallende Strahlung im Brennpunkt F gesammelt wird.

Beweis:

Parabel sei $y = \frac{1}{4c}x^2$. Wir wählen einen Punkt $P(p, \frac{1}{4c}p^2)$auf der Parabel.

Gleichung der Tangente an die Parabel im Punkt P:

$$y = \frac{p}{2c}x + b$$

[Analysis: Anstieg der Tangente = Wert der Ableitung im gegebenen Punkt]

$$y = \frac{1}{4c}x^2 \quad , \quad y' = \frac{1}{4c}2x$$
$$y'(p) = \frac{p}{2c}$$

$$\tan\beta = \frac{p}{2c}$$
$$\tan\beta = \tan(90^\circ - \gamma) = \cot\gamma = \frac{1}{\tan\gamma}$$

Also: $\boxed{\tan\gamma = \frac{2c}{p}}$

Dann ist: $\alpha = \alpha_1 - \alpha_2$
Wir wissen, dass $\tan\alpha_1$ der Anstieg der Tangente ist:

$$\tan\alpha_1 = \tan\beta = \frac{p}{2c}$$

Gleichung der Geraden durch F und P:

$$y = kx + c$$
$$\frac{1}{4c}p^2 = kp + c$$
$$k = \frac{1}{4c}p - \frac{c}{p} = \frac{1}{p}\left(\frac{p^2}{4c} - c\right)$$

Also: $\tan\alpha_2$ = Anstieg der Geraden durch F und P

$$= k = \frac{1}{p}\left(\frac{p^2}{4c} - c\right)$$

Somit ist:

$$\tan\alpha = \tan(\alpha_1 - \alpha_2) = \frac{\tan\alpha_1 - \tan\alpha_2}{1 + \tan\alpha_1 \tan\alpha_2}$$

$$= \frac{\frac{p}{2c} - \frac{1}{p}\left(\frac{p^2}{4c} - c\right)}{1 + \frac{p}{2c} * \frac{1}{p}\left(\frac{p^2}{4c} - c\right)} = \frac{\frac{p}{2c} - \frac{p}{4c} + \frac{c}{p}}{1 + \frac{p^2}{8c^2} - \frac{1}{2}}$$

$$= \frac{\frac{p}{4c} + \frac{c}{p}}{\frac{1}{2} + \frac{p^2}{8c^2}} = \frac{\frac{p^2+4c^2}{4cp}\frac{8c^2+2p^2}{16c^2}}{=} \frac{(p^2+4c^2)16c^2}{4cp(8c^2+2p^2)}$$

$$= \frac{(p^2+4c^2)16c^2}{8cp(p^2+4c^2)} = \underline{\frac{2c}{p}}$$

$$\Rightarrow \tan\gamma = \tan\alpha \Rightarrow \alpha = \gamma\#$$

Gemeinsame Eigenschaften von Ellipse, Hyberbel und Parabel

In den obigen Definitionen spielte die Definition der Parabel eine gewisse Sonderrolle. Es gibt mehrere Definitionen von Ellipse, Hyperbel und Parabel, die alle mehr oder weniger auf dem gleichen Prinzip beruhen.

Hier sind 2 Beispiele:

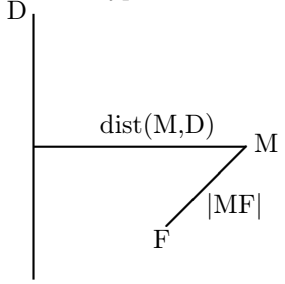
Als erstes sei in der Ebene eine Gerade D und ein Punkt F außerhalb von D gegeben. Sei desweitere $\epsilon > 0$ eine gegebene Zahl. Die Menge der Punkte M der Ebene mit

$$\frac{|MF|}{dist(M,D)} = \epsilon$$

ist eine Ellipse für $0 < \epsilon < 1$,

eine Parabel für $\epsilon = 1$,

eine Hyperbel für $1 < \epsilon < \infty$.



Das können wir beweisen.

Als zweites kann man Kegelschnitte betrachten.

Das können wir noch nicht beweisen.

Algebraische Kurven

Eine algebraische Kurve der Ordnung n in der Ebene ist gegeben als die Menge aller Punkte (x,y), die eine Gleichung der Form

$$\sum_{k+l\leqq n,\ k,l\geqq 0} a_{kl}x^k y^l = 0$$

erfüllen.

Kurven der Ordnung 1

$$\sum_{k+l\leqq 1,\ k,l\geqq 0} a_{kl}x^k y^l = a_{00}x^0y^0 + a_{01}x^0y^1 + a_{10}x^1y^0$$

Gleichung einer Kurve 1. Ordnung ist also: $\boxed{a_{10}x + a_{01}y + a_{00} = 0}$

→ drei Koeffizienten werden beliebig gewählt

Wir wissen, dass dies für $a_{10}^2 + a_{01}^2 > 0$ die Gleichung einer Geraden ist. Für $a_{10} = a_{01} = 0$ entsteht die leere Menge ($a_{00} \neq 0$) oder die ganze Ebene ($a_{00} = 0$). 28.10.05

Kurven der Ordnung 2

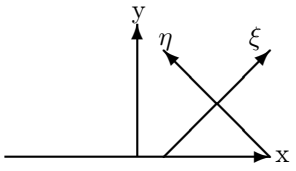
Diese werden gegeben durch eine Gleichung der Form

$$\sum_{k+l\leqq 2,\ k,l\geqq 0} = a_{kl}x^k y^l$$

$= \boxed{a_{20}x^2 + a_{11}xy + a_{02}y^2 + a_{10}x + a_{01}y + a_{00}}$

Für $a_{20} = a_{11} = a_{02} = 0$ wird dies eine Kurve 1. Ordnung. Sei also $a_{20}^2 + a_{11}^2 + a_{02}^2 > 0$.

Man kann zeigen, dass man ein neues Koordinatensystem so wählen kann, dass die durch obige Gleichung beschriebene Menge im neuen System genau eine der folgenden Gleichungen hat:



1. $\boxed{\frac{\xi^2}{a^2} + \frac{\eta^2}{b^2} = 1 \qquad a, b \neq 0}$

Ellipse

2. $\boxed{\frac{\xi^2}{a^2} + \frac{\eta^2}{b^2} = -1 \qquad a, b \neq 0}$

leere Menge $\varnothing$

3. $\boxed{\frac{\xi^2}{a^2} + \frac{\eta^2}{b^2} = 0 \qquad a, b \neq 0}$

Punkt (0,0)

4. $\boxed{\frac{\xi^2}{a^2} - \frac{\eta^2}{b^2} = 1 \qquad a, b \neq 0}$

Hyperbel

5. $\boxed{\frac{\xi^2}{a^2} - \frac{\eta^2}{b^2} = 0 \qquad a, b \neq 0}$

$$\Leftrightarrow \left(\frac{\xi}{a} + \frac{\eta}{b}\right)\left(\frac{\xi}{a} - \frac{\eta}{b}\right) = 0$$

Paar sich schneidender Geraden

6. $\boxed{\xi = c\eta^2 \qquad c \neq 0}$

Parabel

7. $\boxed{\xi^2 + b^2 = 0 \qquad b \neq 0}$

leere Menge

8. $\boxed{\xi^2 - b^2 = 0 \qquad b \neq 0}$

$$\Leftrightarrow \xi = \pm b$$

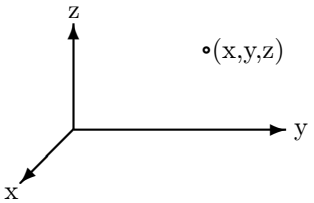
Paar paralleler Geraden

9. $\boxed{\xi^2 = 0}$

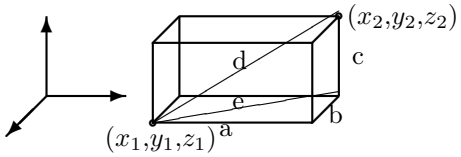
eine Gerade (zwei zusammenfallende Geraden)

1.3 Ebenen und Geraden im Raum

Wir wählen im Raum ein rechtwinkliges Koordinatensystem und können dann den Raum mit der Menge aller geordneten Tripel (x,y,z) von reellen Zahlen identifizieren.



Beispiel: Wie groß ist der Abstand zweier gegebener Punkte (x_1,y_1,z_1) und (x_2,y_2,z_2)?



$$e^2 = a^2 + b^2$$
$$d^2 = e^2 + c^2$$
$$\Rightarrow d^2 = a^2 + b^2 + c^2$$

und $a = |y_2 - y_1|$, $b = |x_2 - x_1|$, $c = |z_2 - z_1|$ ergibt

$$\boxed{d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}}$$

→ Hier ist das Analogon zu Satz 4 in 1.1.:

Satz 1 *Die Menge aller Punkte des Raumes, deren Koordinaten die Gleichung der Form*

$$ax + by + cz + d = 0$$

erfüllen, ist
eine Ebene für $a^2 + b^2 + c^2 > 0$,
die leere Menge für $a = b = c = 0, d \neq 0$,
der ganze Raum für $a = b = c = d = 0$.

Jede Ebene im Raum wird durch eine Gleichung der obigen Form mit $a^2 + b^2 + c^2 > 0$ gegeben.

Beweis: Wir wählen einen Punkt (x_0,y_0,z_0) mit

$$ax_0 + by_0 + cz_0 + d = 0$$

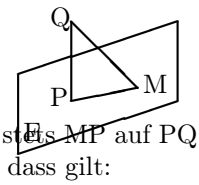
(setzen $a^2 + b^2 + c^2 > 0$ voraus)

Wir bezeichnen die durch die Gleichung

$$ax + by + cz + d = 0$$

beschriebene Menge E. Dann gehört also P = (x_0,y_0,z_0) zu E.

Wir betrachten den Punkt Q = ($x_0 + a$,$y_0 + b$,$z_0 + c$) und einen beliebigen Punkt M = (x,y,z) ∈ E.



Wir zeigen, dass stets MP auf PQ senkrecht steht. Daraus folgt, dass E eine Ebene ist. Dazu zeigen wir, dass gilt:

$$|MQ|^2 = |PM|^2 + |PQ|^2$$

(Umkehrung des Pythagoras liefert dann MP ⊥ QP).

Aus obigem Beispiel folgt

$$|MQ|^2 = (x - x_0 - a)^2 + (y - y_0 - b)^2 + (z - z_0 - c)^2$$
$$|PM|^2 = (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2$$
$$|PQ|^2 = a^2 + b^2 + c^2$$

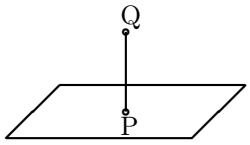
und damit

$$(x - x_0 - a)^2 + \ldots = (x - x_0)^2 + \ldots + a^2 + \ldots \quad (?)$$
$$-2a(x - x_0) - \ldots = 0$$

$$2a(x - x_0) + 2b(y - y_0) + 2c(z - z_0) = 0$$
$$ax + by + cz - \underbrace{(ax_0 + by_0 + cz_0)}_{d} = 0$$

Letztere Gleichung ist aber genau für (x,y,z) ∈ E erfüllt.

Sei nun E eine beliebige Ebene im Raum. Wir wählen einen Punkt P = (x_0, y_0, z_0) in der Ebene und bezeichnen mit Q einen von P verschiedenen Punkt auf der Senkrechten zur Ebene in P.



Die Koordinaten von Q seien ($x_0 + a$,$y_0 + b$,$z_o + c$).

Dann ist E die Menge aller Punkte M = (x,y,z) mit

$$|MQ|^2 = |PM|^2 + |PQ|^2$$

Gleiche Rechnung wie oben ergibt

$$ax + by + cz - (ax_0 + by_0 + cz_0) = 0.$$

Bezeichnet man $-(ax_0 + by_0 + cz_0)$ mit d, so folgt die Behauptung.

\# 04.11.05

Satz 2 *Die Menge aller Lösungen (x,y,z) von m linearen Gleichungen mit drei Unbekannten ist leer, eine Gerade, eine Ebene oder der ganze Raum.*

Beweis: Wir nehmen an, dass alle Gleichungen nichttrivial sind (d.h.$a_i^2 + b_i^2 + c_i^2 > 0$).

m = 1

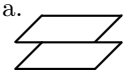
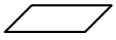
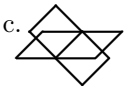
$ax + by + cz + d = 0$

Behauptung folgt aus Satz 1 (→ Ebene)

m = 2

Nach Satz 1 beschreibt jede Gleichung eine Ebene. Lösungsmenge beider Gleichungen ist der Durchschnitt dieser beiden Ebenen.

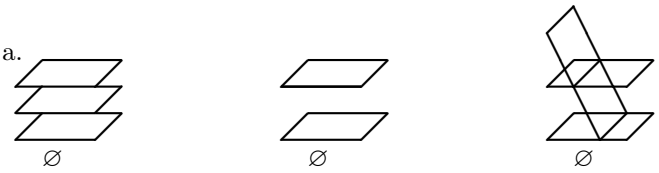
Wir haben folgende Fälle:

a.	b.	c.
		
parallele Ebenen	zusammenfallende Ebenen	zwei sich schneidende Ebenen
∅	Ebene	Gerade

m = 3

Wir nehmen eine dritte Ebene hinzu:

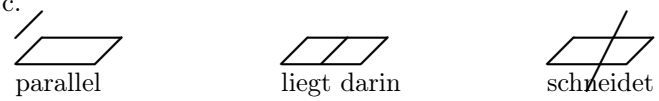
a.



∅ ∅ ∅

b. wird einer der 3 Fälle aus „m = 2“.

c.



parallel	liegt darin	schneidet
∅	Gerade	Punkt

$\boxed{m \geqq 4}$

analog #

Einige Grundaufgaben

Gegeben seien zwei verschiedene Punkte im Raum.Wir suchen eine Beschreibung der Geraden durch diese Punkte. Man könnte diese Gerade als Schnitt zweier Ebenen

$$a_1x + b_1y + c_1z + d_1 = 0$$
$$a_2x + b_2y + c_2z + d_2 = 0$$

darstellen.

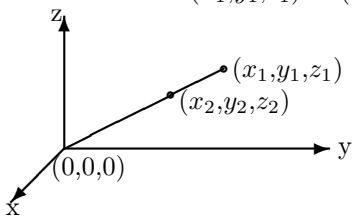
Günstiger ist aber die durch folgenden Satz gelieferte Beschreibung, die Parameterdarstellung genannt wird.

Satz 3 *Die Gerade durch zwei verschiedene Punkte* (x_1,y_1,z_1) *und* (x_2,y_2,z_2) *ist gleich der Menge aller Punkte (x,y,z), für die eine reelle Zahl t mit*

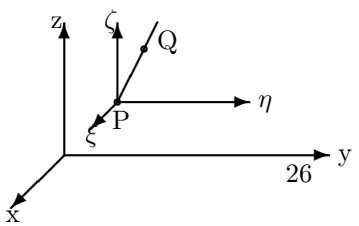
$$x = x_1 + t(x_2 - x_1) = (1-t)x_1 + tx_2$$
$$y = y_1 + t(y_2 - y_1) = (1-t)y_1 + ty_2$$
$$z = z_1 + t(z_2 - z_1) = (1-t)z_1 + tz_2$$

existiert.

Beweis: Dies ist offensichtlich für $(x_1,y_1,z_1) = (0,0,0)$:



Andernfalls wählen wir ein neues, achsenparalleles Koordinatensystem mit Ursprung in (x_1,y_1,z_1):



In den neuen Koordinaten ist die Gerade durch die Gleichung

$$\xi = t(x_2 - x_1)$$

$$\eta = t(y_2 - y_1)$$

$$\zeta = t(z_2 - z_1)$$

gegeben, da Q im neuen Koordinatensystem die Koordinaten ($x_2 - x_1$,$y_2 - y_1$,$z_2 - z_1$) hat.

Der Zusammenhang zwischen alten und neuen Koordinaten ist gegeben durch

$$x = x_1 + \xi$$

$$y = y_1 + \eta$$

$$z = z_1 + \zeta$$

Dies liefert

$$x - x_1 = t(x_2 - x_1) \rightarrow x_1 + t(x_2 - x_1)$$

$$y - y_1 = t(y_2 - y_1) \rightarrow y_1 + t(y_2 - y_1)$$

$$z - z_1 = t(z_2 - z_1) \rightarrow z_1 + t(z_2 - z_1)$$

#

Satz 4 *Die Ebene durch 3 Punkte P_1 = (x_1,y_1,z_1), P_2 = (x_2,y_2,z_2), P_3 = (x_3,y_3,z_3), die nicht auf einer Geraden liegen, ist gegeben durch die Gleichung*

$$(x - x_1)\begin{vmatrix} y_2 - y_1 & z_2 - z_1 \\ y_3 - y_1 & z_3 - z_1 \end{vmatrix} - (y - y_1)\begin{vmatrix} x_2 - x_1 & z_2 - z_1 \\ x_3 - x_1 & z_3 - z_1 \end{vmatrix}$$

$$+(z - z_1)\begin{vmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{vmatrix} = 0,$$

und außerdem ist diese Ebene gleich der Menge aller Punkte (x,y,z), für die reelle Zahlen λ und μ existieren, sodass gilt:

$$x = x_1 + \lambda(x_2 - x_1) + \mu(x_3 - x_2)$$

$$y = y_1 + \lambda(y_2 - y_1) + \mu(y_3 - y_2)$$

$$z = z_1 + \lambda(z_2 - z_1) + \mu(z_3 - z_2)$$

(Parameterdarstellung)

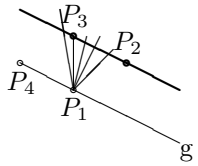
Beweis: Die Gleichung 1 ist von der Form

$$ax + by + cz + d = 0.$$

Die Gleichung ist erfüllt für (x,y,z)=(x_i,y_i,z_i) (i=1,2,3).

Die Gleichung beschreibt also entweder eine Ebene (auf der P_1,P_2,P_3 liegen) oder den ganzen Raum. Der ganze Raum liegt genau dann vor, wenn alle 3 Determinanten Null sind. Man kann zeigen (tun wir später), dass dies wiederum genau dann der Fall ist,

wenn P_1,P_2,P_3 auf einer Geraden liegen. Diesen Fall haben wir ausgeschlossen, also ist die Gleichung die Ebene durch P_1,P_2,P_3.



Gerade durch P_2 und P_3 ist gegeben durch

$$x = x_2 + t(x_3 - x_2)$$

$$y = y_2 + t(y_3 - y_2)$$

$$z = z_2 + t(z_3 - z_2)$$

(nach Satz 3)

Wir verbinden P_1 mit der Geraden durch P_2 und P_3 über einzelne Geraden. Die entstehende Menge ist eine Ebene minus g und ist gegeben durch die Parameterdarstellung (erneut Satz 3)

$$x = x_1 + \lambda(x_2 + t(x_3 - x_2) - x_1) = x_1 + \lambda(x_2 - x_1) + \lambda t(x_3 - x_2)$$

analog für y und z

Wir setzen $\mu = \lambda t$ und erhalten

$$x = x_1 + \lambda(x_1 - x_1) + \mu(x_3 - x_2) \quad \textit{analog für y und z}$$

Der Punkt P_4 hat die Koordinaten $(x_3 - x_2,y_3 - y_2,z_3 - z_2)$. Der Fall $(\lambda,\mu) = (0,\mu)$ liefert

$$x = x_1 + \lambda(x_3 - x_2) \quad \textit{analog für y und z}$$

und nach Satz 3 ist dies gerade die Gerade g.

Somit wird duch die Gleichung 2 die Ebene durch P_1,P_2,P_3 beschrieben. #

Wir haben

$$\begin{aligned} x &= x_1 + \lambda(x_2 - x_1) + \mu(x_3 - x_2) \\ &= x_1 + \lambda(x_2 - x_1) + \mu(x_3 - x_1 + x_1 - x_2) \\ &= x_1 + (\lambda - \mu)(x_2 - x_1) + \mu(x_3 - x_1) \\ &= x_1 + \lambda'(x_2 - x_1) + \mu'(x_3 - x_1) \end{aligned}$$

und (λ,μ) durchläuft die ganze Ebene genau dann, wenn $(\lambda',\mu') = (\lambda - \mu,\mu)$ die ganze Ebene durchläuft.

Statt 2 kann man also auch mit der Parameterdarstellung

$$x = x_1 + \lambda(x_2 - x_1) + \mu(x_3 - x_1)$$

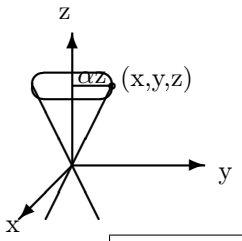
$$y = y_1 + \lambda(y_2 - y_1) + \mu(y_3 - y_1)$$

$$z = z_1 + \lambda(z_2 - z_1) + \mu(z_3 - z_1)$$

arbeiten.

Nochmals zu Kegelschnitten:

Wir betrachten einen Doppelkegel



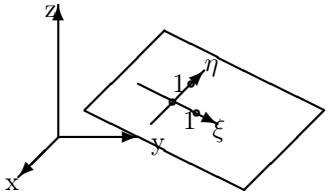
$\boxed{\alpha z = \sqrt{x^2 + y^2}}$ oberer Kegel

$\boxed{x^2 + y^2 = \alpha^2 z^2}$ Doppelkegel

= Gleichung eines Kegels im Raum

(durch das Quadrieren erhält man auch den unteren Kegel)

Ebene im Raum:



Wir wählen P_1,P_2,P_3 wie in der Abbildung. Dann hat P_i die Koordinaten (x_i,y_i,z_i). Ebene im Raum wird gegeben durch

$$x = x_1 + \lambda(x_2 - x_1) + \mu(x_3 - x_1)$$

$$y = y_1 + \lambda(y_2 - y_1) + \mu(y_3 - y_1)$$

$$z = z_1 + \lambda(z_2 - z_1) + \mu(z_3 - z_1)$$

Man kann sich unschwer überlegen, dass λ gerade die ξ-Koordinate und μ gerade die η-Koordinate ist. Wir ersetzen also λ und μ durch ξ und η und setzen alles in die Kegelgleichung ein:

$$(x_1 + \xi(x_2 - x_1) + \eta(x_3 - x_1))^2 + (y_1 + \xi(y_2 - y_1) + \eta(y_3 - y_1))^2$$

$$= \alpha^2(z_1 + \xi(z_2 - z_1) + \eta(z_3 - z_1))^2$$

Dies ergibt

$$\sum_{j+k \leqq 2,\ j,k \geqq 0} a_{jk}\xi^j\eta^k = 0 \quad (Kurve\ 2.\ Ordnung)$$

Wir erhalten also eine Gleichung der Ordnung 2 und damit Ellipse, Hyperbel, Parabel,...

#

1.4 Äquivalenzrelationen

07.11.05

Sei $X_1, \ldots, X_n$ eine endliche Anzahl von Teilmengen einer Menge X.

Vereinigung und Durchschnitt dieser Mengen sind definiert über

$$\bigcup_{\alpha=1}^{n} X_\alpha = \bigcup_{\alpha\in\{1,\ldots,n\}} \qquad X_\alpha = \{x \in X : \exists\alpha \in \{1,\ldots,n\}\ mit\ x \in X_\alpha\}$$

$$\bigcap_{\alpha=1}^{n} X_\alpha = \bigcap_{\alpha\in\{1,\ldots,n\}} \qquad X_\alpha = \{x \in X : x \in X_\alpha \forall\alpha \in \{1,\ldots,n\}\}$$

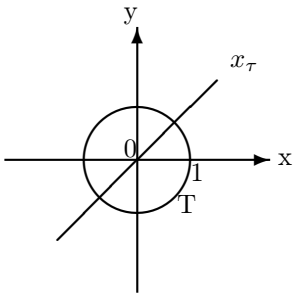
Man kann Vereinigung und Durchschnitt auch für beliebig viele (nicht notwendigerweise endilch viele) Mengen definieren. Sei I eine Menge (sogenannte Indexmenge) und für jedes $\alpha \in$ I sei eine Teilmenge X_α von X gegeben. Dann definiert man

$$\bigcup_{\alpha\in\ I} X_\alpha = \{x \in X : \exists\alpha \in I\ mit\ x \in X_\alpha\}$$

$$\bigcap_{\alpha\in\ I} X_\alpha = \{x \in X : x \in X_\alpha \forall\alpha \in\ I\}$$

1.Beispiel:

Sei T der Einheitskreis in der Ebene, und für jeden Punkt $\tau \in$ T sei x_τ die Gerade durch 0 und τ.



$$\bigcup_{\tau\in\ T} X_\tau \ = \ ganze\ Ebene$$

$$\bigcap_{\tau\in\ T} X_\tau \ = \ \{0\}.$$

Für eine Menge X ist das direkte Produkt X × X definiert als die Menge aller geordneten Paare (x,y) mit $x \in X,\ y \in X$.

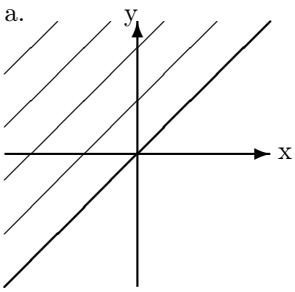
Definition 1 *Eine Relation R auf einer Menge X ist eine Teilmenge $R \subset X \times X$. Man sat, dass zwei Elemente $x \in X$, $y \in X$ in der Relation R zueinander stehen und schreibt dann x R y, wenn (x,y) $\in$ R gilt.*

Bemerkung: Bei uns ist $\subset = \subseteq$ (Teilmenge muss nicht echte Teilmenge sein.) Soll betont werden, dass A eine echte Teilmenge von X ist, schreiben wir A $\subsetneq$ X.

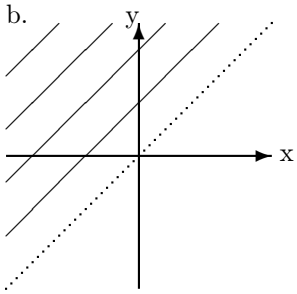
2.Beispiel:

Sei X = $\mathcal{R}$. 4 Teilmengen von X × X:

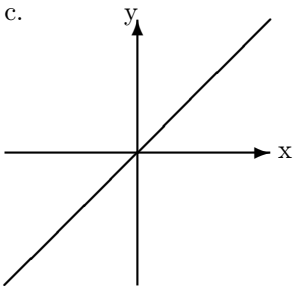
a.



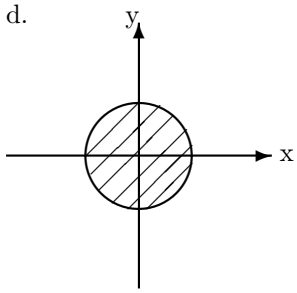
b.



c.



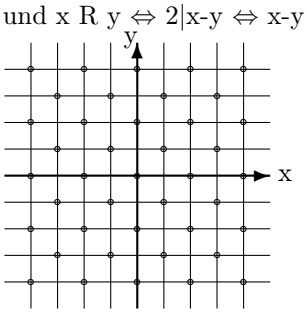
d.



a.	b.	c.	d.
x R y	x R y	x R y	x R y
$\Leftrightarrow x \leqq y$	$\Leftrightarrow x < y$	$\Leftrightarrow x = y$	$\Leftrightarrow x^2 + y^2 \leqq 1$

3.Beispiel:

Sei X = $\mathcal{Z}$ und x R y $\Leftrightarrow$ 2|x-y $\Leftrightarrow$ x-y ist gerade $\Leftrightarrow$ x $\equiv$ y (mod 2)



Definition 2 *Eine Relation R auf X heißt Äquivalenzrelation, wenn R folgende drei Eigenschaften hat:*
(r) (Reflexivität) x R x $\forall$ x $\in$ X; (s) (Symmetrie) x R y $\Rightarrow$ y R x; (t) (Transitivität) x R y, y R z $\Rightarrow$ x R z.

Bei Äquivalenzrelationen schreibt man üblicherweise x $\sim$ y statt x R y.

4.Beispiel:

(a) Beispiel 2a: r, nicht s, t

Beispiel 2b: nicht r, nicht s, t

Beispiel 2c: r, s, t $\rightarrow$ Äquivalenzrelation

Beispiel 2d: nicht r, s, nicht t

Beispiel 3: r, s, t $\rightarrow$ Äquivalenzrelation

(b) X sei Menge der einfarbigen Artikel in einem Kaufhaus, x R y $\Leftrightarrow$ x, y haben die gleiche Farbe.

r,s,t $\Rightarrow$ ist eine Äquivalenzrelation

(c) X sei Menge aller Menschen, x R y $\Leftrightarrow$ x, y haben die gleiche Hautfarbe

$\Rightarrow$ ist eine Äquivalenzrelation

(d) X sei Menge aller konvexen Vielecke in der Ebene, x R y $\Leftrightarrow$ x, y haben die gleiche Anzahl von Ecken

$\Rightarrow$ ist eine Äquivalenzrelation

(e) X sei Menge aller Dreiecke in der Ebene, x R_1 y $\Leftrightarrow$ x und y sind ähnlich, x R_2 y $\Leftrightarrow$ x und y sind kongruent

$\rightarrow$ beides sind Äquivalenzrelationen

Definition 3 *Sei $\sim$ eine Äquivalenzrelation auf X. Eine nichtleere Teilmenge M von X heißt Äquivalenzklasse (bezüglich $\sim$), wenn M folgende zwei Eigenschaften hat:*
(i) $x \in M$, $y \in M \Rightarrow x \sim y$, (ii) $x \in M$, $y \notin M \Rightarrow x \not\sim y$.

Definition 4 *Ein System $\{X_\alpha\}$ von Teilmengen einer Menge X heißt Klasseneinteilung, wenn folgende drei Eigenschaften gelten:*
(iii) $X_\alpha \neq \varnothing \quad \forall \alpha \in I$ (iv) $X_\alpha \cap X_\beta = \varnothing \forall \alpha, \beta \in I$ mit $\alpha \neq \beta$
(v) $X = \bigcup_{\alpha \in I} X_\alpha$.

5.Beispiel:

Beispiel 3: x $\sim$ y $\Leftrightarrow$ 2|x-y

M_1 = Menge der geraden Zahlen

M_2 = Menge der ungeraden Zahlen

$\mathcal{Z} = M_1 \cup M_2$

M_1 und M_2 sind Äquivalenklassen

$\{M_1, M_1\}$ ist eine Klasseneinteilung.

Beispiel 4b:

Menge aller grünen Artikel ist eine Äquivalenklasse, falls mindestens ein grüner Artikel vorhanden ist.

Artikel = Menge der grünen Artikel $\cup$ Menge der roten Artikel ...

$\rightarrow$ ist eine Klasseneinteilung

Beispiel 4c:

Menge aller Menschen = Weiße $\cup$ Indianer $\cup \ldots$

ist Klasseneinteilung in Äquivalenzklassen

Beispiel 4d:

Menge aller konvexen Vielecke = Menge der Dreiecke $\cup$ Menge der Vierecke $\cup$ Menge der Fünfecke ...

$\rightarrow$ Klasseneinteilung in Äquivalenzklassen

Beispiel 4e:

$$X = \bigcup_{\alpha>0,\beta>0,\ \alpha+\beta<180^\circ} X_{\alpha,\beta} \text{ mit } X_{\alpha,\beta} = \text{ Menge aller Dreiecke mit}$$

$$\text{Innenwinkeln } \alpha, \beta, 180^\circ - \alpha - \beta$$

→ Klasseneinteilung in Äquivalenzklassen

$$X = \bigcup_{a,b,c>0,\ a+b>c,\ a+c>b,\ b+c>a} X_{a,b,c} \text{ mit } X_{a,b,c} = \text{ Menge aller Dreiecke mit}$$

$$\text{den Seitenlängen } a, b, c$$

→ Klasseneinteilung in Äquivalenzklassen

Satz 1 *Das System der Äquivalenzklassen einer Äquivalenzrelation auf X ist stets eine Klasseneinteilung von X.*

Beweis:

Sei auf X eine Äquivalenzrelation $\sim$ gegeben. Sei $\{X_\alpha\}_{\alpha\in\ I}$ das *System (= Menge = Familie)* aller Äquivalenzklassen.

Wir müssen die Eigenschaften (iii),(iv) und (v) überprüfen:

(iii) Jede Äquivalenzklasse ist nach Definition nichtleer.

(iv) Seien X_α und X_β verschiedene Äquivalenzklassen. Dann ist $X_\alpha \setminus X_\beta$ oder $X_\alpha \setminus X_\beta$ nichtleer.

Sei etwa $X_\alpha \setminus X_\beta \neq \varnothing$.

Dann existiert ein b $\in\ X_\alpha \setminus X_\beta$.

Wir wollen $X_\alpha \cap X_\beta = \varnothing$ beweisen. Wir nehmen das Gegenteil an, d.h. dass ein a $\in X_\alpha \cap X_\beta$ existiert. Wir haben dann a $\in X_\alpha$, b $\in X_\alpha \overset{(i)}{\Rightarrow} a \sim b$

a $\in X_\beta$, b $\notin X_\beta \overset{(ii)}{\Rightarrow}$ a $\not\sim$ b $\lightning$

(v) Sei a $\in$ X beliebig. Wir definieren

$$M = \{x \in\ X : x \sim\ a\}.$$

Die Menge M ist eine Äquivalenzklasse:

M $\neq \varnothing$, da a $\in$ M (wegen (r))

(i): x $\in$ M, y $\in$ M $\Rightarrow$ x $\sim$ a, y $\sim$ a $\overset{(s)}{\Rightarrow}$ x $\sim$ a, a $\sim$ y $\overset{(t)}{\Rightarrow}$ x $\sim$ y

(ii): x $\in$ M, y $\notin$ M

Wir nehmen das Gegenteil an:

x $\sim$ y

x $\in$ M $\Rightarrow$ x $\sim$ a (nach Definition von M)

x $\sim$ y $\overset{(s)}{\Rightarrow}$ y $\sim$ x $\overset{(t)}{\Rightarrow}$ y $\sim$ a $\Rightarrow$ y $\in$ M (nach Definition von M) $\lightning$

Da M eine Äquivalenzklasse ist, existiert ein $\alpha \in I$ mit $M = X_\alpha$.
Weil $a \sim a$ (wegen (r)) ist $a \in M = X_\alpha$. Also gilt:

$$X = \bigcup_{\alpha \in I} X_\alpha$$

#

Definition 5 *Sei $\sim$ eine Äquivalenzrelation auf X. Die (eindeutig bestimmte) Äquivalenzklasse, die ein Element $a \in X$ enthält, wird mit $\hat{a}$ bezeichnet und Äquivalenzklasse, die a enthält, oder Restklasse, die a enthält, genannt. Ist M eine Äquivalenzklasse, so heißt jedes $a \in M$ Repräsentant.*

11.11.05

Satz 2 *Sei $\{X_\alpha\}_{\alpha \in I}$ eine Klasseneinteilung von X. Dann ist die durch*

$$x \sim y \Leftrightarrow \exists \alpha \in I : x \in X_\alpha \text{ und } y \in X_\alpha$$

auf X definierte Relation eine Äquivalenzrelaion.

Beweis:

(R) $x \sim x$, da $x \in X_\alpha$, $x \in X_\alpha$.

(S) $x \sim y \Leftrightarrow \exists \alpha : x \in X_\alpha$, $y \in X_\alpha \Rightarrow \exists \alpha$: $y \in X_\alpha$, $x \in X_\alpha \Rightarrow y \sim x$.

(T) $x \sim y$, $y \sim z \Rightarrow \exists \alpha$: $x \in X_\alpha$, $y \in X_\alpha$

$\exists \beta$: $y \in X_\beta$, $z \in X_\beta$

$\Rightarrow \alpha = \beta$ (da $y \in X_\alpha \cap X_\beta$)

$\Rightarrow x \in X_\alpha$, $z \in X_\alpha = X_\beta$

$\Rightarrow x \sim z$ #

Satz 3 *(a) Sei $\sim$ eine Äquivalenzrelation auf X und $\{X_\alpha\}_{\alpha \in I}$ die Klasseneinteilung der Äquivalenzklassen von $\sim$. Dann stimmt die durch*

$$x \approx y \Leftrightarrow \exists \alpha \in I : x \in X_\alpha, y \in X_\alpha$$

auf X definierte Relation mit $\sim$ überein.
(b) Sei $\{X_\alpha\}_{\alpha \in I}$ eine Klasseneinteilung von X und $\sim$ die durch

$$x \sim y \Leftrightarrow \exists \alpha \in I : x \in X_\alpha, y \in X_\alpha$$

auf X definierte Äquivalenzrelation. Dann stimmt die durch die Äquivalenzklasse von $\sim$ auf X gegebene Klasseneinteilung mit $\{X_\alpha\}_{\alpha \in I}$ überein.

Definition 6 *Sei $\sim$ eine Äquivalenzrelation auf X. Die Menge der Äquivalenzklassen heißt dann Faktormenge von X bezüglich $\sim$ und wird mit $X/\sim$ oder $\hat{X}$ bezeichnet.*

Die Faktormenge ist eine Teilmenge der Potenzmenge $\mathcal{P}(X)$ (= Menge aller Teilmengen von X). Die Konstruktion von Klasseneinteilung über Äquivalenzklassen liegt dem Prinzip der Begriffsbildung zugrunde (z.B. „der Hund“, „grün“, „Indianer“, „das Dreieck“).

Von Bedeutung ist die Suche nach *Invarianten* und *Charakteristiken* von Äquivalenzklassen.

Beispiel:

Innenwinkel sind Invariante von (den Äquivalenzklassen) der Äquivalenzrelation der Kongruenz von Dreiecken.

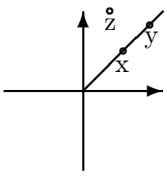
Innenwinkel sind Charakteristik von (den Äquivalenzklassen) der Äquivalenzrelation der Ähnlichkeit von Dreiecken.

Von höchstem Interesse ist das Problem der *Identifikation* von Äquivalenzklassen, d.h. das Problem der Suche nach bijektiven Abbildungen von X/$\sim$ auf andere, leichter handhabbare Mengen.

Beispiel:

Sei X=$\mathcal{R}^{\in} \setminus \{\varnothing\}$ und

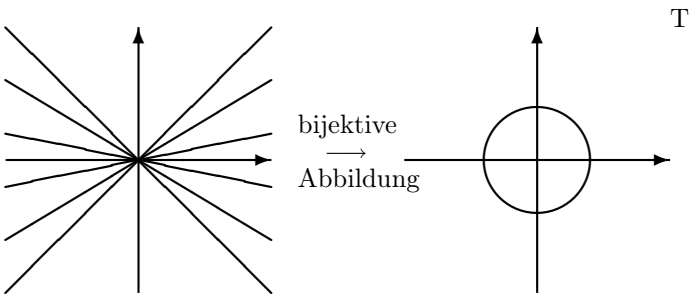
$x \sim y \Leftrightarrow \exists$ ein in 0 beginnender Strahl, der x und y enthält



$$x \sim y \quad x \not\sim z$$

$\rightarrow$ ist eine Äquivalenzrelation

Klasseneinteilung ist Menge aller in 0 beginnender Strahlen



Faktormenge ist

Menge dieser Strahlen

$$X/\sim \cong T$$

Man wählt sich von jedem Strahl nur einen Repräsentanten (z.B. den mit Abstand 1 zum Ursprung).

Alle Punkte mit Abstand 1 zum Ursprung bilden den Einheitskreis T.

$\rightarrow$ X/$\sim$ lässt sich auf den Einheitskreis abbilden

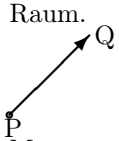
1.5 Vektoralgebra

Es gibt physikalische Größen, die durch einen einzigen Zahlenwert charakteristiert werden (z.B. Druck, Temperatur), sogenannte *Skalare*, und es gibt physikalische Größen, die durch eine Richtung und einen Zahlenwert charakterisiert werden (z.B. Geschwindigkeit, Kraft, Beschleunigung), sogenannte *Vektoren*.

Wir kommen zur mathematisch präzisen Definition des Vektors.

Wir betrachten Vektoren im dreidimensionalen Raum.

Ein gebundener Vektor ist ein geordnetes Paar (P,Q) von zwei Punkten P und Q im Raum.



Man nennt P den Anfangspunkt, Q den Endpunkt und schreibt $\vec{PQ}$ statt (P,Q).

Die Menge aller gebundenen Vektoren bezeichnen wir mit GV^3 (gebundene Vektoren im Raum)

Wir betrachten in GV^3 die folgende Relation: 14.11.05

$\vec{PQ} \sim \vec{RS} \Leftrightarrow \vec{RS}$ lässt sich aus $\vec{PQ}$ durch eine Verschiebung (*Translation*) im Raum erhalten.

In analytischer Sprache lautet dies wie folgt:

Wir wählen uns ein Koordinatensystem und können dann $\vec{PQ}$ und $\vec{RS}$ als $(p_1,p_2,p_3;q_1,q_2,q_3)$ und $(r_1,r_2,r_3;s_1,s_2,s_3)$ auffassen. Dann ist $\vec{PQ} \sim \vec{RS}$ genau dann, wenn t_1,t_2,t_3 mit

$$p_1 + t_1 = r_1 \quad q_1 + t_1 = s_1$$

$$p_2 + t_2 = r_2 \quad q_2 + t_2 = s_2$$

$$p_3 + t_3 = r_3 \quad q_3 + t_3 = s_3$$

existieren. Dies ist äquivalent zu

$$r_1 - p_1 = s_1 - q_1$$

$$r_2 - p_2 = s_2 - q_2$$

$$r_3 - p_3 = s_3 - q_3$$

Die so definierte Relation ist eine Äquivalenzrelation:

$$\vec{PQ} \sim \vec{PQ}$$

$$\vec{PQ} \sim \vec{RS} \Leftrightarrow \vec{RS} \sim \vec{PQ}$$

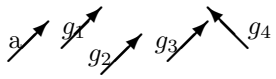
$$\vec{PQ} \sim \vec{RS}, \vec{RS} \sim \vec{TU} \Rightarrow \vec{PQ} \sim \vec{TU}$$

(alle Eigenschaften einer Äquivalenzrelation sind erfüllt)

Die Äquivalenzrelation auf GV^3 erzeugt eine Klasseneinteilung von GV^3. Die Menge aller Äquivalenzklassen ist die Faktormenge $GV^3/\sim$. Die Elemente dieser Faktormenge heißen *freie Vektoren* oder einfach *Vektoren*.

Ein Vektor ist also eine Äquivalenzklasse von gebundenen Vektoren.

Jeder Vektor a lässt sich also als a = $\hat{g}$ schreiben, wobei g ein gebundener Vektor ist. Das g ist nicht eindeutig bestimmt:

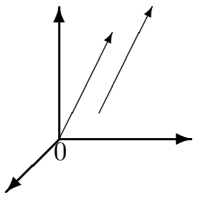


a=$\hat{g_1} = \hat{g_2} = \hat{g_3} \neq \hat{g_4}$

Die Menge aller Vektoren bezeichnen wir mit V^3.

$V^3 = GV^3/\sim = G\hat{V}^3$

Wir fixieren im Raum ein rechtwinkliges Koordinatensystem mit dem Ursprung 0.



Ist a ein Vektor, so gibt es genau einen gebundenen Vektor $\vec{OP}$ mit a = $\hat{\vec{OP}}$.

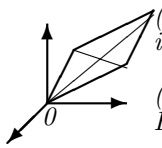
Gebundene Vektoren mit dem Anfangspunkt in 0 heißen *Ortsvektoren*. Der eindeutig bestimmte Ortsvektor $\vec{OP}$ mit a = $\hat{\vec{OP}}$ wird mit o(a) bezeichnet.

Wir haben also a = $\hat{o(a)}$.

Definition 1 *Die Koordinaten des Endpunktes des Ortsvektors o(a) werden die Koordinaten des Vektors a genannt.*

Ist a = $\hat{o(a)}$ und o(a) = $\vec{OP}$ mit P(p_1,p_2,p_3), so sind p_1,p_2,p_3 also die Koordinaten von a und man schreibt a = (p_1,p_2,p_3).

Definition 2 *Unter der Summe zweier Ortsvektoren $\vec{OP}$ und $\vec{OQ}$ ist der durch folgende Forderungen eindeutig bestimmte Ortsvektor $\vec{OR}$ zu verstehen:*



(i) die vier Punkte O,P,Q,R liegen in einer Ebene

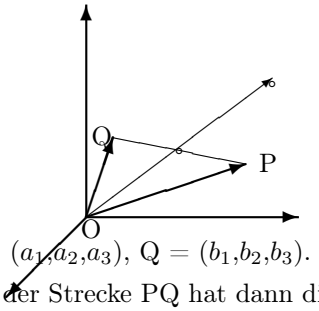
(ii) die vier Punkte O,P,Q,R bilden die Ecken eines Parallelogramms.

Die Summe zweier Vektoren a und b ist definiert als

$$a + b = o(a) \hat{+} o(b) = (o(a) + o(b))^{\cdot}$$

Satz 1 *Haben a,b $\in$ V^3 die Koordinaten (a_1,a_2,a_3) und (b_1,b_2,b_3), so hat a + b $\in$ V^3 die Koordinaten ($a_1 + b_1$,$a_2 + b_2$,$a_3 + b_3$).*

Beweis:



Wir haben P = (a_1,a_2,a_3), Q = (b_1,b_2,b_3).
Mittelpunkt M der Strecke PQ hat dann die Koordinaten

$$M = \left(\frac{a_1 + b_1}{2}, \frac{a_2 + b_2}{2}, \frac{a_3 + b_3}{2} \right),$$

und R liegt auf der Geraden durch 0 und M und hat von den 0 den doppelten Abstand wie M.

Also R = $2 * \left(\frac{a_1+b_1}{2}, \frac{a_2+b_2}{2}, \frac{a_3+b_3}{2} \right) = (a_1 + b_1, a_2 + b_2, a_3 + b_3)$. #

Satz 2 *Für beliebige a, b, c* $\in V^3$ *gilt*
$a + b = b + a$ *(Kommutativität der Addition)*
$(a + b) + c = a + (b + c)$ *(Assoziativität der Addition)*

Beweis: folgt aus Satz 1.

Haben bisher nur definiert, was man unter der Summe von 2 Vektoren versteht. Durch die Assoziativität ist klar, was unter der Summe a + b + c von 3 Vektoren, a + b + c + d von 4 Vektoren usw. zu verstehen ist.

Definition 3 *Der Vektor* $a \in V^3$ *mit* $o(a) = \vec{0}\vec{0}$ *heißt Nullvektor und wird mit 0 bezeichnet.*

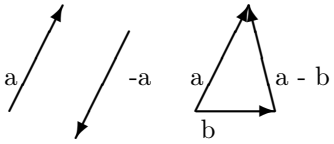
Satz 3 *Für jeden Vektor* $a \in V^3$ *gibt es genau einen Vektor* $b \in V^3$ *mit* $a + b = b + a = 0$.

Beweis: folgt aus Satz 1.

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

#

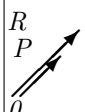
Definition 4 *Den nach Satz 3 eindeutig bestimmten Vektor b mit* $a + b = b + a = 0$ *nennt man den zu a entgegengesetzten Vektor und bezeichnet ihn mit -a.*
Die Differenz zweier Vektoren a und b ist definiert als $a - b = a + (\overline{-b})$.



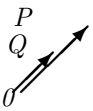
Satz 4 *Wenn a, b ∈ V^3 die Koordinaten (a_1,a_2,a_3) und (b_1,b_2,b_3) haben, so hat -a die Koordinaten ($-a_1$,$-a_2$,$-a_3$) und a - b hat die Koordinaten ($a_1 - b_1$,$a_2 - b_2$,$a_3 - b_3$).*

Beweis: folgt aus Satz 1. #

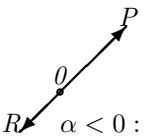
Definition 5 *Sei $\alpha \in \mathcal{R} \setminus \{0\}$ und $\vec{OP}$ ein vom Nullvektor verschiedener Ortsvektor. Dann ist $\alpha\ \vec{OP}$ wie folgt definiert:*
$\alpha\ \vec{OP} = \vec{OR}$ mit



$\alpha > 1$:
$|OR| = \alpha\ |OP|$



$0 < \alpha \leqq 1$:
$|OR| = \alpha\ |OP|$



$\alpha < 0$:
$|OR| = \alpha|OP|$

Man definiert desweiteren $0\ \vec{OP} = \vec{00}$, $\alpha\ \vec{00} = \vec{00}$.

Ist $\alpha\ \in \mathcal{R}$ und a $\in\ V^3$, so definiert man $\alpha_a = (\alpha\ o(a))\wedge$.

Satz 5 *Hat a $\in\ V^3$ die Koordinaten (a_1,a_2,a_3), so hat α_a die Koordinaten (αa_1,αa_2,αa_3).*

Beweis: Definition 5 und Strahlensatz. #

Bemerkung:

Es gilt -a= (-1)a, wobei -a aus Definition 4 und (-1)a aus Definition 5 ist. Beweis: Satz 4 und Satz 5 #

Satz 6 *Es gilt*

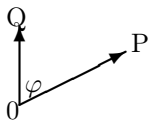
$$\alpha(\beta_a) = (\alpha\beta)a \quad \forall\alpha,\ \beta \in \mathcal{R}\ \forall\ \dashv \in\ \mathcal{V}^\ni$$

$$\alpha(a+b) = \alpha a + \alpha b \quad \forall\alpha \in \mathcal{R}\ \forall\ \dashv,\ \lfloor \in\ \mathcal{V}^\ni$$

$$(\alpha+\beta)a = \alpha a + \beta a \quad \forall\alpha,\ \beta \in \mathcal{R}\ \forall\ \dashv \in\ \mathcal{V}^\ni$$

Beweis: Satz 1 und Satz 5. #

Definition 6 *Das Skalarprodukt $\vec{OP} * \vec{OQ}$ zweier vom Nullvektor verschiedener Ortsvektoren ist die Zahl $|OP| * |OQ| * \cos\varphi$, wobei φ der Winkel zwischen $\vec{OP}$ und $\vec{OQ}$ ist.*



$$\cos\varphi = \cos(360^\circ - \varphi)$$

Ist $\vec{OP}$ oder $\vec{OQ}$ der Nullvektor, so definiert man $0\vec{P} * 0\vec{0} = 0$.

Das Skalarprodukt zweier Vektoren a, b $\in V^3$ ist definiert als die Zahl o(a) * o(b) und wird mit (a,b) (manchmal auch mit a*b oder <a,b>) bezeichnet.

18.11.05

Der Betrag eines Vektors a ist definiert als die Länge von o(a) und wird mit |a| bezeichnet. Haben also

$$(a, b) = |a||b| \cos\varphi.$$

Zwei Vektoren a und b heißen senkrecht, a $\perp$ b, wenn o(a) und o(b) senkrecht sind (ist a = 0 oder b = 0, so werden a und b als senkrecht angesehen).

Definition 7 *Im Raum sei ein rechtwinkliges Koordinatensystem gegeben. Die Vektoren i = (1,0,0), j = (0,1,0), k = (0,0,1) heißen Basisvektoren (bezüglich des Koordinatensystem).*

Satz 7 *Ist $a \in V^3$ ein Vektor mit den Koordinaten (a_1,a_2,a_3), so gilt $a = a_1 i + a_2 j + a_3 k$.*

Sind b_1,b_2,b_3 Zahlen mit $a = b_1 i + b_2 j + b_3 k$, so gilt $b_1 = a_1$, $b_2 = a_2$, $b_3 = a_3$.

Beweis: folgt aus Satz 1, Satz 5, Definition 7. #

$[c_1 i + c_2 j + c_3 k = c_1(1,0,0) + c_2(0,1,0) + c_3(0,0,1) = (c_1,0,0) + (0,c_2,0) + (0,0,c_3) = (c_1,c_2,c_3).]$

Satz 8 *(a) Es gilt $a \perp b \Leftrightarrow (a,b) = 0$.*
(b) Es gilt $(a+b,c) = (a,c) + (b,c)$
$(a,b+c) = (a,b) + (a,c)$
$(\alpha a,b) = \alpha(a,b)$
$(a,\alpha b) = \alpha(a,b)$
$\forall a,b,c \in V^3\ \forall \alpha \in \mathcal{R}$
(Bilinearität des reellen Skalarprodukts)
(c) Ist $a = (a_1,a_2,a_3)$ und $b = (b_1,b_2,b_3)$, so gilt $(a,b) = a_1b_1 + a_2b_2 + a_3b_3$.

Beweis:

(a) $a \perp b \Leftrightarrow \varphi \in \{90^\circ, 270^\circ\} \Leftrightarrow \cos\varphi = 0$

$(\varphi \in [0^\circ, 360^\circ]) \Leftrightarrow (a,b) = 0.$

(b) Haben $(a,b) = |a| \underbrace{|b| \cos\varphi}_{Pr_a b} = |a| * Pr_a b = |b||a| \cos\varphi = |b| Pr_b a$,

wobei $Pr_x y$ die Projektion von y auf x ist.

Daraus folgt $(a, b+c) = |a| * Pr_a(b+c) = |a|(Pr_a b + Pr_a c) = |a| * Pr_a b + |a| * Pr_a c = (a,b) + (a,c)$

Analog zeigt man die erste Formel (a+b,c) = (a,c) + (b,c). Die letzten beiden Formeln sind offensichtlich.

(c) Haben $(a,b) = (a_1 i + a_2 j + a_3 k, b_1 i + b_2 j + b_3 k) = (a_1 i, b_1 i) + (a_1 i, b_2 j) + (a_1 i, b_3 k) + (a_2 j, b_1 i) + (a_2 j, b_2 j) + (a_2 j, b_3 k) + (a_3 k, b_1 i) + (a_3 k, b_2 j) + (a_3 k, b_3 k)$ (nach (b)) $= a_1 b_1 (i,i) + a_1 b_2 (i,j) + a_1 b_3 (i,k) + a_2 b_1 (j,i) + a_2 b_2 (j,j) + a_2 b_3 (j,k) + a_3 b_1 (k,i) + a_3 b_2 (k,j) + a_3 b_3 (k,k)$ (nach (b)) $= a_1 b_1 * 1 + a_1 b_2 * 0 + a_1 b_3 * 0 + a_2 b_1 * 0 + a_2 b_2 * 1 + a_2 b_3 * 0 + a_3 b_1 * 0 + a_3 b_2 * 0 + a_3 b_3 * 1 = a_1 b_1 + a_2 b_2 + a_3 b_3$. #

Bemerkung:

Betrachten im Raum die Menge E aller Punkte (x,y,z) mit

$$ax + by + cz + d = 0.$$

$(a^2 + b^2 + c^2 > 0)$.

Sei $(x_0, y_0, z_0) \in$ E, d.h.

$$ax_0 + by_0 + cz_0 + d = 0.$$

Dann gilt: (x,y,z) $\in$ E $\Leftrightarrow$ $a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$.

Betrachten die Vektoren n = (a,b,c), p = (x,y,z), $p_0 = (x_0, y_0, z_0)$

Dann ist (x,y,z) $\in$ E $\Leftrightarrow$ $(n, p - p_0) = 0 \Leftrightarrow$ $n \perp p - p_0$

$\Leftrightarrow$ (x,y,z) gehört zu der Ebene, auf der (a,b,c) senkrecht steht und die durch den fixierten Punkt (x_0, y_0, z_0) geht.

Sehen auch, dass (a,b,c) ein Normalenvektor zu der Ebene ist.

Definition 8 *Seien P und Q zwei Punkte im Raum, so dass O,P,Q nicht auf einer Geraden liegen. Das Vektorprodukt der Ortsvektoren $\vec{OP}$ und $\vec{OQ}$ ist der durch die folgenden 3 Bedingungen eindeutig bestimmte Vektor $\vec{OR}$:*

(a) $\vec{OR}$ steht auf $\vec{OP}$ und $\vec{OQ}$ senkrecht.

(b) $\vec{OP}$, $\vec{OQ}$, $\vec{OR}$ bilden ein Rechtssystem, d.h. $\vec{OP}$,$\vec{OQ}$,$\vec{OR}$ sind die Daumen, Zeigefinger, Mittelfinger der rechten Hand.

(c) Die Länge /OR/ ist gleich dem Flächeninhalt des von OP und OQ aufgespannten Parallelogramms, d.h. gleich

$$|OP| * |OQ| * \sin\varphi$$

Wenn O,P,Q auf einer Geraden liegen, so ist ihr Vektorprodukt als $\vec{0}\vec{0}$ definiert.

Für a,b $\in$ V^3 definiert man das Vektorprodukt als die Äquivalenzklasse, die das Vektorprodukt von o(a) und o(b) enthält. Bezeichnung für das Vektorprodukt:

$$a \times b = a \wedge b = \ldots$$

Also $a \times b = (o(a) \times o(b)) \wedge$.

Sind a_1,a_2,a_3,b_1,b_2,b_3 Zazhlen, so ist

$$\begin{vmatrix} i & j & k \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

definiert als der Vektor

$$\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} i - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} j + \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} k$$

$$= \left(\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}, - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}, \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \right).$$

Hierbei ist $\begin{vmatrix} x & y \\ u & v \end{vmatrix} := xv - yu$.

Satz 9 *(a) Es gilt*

$$a \times b = -b \times a,$$

$$a \times (b + c) = a \times b + a \times c$$

$$(a + b) \times c = a \times c + b \times c$$

$$(\alpha a) \times b = \alpha(a \times b) = a \times (\alpha b)$$

(b) Sind a,b $\in V^3$ Vektoren mit den Koordinaten (a_1,a_2,a_3) und (b_1,b_2,b_3), so ist

$$a \times b = \begin{vmatrix} i & j & k \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}.$$

Beweis:

(a) Formeln $a \times b = -b \times a, (\alpha a) \times b = \alpha(a \times b) = a \times (\alpha b)$ folgen unmittelbar aus der Definition.

Der Beweis von $a \times (b + c) = a \times b + a \times c$ ist machbar, aber mühsam, und wird deshalb weggelassen.

(b) Haben nach Satz 7

$$a \times b = (a_1 i + a_2 j + a_3 k) \times (b_1 i + b_2 j + b_3 k)$$

$$= a_1 b_1 (i \times i) + a_1 b_2 (i \times j) + a_1 b_3 (i \times k) + a_2 b_1 (j \times i)$$

$$+ a_2 b_2 (j \times j) + a_2 b_3 (j \times k) + a_3 b_1 (k \times i) + a_3 b_2 (k \times j) + a_3 b_3 (k \times k)$$

$$= i(a_2 b_3 - a_3 b_2) + j(a_3 b_1 - a_1 b_3) + k(a_1 b_2 - a_2 b_1)$$

$$= i \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} - j \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} + k \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$$

#

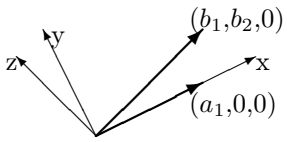
Bemerkung:

Man kann auch anders vorgehen, nämlich $a \times b$ über

$$a \times b = \begin{vmatrix} i & j & k \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

definieren und dann Definition 8 als Satz betrachten. Dies erspart den mühsamen Teil des Beweises von Satz 9 (a).

Der Beweis von Definition 8 als Satz geht dann wie folgt:

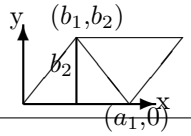


Koordinatensystem wie in der Abbildung wählen

Haben dann

$$a \times b = \begin{vmatrix} i & j & k \\ a_1 & 0 & 0 \\ b_1 & b_2 & 0 \end{vmatrix} = (0, 0, a_1 b_2),$$

und Eigenschaften (a) und (b) sind erfüllt. Um (c) zu zeigen, müssen wir beweisen, dass das von a und b aufgespannte Paralellogramm den Flächeninhalt $a_1 b_2$ hat.



21.11.05

Satz 10 *Seien A,B,C Punkte im Raum mit den Koordinaten* (a_1,a_2,a_3), (b_1,b_2,b_3), (c_1,c_2,c_3), *seien a,b,c Vektoren mit denselben Koordinaten. Dann sind folgende Bedingungen äquivalent:*

(i) A,B,C liegen auf einer Geraden,
(ii) $(b-a)x(c-a) = 0$,
(iii) $\begin{vmatrix} i & j & k \\ b_1 - a_1 & b_2 - a_2 & b_3 - a_3 \\ c_1 - a_1 & c_2 - a_2 & c_3 - a_3 \end{vmatrix} = 0$, *d.h.*

$$\begin{vmatrix} b_2 - a_2 & b_3 - a_3 \\ c_2 - a_2 & c_3 - a_3 \end{vmatrix} = \begin{vmatrix} b_1 - a_1 & b_3 - a_3 \\ c_1 - a_1 & c_3 - a_3 \end{vmatrix} = \begin{vmatrix} b_1 - a_1 & b_2 - a_2 \\ c_1 - a_1 & c_2 - a_2 \end{vmatrix} = 0.$$

[(i) $\Leftrightarrow$ (iii) hatten wir schon in 1.3., damals aber mit Krampfbeweis.]

Beweis:

(i) $\Leftrightarrow$ (ii):

A,B,C auf einer Geraden $\Leftrightarrow \vec{AB} \times \vec{AC} = 0$.

$\Leftrightarrow (b-a) \times (c-a) = 0$.

(ii)⇔ (iii)

Satz 9(b). #

Bemerkung:

Hatten in 1.3. auch die Gleichung der Ebene durch 3 gegebene Punkte bestimmt. Auch dies geht natürlicher mit Vektorrechnung:

Seien A,B,C,a,b,c wie in Satz 10.

Sei desweiteren P ein beliebiger Punkt im Raum mit den Koordinaten (x,y,z) und p ein Vektor mit den gleichen Koordinaten.

Offenbar gilt:

P liegt in der von A,B,C, aufgespannten Ebene $\Leftrightarrow p - a \perp (b-a) \times (c-a)$

$\Leftrightarrow (p-a, (b-a) \times (c-a)) = 0$

$$\Leftrightarrow (x-a_1) \begin{vmatrix} b_2 - a_2 & b_3 - a_3 \\ c_2 - a_2 & c_3 - a_3 \end{vmatrix} + (y-a_2) \left|\cdots\right| + (z-a_3) \left|\cdots\right| = 0,$$

d.h. erhalten im Handumdrehen die Ebenengleichung aus 1.3.

Die Parameterdarstellung dieser Ebene (siehe 1.3.) ist jetzt auch anschaulich klar:

$$p - a = \lambda(b-a) + \mu(c-a) \quad , \ \lambda, \mu \in \mathcal{R}$$
$$p = a + \lambda(b-a) + \mu(c-a).$$

1.6 Körper und komplexe Zahlen

In der Menge $\mathcal{R}$ der reellen Zahlen sind eine Addition a + b und eine Multiplikation ab erklärt, die folgende 9 Eigenschaften haben:

$(K+)$ $a + b = b + a = \forall\ a, b \in \mathcal{R}$ (Kommutativität der Addition)
$(A+)$ $(a + b) + c = a + (b + c) \forall\ a, b, c \in \mathcal{R}$ (Assoziativität der Addition)
(0) $a + 0 = 0 + a = a \forall\ a \in \mathcal{R}$ (Existenz der Null)
$(-)$ $\forall\ a \in \mathcal{R} \exists\ \in \mathcal{R} : \dashv + \lfloor = \lfloor + \dashv = \prime$ (Durchführbarkeit der Subtraktion)
$(K*)$ $a * b = b * a \forall\ a, b \in \mathcal{R}$ (Kommutativität der Multipliktaion)
$(A*)$ $(a * b) * c = a * (b * c) \forall\ a, b, c \in \mathcal{R}$ (Assoziativität der Multiplikation)
(E) $a * 1 = 1 * a = a \forall\ a \in \mathcal{R}$ (Existenz der Eins)
$(/)$ $a * (b + c) = a * b + a * c$
$(a + b) * c = a * c + b * c \forall\ a, b, c \in \mathcal{R}$ (Distributivgesetz)

Man möchte nun auch in beliebigen Mengen wie in der Menge der reellen Zahlen rechnen, d.h. man möchte eine Addition und eine Multiplikation einführen, sodass möglichst viele der obigen Eigenschaften gelten. Wenn es gelingt, alle 9 Eigenschaften zu erreichen, sagt man, dass man die Menge zu einem Körper gemacht hat.

Definition 1 *Eine Menge heißt Körper (engl.field), wenn in K eine Addition und eine Multiplikation, d.h. zwei Vorschriften, die jedem geordneten Paar (a,b)* $\in K^2$ $(= K \times K)$ *ein eindeutig bstimmtes Element* $a + b \in K$ *bzw.* $ab \in K$ *zuordnen, erklärt sind, sodass dabei folgende 9 Axiome gelten:*

$$(K+) a + b = b + a \forall\ a, b \in K$$

$$(A+)(a + b) + c = a + (b + c) \forall\ a, b, c \in K$$

$(0) \exists$ *ein Element aus K, das mit 0 bezeichnet und Nullelement genannt wird,*

$$\text{sodass } a + 0 = 0 + a = a \forall\ a \in K$$

$$(-) \forall\ a \in K \exists\ b \in K : a + b = b + a = 0$$

$$(K*) a * b = b * a \forall\ a, b \in K$$

$$(A*)(ab)c = a(bc) \forall\ a, b, c \in K$$

(E) *es existiert ein Element aus K, das mit e (oder* 1 *oder I oder ...) bezeichnet*

$$\text{und Einselement genannt wird, sodassae} = ea = a \forall\ a \in K$$

$$(/)\ \forall\ a \in K \setminus \{\varnothing\} \exists\ b \in K : ab = ba = 1$$

$$a(b + c) = ab + ac, (a + b)c = ac + bc \forall\ a, b \in K$$

Beispiel 1:

$\mathcal{R}$ (reelle Zahlen) und $\mathcal{Q}$ (rationale Zahen) sind Körper mit der üblichen Addition und Multiplikation. Die Menge $\mathcal{Z}$ (ganze Zahlen) ist hingegen kein Körper mit der üblichen Addition und Multiplikation.

[Axiom $(/)$ ist nicht erfüllt]

Beispiel 2:

Gegeben sei eine natürliche Zahl n $\geqq 2$. Definieren in $\mathcal{Z}$ eine Äquivalenzrelation durch

$$x \sim y \Leftrightarrow x - y \text{ ist durch } n \text{ teilbar} (\Leftrightarrow x \equiv y (mod\ n))$$

Die Äquivalenzklassen sind

$$\overline{0} = \{0, n, -n, 2n, -2n, \ldots\}$$
$$\overline{1} = \{0, n + 1, -n + 1, 2n + 1, -2n + 1, \ldots\}$$
$$\vdots$$
$$\overline{n - 1} = \{n - 1, n + n - 1, -n + n - 1, 2n + n - 1, -2n + n - 1, \ldots\}$$

Für n = 3 gibt es dann beispielsweise 3 Äquivalenzklassen:

$$\overline{0} = \{0, 3, 6, \ldots\}$$
$$\overline{1} = \{1, 4, 7, \ldots\}$$
$$\overline{2} = \{2, 5, 8, \ldots\}$$

Man bezeichnet die Menge dieser Äquivalenzklassen mit $\mathcal{Z} \setminus \setminus \mathcal{Z} =: \mathcal{Z}_\setminus$

Definieren in $\mathcal{Z}_\setminus$ Addition und Multiplikation wie folgt:

$\overline{a} + \overline{b} = \overline{c}$, wobei c $\in \{0, 1, \ldots, n-1\}$ der Rest von a + b bei Division durch n ist

$\overline{a} * \overline{b} = \overline{c}$, wobei c $\in \{0, 1, \ldots, n-1\}$ der Rest von ab bei Division durch n ist

Ist dies ein Körper?

(K+),(A+),(0) [$\overline{0}$ ist Nullelement], (-), (K*),(A*),(E) [e = $\overline{1}$ ist Einselement], (D) sind erfüllt. Behaupten, dass (/) genau dann gilt, wenn n eine Primzahl ist.

Sei n keine Primzahl. Dann ist n =pq mit p $\geqq$ 2, q $\geqq$ 2. Es ist $\overline{p} \neq \overline{0}$ und die Gleichung $\overline{p} * \overline{b} = \overline{1}$ ist nicht lösbar.

Wäre nämlich $\overline{pb} = \overline{1} \Rightarrow n|pb-1 \Rightarrow p|pb-1 \Rightarrow p|1$ ↙.

Sei nun n eine Primzahl. Nehmen $\overline{a} \neq \overline{0}$ (d.h. a $\in \{1, 2, \ldots, n-1\}$) und betrachten $\overline{a} * \overline{0}, \overline{a} * \overline{1}, \ldots, \overline{a} * \overline{n-1}$. Keine zwei dieser n Elemente sind gleich:

$$\overline{a} * \overline{k} = \overline{a} * \overline{l}$$

$$\Rightarrow ak \equiv al (mod\ n)$$

$$\Rightarrow n|a(k-l) \overset{Primzahl}{\Rightarrow} n|a\ oder\ n|k-l$$

$$\swarrow,\ da\ a \in \{1, \ldots, n-1\}\ und\ k-l \in \{-(n-1), \ldots, (n-1)\}$$

Also ist $\{\overline{a} * \overline{0}, \overline{a} * \overline{1}, \ldots, \overline{a} * \overline{n-1}\} = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$

(Gleichheit von Mengen). Somit muss es irgendein $\overline{b}$ geben mit $\overline{a} * \overline{b} = \overline{1}$. Also ist (/) bewiesen.

Beispiel 3:

Betrachten $\mathcal{R}^\in$, d.h. Menge aller geordneten Paare (a,b) mit a,b $\in \mathcal{R}$. Können uns $\mathcal{R}^\in$ als Punkte der Ebene, als Spitzen von ebenen Ortsvektoren, als ebene freie Vektoren usw. vorstellen.

Definieren eine Addition in $\mathcal{R}^\in$ durch

$$(a, b) + (c, d) = (a + c, b + d)$$

Axiome (K+),(A+),(0) [Nullelement ist (0,0)], (-) sind erfüllt.

Versuchen es mit der Multiplikation

$$(a, b) * (c, d) = (ac, bd)$$

(K*),(A*),(E) [Einselement ist (1,1)] sind erfüllt. Aber (/) ist nicht erfüllt: (1,0)(c,d) = (1,1) lässt sich nicht lösen.

Versuch ist also gescheitert. Definieren Multiplikation durch

$$(a, b)(c, d) = (ac - bc, ad + bc).$$

(K*),(A*),(E) [Einselement ist (1,0)] sind erfüllt.

Axiom (/) gilt auch: für (a,b)$\neq$ (0,0) ist

$$(a, b)\left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right) = \left(\frac{a^2}{a^2+b^2} + \frac{b^2}{a^2+b^2}, \frac{-ab}{a^2+b^2} + \frac{ba}{a^2+b^2}\right) = (1, 0).$$

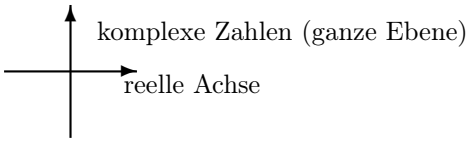
Also ist $\mathcal{R}^{\in}$ mit obiger Addition und Multiplikation ein Körper. Dieser Körper heißt *Körper der komplexen Zahlen* und wird mit $\mathcal{C}$ bezeichnet. 25.11.05

Man kann sich die komplexen Zahlen also als $\mathcal{R}^{\in}$ (= Menge der geordneten Paare (a,b) $\in \mathcal{R}$), als Punkte einer Ebene, als Spitzen von Ortsvektoren oder als freie Vektoren (=Elemente von V^2) vorstellen, mit denne nach obigen Regeln gerechnet wird.

Identifizieren Paare der Form (a,0) mit der reellen Zahl a, d.h. setzen (a,0)=a. Damit wird $\mathcal{R}$ eine Teilmenge von $\mathcal{C}$, und die Rechenoperationen von $\mathcal{C}$, eingeschränkt auf diese Teilmenge, sind die üblichen Operationen von $\mathcal{R}$:

$$(a,0)+(b,0)=(a+b,0)$$
$$a+b=a+b$$
$$(a,0)(b,0)=(ab-00,a0+0b)=(ab,0)$$
$$ab=ab$$

Geometrisch:



Man setzt i:=(0,1). Haben dann

$$(a,b)=(a,0)+\overbrace{(b,0)(0,1)}^{(0,b)}=a+bi$$

und analog

$$(a,b)=a+ib.$$

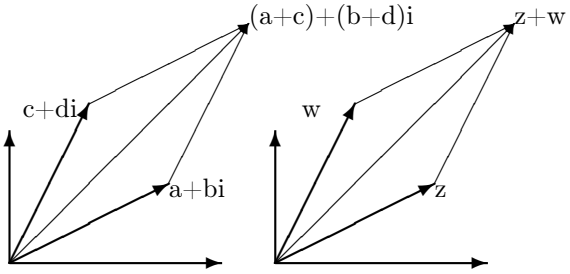
Man schreibt komplexe Zahlen daher in der Form $a+bi$ oder $a+ib$.

Rechenregeln sind in dieser Schreibweise wie folgt:

$$(a+bi)+(c+di)=(a+c)+(b+d)i,$$
$$(a+bi)(c+di)=ac+adi+bci+bdi^2$$
$$[i^2=(0,1)(0,1)=(-1,0)=-1]$$
$$=(ac-bd)+(ad+bc)i$$

Man braucht sich also bloß zu merken, dass $i^2=-1$ ist.

Wie kann man sich Addition und Multiplikation komplexer Zahlen anschaulich vorstellen?



Man nennt a bzw. b den *Real- bzw. Imaginärteil* von a+bi. Bei Addition komplexer Zahlen addieren sich also Real- und Imaginärteil.

Definition 2 *Sei $z = a + bi$ eine komplexe Zahl. Der Betrag $|z|$ von z ist definiert als $|z| = \sqrt{a^2 + b^2}$, und falls $z \neq 0$ ist, so ist das Argument von z jeder Winkel φ mit*

$$\cos\varphi = \frac{a}{\sqrt{a^2+b^2}}, \sin\varphi = \frac{b}{\sqrt{a^2+b^2}}.$$

Der Winkel φ aus $[0°, 360°] = [0, 2\pi)$ heißt Hauptargument. Argument und Hauptargument werden mit arg z und Arg z bezeichnet.

Bemerkung:

$\tan\varphi = \frac{b}{a}$ bestimmt φ nicht eindeutig!

Haben

$$z = a + bi = \sqrt{a^2+b^2}\left(\frac{a}{\sqrt{a^2+b^2}} + \frac{b}{\sqrt{a^2+b^2}}i\right)$$
$$= |z|(\cos\varphi + i\sin\varphi),$$

dies heißt *trigonometrische Form* von z.

Satz 1 *Ist $z = |z|(\cos\varphi + i\sin\varphi), w = |w|(\cos\psi + i\sin\psi)$, so ist*

$$zw = |z||w|(\cos(\varphi+\psi) + i\sin(\varphi+\psi))$$

.

Beweis:

$zw = |z||w|(\cos\varphi\cos\psi - \sin\varphi\sin\psi + i(\cos\varphi\sin\psi + \sin\varphi\cos\psi)) = |z||w|((\cos(\varphi+\psi) + i\sin(\varphi+\psi))$

#

Verschiedene Sorten von Zahlen entstanden aus der Notwendigkeit, sogenannte *algebraische Gleichungen*, d.h. Gleichungen der Form

$$x^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0 = 0$$

zu lösen (Koeffizienten $a_0, a_1, \ldots, a_{n-1}$ sind gegeben, x ist gesucht).

$$x + 2 = 0 \;\rightarrow\; Einführung\ \mathcal{Z}$$
$$3x - 4 = 0 \;\rightarrow\; \mathcal{Q}$$
$$x^2 - 2 = 0 \;\rightarrow\; \mathcal{R}$$
$$x^2 + 1 = 0 \;\rightarrow\; \mathcal{C}$$

Definition 3 *Seien K und L Körper. Der Körper L heißt Körpererweiterung von K, wenn K mit einer Teilmenge von L identifiziert werden kann, d.h. wenn es eine injektive Abbildung f: $K \to L$ mit*

$$f(a+b) = f(a) + f(b), f(ab) = f(a)f(b) \ \forall \ a, b \in \mathcal{K}$$

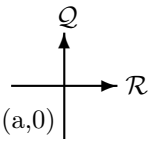
gibt. Jede Abbildung f heißt dann Einbettung von K in L.

Beispiel:

$\mathcal{R}$ ist Körpererweiterung von $\mathcal{Q}$:

$f : \mathcal{Q} \to \mathcal{R}$, $f(a) = a$ identische Abbildung

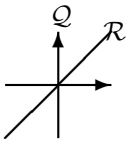
$\mathcal{C}$ ist Körpererweiterung von $\mathcal{R}$:



$f : \mathcal{R} \to \mathcal{Q}$, f(a)=(a,0)

$f : \mathcal{R} \to \mathcal{Q}$, f(a)=(a,a)

f ist injektiv:



$$f(a+b) = (a+b, a+b) = (a,a) + (b,b) = f(a) + f(b)$$
$$f(ab) = (ab, ab), f(a)f(b) = (a,a) * (b,b) = (ab - ab, ab + ab) = (0, 2ab)$$

Das ist keine Körpererweiterung (Multiplikation zerstört).

Die komplexen Zahlen wurden eingeführt, um die Gleichung $x^2 + 1 = 0$ lösbar zu machen.

Frage: Geht das auch anders? Der folgende Satz zeigt, dass die Antwort „nein“ ist.

Satz 2 *Sei K eine Körpererweiterung von $\mathcal{R}$ mit der Eigenschaft, dass in K ein Element j mit $j^2 = -1$ existiert. Dann ist K automatisch eine Körpererweiterung von $\mathcal{C}$.*

Mit anderen Worten:

$\mathcal{C}$ ist die kleinste Körpererweiterung von $\mathcal{R}$, in der die Gleichung $x^2 + 1 = 0$ lösbar ist.

Beweis:

Sei $f : \mathcal{R} \to$ K eine Einbettung und sei $g : \mathcal{R} \to \mathcal{C}$ die übliche Einbettung.

$$g(a) = (a, 0) = a + 0i$$

$$h(\alpha + \beta i) = f(\alpha) + f(\beta)j$$

Satz ist bewiesen, wenn wir gezeigt haben, dass h eine Einbettung ist.

(a)

$$h(a+b) = h((\alpha+\beta i+(\gamma+\delta i)) = h((\alpha+\gamma)+(\beta+\delta)i))$$
$$= f(\alpha+\gamma)+f(\beta+\delta)j = f(\alpha)+f(\gamma)+(f(\beta)+f(\delta))j$$
$$= (f(\alpha)+f(\beta)j)+(f(\gamma)+f(\delta)j)) = h(a)+h(b).$$

(b)

$$h(ab) = h((\alpha+\beta i)(\gamma+\delta i)) = h((\alpha\gamma\beta\delta)+(\alpha\delta+\beta\gamma)i)$$
$$= f(\alpha\gamma-\beta\delta)+f(\alpha\delta+\beta\gamma)j,$$
$$h(a)h(b) = h(\alpha+\beta i)h(\gamma+\delta i) = (f(\alpha)+f(\beta)j)(f(\gamma)+f(\delta)j)$$
$$= f(\alpha)f(\gamma)+f(\alpha)f(\delta)j+f(\beta)f(\gamma)j+f(\beta)f(\delta)j^2,$$

d.h. $h(ab) = h(a)h(b)$.

(c) Zeigen noch, dass h injektiv ist. Sei dazu h(a) = h(b). Setzen c = a-b und haben dann h(c) = 0. Müssen zeigen, dass c = 0 ist.

Nehmen c $\neq$ 0 an. Dann existiert ein d $\in \mathcal{C}$ mit cd = 1. Also gilt:

$$0 = h(c) * h(d) = h(cd) = h(1) = 1. \swarrow$$

[Warum $h(1) = 1$? Beweis: $a * 1 = a \ \forall \ a \in \mathcal{C} \Rightarrow \langle(\dashv)\langle(\infty) = \langle(\dashv) \ \forall \ \dashv \in \mathcal{C}$
$\exists \ a \in \mathcal{C}$ mit $h(a) \neq 0 \Rightarrow h(1) = 1$.] # 28.11.05

Haben zu $\mathcal{Q}$ neue Zahlen hinzugefügt, um neue Gleichungen lösbar zu machen, z.B. kann man Lösungen von $x^2 - 2 = 0$ mit $\sqrt{2}$ bezeichnen und $\sqrt{2}$ zu $\mathcal{Q}$ hinzufügen. Möchten dabei, dass wieder ein Körper herauskommt. Der kleinste Körper, der $\mathcal{Q}$ und $\sqrt{2}$ enthält, ist $\mathcal{Q}(\sqrt{2})$:={a+b$\sqrt{2}$: a,b $\in \mathcal{Q}$}

[Haben $(a+b\sqrt{2})(c+d\sqrt{2}) = ac+ad\sqrt{2}+bc\sqrt{2}+2bd = ac+2bd+(ad+bc)\sqrt{2}$,

$\frac{a+b\sqrt{2}}{c+d\sqrt{2}} = \frac{(a+b\sqrt{2})(c-d\sqrt{2})}{(c+d\sqrt{2})(c-d\sqrt{2}} = \frac{ac-2bd}{c^2-2d^2} + \frac{bc-ad}{c^2-2d^2}\sqrt{2}$.]

Wieder gibt es neue Gleichungen, z.B. $x^2 - 3 = 0$, die in $\mathcal{Q}(\sqrt{\in})$ nicht lösbar sind. Man müsste also $\mathcal{Q}(\sqrt{\in})(\sqrt{\ni})$:={a+b$\sqrt{2}$+c$\sqrt{3}$+d$\sqrt{6}$:a,b,c,d $\in \mathcal{Q}$} usw. bilden.

Der Körper $\mathcal{R}$ erledigt alle diese Erweiterungen mit einem Schlag: Ist p(x) ein Polynom, das sowohl nichtnegative als auch positive Werte annimmt, so hat die Gleichung p(x) = 0 stets eine Lösung in $\mathcal{R}$. Die komplexe Zahl i wurde eingeführt, um die Gleichung $x^2 + 1 = 0$ lösbar zu machen.

Der kleinste Körper, der $\mathcal{R}$ und i enthält, ist $\mathcal{C} = \mathcal{R}$(i):={a + bi: a, b $\in \mathcal{R}$}.

Wieder gibt es neue Gleichungen, z.B. $x^{28} + x^{11} + 1 = 0$, bei denen unklar ist, ob sie eine Lösung in $\mathcal{C}$ haben. Die Frage ist, ob wir wieder uferlos einzelne Körpererweiterungen machen müssen, damit alle algebraischen Gleichungen lösbar werden. Das Überraschende ist, dass die Antwort auf diese Frage „nein“ ist. Dies ist der Inhalt des folgenden Satzes:

Satz 3 Fundamentalsatz der Algebra
Für beliebige komplexe Zahlen a_0, a_1, …, a_{n-1} ($n \geq 1$) hat die algebraische Gleichung

$$x^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0 = 0$$

stets eine Lösung in $\mathcal{C}$.

Der Satz wurde erstmals von Gauß streng bewiesen (der insgesamt 7 verschiedene Beweise gefunden hat). Alle Beweise des Satzes sind kompliziert oder nicht elementar. Verschieben den Beweis daher auf später.

Vermerken, dass im Satz 3 für die Koeffizienten a_0, a_1, ..., a_{n-1} beliebige komplexe Zahlen zugelassen sind.

Definition 4 *Ein Körper K heißt algebraisch abgeschlossen, wenn für beliebige a_0, a_1, ..., a_{n-1} ($n \geq 1$) die Gleichung*

$$x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 = 0$$

stets eine Lösung hat.

Satz 3 lässt sich damit wie folgt umformulieren:

Satz 4 *Der Körper $\mathcal{C}$ ist algebraisch abgeschlossen.*

Vermerken, dass weder $\mathcal{Q}$ noch $\mathcal{R}$ algebraisch abgeschlossen sind.

Im Jahre 1833 stellte sich Hamilton (1805 - 1865) die Frage, wie man $\mathcal{R}^3$ zu einem Körper machen kann. Die Addition sollte durch

$$(a_1, b_1, c_1) + (a_2, b_2, c_2) = (a_1 + a_2, b_1 + b_2, c_1 + c_2)$$

gegeben sein. Dann gilt (K+), (A+), (0) [0 = (0,0,0)], (-).

Das Problem ist die Einführung der Multiplikation.

1. Versuch:

$$(a_1, b_1, c_1)(a_2, b_2, c_2) = (a_1 a_2, b_1 b_2, c_1 c_2)$$

(0,1,0)(x,y,z) = (0,y,0) wird niemals (1,1,1) → Axiom (/) gilt nicht.

2. Versuch:

$$(a_1, b_1, c_1)(a_2 b_2 c_2) = \begin{vmatrix} i & j & k \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix} = (b_1 c_2 - b_2 c_1, c_1 a_2 - a_1 c_2, a_1 b_2 - b_1 a_2)$$

(1,0,0)(x,y,z) = (0,-z,y) wird niemals (1,0,0) → wieder gilt (/) nicht.

Hamilton hat 10 Jahre lang vergeblich versucht, eine Multiplikation im $\mathcal{R}^3$ zu finden, sodass wenigstens ein Schiefkörper entsteht.

Definition 5 *Eine Menge K mit einer Addition und einer Multiplikation heißt Schiefkörper, wenn bis auf (K*) (= Kommutativität der Multiplikation) alle Axiome eines Körpers gelten.*

Am 16.10.1843 hat Hamilton entdeckt, dass $\mathcal{R}^4$ zu einem Schiefkörper gemacht werden kann. Dieser Körper wird Körper der Quaternionen genannt.

Beispiel:

∞=(1,0,0,0), i=(0,1,0,0), j=(0,0,1,0), k=(0,0,0,1)

Ein beliebiges Element $(x_0,x_1,x_2,x_3) \in \mathcal{R}^4$ schreiben wir dann in der Form $x_0\infty+x_1\mathrm{i}+x_2\mathrm{j}+x_3\mathrm{k}$.

Die Addition geschieht komponentenweise.
Multiplikation:
$(x_0\infty+x_1\mathrm{i}+x_2\mathrm{j}+x_3\mathrm{k})(y_0\infty+y_1\mathrm{i}+y_2\mathrm{j}+y_3\mathrm{k})$
mit folgender Multiplikationstabelle:

	∞	i	j	k
∞	∞	i	j	k
i	i	-∞	k	-j
j	j	-k	-∞	i
k	k	j	-i	-∞

und anschließendem Zusammenfassen in der Form $()\infty+()\mathrm{i}+()\mathrm{j}+()\mathrm{k}$

Hier gilt auch das Axiom (/).

Die Menge $\mathcal{R}^4$ mit obiger Addition und Multiplikation heißt *Quaternionenschiefkörper* und wird mit $\mathcal{H}$ bezeichnet.

Die Frage, für welche n sich $\mathcal{R}^n$ zu eine Schiefkörper machen lässt, wurde 1880 von Frobenius (1849 - 1917) gelöst.

Satz 5 *Wenn sich $\mathcal{R}^\backslash$ mit der Addition*

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

und irgendeiner Multiplikation zu einem Schiefkörper machen lässt, so ist entweder
$n = 1$ und $\mathcal{R}$ der Körper der reellen Zahlen,
$n = 2$ und $\mathcal{C}$ der Körper der komplexen Zahlen oder
$n = 4$ und $\mathcal{H}$ der Schiefkörper der Quaternionen.

Beweis: Literatur.

2 Lineare Räume und ihre Operatoren

2.1 Definitionen und Beispiele

Im folgenden sei stets $\mathcal{K} = \mathcal{R}$ oder $\mathcal{K} = \mathcal{C}$.

Definition 1 *Ein linearer Raum ist eine Menge X, in der eine Addition definiert ist, d.h. eine Vorschrift, die jedem Paar (x,y) $\in$ X x X ein eindeutig bestimmtes Element x + y $\in$ X zuordnet, in der eine Multiplikation mit Skalaren definiert ist, d.h. eine Vorschrift, die jedem Paar (α,x) $\in \mathcal{K}$ x X ein eindeutig bestimmtes Element αx $\in$ X zuordnet, und in der die folgenden Axiome gelten:*

(K+) $x + y = y + x \; \forall \; x,y \in X$,
(A+) $(x + y) + z = x + (y + z) \; \forall \; x,y,z \in X$,
(0) $\exists$ *ein Element* $0 \in X$ *(Nullelement:* $x + 0 = 0 + x = x \; \forall \; x \in X$,
(-) $\forall \; x \in X \; \exists \; y \in X$: $x + y = y + x = 0$,
(1)* $1 * x = x \; \forall \; x \in X$,
($\alpha\beta$) $\alpha(\beta x) = (\alpha\beta)x \; \forall \alpha, \beta \in \mathcal{K} \; \forall \; x \in X$,
(α+β) $(\alpha+\beta)x = \alpha x + \beta x \; \forall \alpha, \beta \in \mathcal{K}, \; \forall \; x \in X$,
(x+y) $\alpha(x+y) = \alpha x + \alpha y \; \forall \alpha \in \mathcal{K} \; \forall \; x,y \in X$.

Bemerkung:
Lineare Räume werden manchmal auch *Vektorräume* oder lineare Vektorräume genannt. Der Körper $\mathcal{K}$ heißt *Skalarkörper* des linearen Raumes.

Ist $\mathcal{K} = \mathcal{R}$, so nennt man X einen *reellen linearen Raum*, ist $\mathcal{K} = \mathcal{C}$, so heißt X *komplexer linearer Raum*.

Die Elemente von X nennt man häufig *Vektoren*, die aus $\mathcal{K}$ heißen *Skalare*.

1. Beispiel:
Die Menge V^3 aller freien Vektoren im Raum ist ein linearer Raum mit der üblichen Addition von Vektoren und der üblichen Multiplikation von Vektoren mit Zahlen. Das Nullelement ist der Nullvektor.

2. Beispiel:
Die Menge $\mathcal{K}^n$ aller geordneten n-Tupel (x_1,...,x_n) mit x_1,...,x_n ist ein linearer Raum über $\mathcal{K}$ mit
(x_1,...,x_n) + (y_1,...,y_n),
$\alpha(x_1,...,x_n) := (\alpha x_1,...,\alpha x_n)$
Das Nullelement ist (0,...,0). Haben insbesondere
$\mathcal{R} = \mathcal{R}^1, \mathcal{R}^2, \mathcal{R}^3, \mathcal{R}^n$ (über $\mathcal{R}$) und
$\mathcal{C} = \mathcal{C}^1, \mathcal{C}^2, \mathcal{C}^n$ (sowohl über $\mathcal{R}$ als auch über $\mathcal{C}$).

3. Beispiel:
Die Menge F[a,b] aller Funktionen f: [a,b] $\rightarrow \mathcal{R}$ ist ein linearer Raum über $\mathcal{R}$ mit
(f+g)(x):= f(x) + g(x)

$(\alpha f)(x) := \alpha f(x)$.
Das Nullelement ist die Funktion f mit $f(x) = 0 \; \forall \; x \in [a,b]$.

4. Beispiel:
Die Menge $K_n[x]$ aller Polynome der Form $a_0+a_1x+\ldots+a_nx^n$ bildet einen linearen Raum über $\mathcal{K}$ mit
$(a_0+a_1+\ldots+a_nx^n)+(b_0+\ldots+b_nx^n) := (a_0+b_0)+\ldots+(a_n+b_n)x^n$,
$\alpha(a_0+\ldots+a_nx^n) := \alpha a_0 + \ldots + \alpha a_n x^n$
Das Nullelement ist das Polynom $0 + 0x + \ldots + 0\, x^n$.

5. Beispiel:
Die Menge $\mathcal{K}^\infty$ aller Folgen $(x_1,x_2,x_3,\ldots)$ mit
$(x_1,x_2,\ldots)+(y_1,y_2,\ldots) := (x_1+y_1, x_2, y_2+\ldots)$
$\alpha(x_1,x_2,\ldots) := (\alpha x_1, \alpha x_2, \ldots)$

6. Beispiel:
Die Menge aller Lösungen $(x_1,\ldots,x_n)$ des Gleichungssystems

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n &= 0 \end{aligned}$$

mit $a_{ij} \in \mathcal{K}$ ist mit den Operationen wie im 2. Beispiel ein linearer Raum über $\mathcal{K}$. Es reicht zu zeigen, dass Summe und Vielfaches von Lösungen wieder Lösungen sind und dass (0,...,0) eine Lösung ist. Axiome folgen aus Beispiel 2.

So ist also z.B. $\{(x,y) \in \mathcal{R}^2 : x + y = 0\}$ ein linearer Raum.
Keine linearen Räume sind hingegen $\{(x,y) \in \mathcal{R}^2 : x^2 + y^2 = 1\}$,
$\{(x,y) \in \mathcal{R}^2 : x + y = 1\}$.

Einige einfache Eigenschaften von linearen Räumen

(a) Eindeutigkeit der Null
Es gibt genau ein Element $0 \in X$ mit $a + 0 = 0 + a = a \; \forall \; a \in X$.

Beweis: Existenz folgt aus Axiom. Zum Beweis der Eindeutigkeit nehmen wir an, dass es zwei Nullelemente 0_1 und 0_2 gibt. Dann ist $0_1 = 0_1 + 0_2 = 0_2 + 0_1 = 0_2$.

(b) Eindeutigkeit des negativen Elements
Zu jedem $a \in X$ gibt es genau ein $b \in X$ mit $a + b = b + a = 0$.

Beweis: Existenz folgt aus Axiomen. Seien b_1,b_2 zwei Elemente mit obiger Eigenschaft, dann gilt:
$b_1 = b_1 + 0 = b_1 + (a + b_2) = (b_1 + a) + b_2 = 0 + b_2 = b_2$.

(c) Es gilt $0x = 0 \; \forall \; x \in X$
$\alpha 0 = 0 \; \forall \alpha \in \mathcal{K}$

Beweis: $0x = (0 + 0)x = 0x + 0x$
Es existiert ein Element b mit $0x + b = b + 0x = 0$, und dies ergibt
$0 = 0x + b = (0x + 0x) + b = 0x + (0x + b) = 0x$.

Gleichung $\alpha 0 = 0$ zeigt man analog.

Das nach (b) eindeutig bestimmte Element b wird mit -a bezeichnet. Subtraktion in einem linearen Raum wird dann durch x - y := x + (-y) definiert.

(d) Es gilt (-1)x = -x $\forall$ x $\in$ X.

Beweis: Haben 1x = x (Axiom) und damit
0 = 0x = (1 + (-1))x = 1x + (-1)x = x + (-1)x und somit ist (-1)x = -x.

(e) Für beliebige a,b $\in$ X existiert genau ein x $\in$ X mit a + x = b.

Beweis: Sei a + x = b. Dann folgt
x = 0 + x = (-a + a) + x = -a + (a + x) = -a + b,
d.h. erhalten Eindeutigkeit. Andererseits ist
a + (-a + b) = (a - a) + b = 0 + b = b, d.h. x = -a + b ist eine Lösung.

Nach (d) ist die eindeutige Lösung von a + x = b gegeben durch
x = b - a = b + (-1)a.

Definition 2 *Seien X und Y lineare Räume mit dem gleichen Skalarkörper $\mathcal{K}$. Eine Abbildung A: X $\rightarrow$ Y heißt linearer Operator (= Vektorraumhomomorphismus), wenn gilt*
A(x+y) = A(x) + A(y) $\forall$ x,y $\in$ X,
A(αx) = α A(x) $\forall \alpha \in \mathcal{K} \; \forall$ x $\in$ X.

Bei linearen Abbildungen schreibt man Ax statt A(x).

7. Beispiel:
(1) A: $\mathcal{R}^2 \rightarrow \mathcal{R}^2$, A(x,y) = (x,-y)

A($(x_1,y_1)+(x_2,y_2)$) = A(x_1+x_2,y_1+y_2) = (x_1+x_2,-(y_1+y_2)) = (x_1-y_1)+(x_2,-y_2) = A(x_1,y_1)+A(x_2,y_2)

analog A(α(x,y)) = α A(x,y)
Also ist A linear: Es stellt die Spiegelung an der x-Achse dar.

(2) A: $\mathcal{R}^2 \rightarrow \mathcal{R}^2$, A(x,y) = (-y,x)
$\rightarrow$ Drehung um 90° entgegengesetzt zum Uhrzeigersinn

(3) A: $\mathcal{R}^2 \rightarrow \mathcal{R}^2$, A(x,y) = (x,0)
$\rightarrow$ Projektion auf die x-Achse

(4) A: $\mathcal{R}^2 \rightarrow \mathcal{R}^2$, A(x,y) = (3x,3y)
$\rightarrow$ Streckung um den Faktor 3

Alle diese Abbildungen sind von der Form:
A(x,y)=(αx+βy, γx+δy)

Jeder Abbildung dieser Form ist linear:
A($(x_1,y_1)+(x_2,y_2)$)=A(x_1+x_2,y_1+y_2)=($\alpha(x_1+x_2)+\beta(y_1+y_2),\gamma(x_1+x_2)+\delta(y_1+y_2)$)

=A(x_1,y_1) + A(x_2,y_2),
A(μ(x,y)) = A(μx,μy)
= ($\alpha\mu$x + $\beta\mu$y,$\gamma\mu$x+$\delta\mu$y)
= μ(αx+βy,γx+δy)
= μ A(x,y).

8. Beispiel:
Eine Abbildung A: $\mathcal{R} \to \mathcal{R}$ ist genau dann linear, wenn es ein a $\in \mathcal{R}$ mit Ax = ax $\forall$ x $\in \mathcal{R}$ gibt.
Beweis: Offenbar ist Ax = ax linear:
A(x+y) = a(x+y) = ax + ay = Ax + Ay,
A(αx) = a(αx) = α(ax) = αAx.
Sei nun A linear. Setzten a = A1. Für beliebiges x $\in \mathcal{R}$ ist dann
Ax = A(x*1) = x A1 = xa = ax.

9. Beispiel:
Für einen beliebigen Operator gilt stets A0 = 0.
Dies (oder Beispiel 8) impliziert, dass eine Abbildung A: $\mathcal{R} \to \mathcal{R}$ der Form Ax = ax + b genau dann linear ist, wenn b = 0 ist. Eine Translation T: $\mathcal{R}^n \to \mathcal{R}^n$, Tx = x + b ist genau dann linear, wenn b = 0 ist, d.h. wenn T die identische Abbildung ist.

Wie in Beispiel 7 kann man vielleicht sehen, dass für beliebige Zahlen $a_{ij} \in \mathcal{R}$ die Abbildung A: $\mathcal{R}^n \to \mathcal{R}^m$, $(x_1,\ldots,x_n) \to (a_{11}x_1+\ldots+a_{1n}x_n,\ldots,a_{m1}x_1+\ldots+a_{mn}x_n)$ linear ist. Man kann zeigen, dass jeder lineare Operator A: $\mathcal{R}^n \to \mathcal{R}^m$ von dieser Form ist. Die Theorie des linearen Gleichungssystems ist also äquivalent zur Theorie der Gleichungen der Form

$$Ax = b$$

mit linearen Operatoren A: $\mathcal{R}^n \to \mathcal{R}^m$.

Vermerken ausdrücklich, dass solche Operatoren wie
A: $\mathcal{R} \to \mathcal{R}$, Ax = x^2,
A: $\mathcal{R}^2 \to \mathcal{R}^2$ A(x,y) = sin (x+y),
A: $\mathcal{R}^2 \to \mathcal{R}^2$ A(x,y) = (e^x,e^y)
nicht linear sind.

Definition 3 *Seien X und Y lineare Räume über demselben Skalarkörper $\mathcal{K}$. Mit $\mathcal{L}(X,Y)$ bezeichnet man die Menge aller linearer Operatoren von X nach Y. Für A,B $\in \mathcal{L}(X,Y)$ und $\alpha \in \mathcal{K}$ definiert man die Summe A + B und das skalare Vielfache αA durch (A+B)(x):=Ax+Bx, (αA)(x):=α(Ax).*
Das Produkt AB zweier Operatoren A $\in \mathcal{L}(Y,Z)$, B $\in \mathcal{L}(X,Y,)$ ist definiert durch
(AB)(x):= A(Bx)

Satz 1 *Die Menge $\mathcal{L}(X,Y)$ mit der Additon A + B und dem skalaren Vielfachen αA ist ein linearer Raum.*

Beweis:
Man muss zeigen:
A,B $\in \mathcal{L}$(X,Y) $\Rightarrow$ A + B, αA $\in \mathcal{L}$(X,Y).
So ist z.B.
(A+B(x+y)=A(x+y)+B(x+y)=Ax+Ay+Bx+By=Ax+Bx+Ay+By=(A+B)x + (A+B)y.
Man überprüft leicht, dass alle Axiome eines linearen Raumes gelten. Nullelement ist der *Nulloperator* A: X $\rightarrow$ Y, Ax = 0 $\forall$ x $\in$ X.

Satz 2 *Ist A $\in \mathcal{L}$(Y,Z) und B $\in \mathcal{L}$(X,Y), so ist AB $\in \mathcal{L}$(X,Z).*

Beweis: Haben
(AB)(x+y)=A(B(x+y))=A(Bx+By)=A(Bx) + A(By) = (AB)(x) + (AB)(y),
analog zeigt man (AB)(αx)=α(AB(x)).

Bemerkung:
Man setzt $\mathcal{L}$(X) := $\mathcal{L}$(X,X). In $\mathcal{L}$(X) sind Summe, skalare Vielfache und Produkt definiert. Betrachtet man bloß Summe und Produkt, so entsteht die Frage, wann $\mathcal{L}$(X) ein Körper ist.
Es zeigt sich, dass dies im Wesentlichen nur für X = $\mathcal{K}$ der Fall ist. In allen anderen Situationen ist $\mathcal{L}$(X) nicht mal ein Schiefkörper.

Definition 4 *Zwei lineare Räume X und Y über demselben Skalarkörper $\mathcal{K}$ heißen isomorph, wenn es einen bijekten linearen Operator A: X $\rightarrow$ Y gibt. Jeder solche Operator heißt dann Isomorphismus (oder Vektorraumisomorphismus.*

In der Mengenlehre werden zwei Mengen als gleich angesehen, wenn sie gleichmächtig sind, d.h. wenn eine bijektive Abbildung zwischen ihnen existiert. In der Algebra betrachtet man Mengen mit algebraischen Strukturen und in der Analysis geht es um Mengen mit topologischen Strukturen, und zwei solche strukturierte Mengen sieht man als gleich an, wenn es eine bijektive Abbildung zwischen ihnen gibt, die die Struktur invariant lässt.

Satz 3 *Isomorphie $\cong$ von linearen Räumen ist eine Äquivalenzrelation.*

Beweis:
(R) A: X $\rightarrow$ X, Ax = x ist bijektiver linearer Operator, d.h. X $\cong$ X für jeden linearen Raum.
(S) Sei X $\cong$ Y. Dann existiert ein bijektiver linearer Operator A: X $\rightarrow$ Y. Die bijektive Abbildung A besitzt eine bijektive Umkehrabbildung A^{-1}: Y $\rightarrow$ X. Zeigen, dass A^{-1} ein linearer Operator ist.
Seien dazu y_1,y_2 $\in$ Y. Dann gibt es eindeutig bestimmte x_1,x_2 $\in$ X mit Ax_1 = y_1, Ax_2 = y_2 und es gilt

$$A^{-1}(y_1 + y_2) = A^{-1}(Ax_1 + Ax_2) = A^{-1}(A(x_1 + x_2)) = x_1 + x_2 = A^{-1}y_1 + A^{-1}y_2.$$

und analog $A^{1}(\alpha y) = \alpha A^{-1} y$.
Also ist A^{-1}: Y → X ein Isomorphismus.
(T) Sei X ≅ Y, Y ≅ Z. Dann gibt es bijektive lineare Operatoren A: X → Y, B: Y → Z. Dann ist BA: X → Z bijektiv (Hinterinanderausführung zweier bijektiver Abbildungen ist immer bijektiv) und nach Satz 2 ist BA linear. Also ist BA ein Isomorphismus, d.h. X ≅ Z.

10. Beispiel
$V^3 \cong \mathcal{R}^3$
Ein Isomorphismus ist A: $V^3 \to \mathcal{R}^3$, A(xi+yj+zk)=(x,y,z).
Analog ist $\mathcal{K}_n[x] \cong \mathcal{K}^{n+1}$.
Ein Isomorphismus ist
$A(a_0+a_1x+\ldots+a_nx^n) = (a_0,a_1,\ldots,a_n)$
Obige Bemerkung lässt sich wie folgt präzisieren:
$\mathcal{L}(X)$ ist ein Körper ⇔ X ≅ $\mathcal{K}$.

11. Beispiel
$\mathcal{R}^3$ und $\mathcal{R}^2$ sind nicht isomorph.
Nehmen an, dass A: $\mathcal{R}^3 \to \mathcal{R}^2$ ein Isomorphismus ist.
Setzen
$A(1,0,0) = (a_1,a_2)$, $A(0,1,0)=(b_1,b_2)$, $A(0,0,1)=(c_1,c_2)$.
Dann ist:

$$\begin{aligned} A(x,y,z) &= A(x(1,0,0) + y(0,1,0) + z(0,0,1) \\ &= xA(1,0,0) + yA(0,1,0) + zA(0,0,1) \\ &= x(a_1,a_2) + y(b_1,b_2) + z(c_1,c_2) \\ &= (a_1x + b_1y + c_1z, a_2x + b_2y + c_2z) \end{aligned}$$

Betrachten die (x,y,z), die auf (0,0) abgebildet werden:
$(0,0) \Leftrightarrow (x,y,z) \in \{(x,y,z) \in \mathcal{R}^3 : a_1x+b_1y+c_1z=0\} \cap \{(x,y,z) \in \mathcal{R}^3 : a_2x+b_2y+c_2z=0\}$

Gleichung A(x,y,z) = (0,0) wird also mindestens für alle Punkte auf einer Geraden erfüllt. Damit ist A nicht injektiv. Widerspruch!

Gleichmäßigkeit von Mengen wird durch die Kardinalzahl charakterisiert:
Zwei Mengen werden genau dann als gleich angesehen (⇔ liegen in derselben Äquivalenzklasse bezüglich der Gleichmächtigkeit), wenn sie die gleiche Kardinalzahl haben.
Suchen eine ähnliche Charakteristik für die Isomorphie von linearen Räumen, d.h. eine Eigenschaft, sodass gilt:

Zwei lineare Räume sind isomorph (⇔ liegen in derselben Äquivalenzklasse bezüglich Isomorphie), wenn jene Eigenschaft bei beiden erfüllt ist, d.h. wenn ihre Charakteristiken übereinstimmen.

Wir werden sehen, dass diese Charakteristik die Dimension ist.

2.2 Basen und Dimensionen

Im folgenden sei X ein linearer Raum über $\mathcal{K}$.

Definition 1 *Eine endliche Menge $x_1,\ldots,x_n$ von Elementen aus X heißt linear abhängig, wenn es Zahlen $\alpha_1, \ldots, \alpha_n \in \mathcal{K}$ gibt, sodass wenigstens eine der Zahlen $\alpha_1, \ldots, \alpha_n$ von Null verschieden ist und $\alpha_1 x_1 + \ldots + \alpha_n x_n = 0$ gilt.*

Sind $x_1, \ldots, x_n$ nicht linear abhängig, so nennt man sie linear unabhängig.

Es gilt also folgendes: Die Elemente $x_1, \ldots, x_n$ sind linear unabhängig genau dann, wenn gilt

$$\alpha_1 x_1 + \ldots + \alpha_n x_n = 0 \Rightarrow \alpha_1 = \ldots = \alpha_n = 0$$

Die Elemente $x_1,\ldots,x_n$ sind linear abhängig genau dann, wenn gilt

$$\exists(\alpha_1, \ldots, \alpha_n) \neq (0, \ldots, 0) : \alpha_1 x_1 + \ldots + \alpha_n x_n = 0.$$

$$x_1 = -\frac{\alpha_2}{\alpha_1} x_2 - \ldots - \frac{\alpha_n}{\alpha_1} x_n$$

09.12.05

1. Beispiel:
Zwei Vektoren a, b $\in V^3$ sind genau dann linear abhängig, wenn sie gleiche oder entgegengesetzte Richtung haben. In der Tat, ist $\alpha a + \beta b = 0$ mit $(\alpha,\beta)\neq(0,0)$, so ist $a = \frac{-\beta}{\alpha} b$ $(\alpha \neq 0)$ oder $b = \frac{-\alpha}{beta} a$ $(\beta \neq 0)$, und haben umgekehrt a, b die gleiche bzw. entgegengesetzte Richtung, so ist $a = \lambda b$ $(\lambda \geq 0)$ oder $b = -\lambda a$ $(\lambda < 0)$, d.h. $b - \lambda a = 0$ bzw. $b + \lambda a = 0$.

Drei Vektoren a,b,c $\in V^3$ sind genau dann linear abhängig, wenn sie *komplanar* sind, d.h. wenn sie (präziser: die entsprechenden Ortsvektoren) in einer Ebene liegen.
Vier Vektoren aus V^3 sind stets linear abhängig.

Definition 2 *Sind $x_1,\ldots,x_n$ und $\alpha_1,\ldots,\alpha_n \in \mathcal{K}$, so heißt $\alpha_1 x_1 + \ldots + \alpha_n x_n$ Linearkombination von $x_1,\ldots,x_n$.*

2. Beispiel:
Ein einzelnes Element x $\in$ X ist genau dann linear unabhängig, wenn x $\neq 0$ ist.
Ist x = 0, so ist 1x = x = 0.
Umgekehrt sei x linear unabhängig, dann $\exists\alpha \neq 0$: αx=0. Daraus folgt

$$\frac{1}{\alpha} \underbrace{\alpha x}_{=0} = (\frac{1}{\alpha}\alpha)x = 1x = x = 0.$$

3. Beispiel:
Zwei Elemente (a,b), (c,d) $\in \mathcal{R}^2$ sind genau dann linear abhängig, wenn es (x,y)$\neq$(0,0) mit x(a,b) + y(c,d) = (0,0) gibt. Dies ist genau dann der Fall, wenn das Gleichungssystem

$$\begin{array}{lllll} ax & + & cy & = & 0 \\ bx & + & dy & = & 0 \end{array}$$

eine nichttriviale Lösung (x,y) besitzt. Wissen, dass dies zu $|\,a \quad cb \quad d\,| = ad - bc = 0$ äquivalent ist.

4. Beispiel:
Die drei Funktionen sin x, cos x, sin (x+1) sind als Elemente von F[0,2π] linear abhängig.
In der Tat,
sin(x+1) = sin x cos 1 + cos x sin 1,
d.h. (cos 1)sin x + (sin 1)cos x + $\underbrace{(-1)}_{\neq 0}$ sin (x+1) = 0.
Die Funktionen sin x und cos x sind hingegen linear unabhängig.
$\alpha \sin x + \beta \cos x = 0 \forall x \in [0, 2\pi$
Setzen x = 0 $\Rightarrow \beta = 0$,
Setzen x = $\frac{\pi}{2} \Rightarrow \alpha = 0$.

5. Beispiel:
Die Elemente e_1 = (1,0,0,...), e_2=(0,1,0,0,...), ..., e_n = (0,...,0,1,0,...) von $\mathcal{R}^\infty$ sind für beliebiges n $\geq$ 1 linear unabhängig.
$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n = 0 = (0, 0, \ldots)$
$\Rightarrow \alpha_1$(1,0,...)+ α_2(0,1,0,...)+...+α_n(0,...,0,1,0,...) = (0,0,...)
$\Rightarrow$ (α_1,α_2,...,α_n) = (0,0,...,0)

Definition 3 *Sei X ein linearer Raum über $\mathcal{K}$. Man nennt X n-dimensional (über $\mathcal{K}$) und schreibt dim X = n oder $dim_{\mathcal{K}}$ X = n, wenn es in X n linear unabhängige Elemente gibt und beliebige n+1 Elemente aus X stets linear abhängig sind.*
Der Raum heißt unendlichdimensional dim X = ∞, wenn sich für jedes n in X n linear unabhängige Elemente finden lassen.
Der Raum X = {0} hat nach Definition die Dimension Null.

6. Beispiel:
Haben dim $V^3 = 3$, denn es gibt 3 linear unabhängige Vektoren (z.B. i,j,k) und beliebige 4 Vektoren sind linear abhängig. Aus Beispiel 5 folgt, dass dim $\mathcal{K}^\infty = \infty$ ist.

7. Beispiel:

$$\boxed{dim_{\mathcal{R}} R = 1}$$

Element 1 $\in \mathcal{R}$ ist linear unabhängig (siehe Beispiel 2).
Sind x,y $\in \mathcal{R}$, so ist $(-\frac{y}{x})x + y = 0$ für x $\leq$ 0 und $x + (-\frac{x}{y})y = 0$ für y *neq* 0.

$$\boxed{dim_{\mathcal{R}} R^2 = 2}$$

Elemente (1,0), (0,1) sind linear unabhängig.
α(1,0) + β(0,1) = (0,0)
$\Rightarrow$ (α,0) + (0,β) = (0,0) $\Rightarrow$ (α,β) = (0,0)

Seien (x_1,y_1), (x_2,y_2), (x_3,y_3) drei beliebige Elemente aus $\mathcal{R}^2$: Suchen α,β,$\gamma \in \mathcal{R}$ mit

$$\alpha(x_1, y_1) + \beta(x_2, y_2) + \gamma(x_3, y_3) = (0, 0)$$

Das gilt genau dann, wenn

$$\alpha x_1 + \beta x_2 + \gamma x_3,$$

$$\alpha y_1 + \beta y_2 + \gamma y_3,$$

d.h. (α,β,γ) $\in$ {(u,v,w): x_1u+x_2v+x_3w = 0} $\cap$ {(u,v,w): y_1u+y_2v+y_3w = 0}
und der Durchschnitt enthält mindestens eine Gerade, d.h. (α,β,γ) $\neq$ (0,0,0).

$$\boxed{dim_{\mathcal{C}} C = 1}$$

$1 \in \mathcal{C}$ ist linear unabhängig.
Sind z,w $\in \mathcal{C}$, so ist
$(-\frac{w}{z})z + w = 0$ für z $\neq$ 0 und
$z + (-\frac{z}{w})w = 0$ für w $\neq$ 0.

$$\boxed{dim_{\mathcal{R}} C = 2}$$

Elemente 1, i $\in \mathcal{C}$ sind linear unabhängig:
α 1 + β i = 0 mit $\alpha, \beta \in \mathcal{R}$.
$\Rightarrow \alpha = 0$, $\beta = 0$.

Sind $x_1 + y_1 i$, $x_2 + y_2 i$, $x_3 + y_3 i$ drei beliebige komplexe Zahlen, so zeigt man wie oben für den $\mathcal{R}^2$, dass diese linear abhängig sind.

Definition 4 *Jedes System von n linear unabhängigen Elementen eines n-dimensionalen linearen Raumes heißt Basis des Raumes.*

8. Beispiel:
Die Standardvektoren i,j,k bilden eine Basis im V^3.
Die beiden Elemente e_1 = (1,0) und e_2 = (0,1) bilden eine Basis im $\mathcal{R}^2$. Nach Beispiel 3 bilden (a,b) und (c,d) eine Basis im $\mathcal{R}^2$ genau dann, wenn ad - bc $\neq$ 0 ist.
Eine Basis in $\mathcal{C}$ über $\mathcal{C}$ ist {1}, eine Basis in $\mathcal{C}$ über $\mathcal{R}$ ist {1,i}.

Satz 1 *Sei X ein n-dimensionaler linearer Raum über $\mathcal{K}$ und $\{e_1,\ldots,e_n\}$ eine Basis in X. Dann lässt sich jedes $x \in X$ eindeutig in der Form $x = x_1e_1 + \ldots + x_ne_n$ mit $x_1,\ldots,x_n \in \mathcal{K}$ darstellen.*

Beweis:
Wegen dim X = n sind die n+1 Elemente x, $e_1,\ldots,e_n$ linear abhängig. Also gibt es $\alpha_0,\ldots,\alpha_n \in \mathcal{K}$ mit $(\alpha_0,\ldots,\alpha_n) \neq (0,\ldots,0)$ und $\alpha_0 x + \alpha_1 e_1 + \ldots + \alpha_n e_n = 0$.

Haben $\alpha_0 \neq 0$, da ansonsten $\alpha_1 e_1 + \ldots + \alpha_n e_n = 0$ mit $(\alpha_1,\ldots,\alpha_n) \neq (0,\ldots,0)$ wäre und somit $e_1,\ldots,e_n$ linear abhängig wären. Also ist

$$x = (-\frac{\alpha_1}{\alpha_n})e_1 + \ldots + (-\frac{\alpha_n}{\alpha_0})e_n$$

eine Darstellung der gewünschten Form.
Zum Beweis der Eindeutigkeit sei

$$x_1e_1 + \ldots + x_ne_n = y_1e_1 + \ldots + y_ne_n.$$

Dann ist

$$(x_1 - y_1)e_1 + \ldots + (x_n - y_n)e_n = 0,$$

und da $e_1,\ldots,e_n$ linear unabhängig sind, folgt

$$x_1 - y_1 = 0, \ldots, x_n - y_n = 0.$$

#

Definition 5 *Die durch Satz 1 eindeutig bestimmten Zahlen $x_1,\ldots,x_n \in \mathcal{K}$ heißen Koordinaten von x in der Basis $\{e_1,\ldots,e_n\}$*

Haben x und y die Koordinaten $(x_1,\ldots,x_n)$ und $(y_1,\ldots,y_n)$, so haben x + y, x - y, αx, 0 die Koordinaten

$$(x_1 + y_1, \ldots, x_n + y_n)$$
$$(x_1 - y_1, \ldots, x_n - y_n)$$
$$(\alpha x_1, \ldots, \alpha x_n)$$
$$(0, \ldots, 0).$$

Der folgende Satz ist die Umkehrung von Satz 1.

Satz 2 *Sei X ein linearer Raum über $\mathcal{K}$ und seinen $e_1,\ldots,e_n$ Elemente aus X. Wenn sich jedes $x \in X$ eindeutig in der Form $x = x_1e_1 + \ldots + x_ne_n$ mit $x_1,\ldots,x_n \in \mathcal{K}$ darstellen lässt, dann ist dim X = n und $\{e_1,\ldots,e_n\}$ eine Basis in X.*

Lemma: Für m < n hat das Gleichungssystem

$$a_{11}x_1 + \ldots + a_{1n}x_n = 0$$

$$\vdots$$

$$a_{m1}x_1 + \ldots + a_{mn}x_n = 0$$

(m Gleichungen mit n Unbekannten) stets eine nichttriviale Lösung.

12.12.05

Beweis:
Beweisen das Lemma durch vollständige Induktion nach m.
Für m = 1 (n ≥ 2) ist die Behauptung offensichtlich. Sei die Behauptung für m = k bewiesen. Zeigen Sie für m = k+1, d.h. für das System (*):

$$\begin{array}{lcl} a_{11}x_1 + \ldots + a_{1n}x_n & = & 0 \\ \vdots & & \\ a_{k1}x_1 + \ldots + a_{kn}x_n & = & 0 \\ a_{k+1,1}x_1 + \ldots + a_{k+1,n}x_n & = & 0 \end{array}$$

Betrachten die letzte Gleichung:
Ist $a_{k+1,1} = \ldots = a_{k+1,n} = 0$, so folgt die Behauptung aus der Induktionsvoraussetzung.

Sei also eine der Zahlen $a_{k+1,1}, \ldots, a_{k+1,n}$ von Null verschieden, etwa $a_{k+1,1} \neq 0$.
Nach Induktionsvoraussetzung existiert eine nichttriviale Lösung von

$$a_{11}(-\frac{a_{k+1,2}}{a_{k+1,1}}x_2 - \ldots - \frac{a_{k+1,n}}{a_{k+1,1}}x_n) + a_{12} + \ldots + a_{1n}x_n = 0$$

$$\vdots$$

$$a_{k1}(-\frac{a_{k+1,2}}{a_{k+1,1}}x_2 - \ldots - \frac{a_{k+1,n}}{a_{k+1,1}}x_n) + a_{k2} + \ldots + a_{kn}x_n = 0$$

(Dies ist ein homogenes Gleichungssystem mit k Gleichungen und n-1 > k Unbekannten)

Setzen wir

$$x_1 = -\frac{a_{k+1,2}}{a_{k+1,1}}x_2 - \ldots - \frac{a_{k+1,n}}{a_{k+1,1}}x_n,$$

so erhalten wir eine nichttriviale Lösung vom Ausgangssystem (*).
#
Beweis von Satz 2:
Haben $0 = 0e_1 + \ldots + 0e_n$ und aus Eindeutigkeit der Darstellung folgt, dass $0 = \alpha_1 e_1 + \ldots + \alpha_n e_n$ nur mit $\alpha_1 = \ldots = \alpha_n = 0$ möglich ist.
Also sind $e_1, \ldots, e_n$ linear unabhängig.

Seien $x_0, x_1, \ldots, x_n$ beliebige n+1 Elemente aus X. Haben dann die Darstellungen

$$x_j = \xi_j^{(1)}e_1 + \ldots + \xi_j^{(n)}e_n$$

mit $\xi_j^{(i)} \in \mathcal{K}$.

Dann ist

$$\alpha_0 x_0 + \ldots + \alpha_n x_n = \alpha_0(\xi_0^{(1)} e_1 + \ldots + \xi_0^{(n)} e_n) + \ldots + \alpha_n(\xi_n^{(1)} e_1 + \ldots + \xi_n^{(n)} e_n)$$

$$= (\xi_0^{(1)}\alpha_0 + \ldots + \xi_n^{(1)}\alpha_n)e_1 + \ldots + (\xi_0^{(n)}\alpha_0 + \ldots + \xi_n^{(n)}\alpha_n)e_n$$

Nach dem Lemma gibt es (α_0,*ldots*,α_n) $\neq$ (0,...,0) mit der Eigenschaft, dass

$$\xi_0^{(1)}\alpha_0 + \ldots + \xi_n^{(1)}\alpha_n = 0$$

$$\vdots$$

$$\xi_0^{(n)}\alpha_0 + \ldots + \xi_n^{(n)}\alpha_n = 0$$

gilt. [n homogene mit n+1 Unbekannten α_0, ..., α_n]
Für diese α_0,...,α_n ist dann $\alpha_0 x_0 + \ldots + \alpha_n x_n = 0$.
Also ist dim X = n und $\{e_1,\ldots,e_n\}$ eine Basis in X.
#

Satz 3 *Isomorphiesatz für endlichdimensionale lineare Räume I*
Jeder n-dimensionale lineare Raum über $\mathcal{K}$ ($n \in \mathcal{N}$) ist zu $\mathcal{K}^n$ isomorph, d.h.
$dim_{\mathcal{K}} X = n \Rightarrow X \cong \mathcal{K}^n$.

Beweis:
Wählene eine Basis $\{e_1,\ldots,e_n\}$ in X. Nach Satz 1 lässt sich jedes x $\in$ X eindeutig in der Form x = $x_1e_1 + \ldots + x_ne_n$ darstellen.
Definieren Abbildung A durch
A: X $\rightarrow \mathcal{K}^n$, Ax = $(x_1,\ldots,x_n)$.

Offenbar ist A ein linearer Operator.
Sei Ax = Ay. Dann ist $(x_1,\ldots,x_n) = (y_1,\ldots,y_n)$, d.h. x_1=y_1,...,x_n=y_n, und somit x = $x_1e_1 + \ldots + x_ne_n = y_1e_1 + \ldots + y_ne_n$ = y. Also ist A injektiv.
Wegen A($x_1e_1 + \ldots + x_ne_n$) = $(x_1,\ldots,x_n)$ ist A auch surjektiv.
Also ist A ein bijektiver linearer Operator und damit ein Isomorphismus.
#

Satz 4 *Isomorphiesatz für endlichdimensionale lineare Räume II*
Zwei endlichdimensionale Räume über dem gleichen Skalarkörper sind genau dann isomorph, wenn sie die gleiche Dimension haben, d.h.
dim X = dim $Y \Leftrightarrow X \cong Y$

Beweis:
„$\Rightarrow$“: Folgt aus Satz 3 und der Tatsache, dass Isomorphie eine Äquivalenzrelation ist.
„$\Leftarrow$“: Sei A: X $\rightarrow$ Y ein Isomorphismus und dim X = n. Seien $y_0,\ldots,y_n$ beliebige n+1 Elemente aus Y. Da A surjektiv ist, gibt es $x_0,\ldots,x_n \in$ X mit

$$Ax_0 = y_0, \ldots, Ax_n = y_n$$

Weil A linear ist, haben wir

$$\alpha_0 y_0 + \ldots + \alpha_n y_n = \alpha_0 A x_0 + \ldots + \alpha_n A x_n$$

$$= A(\alpha_0 x_0 + \ldots + \alpha_n x_n).$$

Wegen dim X = n gibt es (α_0,...,α_n) $\neq$ (0,...,0) mit

$$\alpha_0 x_0 + \ldots + \alpha_n x_n = 0.$$

Für diese α_0, ..., α_n ist also $\alpha_0 y_0 + \ldots + \alpha_n y_n$ = A0 = 0.
Somit sind beliebige n+1 Elemente aus Y linear abhängig. Daraus folgt:
dim Y $\leq$ n.
Haben also dim Y $\leq$ dim X gezeigt. Betrachtet man A^{-1}: Y $\to$ X, so ergibt sich analog dim X $\leq$ dim Y. Also ist dim X = dim Y.
#

Folgerung: Es gilt $dim_{\mathcal{R}} R^n$ = n, $dim_{\mathcal{C}} C^n$ = n, $dim_{\mathcal{R}} C^{2n}$ = 2n.

1. Beweis:
Da $\mathcal{K}^n$ zu $\mathcal{K}^n$ isomorph ist, folgt aus den Sätzen 3 und 4, dass $dim_{\mathcal{K}} K^n$ = n ist.
Da A: $\mathcal{C}^n \to \mathcal{R}^{2n}$,
A ($x_1 + iy_1$,...,$x_n + iy_n$) = (x_1,y_1,...,x_n,y_n) ein Isomorphismust ist, folgt wieder aus den Sätzen 3 und 4, dass $dim_{\mathcal{R}} C^n = dim_{\mathcal{R}} R^{2n}$ = 2n gilt.
#
2. Beweis:
Setzen e_1 = (1,0,...,0),..., e_n = (0,...,0,1). Dann lässt sich jedes x $\in \mathcal{K}^n$ eindeutig in der Form x = $x_1 e_1 + \ldots + x_n e_n$ mit $x_j \in \mathcal{K}$ schreiben.
Nach Satz 2 ist also $dim_{\mathcal{K}} K^n$ = n.
Des Weiteren lässt sich jedes z = $(x_1 + iy_1, \ldots, x_n + iy_n) \in \mathcal{C}^n$ eindeutig in der Form z = $x_1 e_1 + \ldots + x_n e_n + y_1(ie_1) + \ldots + y_n(ie_n)$ mit x_j, $y_j \in \mathcal{R}$ schreiben.
Das heißt: $dim_{\mathcal{R}} C^{2n}$ = 2n nach Satz 2.
#
Charakteristik der linearen Räume: Dimension

→ Menge der linearen Räume in Äquivalenzklassen zerlegt:
→ endlichdimensional (0,1,2,...) oder unendlichdimensional
→ vergleichbar mit den Kardinalzahlen der Mengen

2.3 Matrizen

Definition 1 *Seien m,n ∈ $\mathcal{N}$ = {1,2,3,...}. Eine m x n-Matrix ist ein rechteckiges Schema der Form*

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix},$$

wobei a_{ij} aus $\mathcal{K}$ sind. Die Zahlen a_{ij} heißen Elemente, Komponenten oder Einträge der Matrix. Eine m x n-Matrix hat m Zeilen und n Spalten. Für $\mathcal{K} = \mathcal{R}$ spricht man von einer reellen Matrix für $\mathcal{K} = \mathcal{C}$ von einer komplexen Matrix. Die Menge aller m x n-Matrizen bezeichnet man mit $M_{m,n}(\mathcal{K})$, gelegentlich auch mit $\mathcal{K}^{mxn}$.

Eine m x n-Matrix ist nichts weiter als eine Abbildung
a: {1,2,...,m} x {1,2,...,n} → $\mathcal{K}$,
bei der a_{ij} statt a((i,j)) geschrieben wird, und der gesamte Wertevorrat in einem Rechteck tabelliert ist, und zwar so, dass a_{ij} in der i-ten Zeile und der j-ten Spalte steht.

Definition 2 *Seien $A = (a_{ij})_{i,j=1}^{m,n} \in M_{(m,n)}(\mathcal{K})$ und $B = (b_{ij})_{i,j=1}^{m,n} \in M_{(m,n)}(\mathcal{K})$ [Kurzschreibweise definiert sich selbst] zwei Matrizen der gleichen Größe. Die Summe A + B ist definiert als die m x n-Matrix $(a_{ij} + b_{ij})_{i,j=1}^{m,n}$, d.h.*

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} := \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \dots & a_{mn} + b : mn \end{pmatrix}$$

Für $\alpha \in \mathcal{K}$ wird das skalare Vielfache αA definiert als die m x n-Matrix $(\alpha a_{ij})_{i,j=1}^{m,n}$, d.h.

$$\alpha \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} := \begin{pmatrix} \alpha a_{11} & \dots & \alpha a_{1n} \\ \vdots & & \vdots \\ \alpha a_{m1} & \dots & \alpha a_{mn} \end{pmatrix}.$$

Satz 1 *Die Menge $M_{m,n}(\mathcal{K})$ mit den Operationen aus Definition 2 ist ein linearer Raum der Dimension mn über $\mathcal{K}$*

Beweis:
Es ist leicht zu sehen, dass alle Axiome eines linearen Raumes erfüllt sind, das Nullele-

ment ist die *Nullmatrix* $\begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix}$ Bezeichnen mit E_{ij} die Matrix

$$E_{ij} = \begin{matrix} \\ (i) \end{matrix} \begin{matrix} \quad (j) \quad \\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{matrix},$$

d.h. die Matrix, deren (i,j)-Eintrag 1 ist und deren andere Einträge alle 0 sind.
Für eine beliebige Matrix A = $(a_{ij}) \in M_{m,n}\ (\mathcal{K})$ ist dann

$$A = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} E_{ij}$$

und diese Darstellung ist eindeutig.
Nach Satz 2 aus 2.2. hat also $M_{m,n}\ (\mathcal{K})$ die Dimension mn (= Anzahl der möglichen E_{ij}).
#

Definition 3 *Unter dem Produkt einer Matrix*
$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$ *und einer Spalte* $x = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} \in M_{n,1}\ (\mathcal{K})$ *versteht man die Spalte*

$$Ax = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix} \in M_{m,1}(\mathcal{K}).$$

[Dies ist eine Abbildung $M_{m,n}\ (\mathcal{K})$ x $M_{n,1}\ (\mathcal{K}) \to M_{m,1}\ (\mathcal{K})$].

Ein lineares Gleichungssystem lässt sich also in der Form

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

kurz Ax = b, schreiben.

Der folgende Satz beschreibt alle linearen Operatoren von $M_{n,1}\ (\mathcal{K})$ nach $M_{m,1}\ (\mathcal{K})$, d.h. er beschreibt die Menge $\mathcal{L}(M_{n,1}\ (\mathcal{K}), M_{m,1}\ (\mathcal{K}))$.
Vermerken, dass jeder n-dimensionale lineare Raum über $\mathcal{K}$ zu $\mathcal{K}^n$ isomorph ist und dass $\mathcal{K}^n \to M_{n,1}(\mathcal{K})$, $(x_1,\dots,x_n) \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$
ebenfalls ein Isomorphismus ist.
Erinnern an Beispiele 8 und 9 aus 2.1.

Satz 2 *Eine Abbildung A: $M_{n,1}$ ($\mathcal{K}$) $\to$ $M_{m,1}$ ($\mathcal{K}$) ist genau dann ein linearer Operator, wenn es eine Matrix [A] $\in$ $M_{m,n}$ ($\mathcal{K}$) mit*

$$Ax = [A]x \quad \forall x \in M_{n,1}(\mathcal{K})$$

gibt. Die Matrix [A] ist für jedes A eindeutig bestimmt. Die Abbildung
$\mathcal{L}(M_{n,1}$ ($\mathcal{K}$), $M_{m,1}$ ($\mathcal{K}$)) $\to$ $M_{m,n}$ ($\mathcal{K}$), A $\to$ [A]
ist ein Isomorphismus, d.h.
(a) Die Abbildung ist bijektiv,
(b) [A+B] = [A] + [B],
(c) [αA] = α[A].

Beweis:
Die Abbildung x $\mapsto$ [A]x ist linear:

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 + y_1 \\ \dots \\ x_n + y_n \end{pmatrix} = \begin{pmatrix} a_{11}(x_1 + y_1) + \dots + a_{1n}(x_n + y_n) \\ \vdots \\ a_{m1}(x_1 + y_1) + \dots + a_{mn}(x_n + y_n) \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix} + \begin{pmatrix} a_{11}y_1 + \dots + a_{1n}y_n \\ \vdots \\ a_{m1}y_1 + \dots + a_{mn}y_n \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} a_{11} + \dots + a_{1n} & & \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

d.h.[A](x+y) = [A]x + [A]y. Analog zeigt man, dass [A](αx) = α([A]x).
Sei nun umgekehrt A eine lineare Abbildung. Setzen

$$e_i = (i) \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Wenden A auf e_i an und erhalten so eine Spalte aus $M_{m,1}$ ($\mathcal{K}$), die wir mit $Ae_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix}$ bezeichnen.
Aus diesen Spalten für i = 1,. . ., n bilden wir die Matrix

$$[A] = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}.$$

Ein beliebiges x = $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in M_{n,1}(\mathcal{K})$ lässt sich eindeutig in der Form x = $x_1e_1 + \ldots + x_ne_n$ schreiben.
Da A linear ist, folgt

$$Ax = A(x_1e_1 + \ldots + x_ne_n) = x_1Ae_1 + \ldots + x_nAe_n$$

$$= x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \ldots + \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

$$= \begin{pmatrix} x_1a_{11} + \ldots + x_na_{1n} \\ \vdots \\ x_1a_{n1} + \ldots + x_na_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = [A]x.$$

Ist M $\in M_{m,n}(\mathcal{K})$ irgendeine Matrix mit Ax = Mx $\forall$ x, so ist insbesondere

$$Ae_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix} = \begin{pmatrix} m_{11} & \ldots & m_{1n} \\ \vdots & & \vdots \\ m_{1m} & \ldots & m_{mn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ i \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} m_{1i} \\ \vdots \\ m_{mi} \end{pmatrix},$$

d.h. die Spalten von M stimmen mit denen von [A] überein. Daraus folgt M = [A]. Dies beweist die Eindeutigkeit von A.
(a) Offenbar existiert für jede Matrix M $\in M_{m,n}(\mathcal{K})$ ein A mit [A] = M, nämlich Ax := Mx. Also ist die Abbildung A $\mapsto$ [A] surjektiv.
Sei [A] = [B]. Dann ist Ax = [A]x = [B]x = Bx, d.h. A = B. Somit ist A injektiv.
(b) Die i-te Spalte von [A+B] ist (A+B)e_i = Ae_i + Be_i, d.h. ist gleich der i-ten Spalten von [A] plust i-te Spalte von [B]. Damit ist [A+B] = [A] + [B].
(c) analog zu (b).
#
1. Beispiel:
$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix}$
[A] = $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$
da $\begin{pmatrix} -y \\ x \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$
Vorgehensweise wie im Beweis:
Ae_1 = A $\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$,
Ae_2 = A $\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$
$\Rightarrow$ [A] = $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.
A$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$ Spiegelung an der x-Achse
$\begin{pmatrix} x \\ -y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$, [A] = $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}$ Projektion auf die x-Achse

$\begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$, $[A] = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3x \\ 3y \end{pmatrix}$ Streckung um den Faktor 3

$\begin{pmatrix} 3x \\ 3y \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$, $[A] = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$.

Definition 4 *Unter dem Produkt zweier Matrizen $A = (a_{ij}) \in M_{m,n}(\mathcal{K})$ und $B = (b_{ij}) \in M_{n,k}(\mathcal{K})$ versteht man die Matrix $AB = (c_{ij}) \in M_{m,k}(\mathcal{K})$, die gegeben ist durch $c_{ij} = a_{i1}b_{1j} + \ldots + a_{in}b_{nj}$,*
d.h. durch

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & \ldots & b_{1k} \\ \vdots & & \vdots \\ b_{n1} & \ldots & b_{nk} \end{pmatrix} := \begin{pmatrix} a_{11}b_{11} + \ldots + a_{1n}b_{n1} & \ldots & a_{11}b_{1k} + \ldots + a_{1n}b_{nk} \\ \vdots & & \vdots \\ a_{m1}b_{11} + \ldots + a_{mn}b_{n1} & \ldots & a_{m1}b_{1k} + \ldots + a_{mn}b_{nk} \end{pmatrix}$$

[Dies ist eine Abbildung $M_{m,n}(\mathcal{K})$ x $M_{n,k}(\mathcal{K}) \to M_{m,k}(\mathcal{K})$, (A,B) $\mapsto$ AB]

1. Beispiel:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 2+3 & 1+2 \\ 5+6 & 4+5 \end{pmatrix} = \begin{pmatrix} 5 & 3 \\ 11 & 9 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \text{n.d.}$$

$$(a_1 \ldots a_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = (a_1b_1 + \ldots + a_nb_n) = a_1b_1 + \ldots + a_nb_n$$

$$\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} (a_1 \ldots a_n) = \begin{pmatrix} b_1a_1 & \ldots & b_1a_n \\ \vdots & & \vdots \\ b_na_1 & \ldots & b_na_n \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

($\exists$ A $\neq 0$, B $\neq 0$ mit AB $= 0$, sog. Nullteiler und im Allgemeinen ist AB$\neq$BA)

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

Man nennt diese Matrix $\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$ die Einheitsmatrix und bezeichnet sie mit I oder I_n oder $I_{n\times n}$.

Satz 3 *Seien A: $M_{n,1}$ ($\mathcal{K}$) → $M_{m,1}$ ($\mathcal{K}$) und B: $M_{k,1}$ ($\mathcal{K}$) → $M_{n,1}$ ($\mathcal{K}$) zwei lineare Operatoren. Dann gilt für den linearen Operator AB: $M_{k,1}$ ($\mathcal{K}$) → $M_{m,1}$ ($\mathcal{K}$) die Beziehung [AB] = [A][B].*

19.12.05

Beweis:
Sei [A] = (a_{ij}), [B] = (b_{ij}), [AB] = (c_{ij}). Dann ist

$$\begin{pmatrix} c_{1i} \\ \vdots \\ c_{mi} \end{pmatrix} = (AB)e_i = A(Be_i) = A \begin{pmatrix} b_{1i} \\ \vdots \\ b_{mi} \end{pmatrix}$$

$$= A(b_{1i}e_1 + \ldots + b_{ni}e_n) = b_{1i}Ae_1 + \ldots + b_{ni}Ae_n$$

$$= b_{1i} \begin{pmatrix} a_{11} \\ \vdots \\ a_{1m} \end{pmatrix} + \ldots + b_{ni} \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix},$$

d.h. $c_{ji} = b_{1i}a_{j1} + \ldots + b_{ni}a_{jn} = a_{j1}b_{1i} + \ldots + a_{jn}b_{ni}$

Also ist $(c_{ij}) = (a_{ij})(b_{ij})$, d.h. [AB] = [A][B]. #

Definition 5 *Seien X und Y lineare Räume der Dimension n bzw. m über $\mathcal{K}$ und sei $A \in \mathcal{L}(X,Y)$. Sei $E = \{e_1,\ldots,e_n\}$ eine Basis in X und $F = \{f_1,\ldots,f_m\}$ eine Basis in Y. Dann lässt sich Ae_i eindeutig in der Form $Ae_i = a_{1i}f_1 + \ldots + a_{mi}f_m$ mit $a_{ji} \in \mathcal{K}$ schreiben. Die Matrix $(a_{ji})_{j=1,i=1}^{m,n}$ wird mit $[A]_{E,F}$ bezeichnet und Matrixdarstellung von A in den (bezüglich der) Basen E und F genannt.*

Bemerkung:
Nehmen i-tes Basiselement aus E, wenden A darauf an und zerlegen das Resultat in der Basis F. Die Koeffizienten der Zerlegung bilden die i-te Spalte von $[A]_{E,F}$.

Ist X = $M_{n,1}$ ($\mathcal{K}$), Y = $M_{m,1}$ ($\mathcal{K}$), $e_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ (i), $f_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$, so ist $[A]_{E,F}$ = [A], wobei [A] wie in Definition 3 ist.

3. Beispiel:
Sei A: $\mathcal{R}^2 \to \mathcal{R}^2$, (x,y) ↦ (x,-y)
Dies ist Spiegelung an der x-Achse.
Wählen E = {(1,1);(1,-1)} und F = {(1,0);(0,1)}.

Matrixdarstellung bestimmen:
Ae_1 = A(1,1) = (1,-1) = 1(1,0) + (-1)(0,1) = 1 f_1 + (-1) f_2
Ae_2 = A(1,-1) = (1,1) = 1(1,0) + 1(0,1) = 1 f_1 + 1 f_2,

d.h. $[A]_{E,F} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$

Andere Wahl: E = {(1,0);(1,1)}, F = {(-1,0);(0,1)}
Ae_1 = A(1,0) = (1,0) = -1(-1,0) + 0(0,1) = (-1)f_1 + 0f_2
Ae_2 = A(1,1) = (1,-1) = (-1)f_1 + (-1)f_2,
d.h. $[A]_{E,F} = \begin{pmatrix} -1 & -1 \\ 0 & -1 \end{pmatrix}$

Standardbasen:

Ae_1 = (1,0) = 1 f_1 + 0 f_2
Ae_2 = (0,-1) = 0f_1 + (-1) f_2
$[A]_{E,F} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

4. Beispiel:
A: $\mathcal{R} \to \mathcal{R}$, Ax = ax (Multiplikation mit einer Zahl)
Sei E = {e}, F = {f} mit e $\neq$ 0, f $\neq$ 0 (e,f $\in \mathcal{R}$).
Dann ist Ae = ae = $\frac{ae}{f}$ f. Also ist $[A]_{E,F} = (\frac{ae}{f})$. Erhalten insbesondere $[A]_{E,F}$ = (a) genau dann, wenn e = f ist.

5. Beispiel:
Betrachten D: $\mathcal{R}_n$ [X] $\to \mathcal{R}_n$ [X]
D$(a_0 + a_1x + \ldots + a_nx^n) = a_1 + 2a_2x + \ldots + na_nx^{n-1}$
Df = f' = $\frac{df}{dx}$.
Wählen E = F = {1,x,...,x^n}
D1 = 0 = 0*1 + ... + 0*x^n,
Dx = 1 = 1*1 + 0*x + ... + 0*x^n,
$\vdots$
Dx^n = nx^{n-1} = 0*1 + 0*x + ... + n*x^{n-1} + 0*x^n

$$\Rightarrow [A]_{E,F} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 2 & \ldots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & n \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}$$

Satz 4 *Seien X und Y lineare Räume der Dimension n bzw. m über $\mathcal{K}$, sei $E = \{e_1,\ldots,e_n\}$ eine Basis in X, $F = \{f_1,\ldots,f_m\}$ eine Basis in Y, und sei $A \in \mathcal{L}(X,Y)$. Jedes $x \in X$ lässt sich dann eindeutig in der Form $x = x_1e_1 + \ldots + x_ne_n$ schreiben und Ax lässt sich eindeutig in der Form $Ax = y_1f_1 + \ldots + y_mf_m$ schreiben.*
Mit anderen Worten bezeichnen mit $x_1,\ldots,x_n$ die Koordinaten von x in der Basis E und mit $y_1,\ldots,y_m$ die Koordinaten von Ax in der Basis F.
Sei $[A]_{E,F} = (a_{ij})$.
Dann gilt:
(a) $\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$.
(b) Die Abbildung $\mathcal{L}(X,Y) \to M_{m,n}$ ($\mathcal{K}$), $A \mapsto [A]_{E,F}$ ist ein Isomorphismus, d.h.
(i) Abbildung ist bijektiv
(ii) $[A+B]_{E,F} = [A]_{E,F} + [B]_{E,F}$
(iii) $[\alpha A]_{E,F} = \alpha[A]_{E,F}$
*(c) Es gilt dim $\mathcal{L}(X,Y)$ = dim X * dim Y.*

Beweis:
(a) Haben
$Ax = A(x_1e_1 + \ldots + x_ne_n) = x_1Ae_1 + \ldots + x_nAe_n$
$= x_1(a_{11}f_1 + \ldots + a_{m1}f_m) + \ldots + x_n(a_{1n}f_1 + \ldots + a_{mn}f_m)$
$= \underbrace{(a_{11}x_1 + \ldots + a_{1n}x_n)}_{y_1} f_1 + \ldots + \underbrace{(a_{m1}x_1 + \ldots + a_{mn}x_n)}_{y_m} f_m$, und dies ergibt die Behauptung.
(b) Sei M $\in M_{m,n}$ ($\mathcal{K}$) beliebig.
Müssen zeigen, dass ein A $\in \mathcal{L}$(X,Y) mit $[A]_{E,F}$ = M existiert.
Sei M = (c_{ij}). Setzen

$$A(x_1e_1 + \ldots + x_ne_n) = (c_{11}x_1 + \ldots + c_{1n}x_n)f_1 + (c_{m1}x_1 + \ldots + c_{mn}x_n)f_m.$$

Dies ist eine lineare Abbildung und die Rechnung aus Beweis von (a) zeigt, dass $[A]_{E,F} = (c_{ij})$ ist.
Damit ist die Surjektivität gezeigt.
Sei nun $[A]_{E,F} = [B]_{E,F}$.
Aus (a) folgt daraus: Ax = Bx $\forall$ x $\in$ X, d.h. A = B. Also ist die Abbildung injektiv.
(Teile (ii) und (iii) bleiben dem Leser überlassen)
(c) Nach (b) ist $\mathcal{L}$(X,Y) $\cong M_{m,n}$ ($\mathcal{K}$), nach Satz 2 ist dim $M_{m,n}$ ($\mathcal{K}$) = mn, nach Satz 4 aus 2.2. also dim $\mathcal{L}$(X,Y) = mn. #

Satz 5 *Seien X,Y,Z lineare Räume der Dimensionen k,n,m mit Basen $E = \{e_1,\ldots,e_k\}$, $F = \{f_1,\ldots,f_n\}$, $G = \{g_1,\ldots,g_m\}$.*
Für $A \in \mathcal{L}(Y,Z)$ und $B \in \mathcal{L}(X,Y)$ gilt dann:
$[AB]_{E,G} = [A]_{F,G}[B]_{E,F}$.

Beweis:
Sei $[A]_{F,G} = (a_{ij})$, $[B]_{E,F} = (b_{ij})$, $[AB]_{E,G} = (c_{ij})$.

Haben
$(AB)e_i = A(Be_i) = A(b_{1i}f_1 + \ldots + b_{ki}f_k)$
$= b_{1i}Af_1 + \ldots + b_{ki}Af_n$
$= b_{1i}(a_{11}g_1 + \ldots + a_{m1}g_m) + \ldots + b_{ki}(a_{1k}g_1 + \ldots + a_{mk}g_m)$
$= (a_{11}b_{1i} + \ldots + a_{1k}b_{ki})g_1 + \ldots + (a_{m1}b_{1i} + \ldots + a_{mk}b_{ki})g_m$
und andererseits ist
$(AB)e_i = c_{1i}g_1 + \ldots + c_{mi}g_m$
also ist
$c_{li} = a_{l1}b_{1i} + \ldots + a_{lk}b_{ki}$ (l = 1, …, m)
d.h. $(c_{ij}) = (a_{ij})(b_{ij})$. # 06.01.06

2.4 Gruppen und Permutationen

Definition 1 *Ist X eine Menge, so bezeichnen wir mit S(X) die Menge aller bijektiven Abbildungen von X auf X. Für S =({1,2,…,n}) schreibt man kurz S_n. Man nennt S_n die symmetrische Gruppe vom Grad n. Die Elemente von S_n heißen Permutationen.*

Eine Permutation $\sigma \in S_n$ ist also eine bijektive Abbildung von {1,…,n} auf sich selbst. Diese Abbildung ist eindeutig gegeben durch die Werte $\sigma(1)$, $\sigma(2)$, …, σ(n). Man schreibt σ in der Form

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}.$$

Haben
$S_2 = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix},$
$S_3 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 3 & 2 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},$
$S_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$
$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 1 \end{pmatrix}$ ist z.B. keine Permutation (keine injektive Abbildung).
Man kann sich Permutationen auch als Umordnung vorstellen. Offenbar hat S_n genau n! = 1 * 2 * … * n Elemente.

Definition 2 *Das Produkt $\sigma\tau$ zweier Permutationen $\sigma, \tau \in S_n$ ist definiert als die Komposition (= Hintereinanderausführung) der Abbildungen σ und τ, d.h. $\sigma\tau = \sigma \circ \tau$ oder $(\sigma\tau)(j) = (\sigma)(\tau(j))$.*

Zum Beispiel ist
$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$

Satz 1 *(a) Es gilt* $(\sigma\tau)(\mu) = \sigma(\tau\mu) \quad \forall \sigma, \tau, \mu \in S_n$.
(b) $\exists \epsilon \in S_n$*:* $\sigma * \epsilon = \epsilon * \sigma = \sigma \quad \forall \sigma \in S_n$.
(c) $\forall \sigma \in S_n \; \exists \tau \in S_n : \; \sigma\tau = \tau\sigma = \epsilon$.

Beweis:
(a) Hintereinanderausführung von Abbildungen ist assoziativ.
(b) $\epsilon = \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}$
(c) Ist σ: {1,...,n} $\to$ {1,...,n} bijektiv, so hat $\tau = \sigma^{-1}$: {1,...,n} $\to$ {1,...,n} die gewünschte Eigenschaft.
$\left[\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \right]$ #

Definition 3 *Eine Gruppe ist eine Menge G, in der für jedes geordnete Paar (a,b)* $\in G \times G$ *ein eindeutig bestimmtes Element* $a * b \in G$ *definiert ist und dabei die folgenden Axiome gelten:*
(a) $(a * b) * c = a * (b * c) \quad \forall\, a,b,c \in G$,
(b) $\exists \epsilon \in G$*:* $a * \epsilon = \epsilon * a = a \quad a \in G$,
(c) $\forall\, a \in G \; \exists\, b \in G$*:* $a * b = b * a = \epsilon$.
Eine Gruppe G heißt abelsch oder kommutativ, wenn gilt
(d) $a * b = b * a \quad \forall\, a,b \in G$
Ist G als Menge endlich, so nennt man G endliche Gruppe und bezeichnet die Anzahl der Elemente von G als Ordnung von G.

1. Beispiel:
Satz 1 besagt, dass S_n eine endliche Gruppe der Ordnung n! mit der Operation $\sigma * \tau = \sigma \circ \tau$ ist. Für eine beliebige Menge x ist S(x) eine Gruppe bezüglich $\sigma * \tau = \sigma \circ \tau$. Gruppe S(x) ist kommutativ $\Leftrightarrow$ |x| (Anzahl der Elemente von x) ≤ 2

2. Beispiel:
Ist X ein linearer Raum, so ist X mit a * b = a + b eine abelsche Gruppe, d.h. jeder lineare Raum wird zu einer Gruppe, wenn man die Multiplikation mit Skalaren aus $\mathcal{K}$ vergisst. Haben $\epsilon = 0$ und a + (-a) = (-a) + a = 0. Abelsche Gruppen bezüglich der Addition sind also insbesondere $\mathcal{R}^n$, $\mathcal{C}^n$, $\mathcal{R}_n$[x], $\mathcal{C}_n$[x], $M_{n,m}(\mathcal{R})$, $M_{n,m}(\mathcal{C})$, $\mathcal{L}$(X,Y).

Bemerkungen:
Das Element ϵ aus Definition 3 ist eindeutig bestimmt:
$\epsilon_1 = \epsilon_1 * \epsilon_2 = \epsilon_2$.
Es heißt *neutrales Element* der Gruppe.
Das Element b aus Definition 3c) ist ebenfalls eindeutig bestimmt.
a * b_1 = b_1 * a = ϵ,
a * b_2 = b_2 * a = ϵ
$\Rightarrow b_1 = b_1 * \epsilon = b_1 *$ a * $b_2 = (b_1 *$ a) * $b_2 = \epsilon * b_2 = b_2$
Dieses Element bezeichnet man mit a^{-1} und nennt es das zu a *inverse Element.*
Für beliebiges a,b $\in$ G hat die Gleichung ax = b stets eine eindeutig bestimmte Lösung, nämlich x = a^{-1}b.

Abelsche Gruppen schreibt man in der Regel *additiv*, d.h. man schreibt a + b statt

a * b. Das neutrale Element heißt dann *Nullelement* und wird mit 0 bezeichnet. Das inverse Element zu a wird mit -a bezeichnet.
Beliebige Gruppen schreibt man häufig *multiplikativ*, d.h. man schreibt ab statt a * b. In diesem Fall heißt ϵ *Einselement* und wird oft mit e bezeichnet. Das inverse Element a^{-1} bleibt.

3. Beispiel:
$\mathcal{R}$ mit Multiplikation ist keine Gruppe: Es existiert kein b $\in \mathcal{R}$ mit 0b = 1.
Aber $\mathcal{R}^*$ (= $\mathcal{R}\backslash\{0\}$) und (0,$\infty$) sind abelsche Gruppen bezüglich der Multiplikation.
(0,∞) mit der Addition ist keine Gruppe (0 liegt nicht darin), ist aber eine abelsche Gruppe mit x * y = log x + log y [log = ln = log nat].
Ist K ein Schiefkörper, so ist (K,+), d.h. K mit der Addition, eine abelsche Gruppe und K* = K\{0} mit der Multiplikation eine Gruppe.

4. Beispiel:
Die Menge $\{\overline{0},\ldots,\overline{n-1}\}$ mit $\overline{a}+\overline{b}$:= Rest von a + b modulo n ist eine abelsche Gruppe der Ordnung n (n *geq* 2). Diese Gruppe heißt *zyklische Gruppe der Ordnung n* und wird mit $\mathcal{Z}_n$ oder $\mathcal{Z}/n\mathcal{Z}$ bezeichnet.

5. Beispiel:
Das Intervall (-c,c) ist eine abelsche Gruppe mit der Operation $\boxplus$, die gegeben ist durch

$$u \boxplus v = \frac{u+v}{1+\frac{uv}{c^2}}.$$

Dies ist Addition von Geschwindigkeiten in der Relativitätstheorie.

Definition 4 *Seien (G,$\circ$) und (H,*) Gruppen. Eine Abbildung f: $G \to H$ heißt Gruppenhomomorphismus, wenn gilt*
$f(a \circ b) = f(a) * f(b) \quad \forall\ a,b \in G.$
Bijektive Gruppenhomomorphismen heißen Gruppenisomorphismen. Zwei Gruppen heißen isomorph, wenn es einen Gruppenisomorphismus zwischen ihnen gibt.

Isomorphie ist eine Äquivalenzrelation:
(R) f: G $\to$ G, f(x) = x ist Isomorphismus
(S) f: G $\to$ H Isomorphismus $\Rightarrow$ f^{-1}: H $\to$ G Isomorphismus
(T) f: G $\to$ H, g: H $\to$ K Isomorphismus $\Rightarrow$ g $\circ$ f: G $\to$ K Isomorphismus

Sind damit bei einem weiteren Klassifizierungsproblem angelangt: Man erstelle eine Liste von Gruppen, sodass jede Gruppe zu genau einer Gruppe aus der Liste isomorph ist. Für endliche abelsche Gruppen geben wir eine solche Liste später an.

6. Beispiel:
Betrachten folgende Tätigkeiten, die man mit einem angezogenen Socken machen kann:
e: wir machen gar nichts
a: wir ziehen den Socken auf den anderen Fuß
b: wir ziehen den Socken aus, stülpen ihn um, und ziehen ihn wieder auf denselben Fuß
c: wir ziehen den socken aus, stülpen ihn um, und ziehen ihn auf den anderen Fuß

Diese vier Tätigkeiten bilden eine Gruppe bezüglich der Hintereinanderausführung.

Die Gruppentafel ist

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Diese Gruppe und jede zu ihr isomorphe Gruppe nennt man die *Kleinsche Vierergruppe* und bezeichnet sie mit V_4 oder $\mathcal{Z}_2 \times \mathcal{Z}_2$.
Eine isomorphe Gruppe wird z.B. durch Hintereinanderausführung folgender Abbildungen in der Ebene gegeben:
e: identische Abbildung
a: Spiegelung an der x-Achse
b: Spiegelung an der y-Achse
c: Spiegelung am Ursprung

7. Beispiel: 09.01.06

Jede Gruppe der Ordnung 1 ist isomorph zu {e} mit

	e
e	e

Jede Gruppe der Ordnung 2 ist isomorph zu {e,a} mit

	e	a
e	e	a
a	a	e

Konkrete Gruppe der Ordnung 2

- {-1,1} mit Multiplikation
  → Assoziativgesetz gilt, weil es für $\mathcal{Z}$ gilt.
  Gruppentafel:

	1	-1
1	1	-1
-1	-1	1

- $\{\overline{0},\overline{1}\}$ mit Addition modulo 2
  (Zyklische Gruppe der Ordnung 2)

	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{1}$
$\overline{1}$	$\overline{1}$	$\overline{0}$

Gruppen der Ordnung 3:

	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

Dies ist tatsächlich eine Gruppentafel:
→ Assoziativgesetz kann man wieder durch konkrete Gruppe überprüfen:
$\{\overline{0},\overline{1},\overline{2}\}$ mit Addition mod 3 (ist assoziativ)
→ jede beliebige Gruppe der Ordnung 3 ist isomorph zur zyklischen Gruppe der Ordnung 3
Somit existiert bis auf Isomorphie genau eine Gruppe der Ordnung 3.

Jede Gruppe der Ordnung 4 ist isomorph zu $\mathcal{Z}_4$ oder V_4 und $\mathcal{Z}_4$ und V_4 sind nicht isomorph zueinander.

Jede Gruppe der Ordnung 5 ist isomorph zu $\mathcal{Z}_5$.

Alle Gruppen der Ordnung ≤ 5 sind abelsch.

S_3 ist eine nichtabelsche Gruppe der Ordnung 6.

Liste aller Gruppen beginnt also wie folgt:

n	Anzahl der Gruppen	Gruppe
1	1	$\mathcal{Z}_1$
2	1	$\mathcal{Z}_2$
3	1	$\mathcal{Z}_3$
4	2	$\mathcal{Z}_4$,V_4
5	1	$\mathcal{Z}_5$

Haben bisher kennengelernt:

Schiefkörper	Lineare Räume	Gruppen
a + b, ab	a + b, αa	a * b
+ Axiome	+ Axiome	+ Axiome
	Homomorphismen	
φ(a+b)=φ(a)+φ(b)	φ(a+b)=φ(a)+φ(b)	φ(a*b)=φ(a)*φ(b)
φ(ab)=φ(a)φ(b)	φ(αa)=$\alpha\varphi$(a)	

Isomorphismus = bijektiver Homomorphismus

Zum Sprachgebrauch:
Homomorphismus X → X = Endomorphismus
Isomorphismus X → X = Automorphismus
surj.Homomorphismus X → Y = Epimorphismus
inj.Homomorphismus X → Y = Monomorphismus

Definition 5 *Eine Teilmenge H einer Gruppe heißt Untergruppe von G, wenn H mit den Operationen aus G selbst eine Gruppe ist, d.h. $H \subset G$ ist Untergruppe von G genau dann, wenn gilt*
(a) $a,b \in H \Rightarrow a * b \in H$;
(b) $e \in H$;
(c) $a \in H \Rightarrow a^{-1} \in H$.
Untergruppen von S_n heißen Permutationsgruppen.

8. Beispiel:
Jede Gruppe G hat die beiden trivialen Untergruppen {e} und G.
Gruppe $\mathcal{Z}_3$ hat nur die trivialen Untergruppen: Ist H eine Untergruppe mit $\overline{1} \in$ H, so ist $\overline{1} + \overline{1} = \overline{2} \in$ H, d.h. H = $\mathcal{Z}_3$, und ist H eine Untergruppe mit $\overline{2} \in$ H, so ist $\overline{2} + \overline{2} = \overline{1} \in$ H, d.h. H = $\mathcal{Z}_3$.
Gruppe V_4 enthält die Untergruppen {e,a}, {e,b}, {e,c}.
Gruppe $\mathcal{Z}_4$ enthält die Untergruppe $\{\overline{0},\overline{2}\}$.

Satz 2 *Cayley*
Jede endliche Gruppe ist zu einer Permutationsgruppe isomorph. Präziser: Eine Gruppe der Ordnung n ist zu einer Untergruppe von S_n isomorph

Beweis:
Sei G eine Gruppe der Ordnung n. Für jedes a ∈ G betrachten wir die Abbildung

$$\varphi(a) : G \to G, \varphi(a)b = ab.$$

φ(a) ist injektiv:
φ(a)b_1 = φ(a)b_2 ⇒ ab_1=ab_2 ⇒ a^{-1}ab_1=a^{-1}ab_2 ⇒ b_1e=b_2e ⇒ $b_1 = b_2$.

φ(a) ist surjektiv:
φ(a)x=b ⇔ ax = b ⇔ x = a^{-1}b.

Also ist φ(a) ∈ S(G) für jedes a ∈ G. Erhalten damit eine Abbildung

$$\varphi : G \to S(G), a \mapsto \varphi(a).$$

Behaupten, dass φ(G) (:= {φ(a): a ∈ G}) eine Untergruppe von S(G) und φ: G → φ(G) ein Isomorphismus ist.

φ ist Homomorphismus, d.h. φ(cd) = φ(c)φ(d) ∀ c,d ∈ G:
φ(cd)f=(cd)f=c(df)=φ(c)φ(d)f ∀ f ∈ G.

φ ist bijektiv:
φ ist surjektiv, da Wertebereich auf φ(G) eingeschränkt wurde
φ ist injektiv, denn φ(a)=φ(b) ⇒ φ(a)f=φ(b)f ∀ f ∈ G
⇒ af = bf ∀ f ∈ G ⇒ ae = be ⇒ a = b.

φ(G) ist Untergruppe:
(a) a,b ∈ φ(G) ⇒ a = φ(c), b = φ(d) mit c,d ∈ G
⇒ ab = φ(c)φ(d) = φ(cd) ∈ φ(G);
(b) Einselement von S(G) ist φ(e), und φ(e) ∈ φ(G);
(c) a = φ(c) ∈ φ(G) ⇒ $a^{-1} = \varphi\ c^{-1} \in \varphi$(G).

Also ist φ: G → φ(G) ein Isomorphismus. Da S(G) zu S_n = S({1,...,n} isomorph ist, ist φ(G) ⊂ S(G) isomorph zu einer Untergruppe von S_n. #
Es gibt eine Untergruppe A_n von S_n, die für verschiedene Belange von Bedeutung ist. Benötigten zur Definition von A_n noch einige Hilfsmittel.

Permutationen der Form

$$\begin{pmatrix} 1 & \dots & i-1 & i & i+1 & \dots & j-1 & j & j+1 & \dots & n \\ 1 & \dots & i-1 & j & i+1 & \dots & j-1 & i & j+1 & \dots & n \end{pmatrix}$$

heißen *Transpositionen* und werden mit (ij) bezeichnet. Die Permutation (ij) überführt also i in j und j in i und lässt alles andere fest.

Zyklenschreibweise:
(123) :=
$1 \to 2$
$2 \to 3$
$3 \to 1$

Satz 3 *Jede Permutation $\sigma \in S_n$ (n≥2) lässt sich als ein endliches Produkt von Transpositionen schreiben:*

$$\sigma = (i_1 j_1)(i_2 j_2) \dots (i_k j_k)$$

Beweis : Vollständige Induktion

Für n=2 ist alles klar:

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = (12)$$

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = (12)(12)$$

Nehmen an, dass Behauptung für n-1 richtig ist und zeigen sie für n:

Sei also $\sigma \in S_n$. Unterscheiden 2 Fälle:

- a) $\sigma(n) = n$
  $\to \begin{pmatrix} 1 & 2 & \dots & n-1 \\ \sigma(1) & \sigma(2) & \dots & \sigma(n-1) \end{pmatrix}$
  lässt sich nach Induktionsvoraussetzung in der Form $(i_1 j_1) \dots (i_k j_k)$ schreiben.

  Dies ergibt: $\sigma = (i_1 j_1) \dots (i_k j_k)$

- b) $\sigma(n) \neq n$

  Sei $\tau = (\sigma(n) n)\sigma$

  Dann ist $\tau(n) = n$. Nach (a) ist

  $\tau = (i_1 j_1) \dots (i_k j_k)$, d.h.

  $(\sigma(n) n)\sigma = (i_1 j_1) \dots (i_k j_k)$

  $\Rightarrow \underbrace{(\sigma(n) n)(\sigma(n) n)}_{\epsilon} \sigma = (\sigma(n) n)(i_1 j_1) \dots (i_k j_k)$

  $[(ij)^{-1} = (ij)]$

  $\Rightarrow \sigma = (\sigma(n) n)(i_1 j_1) \dots (i_k j_k).$ #

Die Darstellung von Satz 3 ist nicht eindeutig. Zeigen aber, dass für jedes $\sigma \in S_n$ (n≥2) die Anzahl der Faktoren entweder stets gerade oder stets ungerade ist.

„Anzahl der Unordnungen“

Definition 6 *Definition 6: Für eine Permutation $\sigma \in S_n$ (n≥2) bezeichnen wir mit $N(\sigma)$ die Anzahl aller geordneten Paare*

$$(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,n\}$$

mit $i < j$ und $\sigma(i) > \sigma(j)$.

Die Zahl $(-1)^{N(\sigma)}$ heißt Signum *der Permutation σ und wird mit sgn σ bezeichnet.*

Eine Permutation σ heißt gerade *bzw.* ungerade*, wenn sgn $\sigma = 1$ bzw. sgn $\sigma = -1$ ist.*

Satz 4 *Satz 4: Eine Permutation ist gerade (bzw. ungerade) genau dann, wenn bei jeder Zerlegung in ein Produkt von Transpositionen die Anzahl der Faktoren gerade (bzw. ungerade) ist.*

Beispiele zu Definition 6:

$$N\begin{pmatrix}1&2\\2&1\end{pmatrix}=0 \qquad N\begin{pmatrix}1&2\\2&1\end{pmatrix}=1$$
$$N\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}=1 \qquad N\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}=1$$
$$N\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}=2 \qquad N\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}=0$$
$$N\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}=2 \qquad N\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}=3$$

Beweis von Satz 4:

Zeigen zunächst, dass gilt:

$$sgn\ \sigma(i\ i+1) = -sgn\ \sigma \qquad (1)$$

In der Tat, haben

$$\sigma(i\ i+1) = \begin{pmatrix}1&\ldots&i&\ldots&i+1&\ldots&n\\ \sigma(1)&\ldots&\sigma(i)&\ldots&\sigma(i+1)&\ldots&\sigma(n)\end{pmatrix}(i\ i+1)$$
$$= \begin{pmatrix}1&\ldots&i&\ldots&i+1&\ldots&n\\ \sigma(1)&\ldots&\sigma(i+1)&\ldots&\sigma(i)&\ldots&\sigma(n)\end{pmatrix},$$

d.h. N$(\sigma(i\ i+1))$ = N$(\sigma)+1$ für $\sigma(i) < \sigma(i+1)$ und N$(\sigma(i\ i+1))$ = N(σ)-1 für $\sigma(i) > \sigma(i+1)$.

Somit ist $sgn\ \sigma(i\ i+1) = (-1)^{N(\sigma(i\ i+1))} = (-1)^{N(\sigma)\pm 1} = -(-1)^{N(\sigma)} = -sgn\ \sigma$.

Damit ist (1) bewiesen. #

Zeigen als nächstes, dass

$$sgn\ \sigma(ij) = -sgn\ \sigma \qquad (2)$$

Dies folgt aus (1) und der Tatsache, dass (ij) als Produkt einer ungeraden Anzahl von Transpositionen der Form (k k+1) geschrieben werden kann.

Beispiel: n=5, i=2, j=5 → Betrachten (25) = $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 5 & 3 & 4 & 2 \end{pmatrix}$

12345 → 13245 → 13425 → 13452 → 13542 → 15342

nur benachbarte Elemente vertauschen

(25) = (23)(34)(45)(34)(23)

→ ungerade Anzahl

Allgemein ist

$$(ij) = \underbrace{(i\ i+1)\dots(j-2\ j-1)}_{j-i-1\ Faktoren}\underbrace{(j-1\ j)}_{1\ Faktor}\underbrace{(j-2\ j-1)\dots(i\ i+1)}_{j-i-1\ Faktoren}$$

Haben also 2(j-i-1)+1 Faktoren. Damit ist (2) bewiesen. #

Nun zu Satz 4. Sei $\sigma = (i_1 j_1)\dots(i_k j_k)$.

Dann ist

$$\begin{aligned} sgn\ \sigma &= sgn\ (i_1 j_1)\dots(i_{k-1} j_{k-1})(i_k j_k) \\ &\overset{(1)}{=} -sgn\ (i_1 j_1)\dots(i_{k-1} j_{k-1}) \\ &\ \ \vdots \\ &= (-1)^k \end{aligned}$$

Ist $sgn\ \sigma = 1$, so ist also k gerade in jeder Zerlegung.

Und ist $sgn\ \sigma = -1$, so ist k ungerade in jeder Zerlegung. #

Satz 5 *Satz 5: Die Abbildung*

$$sgn: \ S_n \to \{-1, 1\}, \sigma \mapsto sgn\ \sigma$$

ist ein Homomorphismus von S_n auf die Gruppe {-1,1} mit der Multiplikation. Die Menge

$$A_n := \{\sigma \in S_n : sgn\ \sigma = 1\}$$

ist eine Untergruppe von S_n, die sogenannte alternierende Gruppe vom Grad n. *Die Gruppe A_n hat für n≥ 2 die Ordnung $\frac{n!}{2}$.*

Beweis: Müssen zeigen,

$$(3) \qquad sgn\ (\sigma\tau) = sgn\ \sigma\ sgn\ \tau$$

(Homomorphismus)

Nach Satz 3 und Satz 4 ist

$$\sigma = (i_1 j_1) \dots (i_m j_m),\ sgn\ \sigma = (-1)^m$$

$$\tau = (k_1 l_1) \dots (k_p l_p),\ sgn\ \tau = (-1)^p$$

Dann ist

$$\sigma\tau = (i_1 j_1) \dots (i_m j_m)(k_1 l_1) \dots (k_p l_p)$$

und Satz 4 liefert

$$sgn\ (\sigma\tau) = (-1)^{m+p}$$

Aus $(-1)^{m+p} = (-1)^m * (-1)^p$ folgt also, dass sgn ein Homomorphismus ist.

Zeigen nun, dass A_n eine Untergruppe ist.

Sind $\sigma\ , \tau \in A_n$, so ist $sgn\ \sigma = sgn\ \tau = 1$, und aus (3) folgt $sgn\ (\sigma\tau) = 1$, d.h. $\sigma\tau \in A_n$.

Wegen $sgn\ \epsilon = 1$, ist $\epsilon \in A_n$.

Ist $\sigma \in A_n$, so ist $sgn\ \sigma = 1$. Aus $\sigma\sigma^{-1} = \epsilon$ und (3) ergibt sich $sgn\ \sigma\ sgn\ (\sigma^{-1}) = sgn\ \epsilon = 1$, d.h. $sgn\ \sigma^{-1} = 1$, d.h. $\sigma^{-1} \in A_n$.

Also ist A_n eine Untergruppe von S_n.

Sei schließlich p die Anzahl der geraden und q die Anzahl der ungeraden Permutationen. (p + q = n!). Die Abbildung

$$\varphi : S_n \to S_n,\ \sigma \mapsto \sigma(12)$$

ist injektiv: $\varphi(\sigma) = \varphi(\tau) \Rightarrow \sigma(12) = \tau(12) \Rightarrow \sigma(12)(12) = \tau(12)(12) \Rightarrow \sigma = \tau$.

Nach Satz 4 (oder nur (1)) ändert φ die *Signatur* (= das Signum) von σ.

Daraus folgt p ≤ q und q ≤ p, d.h. p = q. [Die mit Signum 1 werden injekiv in die mit -1 abgebildet und andersherum.] #

Gruppenhomomorphismen liefern vereinfachte „Fotografien“ von Gruppen, d.h. sie verwischen Details (außer bei Isomorphismen), geben aber die Struktur richtig wider.

Definition 7 *Ein Gruppenhomomorphismus $\varphi : G \to H$ heißt* vereinfachend*, wenn $1 < |\varphi(G)| < |G|$ gilt. ($|X|$ = Anzahl der Elemente von X).*
Eine endliche Gruppe heißt einfach*, wenn es keinen vereinfachenden Gruppenhomomorphismus gibt.*

G einfache Gruppe $\Leftrightarrow$ es existiert kein vereinfachender Homomorphismus
f: G $\to$ H
$1 < |\text{Im f}| = |\text{f(G)}| < |\text{G}|$

Satz 6 *Die zyklische Gruppe $\mathcal{Z}_n$ ist genau dann einfach, wenn n eine Primzahl ist*

Beweis:
Sei n Primzahl. Nehmen an, dass $\mathcal{Z}_n$ nicht einfach ist. Dann existiert ein Homomorphismus f: $\mathcal{Z}_n \to$ H mit $|\text{f}(\mathcal{Z}_n)| <$ n. Damit müssen zwei verschiedene Elemente von $\mathcal{Z}_n$ = $\{\overline{0},\ldots,\overline{n}\}$ das gleiche Bild haben, d.h. es existiert a,b $\in \{0,\ldots,\text{n-1}\}$ mit a $\neq$ b und f($\overline{a}$) = f($\overline{b}$).
Schreiben H additiv. Da f ein Homomorphismus ist, gilt
f($\overline{a}$-$\overline{b}$)=f($\overline{a}$)-f($\overline{b}$)=0

O.B.d.A. sei a > b. Für w = a - b gilt also f($\overline{w}$) = 0.
Die Zahlen 0 * w, …, (n-1) * w lassen bei Division durch n alle verschiedene Reste (weil n Primzahl: n | k*w - l*w = (k-l)*w $\Rightarrow$ n|k-l oder n|w, da |k-l|$\leq$n-1 und |w| $\leq$n-1 und k-l $\neq$ 0, w $\neq$ 0).
Damit lässt sich jedes c $\in \{0,\ldots, \text{n-1}\}$ in der Form c = m*n+r*w mit r $\in \{0,\ldots,\text{n-1}\}$ schreiben.

Dies ergibt f($\overline{c}$)=f(r$\overline{w}$)=f($\underbrace{\overline{w}+\ldots+\overline{w}}_{r}$=f($\overline{w}$)+…+f($\overline{w}$) = 0+…+0.

D.h. jedes Element von $\mathcal{Z}_n$ wird auf 0 abgebildet. Also ist $|\text{f}(\mathcal{Z}_n)| = 1$ Widerspruch (kein vereinfachender Homomorphismus)!!
Ist n keine Primzahl und k ein nichttrivialer Teiler von n, so ist
f: $\mathcal{Z}_n \to \mathcal{Z}_k$, $\overline{a} \mapsto \overline{a \bmod k}$
ein vereinfachender Homomorphismus. #
Erhalten also eine unendliche Serie $\mathcal{Z}_1$,$\mathcal{Z}_2$, $\mathcal{Z}_3$, $\mathcal{Z}_5$, $\mathcal{Z}_7$, $\mathcal{Z}_{11}$, … von einfachen Gruppen.
Man kann zeigen, dass auch A_5,A_6,A_7,A_8,… eine unendliche Serie von einfachen Gruppen ist. [Beweis nicht trivial]

Man kann des Weiteren zeigen, dass es noch 16 weitere solche unendlicher Serien von einfachen Gruppen gibt. Die Frage ist, ob mit diesen 18 unendlichen Serien alle einfachen Gruppen erschöpft sind. Überraschenderweise ist die Antwort „nein“.
Es gibt noch sogenannte *sporadische Gruppen*. Bis 1965 waren 5 solcher sporadischer Gruppen (die sog. Mathieu-Gruppen) bekannt. Dies sind sehr komplizierte Gruppen, die mit Symmetrietransformationen im $\mathcal{R}^{12}$ und $\mathcal{R}^{24}$ zusammenhängen.
Janko hat 1965 eine sechste sporadische Gruppe entdeckt, woraufhin ein Boom bei der Suche nach weiteren sporadischen Gruppen einsetzte. Bis 1975 wurden noch 20 weitere gefunden. Diese sind alle hochgradig kompliziert. Die größte dieser 26 gruppen heißt „big monster“ und hat eine Ordnung von $\approx 10^{54}$. Die zweitgrößte heißt „baby monster“.

Man nimmt heute als bewiesen an, dass es außer den 18 unendlichen Serien und den 26 sporadischen Gruppen keine weiteren (endlichen) einfachen Gruppen gibt. Beweis füllt mehrere Bände (Gorenstein).
Die Konstruktion von sporadischen Gruppen ist kompliziert und clever.

Conway (der 3 sporadische Gruppen entdeckt hat) benutzt z.B. die Theorie der Kugelpackungen in höherdimensionalen Räumen. (Wie viele gleichgroße Kugeln passen in n-dimensionalen Räumen an eine?)

Bei solchen Fragen versagt unsere Intuition oft. Hier ist ein Beispiel:
Vier Kreise werden in ein Quadrat einbeschrieben. Dann wird ein kleinerer Kreis in die Mitte positioniert. Gesucht ist dessen Radius.
Nehmen Seitenlänge des Quadrates an = 4
$d = \sqrt{(1-0)^2 + (1-0)^2} = \sqrt{2}$
$\underline{\underline{r = \sqrt{2} - q}}$
Haben nun 8 Kugeln in einem Würfel. Gesucht ist der Radius der Kugel in der Mitte.
$d = \sqrt{(1-0)^2 + (1-0)^2 + (1-0)^2} = \sqrt{3}$
$\underline{\underline{r = \sqrt{3} - 1}}$
In höheren Dimensionen:
2^n Kugeln im n-dimensionalen Würfe. Gesucht ist Radius der Kugel in der Mitte. Eine Ecke ist (2,2,...,2). Mittelpunkt der Kugel (1,1,...,1). Koordinaten des Ursprungs (0,0,...,0).
$$d = \sqrt{\underbrace{(1-0)^2 + \ldots + (1-0)^2}_{n}} = \sqrt{n}$$
$\underline{\underline{r = \sqrt{n} - 1}}$
Für n $\geq$ 10 guckt die Kugel in der Mitte also aus dem Würfel heraus.

2.5 Determinanten

Definition 1 *Eine Matrix $A \in M_n(\mathcal{K})$ heißt invertierbar (= nichtsingulär = regulär, wenn eine Matrix $B \in M_n(\mathcal{K})$ mit $AB = BA = I$ existiert, wobei $I \in M_n(\mathcal{K})$ die Einheitsmatrix (= identische Matrix) ist.*

Wenn eine Matrix B wie in Definition 1 exisitiert, so ist diese eindeutig bestimmt.
Beweis: $AB_1 = B_1A = I$, $AB_2 = B_2A = I$
$\Rightarrow B_1I = IB_1 = (B_2A)B_1 = B_2(AB_1) = B_2I$.

Definition 2 *Wenn $A \in M_n(\mathcal{K})$ invertierbar ist, so wird die eindeutig bestimmte Matrix $B \in M_n(\mathcal{K})$ mit $AB = BA = I$ mit A^{-1} bezeichnet und die Inverse (= inverse Matrix) von / zu A genannt.*

Seien X und Y lineare Räume über $\mathcal{K}$ und sei dim X = dim Y.
Für eine lineare Abbildung A: X $\rightarrow$ Y sind dann folgende Aussagen äquivalent:
(i) A ist ein Isomorphismus;
(ii) es existieren Basen E in X und F in Y, sodass $[A]_{E,F}$ invertierbar ist;
(iii) für alle Basen E in X und F in Y ist $[A]_{E,F}$ invertierbar.
Beweis: Hausaufgabe

n = 1 A = (a) ist invertierbar $\Leftrightarrow$ a $\neq 0$.
Die Inverse ist B = (a^{-1}).

n = 2 A = $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$
Behaupten: A ist invertierbar $\Leftrightarrow$ ad - bc $\neq 0$
Beweis:
Sei ad - bc $\neq 0$. Setzen B = $\frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Dann ist AB = $\begin{pmatrix} a & b \\ c & d \end{pmatrix}\frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$
$= \frac{1}{ad-bc}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$
$= \frac{1}{ad-bc}\begin{pmatrix} ad-bc & -ab+ba \\ cd-dc & -bc+da \end{pmatrix}$
$= \frac{1}{ad-bc}\begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.
und analog BA = I.
Umgekehrt sei A invertierbar und B = $\begin{pmatrix} u & v \\ x & y \end{pmatrix}$ die Inverse.
Dann gilt also

$$AB = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} u & v \\ x & y \end{pmatrix} = \begin{pmatrix} au+bx & av+by \\ cu+dx & cv+dy \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Aus der Identität

$$(ad-bc)(uy-vx) = (au+bx)(cv+dy) - (av+by)(cu+dx)$$

folgt

$$(ad-bc)(uy-vx) = 1*1+0*0 = 1,$$

d.h. ad - bc $\neq 0$. #
Die Zahlen det(a) := a und det $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$:= ad-bc „determinieren“ also, ob (a) bzw. $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ invertierbar sind.
Im Allgemeinen Fall hat man folgendes:

Definition 3 *Sei $A = (a_{ij})_{i,j=1}^{n} \in M_n(\mathcal{K})$. Die Zahl*

$\det A := \sum_{\sigma \in S_n} (sgn\ \sigma) a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)}$

heißt Determinante von A. Die Summe erstreckt sich also über alle n! Permutationen $\sigma \in S_n$; sgn σ ist die Signatur von σ. Statt det A schreibt man häufig auch

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

20.01.06

n = 1 A = (a_{11}) $\quad s_1 = \{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\}$
det A = sgn $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ * $a_{11} = \underline{\underline{a_{11}}}$

$\underline{n = 2}$ A = $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$

$s_2 = \{ \underbrace{\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}}_{\sigma}, \underbrace{\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}}_{\tau} \}$

det A = sgn σ * $a_{1\sigma(1)}$ * $a_{2\sigma(2)}$ + sgn τ * $a_{1\tau(1)}$ * $a_{2\tau(2)}$

= $\underline{\underline{a_{11}a_{22} - a_{12}a_{21}}}$

$\underline{n = 3}$ A = $\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$

σ	sgn σ	det A
123	1	$= a_{11}a_{22}a_{33}$
132	-1	- $a_{11}a_{23}a_{32}$
213	-1	- $a_{12}a_{21}a_{33}$
231	1	+ $a_{12}a_{23}a_{13}$
312	1	+ $a_{13}a_{21}a_{31}$
321	-1	- $a_{13}a_{22}a_{31}$

Sarrus'sche Regel (nur für n = 3)
Diagonalen von links oben nach rechts unten durchmultiplizieren und einzelne Diagonalen addieren, dann die Produkte der Diagonalenwerte (von rechts oben nach links unten) davon abziehen

$\underline{\text{Beispiel:}}$

$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} \begin{matrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{matrix}$

$= 1*5*9 + 2*6*7 + 3*4*8 - 7*5*3 - 8*6*1 - 9*4*2$

$= 45 + 84 + 96 - 105 - 48 - 72 = \underline{\underline{0}}$

Die Determinante det A ist eine Summe von n! Summanden der Form $\pm a_{i_1 j_1} a_{i_2 j_2} \dots a_{i_n j_n}$ mit $\{i_1, \dots, i_n\} = \{j_1, \dots, j_n\} = \{1,\dots,n\}$.
Das Vorzeichen ist
sgn $\begin{pmatrix} i_1 & i_2 & \dots & i_n \\ j_1 & j_2 & \dots & j_n \end{pmatrix}$.

Vorzeichenregel:
Die Vorzeichenregel besagt, dass dieses Vorzeichen auch über

$$sgn \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix} sgn \begin{pmatrix} 1 & 2 & \dots & n \\ j_1 & j_2 & \dots & j_n \end{pmatrix}$$

bestimmt werden kann.

$\underline{\text{1. Beweis:}}$

$\sigma = \begin{pmatrix} i_1 & i_2 & \dots & i_n \\ j_1 & j_2 & \dots & j_n \end{pmatrix}$, $\mu = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}$, $\nu = \begin{pmatrix} 1 & 2 & \dots & n \\ j_1 & j_2 & \dots & j_n \end{pmatrix}$

Müssen zeigen

$$(-1)^{N(\sigma)} = (-1)^{N(\mu)+N(\nu)}$$

Für jedes Paar (k,l) mit $1 \le k < l \le n$ erhalten wir folgende Beiträge zur Anzahl der

Unordnungen:

Fall	Beitrag zu		
	N(σ)	N(μ)	N(ν)
$i_k < i_l$, $j_k < j_l$	0	0	0
$i_k < i_l$, $j_k > j_l$	1	0	1
$i_k > i_l$, $j_k < j_l$	1	1	0
$i_k > i_l$, $j_k > j_l$	0	1	1

Erhalten also

$$N(\sigma) = N(\mu) + N(\nu) mod 2.$$

#

2. Beweis:
Haben

$$\underbrace{\begin{pmatrix} 1 & 2 & \dots & n \\ j_1 & j_2 & \dots & j_n \end{pmatrix}}_{\nu} = \underbrace{\begin{pmatrix} i_1 & i_2 & \dots & i_n \\ j_1 & j_2 & \dots & j_n \end{pmatrix}}_{\sigma} \underbrace{\begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}}_{\mu}$$

d.h. $\nu = \sigma * \mu$ und somit
sgn ν = sgn σ sgn μ (Satz 5 aus 2.4.)
$\Rightarrow$ sgn μ sgn ν = sgn σ $\underbrace{sgn\ \mu sgn\ \mu}_{=1}$ = sgn σ. #

Beispiel:
n = 4
$a_{21}a_{34}a_{13}a_{42}$
Was hat dieses Produkt für ein Vorzeichen?
Def. 3 $\Rightarrow$ sgn $\begin{pmatrix} 2 & 3 & 1 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix} = (-1)^3$ = -1
Vorzeichenregel
$\Rightarrow$ sgn $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}$ sgn $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}$
= 1 * (-1) = -1

Das Vorzeichen von $a_{\sigma(1)1}a_{\sigma(2)2}\dots a_{\sigma(n)n}$ ist nach Vorzeichenregel gleich

$$\underbrace{sgn \begin{pmatrix} 1 & \dots & n \\ \sigma(1) & \dots & \sigma(n) \end{pmatrix}}_{sgn\ \sigma} \underbrace{sgn \begin{pmatrix} 1 & \dots & n \\ 1 & \dots & n \end{pmatrix}}_{=1}$$

d.h. gleich dem Vorzeichen von $a_{1\sigma(1)}a_{2\sigma(2)}\dots a_{n\sigma(n)}$.
Also ist die Determinante gleich, egal, ob man sie nach aufsteigenden Zeilenindizes oder nach aufsteigenden Spaltenindizes entwickelt.

Für praktische Zwecke sind diese Formeln zur Berechnung einer Determinante ungeeignet.
Insbesondere die Differenzen sind numerisch kritisch (viele Rundungsfehler).
($\sqrt{10^{14}+1} - \sqrt{10^{14}} \approx 10^7 - 10^7 = 0$
$\sqrt{10^{14}+1} - \sqrt{10^{14}} = \frac{1}{\sqrt{10^{14}+1}+\sqrt{10^{14}}} \approx \frac{1}{10^7+10^7} = \frac{1}{2} * 10^{-7}$ $\rightarrow$ wesentlich genauer)

Man geht daher anders vor.

Definition 4 *Sei $A = (a_{ij})_{i,j=1}^n \in M_n(\mathcal{K})$. Die Matrix $B = (b_{ij})_{i,j=1}^n \in M_n(\mathcal{K})$ mit $b_{ij} = a_{ji}$ heißt die zu A transponierte Matrix und wird mit A^T bezeichnet.*

23.01.06

Haben also

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} a_{11} & \dots & a_{n1} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{nn} \end{pmatrix}.$$

A^T entsteht aus A durch Vertauschen von Zeilen und Spalten oder, was dazu äquivalent ist, durch Spiegelung an der *Hauptdiagonalen* (das ist die von links oben nach rechts unten).

Satz 1 *Es gilt det A = det A^T.*

Beweis:
Sei A = (a_{ij}), $A^T = (b_{ij})$. Dann ist

$$\begin{aligned} det\ A &= \sum_{\sigma \in S_n} (sgn\ \sigma) a_{\sigma(1)1} \dots a_{\sigma(n)n} \\ &= \sum_{\sigma \in S_n} (sgn\ \sigma) b_{1\sigma(1)} \dots b_{n\sigma(n)} \\ &= det A^T. \end{aligned}$$

#

Sei A = $(a_{ij}) \in M_n(\mathcal{K})$. Bezeichnen die Spalten von A mit $s_1, \dots, s_n$ und schreiben A = $(s_1 \dots s_n)$.

Satz 2 *Die Determinante ist eine alternierende Multilinearform ihrer Spalten, d.h. es gilt*
(a) det $(s_1, \dots, s_i' + s_i'', \dots, s_n)$ = det $(s_1, \dots, s_i', \dots, s_n)$ + det $(s_1, \dots, s_i'', \dots, s_n)$;
(b) det $(s_1, \dots, \lambda s_i, \dots, s_n) = \lambda$ det $(s_1, \dots, s_i, \dots, s_n)$
(c) det $(s_{\sigma(1)} \dots s_{\sigma(n)}) = sgn(\sigma)$ det $(s_1 \dots s_n)$.

Beweis:
(a) det $(s_1, \dots, s_i' + s_i'', \dots, s_n)$

$$\begin{aligned} &= \sum_{\sigma \in S_n} (sgn\ \sigma) a_{\sigma(1)1} \dots (a'_{\sigma(i)i} + a''_{\sigma(i)i}) \dots a_{\sigma(n)n} \\ &= \sum_{\sigma \in S_n} (sgn\ \sigma) a_{\sigma(1)1} \dots a'_{\sigma(i)i} \dots a_{\sigma(n)n} \\ &+ \sum_{\sigma \in S_n} (sgn\ \sigma) a_{\sigma(1)1} \dots a''_{\sigma(i)i} \dots a_{\sigma(n)n} \end{aligned}$$

$= \det(s_1 \dots s_i' \dots s_n) + \det(s_1 \dots s_i'' \dots s_n)$.

(b) $\det(s_1 \dots \lambda s_i \dots s_n)$

$$= \sum_{\sigma \in S_n} (sgn\ \sigma) a_{\sigma(1)1} \dots (\lambda a_{\sigma(i)i}) \dots a_{\sigma(n)n}$$

$$= \lambda \sum_{\sigma \in S_n} (sgn\ \sigma) a_{\sigma(1)1} \dots a_{\sigma(i)i} \dots a_{\sigma(n)n} = \lambda det(s_1 \dots s_i \dots s_n).$$

(c) Sei A $= (s_1 \dots s_n) = (a_{ij})$, B $= (s_{\sigma(1)} \dots s_{\sigma(n)}) = (b_{ij})$.
Dann ist $b_{ij} = a_{i\sigma(j)}$ und somit

$\det B = \sum_{\tau \in S_n} (sgn\ \tau) b_{1\tau(1)} \dots b_{n\tau(n)}$

$$= \sum_{\tau \in S_n} (sgn\ \tau) a_{1\sigma(\tau(1))} \dots a_{n\sigma(\tau(n))}$$

$$= \sum_{\tau \in S_n} (sgn\ \tau) a_{1(\sigma\tau)(1)} \dots a_{n(\sigma\tau)(n)}$$

$$= sgn\ \sigma \sum_{\tau \in S_n} (sgn\ \sigma)(sgn\ \tau) a_{1(\sigma\tau)(1)} \dots a_{n(\sigma\tau)(n)}$$

$$= sgn\ \sigma \sum_{\tau \in S_n} (sgn\ (\sigma\tau)) a_{1(\sigma\tau)(1)} \dots a_{n(\sigma\tau)(n)}$$

$$= sgn\ \sigma \sum_{\mu \in S_n} (sgn\ \mu) a_{1\mu(1)} \dots a_{n\mu(n)}$$

$= \text{sgn}\ \sigma * \det A.$ #

Bemerkung: 1. Eigenschaften (a) und (b) gesagen, dass det eine Multilinearform ist. Es gilt insbesondere

$$\begin{vmatrix} a_1 + a_2 & b \\ c_1 + c_2 & d \end{vmatrix} = \begin{vmatrix} a_1 & b \\ c_1 & d \end{vmatrix} + \begin{vmatrix} a_2 & b \\ c_2 & d \end{vmatrix},$$

$$\begin{vmatrix} a_1 + a_2 & b_1 + b_2 \\ c_1 + c_2 & d_1 + d_2 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 + b_2 \\ c_1 & d_1 + d_2 \end{vmatrix} + \begin{vmatrix} a_2 & b_1 + b_2 \\ c_2 & d_1 + d_2 \end{vmatrix}$$

$$= \begin{vmatrix} a_1 & b_1 \\ c_1 & d_1 \end{vmatrix} + \begin{vmatrix} a_1 & b_2 \\ c_1 & d_2 \end{vmatrix} + \begin{vmatrix} a_2 & b_1 \\ c_2 & d_1 \end{vmatrix} + \begin{vmatrix} a_2 & b_2 \\ c_2 & d_2 \end{vmatrix},$$

$$\lambda \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{pmatrix}$$

$$\lambda \begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} \lambda a & b \\ \lambda c & d \end{vmatrix} = \begin{vmatrix} a & \lambda b \\ c & \lambda d \end{vmatrix},$$

$$\begin{vmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{vmatrix} = \lambda^2 \begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

Aus (b) folgt auch, dass det a $= 0$ ist, wenn eine Spalte von A Null ist:
$det(s_1 \dots 0(i) \dots s_n) = det(s_1 \dots 0(i) * \mathcal{O} \dots s_n)$
$= 0 * det(s_1 \dots \mathcal{O} \dots s_n) = 0.$

2. Eigenschaft (c) besagt, dass die Determinante alternierend ist. Es gilt insbesondere:

$$det(s_1 \dots s_i \dots s_i \dots s_n) = -det(s_1 \dots s_i \dots s_i \dots s_n)$$

$$\Rightarrow det(s_1 \dots s_i \dots s_i \dots s_n) = 0.$$

Es folgt auch, dass sich die Determinante nicht ändert, wenn man zu einer Spalte ein beliebiges Vielfaches einer anderen Spalte addiert:

$$det(s_1 \dots s_i + \lambda s_j(i) \dots s_j(j) \dots s_n)$$

$$= det(s_1 \dots s_i(i) \dots s_j(j) \dots s_n) + \underbrace{\lambda det(s_1 \dots s_j(i) \dots s_j(j) \dots s_n)}_{=0}.$$

3. Wegen Satz 1 gilt alles, was oben gesagt wurde, auch für Zeilen.
Eine Determinante ist also eine alternierende Multilinearform ihrer Zeilen. Insbesondere gilt:
- Determinante ist Null, wenn eine Zeile null ist.
- Determinante ist Null, wenn sie zwei gleiche Zeilen hat.
- Determinante ändert das Vorzeichen, wenn man zwei Zeilen miteinander vertauscht.
- Determinante ändert sich nicht, wenn man zu einer Zeile ein beliebiges Vielfaches einer anderen Zeile addiert.

4. Man kann folgendes zeigen. Ist:
$\varphi : M_n(\mathcal{K}) \to \mathcal{K}$
eine alternierende Multilinearform der Spalten mit φ(I) = 1 [vermerken, dass det I = ist], so gilt φ(A) = det A $\forall$ A $\in$ M_n ($\mathcal{K}$). Diese Eigenschaft heißt *Eindeutigkeit der Determinante.*

Definition 5 *Matrizen der Gestalt*

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & a_{nn} \end{pmatrix} \text{ bzw. } \begin{pmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ \vdots & \vdots & \ddots & 0 \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

heißen obere bzw. untere Dreiecksmatrizen.

Haben also A = (a_{ij}) obere Dreicksmatrix $\Leftrightarrow$ $a_{ij} = 0$ für i > j;
A = (a_{ij}) untere Dreiecksmatrix $\Leftrightarrow$ $a_{ij} = 0$ für i <j;

Ist A obere oder untere Dreiecksmatrix, so ist det A gleich dem Produkt der Einträge auf der Hauptdiagonalen. In der Tat sei etwa A = $(a_{ij})_{i,j=1}^n$ eine obere Dreiecksmatrix. Dann ist

$$det\ A = \sum_{\sigma \in\ S_n} (sgn\ \sigma) a_{1\sigma(1)} \dots a_{n\sigma(n)}$$

und das Produkt $a_{1\sigma(1)} \dots a_{n\sigma(n)}$ ist von Null verschieden nur für $1 \leq \sigma(1)$, ..., n $\leq$ σ(n), d.h. für $\sigma(1) = 1$, ..., σ(n) = n. Die Summe reduziet sich also auf
det A = $\begin{pmatrix} 1 & \dots & n \\ 1 & \dots & n \end{pmatrix} a_{11} \dots a_{nn} = a_{11} \dots a_{nn}$. [analog für untere Dreiecksmatrizen]

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & 12 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{vmatrix} = 0.$$

Definition 6 *Seien A_{11}, …, A_{nn} quadratische Matrizen (nicht notwendigerweise der gleichen Größe). Eine quadratische Matrix der Form*

$$\begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ 0 & A_{22} & \dots & A_{2n} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & A_{nn} \end{pmatrix} \text{ bzw. } \begin{pmatrix} A_{11} & 0 & 0 & 0 \\ A_{21} & A_{22} & 0 & 0 \\ \vdots & \vdots & \ddots & 0 \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{pmatrix}$$

heißt obere bzw. untere Blockdreiecksmatrix.

Beispiele:

$$\left(\begin{array}{cc|cc} 1 & 2 & 5 & 6 \\ 3 & 4 & 7 & 8 \\ \hline 0 & 0 & 9 & 1 \\ 0 & 0 & 3 & 4 \end{array}\right)$$

$$\left(\begin{array}{c|cc} 1 & 2 & 3 \\ \hline 0 & 4 & 5 \\ 0 & 6 & 7 \end{array}\right)$$

$$\left(\begin{array}{c|c|c} 1 & 2 & 3 \\ \hline 0 & 4 & 5 \\ \hline 0 & 0 & 6 \end{array}\right) = \left(\begin{array}{c|cc} 1 & 2 & 3 \\ \hline 0 & 4 & 5 \\ 0 & 0 & 6 \end{array}\right)$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = (A_{11})$$

⇒ Jede Matrix ist eine Blockdreiecksmatrix.

Satz 3 *Ist A eine Blockdreiecksmatrix mit den Blöcken A_{11},…, A_{nn} auf der Hauptdiagonalen (so wie in Definition 6), so gilt*

det A = (det A_{11})(det A_{12})…(det A_{nn})

27.01.06

Beweis:

Sei zunächst n = 2 und A = $\begin{pmatrix} B & C \\ 0 & D \end{pmatrix} \begin{matrix} m \\ n \end{matrix}$
m n

Setzen A = (a_{ij}), B = (b_{ij}), C = (c_{ij}), D = (d_{ij}).
Haben det A = $\sum_{\sigma \in S_{m+n}} (sgn\ \sigma) a_{1\sigma(1)} \dots a_{m\sigma(m)} a_{m+1\sigma(m+1)} \dots a_{m+n\sigma(m+n)}$.
Wegen $a_{ij} = 0$ für i > m und j ≤ m ist
$\{\sigma(m+1), \dots, \sigma(m+n)\}$ = {m+1, …, m+n}
und damit
$\{\sigma(1), \dots, \sigma(m)\}$ = {1, …, m}

Sei σ(m+1) = m + τ(1), …, σ(m+n) = m + τ(n) mit $\tau \in S_n$.
Erhalten so

$$detA = \sum_{\sigma \in S_n, \tau \in S_n} sgn \begin{pmatrix} 1 & \dots & m & m+1 & \dots & m+n \\ \sigma(1) & \dots & \sigma(m) & m+\tau(1) & \dots & m+\tau(n) \end{pmatrix} b_{1\sigma(1)} \dots b_{m\sigma(m)} d_{1\tau(1)} \dots d_{n\tau(n)}$$

und wegen

$$sgn \begin{pmatrix} 1 & \dots & m & m+1 & \dots & m+n \\ \sigma(1) & \dots & \sigma(m) & m+\tau(1) & \dots & m+\tau(n) \end{pmatrix}$$

$$= sgn \begin{pmatrix} 1 & \dots & m & m+1 & \dots & m+n \\ \sigma(1) & \dots & \sigma(m) & m+1 & \dots & m+n \end{pmatrix} sgn \begin{pmatrix} 1 & \dots & m & m+1 & \dots & m+n \\ 1 & \dots & m & m+\tau(1) & \dots m+\tau(n) & \end{pmatrix}$$

= sgn σ sgn τ
ergibt sich

$$det\ A = \sum_{\sigma \in S_n, \tau \in S_n} (sgn\ \sigma)(sgn\ \tau) b_{1\sigma(1)} \dots b_{m\sigma(m)} d_{1\tau(1)} \dots d_{n\tau(n)}$$

$$= \left(\sum_{\sigma \in S_n} (sgn\ \sigma) b_{1\sigma(1)} \dots b_{m\sigma(m)} \right) \left(\sum_{\tau \in S_n} (sgn\ \tau) d_{1\tau(1)} \dots d_{n\tau(n)} \right)$$

= det B * det B.
Für allgemeines n ergibt sich die Behauptung durch sukzessive Anwendung des Resultates für n = 2:

$$\det \left(\begin{array}{c|ccc} A_{11} & A_{12} & \dots & A_{1n} \\ \hline & A_{22} & \dots & A_{2n} \\ & & \ddots & \vdots \\ & & & A_{nn} \end{array} \right) = \det\ A_{11} * \det \begin{pmatrix} A_{22} & \dots & A_{2n} \\ & \ddots & \vdots \\ & & A_{nn} \end{pmatrix}$$

$= \dots = (\det\ A_{11}) \dots (\det\ A_{nn})$. #

Gaußscher Algorithmus zur Berechnung von Determinanten

Sei A = $\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$.
Ist die erste Zeile Null, so ist det A = 0.
Möge die erste Zeile also ein von Null verschiedenes Element enthalten. Ist $a_{11} \neq 0$, so lassen wir A unverändert. Ist $a_{11} = 0$, so nehmen wir ein $a_{1j} \neq 0$ und vertauschen die erste und die j-te Spalte.
Erhalten so

$$\det A = \pm \det \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nn} \end{pmatrix} \text{ mit } b_{11} \neq 0.$$

Multiplizieren dann für j = 2, ..., n die erste Zeile mit $-\frac{b_{j1}}{b_{11}}$ und addieren das Resultat zur j-ten Zeile.

Wegen $b_{j1} + (-\frac{b_{j1}}{b_{11}}) b_{11} = 0$
erhalten wir

$$\det \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nn} \end{pmatrix}$$

$$= \det \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ 0 & c_{11} & \dots & c_{1,n-1} \\ \vdots & \vdots & & \vdots \\ 0 & c_{n-1,1} & \dots & c_{n-1,n-1} \end{pmatrix}$$

Wiederholen das Obige für det $\begin{pmatrix} c_{11} & \dots & c_{1,n-1} \\ \vdots & & \vdots \\ c_{n-1,1} & \dots & c_{n-1,n-1} \end{pmatrix}$ usw.

Zum Schluss ergibt sich eine Dreiecksmatrix, deren Determinante gleich dem Produkt der Einträge auf der Hauptdiagonalen ist.

Beispiel:

$$\begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{vmatrix} = - \begin{vmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{vmatrix} = - \left| \begin{array}{c|ccc} 1 & 0 & 1 & 1 \\ \hline 0 & 1 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{array} \right| = - \begin{vmatrix} 1 & 1 & 1 \\ -1 & -1 & 0 \\ 1 & 0 & -1 \end{vmatrix}$$

$$= - \left| \begin{array}{c|cc} 1 & 1 & 1 \\ \hline 0 & -2 & -1 \\ 0 & -1 & -2 \end{array} \right| = - \begin{vmatrix} -2 & -1 \\ -1 & -2 \end{vmatrix} = - \left| \begin{array}{c|c} -2 & -1 \\ \hline 0 & \frac{-3}{2} \end{array} \right| = -(-2) * (\frac{3}{2}) = \underline{\underline{-3}}$$

Satz 4 *Produktsatz für Determinanten von Cauchy*
Für A,B ∈ M_n (K) gilt det (AB) = (det A)(det B).

Beweis:
Betrachten die Blockmatrix

$$\mathrm{C} = \begin{pmatrix} A & 0 \\ -I & B \end{pmatrix} \begin{matrix} n \\ n \end{matrix}$$
$$\quad \mathrm{n} \quad \mathrm{n}$$

Nach Satz 3 ist det C = (det A)(det B).
Durch entsprechendes Addieren von Vielfachen der letzten n Zeilen zu den ersten n Zeilen lässt sich C in die Matrix

$$\begin{pmatrix} 0 & AB \\ -I & B \end{pmatrix}$$

überführen.
Somit ist det C = det $\begin{pmatrix} 0 & AB \\ -I & B \end{pmatrix}$ = (-1)^n det $\begin{pmatrix} AB & 0 \\ B & -I \end{pmatrix}$
= (-1)^n det (AB) * det (-I)
= (-1)^n det (AB) * (-1)^n
= det (AB). #

Bemerkung:
Ein anderer Beweis geht wie folgt:
Man zeigt, dass $\varphi(A) = \frac{det(AB)}{det(B)}$ eine alternierende Multilinearform der Spalten in A ist.
Wegen φ(I) = 1 folgt aus der Eindeutigkeit der Determinante φ(A) = det(A).

Formel von Cauchy-Binet

Sei A $\in$ $M_{n,m}$ ($\mathcal{K}$) und B $\in$ $M_{m,n}$ ($\mathcal{K}$). Dann ist AB $\in$ M_n ($\mathcal{K}$). Für m = n ist det (AB) = (det A)(det B) nach Satz 4.
Man kann zeigen, dass det (AB) = 0 für m < n.
Für n > m kann man die Formel von Cauchy-Binet benutzen.
Bezeichnen mit A$\begin{pmatrix} 1 & \dots & n \\ j_1 & \dots & j_n \end{pmatrix}$ die n x n-Matrix, die aus den Spalten j_1, ..., j_n von A gebildet wird, und mit B $\begin{pmatrix} j_1 & \dots & j_n \\ 1 & \dots & n \end{pmatrix}$ die n x n-Matrix, die aus den Zeilen j_1, ..., j_n von B gebildet wird. Die Formel von Cauchy-Binet besagt

$$det(AB) = \sum_{1 \le j_1 < \dots j_n \le m} det\ A \begin{pmatrix} 1 & \dots & n \\ j_1 & \dots & j_n \end{pmatrix} det\ B \begin{pmatrix} j_1 & \dots & j_n \\ 1 & \dots & n \end{pmatrix}$$

Beispiel:

$$det \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \begin{pmatrix} x & y \\ z & u \\ v & x \end{pmatrix} = \begin{vmatrix} a & b \\ d & e \end{vmatrix} \begin{vmatrix} x & y \\ z & u \end{vmatrix} + \begin{vmatrix} a & c \\ d & f \end{vmatrix} \begin{vmatrix} x & y \\ v & w \end{vmatrix} + \begin{vmatrix} b & c \\ e & f \end{vmatrix} \begin{vmatrix} z & u \\ v & w \end{vmatrix}.$$

Definition 7 *Sei $A = (a_{ij})_{i,j=1}^n \in M_n$ ($\mathcal{K}$). Die Determinante der (n-1)x(n-1)-Matrix, die durch Streichen der i-ten Zeile und der j-ten Spalte von A entsteht, heißt der zu a_{ij} gehörige Minor und wird mit M_{ij} bezeichnet. Die Zahl $(-1)^{i+j} * M_{ij}$ heißt die zu a_{ij} gehörende Adjunkte (= das algebraische Komplement von a_{ij}) und wird mit A_{ij} bezeichnet.*

30.01.06

Beispiel:

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & k \end{pmatrix}$$

$$M_{11} = \begin{vmatrix} e & f \\ h & k \end{vmatrix}, M_{12} = \begin{vmatrix} d & f \\ g & k \end{vmatrix}, M_{21} = \begin{vmatrix} b & c \\ h & k \end{vmatrix}, M_{22} = \begin{vmatrix} a & c \\ g & k \end{vmatrix}$$

$$A_{11} = \begin{vmatrix} e & f \\ h & k \end{vmatrix}, A_{12} = -\begin{vmatrix} d & f \\ g & k \end{vmatrix}, A_{21} = -\begin{vmatrix} b & c \\ h & k \end{vmatrix}, A_{22} = \begin{vmatrix} a & c \\ g & k \end{vmatrix}$$

Satz 5 *Entwicklungssatz*
Sei $A = (a_{ij})_{i,j=1}^n \in M_n$ ($\mathcal{K}$). Dann gilt

$$a_{i1}A_{k1} + a_{i2}A_{k2} + \dots + a_{in}A_{kn} = \delta_{ik} det\ A,$$

$$a_{1j}A_{1k} + a_{2j}A_{2k} + \dots + a_{nj}A_{nk} = \delta_{jk} det\ A,$$

wobei δ_{lk} das Kroneckersymbol ist.

Kommentar:
Sei i = k. Dann erhalten wir:

det A $= a_{i1}A_{i1} + \ldots + a_{in}A_{in}$
Also: Die Determinante von A ist gleich der Summe der Produkte der Elemente einer Zeile mit ihren Adjunkten.

n = 2: $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = a_{11}A_{11} + a_{12}A_{12} = ad + b(-c) = ad - bc.$

n = 3: $\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & k \end{vmatrix} = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13} +$

$= a * \begin{vmatrix} e & f \\ h & k \end{vmatrix} - b * \begin{vmatrix} d & f \\ g & k \end{vmatrix} + c * \begin{vmatrix} d & e \\ g & h \end{vmatrix}$

Analog det A = Summe der Produkte der Elemente einer Spalte mit ihren Adjunkten (j = k).

Schachbrettregel für die Vorzeichen

$$\begin{vmatrix} + & - & + & - & \ldots \\ - & + & - & + & \ldots \\ + & - & + & - & \ldots \\ - & + & - & + & \ldots \end{vmatrix}$$

allgemeines n, Entwicklung nach der ersten Zeile

$$det \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}$$

$$= a_{11} \begin{vmatrix} a_{22} & \ldots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \ldots & a_{nn} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n3} & \ldots & a_{nn} \end{vmatrix} + \ldots - \ldots + (-1)^{n+1} a_{1n} \begin{vmatrix} a_{21} & \ldots & a_{2,n-1} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{n,n-1} \end{vmatrix}.$$

Für i $\neq$ k ($\Leftrightarrow \delta_{ik} = 0$) erhalten wir:
Die Summe der Produkte der Elemente einer Zeile mit den Adjunkten der entsprechenden Elemente einer anderen Zeile ist immer Null. Analog für Spalten (j $\neq$ k).

Beispiel:
n = 2 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, i = 1, k = 2,

$$a_{11}A_{21} + a_{12}A_{22} = a(-b) + b(a) = 0.$$

Beweis:
Zeigen

$$a_{1j}A_{1j} + \ldots + a_{nj}A_{nj} = det\ A$$

(Satz 1 liefert dann $a_{i1}A_{i1} + \ldots + a_{in}A_{in} =$ det A).

Nehmen zunächst an, dass in der j-ten Spalte von A nur a_{ij} und ansonsten Nullen stehen.

$$\begin{matrix} \\ \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} (j) \\ \end{matrix}$$

$$\left.\begin{pmatrix} 0 \\ \vdots \\ 0 \\ a_{ij} \\ 0 \\ \vdots \\ 0 \end{pmatrix}\right. (i)$$

Vertauschen Zeilen i mit i-1, dann i-1 und i-2, ..., 2 mit 1.

Erhalten $(-1)^{(i-1)}$ det A = det $\begin{pmatrix} a_{ij} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

Vertauschen nun Spalten j mit j-1, dann j-1 mit j-2, ..., 2 mit 1.

Es ergibt sich $(-1)^{(j-1)}(-1)^{(i-1)}$ det A = det $\left(\begin{array}{c|c} a_{ij} & \\ \hline 0 & \\ \vdots & \tilde{M}_{ij} \\ 0 & \end{array}\right)$

$= a_{ij}$ det $\tilde{M}_{ij}$ (Satz 3)
$= a_{ij}M_{ij}$.
Also det A $= (-1)^{i+j}a_{ij}M_{ij} = a_{ij}A_{ij}$.

Für eine beliebige j-te Spalte haben wir

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix} = \begin{pmatrix} a_{1j} \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ a_{2j} \\ \vdots \\ 0 \end{pmatrix} + \ldots + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ a_{nj} \end{pmatrix}.$$

und Satz 2 und Obiges liefert

$$det\ A = a_{1j} * A_{1j} + \ldots + a_{nj} * A_{nj}.$$

Sei nun j $\neq$ k. Dann ist

$$\begin{matrix} (j) & (k) \\ \end{matrix}$$

$$0 = det \begin{pmatrix} a_{1j} & a_{1j} \\ \vdots & \vdots \\ a_{nj} & a_{nj} \end{pmatrix} = a_{1j}A_{1k} + \ldots + a_{nj}A_{nk}.$$

Entwicklungssatz von Laplace
Dies ist eine Verallgemeinerung von Satz 5. Sei A $\in M_n(\mathcal{K})$.
Für $1 \leq i_1 < \ldots < i_l \leq$ n und für $1 \leq j_1 < \ldots < j_n \leq$ n bezeichnen wir mit $A\begin{pmatrix} i_1 & \ldots & i_l \\ j_1 & \ldots & j_l \end{pmatrix}$ die Determinante der l x l-Matrix, die aus dem Schnitt der Zeilen i_1,...,i_l und der Spalten j_1, ..., j_l gebildet wird.
Definieren desweiteren $i_1^{'} < \ldots < i_{n-l}^{'}$ und $j_1^{'} < \ldots < j_{n-l}^{'}$ durch $\{i_1^{'}, \ldots, i_{n-l}^{'}\} = \{1,\ldots,n\}\backslash\{i_1, \ldots, i_l\}$,
$\{j_1^{'}, \ldots, j_{n-l}^{'}\} = \{1,\ldots,n\} \setminus \{j_1, \ldots, j_l\}$.

Der Entwicklungssatz von Laplace besagt, dass für fixierte $1 \leq i_1 < \ldots < i_l \leq$ n die Formel

$$det\ A = \sum_{1 \leq j_1 < \ldots < j_l \leq n} (-1)^{i_1+\ldots+i_l+j_1+\ldots+j_l} A\begin{pmatrix} i_1 & \ldots & i_l \\ j_1 & \ldots & j_l \end{pmatrix} A\begin{pmatrix} i_1^{'} & \ldots & i_{n-l}^{'} \\ j_1^{'} & \ldots & j_{n-l}^{'} \end{pmatrix}$$

gilt.

Beweis: Literatur (Gantmacher).
Beispiel:

Sei z.B. A $= \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$.
Wählen l = 1, d.h. fixieren ein $i_1 :=$ i.
Laplace lautet dann (n = 4):

$$det\ A = \sum_{i \leq j \leq n} (-1)^{i+j} A\begin{pmatrix} i \end{pmatrix} A\begin{pmatrix} i_1^{'} & i_2^{'} & i_3^{'} \\ j_1^{'} & j_2^{'} & j_3^{'} \end{pmatrix}$$

mit $\{i_1^{'}, i_2^{'}, i_3^{'}\} = \{1,\ldots,4\}\backslash$ {i}
$\{j_1^{'}, j_2^{'}, j_3^{'}\} = \{1,\ldots,4\} \setminus$ {j}

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

d.h. erhalten Satz 5 (Entwicklung nach der i-ten Zeile).

Wählen nun $\{i_1,i_2\} = \{1,2\}$. Erhalten dann

$$det\ A = \sum_{1 \leq j_1 < j_2 \leq 4} (-1)^{1+2+j_1+j_2} A\begin{pmatrix} 1 & 2 \\ j_1 & j_2 \end{pmatrix} A\begin{pmatrix} 3 & 4 \\ j_1^{'} & j_2^{'} \end{pmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \begin{vmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{vmatrix} - \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} \begin{vmatrix} a_{32} & a_{34} \\ a_{42} & a_{44} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{14} \\ a_{41} & a_{44} \end{vmatrix} \begin{vmatrix} a_{32} & a_{33} \\ a_{42} & a_{43} \end{vmatrix}$$

$$+ \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \begin{vmatrix} a_{31} & a_{34} \\ a_{41} & a_{44} \end{vmatrix} - \begin{vmatrix} a_{12} & a_{14} \\ a_{22} & a_{24} \end{vmatrix} \begin{vmatrix} a_{31} & a_{33} \\ a_{41} & a_{43} \end{vmatrix} + \begin{vmatrix} a_{13} & a_{14} \\ a_{23} & a_{24} \end{vmatrix} \begin{vmatrix} a_{31} & a_{32} \\ a_{41} a_{42} & \end{vmatrix}.$$

Satz 6 über die inverse Matrix:
Eine Matrix $A \in M_n(\mathcal{K})$ ist genau dann invertierbar, wenn det $A \neq 0$ ist. In diesem Fall ist

$$A^{-1} = \frac{1}{det\ A}(A_{ji})_{i,j=1}^{n},$$

wobei A_{ji} das algebraische Komplement von a_{ji} ist.

Bemerkung:
Der (i,j)-Eintrag von A^{-1} ist also $\frac{1}{det\ A}\ A_{ji}$ (und nicht $\frac{1}{det\ A}\ A_{ij}$).
n = 2. Haben

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{det\ A}\begin{pmatrix} A_{11} & A_{21} \\ A_{21} & A_{22} \end{pmatrix} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Beweis:
Sei A invertierbar und B die Inverse. Dann ist AB = I und somit det A * det B = det I = 1, d.h. det A $\neq$ 0.
Sei nun det A $\neq$ 0.
Zeigen, dass $\frac{1}{det\ A}\ (A_{ji})_{i,j=1}^{n} =:$ B die Inverse ist.
Der (i,j)-Eintrag von A * B ist

$$(a_{i1} \dots a_{in})\frac{1}{det\ A}\begin{pmatrix} A_{j1} \\ \vdots \\ A_{jn} \end{pmatrix} = \frac{1}{det\ A}(a_{i1}A_{j1} + \dots + a_{in}A_{jn})$$

$$= \frac{1}{det\ A}\delta_{ij}A = \delta_{ij},$$

d.h. AB = I. Analog zeigt man BA = I. #
Betrachten schließlich ncoh das quadratische Gleichungssystem

$$\begin{array}{lllll} a_{11}x_1+ & \dots & +a_{1n}x_n & = & b_1 \\ \vdots & & & & \\ a_{n1}x_1+ & \dots & +a_{nn}x_n & = & b_n \end{array}$$

d.h.

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

oder kompakt Ax = b.

Satz 7 Cramersche Regel
Sei $A \in M_n(\mathcal{K})$ und det $A \neq 0$. Dann hat das System $Ax = b$ für jedes $b \in \mathcal{K}^n$ genau eine Lösung $x \in \mathcal{K}^n$, nämlich $x = A^{-1}b$.
Die k-te Komponente x_k von x ist gegeben durch $x_k = \frac{det\ A_k}{det\ A}$, wobei A_k aus A entsteht, indem man die k-te Spalte durch die Spalte b ersetzt.

Beweis:
Sei Ax = b. Dann ist x = A^{-1}Ax = A^{-1}b, d.h. es gibt höchstens A^{-1}b als Lösung.
Zeigen, dass A^{-1}b wirklich eine Lösung ist:
A(A^{-1}b) = b.
Berechnen die k-te Komponente von A^{-1}b nach Satz 6:
x_k = k-te Zeile von A^{-1} mal b

$$= \frac{1}{det\ A}(A_{1k} \dots A_{nk}) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \frac{1}{det\ A}(b_1 A_{1k} \dots b_n A_{nk}) = \frac{1}{det\ A} \begin{vmatrix} a_{11} & \dots & b_1 & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & b_n & \dots & a_{nn} \end{vmatrix} = \frac{1}{det\ A} det\ A_k.$$

#

Matrizengruppen
Die Menge aller invertierbaren Matrizen aus M_n ($\mathcal{K}$) bilden eine Gruppe bezüglich der Matrizenmultiplikation:
A,B invertierbar $\Leftrightarrow$ AB invertierbar ($B^{-1}A^{-1}$ ist die Inverse von AB),
Matrizenmultiplikation ist assoziativ,
I ist Einselement,
jede invertierbare Matrix ist invertierbar als Element der Gruppe.
Diese Gruppe wird mit GL(n,$\mathcal{K}$) bezeichnet und *allgemeine* (oder generelle oder vollständige) *lineare Gruppe* über n und $\mathcal{K}$ genannt.

Haben also GL(n,$\mathcal{K}$) = {A $\in$ M_n ($\mathcal{K}$): det A $\neq$ 0}
Die Abbildung
det: GL(n,$\mathcal{K}$) $\to$ $\mathcal{K}\backslash\{0\}$, A $\mapsto$ det A
ist ein Gruppenhomomorphismus ist φ: G $\to$ H ein Gruppenhomomorphismus, so ist das vollständige Urbild φ^{-1}(H) stets eine Untergruppe von G [Beweis: HA].

Auf diese Weise gewinnt man interessante Untergruppen von GL(n,$\mathcal{K}$).
SL(n,$\mathcal{K}$) = {A $\in$ GL(n,$\mathcal{K}$): det A = 1} = {A $\in$ M_n ($\mathcal{K}$): det A = 1}
$\to$ *spezielle lineare Gruppe* über n und $\mathcal{K}$

{A $\in$ M_n ($\mathcal{K}$): |det A| = 1} (={A $\in$ GL(n,$\mathcal{K}$): |det A| = 1}) wird zu
{A $\in$ M_n ($\mathcal{R}$): det A = $\pm$ 1}
{A $\in$ M_n ($\mathcal{C}$): det A $\in$ $\mathcal{T}$} ($\mathcal{T}$ = {z $\in$ $\mathcal{C}$: |z| = 1}),
{A $\in$ M_n ($\mathcal{R}$): det A > 0}.
Für die letzten drei Gruppen sind keine besonderen Namen geläufig.

2.6 Basiswechsel, Äquivalenz und Ähnlichkeit von Matrizen

Sei X ein linearer Raum der Dimension n. Seien E = {e_1, ..., e_n} und F = {f_1,...,f_n} zwei Basen in X. Nach Satz 1 aus 2.2. gibt es eindeutig bestimmte Zahlen $\alpha_{ij} \in \mathcal{K}$ mit

$$\begin{aligned} f_1 &= \alpha_{11}e_1 + \dots + \alpha_{1n}e_n \\ &\vdots \\ f_n &= \alpha_{n1}e_1 + \dots + \alpha_{nn}e_n \end{aligned}$$

$$\begin{aligned} e_1 &= \beta_{11}f_1 + \dots + \beta_{1n}f_n \\ &\vdots \\ e_n &= \beta_{n1}f_1 + \dots + \beta_{nn}f_n \end{aligned}$$

Definition 1 *Die Matrizen* $U_{E,F} := (\alpha_{ij})_{i,j=1}^n$ *und* $U_{F,E} := (\beta_{ij})_{i,j=1}^n$ *heißen Übergangsmatrizen.*

Haben also

$$\begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} = U_{E,F} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}, \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = U_{F,E} \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}.$$

Satz 1 *Übergangsmatrizen sind stets invertierbar und es gilt*

$$U_{E,F}^{-1} = U_{F,E}.$$

07.04.06

Beweis:
Haben

$$\begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} = U_{E,F} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \underbrace{U_{E,F}U_{F,E}}_{=I} \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix},$$

d.h. $U_{E,F}U_{F,E} = \mathrm{I}$. Dies ergibt die Invertierbarkeit von $U_{E,F}$ und $U_{E,F}^{-1} = U_{F,E}$. #
Sind $\mathrm{E} = \{e_1, \ldots, e_n\}$ und $\mathrm{F} = \{f_1, \ldots, f_n\}$ Basen in X, so lässt sich jedes Element $\mathrm{x} \in \mathrm{X}$ eindeutig in der Form $\mathrm{x} = \gamma_1 e_1 + \ldots + \gamma_n e_n$, $\mathrm{x} = \delta_1 f_1 + \ldots + \delta_n f_n$ schreiben.
Die aus den Koordinaten gebildeten Spalten bezeichnen wir mit

$$[x]_E = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix}, [x]_F = \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_n \end{pmatrix}.$$

Uns interessiert, wie $[x]_E$ und $[x]_F$ miteinander zusammenhängen. Brauchen dazu ein einfaches, aber wichtiges Lemma.

Lemma:
Für $\mathrm{A,B} \in M_n(\mathcal{K})$ gilt

$$(AB)^T = B^T A^T,$$

und A^T ist genau dann invertierbar, wenn A invertierbar ist, in welchem Falle

$$(A^T)^{-1} = (A^{-1})^T$$

gilt.

Beweis:
Sei $\mathrm{A} = (a_{ij})$, $\mathrm{B} = (b_{ij})$. Dann ist

$$(AB)^T = (a_{i1}b_{1j} + \ldots + a_{in}b_{nj})^T = (a_{j1}b_{1i} + \ldots + a_{jn}b_{ni}),$$

$$B^T A^T = \begin{pmatrix} b_{11} & \ldots & b_{n1} \\ \vdots & & \vdots \\ b_{1n} & \ldots & b_{nn} \end{pmatrix} \begin{pmatrix} a_{11} & \ldots & a_{n1} \\ \vdots & & \vdots \\ a_{1n} & \ldots & a_{nn} \end{pmatrix} = (b_{1i}a_{j1} + \ldots + b_{ni}a_{jn}),$$

d.h. $(\mathrm{AB})^T = \mathrm{B}^T\ \mathrm{A}^T$.

Ist A invertierbar und A^{-1} die Inverse, so folgt also

$$I = I^T = (AA^{-1})^T = (A^{-1})^T A^T = (A^{-1}A)^T = A^T(A^{-1})^T,$$

und somit ist A^T invertierbar und $(A^T)^{-1} = (A^{-1})^T$. Ist A^T invertierbar, so folgt aus obigem, dass $A = (A^T)^T$ invertierbar ist. #

Satz 2 *Es gilt*

$$[x]_F = U_{F,E}^T [x]_E = (U_{E,F}^{-1})^T [x]_E = (U_{E,F}^T)^{-1} [x]_E.$$

Beweis:
Sei $U_{F,E} = (\beta_{ij})$, $[x]_E = (\gamma_i)$, $[x]_F = (\delta_j)$. Dann ist

$$\begin{aligned} x &= \gamma_1 e_1 + \ldots + \gamma_n e_n = \gamma_1(\beta_{11} f_1 + \ldots + \beta_{1n} f_n) + \ldots + \gamma_n(\beta_{n1} f_1 + \ldots + \beta_{nn} f_n) \\ &= (\beta_{11}\gamma_1 + \ldots + \beta_{nn}\gamma_n) f_1 + \ldots + (\beta_{1n}\gamma_1 + \ldots + \beta_{nn}\gamma_n) f_n \end{aligned}$$

und

$$x = \delta_1 f_1 + \ldots \delta_n f_n.$$

Daraus folgt

$$\underbrace{\begin{pmatrix} \delta_1 \\ \vdots \\ \delta_n \end{pmatrix}}_{[x]_F} = \underbrace{\begin{pmatrix} \beta_{11} & \ldots & \beta_{n1} \\ \vdots & & \vdots \\ \beta_{1n} & \ldots & \beta_{nn} \end{pmatrix}}_{U_{F,E}^T} \underbrace{\begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix}}_{[x]_E}.$$

Die anderen beiden Gleichungen ergeben sich nun aus dem Lemma. #
Seien nun X und Y lineare Räume mit $\dim_{\mathcal{K}} X = m$ und $\dim_{\mathcal{K}} Y = n$.
Sei $A \in \mathcal{L}(X,Y)$. Wählen Basen $E = \{e_1, \ldots, e_m\}$ und $G = \{g_1, \ldots, g_m\}$ in X und $F = \{f_1, \ldots, f_n\}$ und $H = \{h_1, \ldots, h_n\}$ in Y.
Der folgende Satz zeigt, wie die Matrixdarstellungen $[A]_{E,F}$ und $[A]_{G,H}$ zusammenhängen.

Satz 3 *Es gilt*

$$[A]_{G,H} = U_{H,F}^T [A]_{E,F} U_{E,G}^T.$$

Beweis:
Nach Definition der Matrixdarstellung ist

$$[Ax]_H = [A]_{G,H} [x]_G.$$

Nach Satz 2 ist
$[Ax]_H = U_{H,F}^T [Ax]_F$,
$[x]_G = U_{G,E}^T [x]_E$.
Einsetzen liefert

$$U_{H,F}^T [Ax]_F = [A]_{G,H} U_{G,E}^T [x]_E,$$

d.h. $[Ax]_F = \underbrace{U_{F,H}^T [A]_{G,H} U_{G,E}^T}_{[A]_{E,F}} [x]_E$

Somit ist

$$U_{F,H}^T [A]_{G,H} U_{G,E}^T = [A]_E, F,$$

d.h. $[A]_{G,H} = U^T_{H,F}[A]_{E,F}U^T_{E,G}$. #

1. Beispiel:

Sei X = Y = $\mathcal{R}^2$ und A sei Spiegelung an der x-Achse.

Erste Basis: E = {(1,0);(0,1)}, F = {(1,0);(0,1)}
Ae_1=A(1,0)=(1,0)=1*(1,0)+0*(0,1) = $1f_1$+$0f_2$
Ae_2=A(0,1)=(0,-1)=0*(1,0)+(-1)*(0,1) = $0f_1$+(-1)f_2

$$[A]_{E,F} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Zweite Basis: G = {(1,1);(-1,0)}, H = {(0,1);(-1,0)}
Ag_1=A(1,1)=(1,-1)=(-1)(0,1)+(-1)(-1,0) = (-1)h_1+(-1)h_2
Ag_2=A(-1,0)=(-1,0)=0(0,1)+1(-1,0) = $0h_1$+$1h_2$

$$[A]_{G,H} = \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix}$$

$U_{E,G}$:
g_1 = (1,1) = 1(1,0)+1(0,1) = $1e_1$ + $1e_2$
g_2 = (-1,0) = (1-)(1,0)+0(0,1) = (-1)e_1 + $0e_2$

$$U_{E,G} = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$$

$U_{H,F}$:
f_1 = (1,0) = 0(0,1)+(-1)(-1,0) = $0h_1$ + (-1)h_2
f_2 = (0,1) = 1(0,1)+0(-1,0) = $1h_1$ + $0h_2$

$$U_{H,F} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

$$[A]_{G,H} = U^T_{H,F}[A]_{E,F}U^T_{E,G}$$

$$= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix}$$

Durch Wahl verschiedener Basenpaare erreicht man verschiedene Matrixdarstellungen ein und desselben Operators. Wie hängen alle diese Matrixdarstellungen untereinander zusammen?

Satz 4 *Seien X und Y lineare Räume mit $dim_{\mathcal{K}}$ X = m, $dim_{\mathcal{K}}$ Y = n und sei A ∈ $\mathcal{L}$(X,Y). Es gelte $[A]_{E,F}$ = B in einem Basenpaar E,F. Für eine Matrix C ∈ $M_{n,m}$ ($\mathcal{K}$) sind dann folgende Bedingungen äquivalent:*
(i) Es existiert ein Basenpaar G,H mit $[A]_{G,H}$ = C;
(ii) Es gibt Matrizen N ∈ M_n ($\mathcal{K}$), die invertierbar sind, und M ∈ M_m ($\mathcal{K}$) mit C = NBM.

Beweis:
(i) $\Rightarrow$ (ii): Folgt aus Sätzen 1 und 3 mit N $= U_{H,F}^T$ und M $= U_{E,G}^T$.
(ii) $\Rightarrow$ (i): Sei C = NBM.
Setzen

$$\begin{pmatrix} g_1 \\ \vdots \\ g_m \end{pmatrix} = M^T \begin{pmatrix} e_1 \\ \vdots \\ e_m \end{pmatrix}, \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = (N^{-1})^T \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}$$

Behaupten, dass G und H Basen sind. eigen dazu folgendes:
Sind $x_1, \ldots, x_k$ linear unabhängig und ist A $\in M_k$ $(\mathcal{K})$ invertierbar, so sind die durch

$$\begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}$$

gegebenen Elemente $y_1, \ldots, y_k$ ebenfalls linear unabhängig.
Sei $0 = \gamma_1 y_1 + \ldots + \gamma_k y_k = \gamma_1(a_{11}x_1 + \ldots + a_{1k}x_k) + \ldots + \gamma_k(a_{k1}x_1 + \ldots + a_{kk}x_k)$
$= (a_{11}\gamma_1 + \ldots + a_{k1}\gamma_k)x_1 + \ldots + (a_{1k}\gamma_1 + \ldots + a_{kk}\gamma_k)x_k$
Da $x_1, \ldots, x_k$ linear unabhängig sind, folgt

$$\begin{pmatrix} a_{11} & \ldots & a_{k1} \\ \vdots & & \vdots \\ a_{1k} & \ldots & a_{kk} \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_k \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

d.h. $A^T\gamma = 0$. Da A invertierbar ist, ergibt dies $\gamma = 0$. Somit ist gezeigt, dass G und H Basen sind.
Haben $U_{E,G} = M^T$, $U_{F,H} = (N^{-1})^T$ und nach Satz 3 ist damit

$$C = [A]_{G,H} = [((N^{-1})^T)^{-1}]^T B (M^T)^T = NBM.$$

#

Definition 2 *Zwei Matrizen B,C $\in M_{n,m}$ $(\mathcal{K})$ heißen äquivalent, wenn es invertierbare Matrizen N $\in M_n$ $(\mathcal{K})$ und M $\in M_m$ $(\mathcal{K})$ mit C = NBM gibt.*

Satz 4 besagt also, dass zwei Matrizen genau dann Darstellungen ein und desselben Operators sind, wenn sie äquivalent sind.

Satz 5 *Äquivalenz von Matrizen ist eine Äquivalenzrelation in $M_{n,m}$ $(\mathcal{K})$.*

Beweis:
(R) B $= I_{n\times n}$B$I_{m\times m}$
(S) C = NBM $\Rightarrow$ B = N^{-1}CM^{-1}.
(T) C = NBM, B = KDL $\Rightarrow$ C = (NK)D(ML). #
Die Menge $M_{n,m}$ $(\mathcal{K})$ zerfällt somit in Äquivalenzklassen. Ist A ein gegebener Operator, so bilden alle möglichen Matrixdarstellungen $[A]_{E,F}$ also (genau) eine Äquivalenzklasse. Werden sehen, dass der Rang eine Charakteristik der Äquivalenzklassen ist, d.h., dass

zwei Matrizen genau dann äquivalent sind, wenn sie den gleichen Rang haben.
Es folgt insbesondere, dass es genau min(n,m) + 1 Äquivalenzklassen gibt.

Sei nun X ein n-dimensionaler linearer Raum und A $\in \mathcal{L}$(X). Betrachten $[A]_{E,E}$, d.h. stellen A in einer einzigen Basis dar.

Satz 6 *Sei $[A]_{E,E} = B$ in einer Basis E. Dann sind folgende Bedingungen äquivalent:*
(i) Es existiert eine Basis F mit $[A]_{F,F} = C$;
(ii) Es existiert eine invertierbare Matrix V mit $C = V^{-1}BV$.

Beweis:
Satz 3 mit $X_n \xrightarrow{A} X_n$ und jeweils Basen E und F:

$$[A]_{F,F} = U^T_{F,E}[A]_{E,E}U^T_{E,F}$$

liefert C = V^{-1}BV mit V = $U^T_{E,F}$, d.h. liefert die Implikation (i) ⇒ (ii).

Um (ii) ⇒ (i) zu zeigen, sei C = V $^{-1}$BV. Definieren F durch

$$\begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} = V^T \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.$$

Wie im Beweis von Satz 4 bereits gezeigt wurde, ist F eine Basis in X.
Satz 3 mit obigem Muster ergibt

$$[A]_{F,F} = U^T_{F,E}[A]_{E,E}U^T_{E,F} = [(V^T)^{-1}]^T B(V^T)^T = V^{-1}BV = C.$$

#

Definition 3 *Zwei (quadratische) Matrizen B,C $\in M_n$ ($\mathcal{K}$) heißen ähnlich, wenn es eine invertierbare Matrix V $\in M_n$ ($\mathcal{K}$) mit $B = V^{-1}CV$ gibt.*

Satz 6 besagt somit, dass zwei Matrizen Darstellungen ein und desselben Operators A $\in \mathcal{L}$(X) in einem Paar gleicher Basen sind genau dann, wenn sie ähnlich sind.

Satz 7 *Ähnlichkeit von Matrizen ist eine Äquivalenzrelation in M_n ($\mathcal{K}$).*

Beweis:
(R) B = IBI.
(S) B = V^{-1}CV ⇒ C = VBV^{-1} = $(V^{-1})^{-1}$BV^{-1}.
(T) B = V^{-1}CV, C = W^{-1}DW ⇒ B = $V^{-1}W^{-1}$DWV = (WV)$^{-1}$D(WV). #
Die Mengen M_n ($\mathcal{K}$) zerfällt somit in Äquivalenzklassen. Für einen Operator A $\in \mathcal{L}$(X) bilden die Matrizen $[A]_{E,E}$ (genau) eine Äquivalenzklasse.

2. Beispiel:
Ähnliche Matrizen sind offenbar äquivalent. Quadratische äquivalente Matrizen müssen aber nicht ähnlich sein.
Ähnliche Matrizen haben die gleiche Determinante, äquivalente nicht notwendigerweise. Auch Gleichheit der Determinante von äquivalenten Matrizen sichert nicht deren Ähnlichkeit. Hier ist ein Beispiel:

$$\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{N}\underbrace{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}}_{B}\underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{M} = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}}_{C}$$

B und C sind äquivalent mit det B = det C = 0. Nehmen an, es gibt ein V mit $V^{-1}BV$ = C, d.h. BV = VC. Mit V = $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist dies

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} c & d \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} a & 0 \\ c & 0 \end{pmatrix}$$

$\Rightarrow$ d = 0, c = 0, a = 0 $\Rightarrow$ V = $\begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}$
Widerspruch, da V nicht invertierbar (det V = 0).

Die Äquivalenzklassen bezüglich Ähnlichkeit sind vollständig in den Äquivalenzklassen bezüglich der Äquivalenz enthalten (n = m vorausgesetzt).
Werden sehen, dass die Anzahlder Äquivalenzklassen bezüglich der Ähnlichkeit unendlich ist.
Die Charakteristik der Ähnlichkeit wird die Jordansche Normalform sein. Der Rang ist eine Charakteristik der Äquivalenz, aber lediglich eine Invarianz der Ähnlichkeit.

Definition 4 *Sei X ein n-dimensionaler linearer Raum und A* $\in \mathcal{L}(X)$*. Die Determinante von A ist definiert als det A = det* $[A]_{E,E}$*, wobei E irgendeine Basis in X ist.*

Diese Definition ist korrekt in dem Sinne, dass det A nicht von der speziellen Wahl von E abhängt. Nach Satz 6 ist nämlich $[A]_{F,F} = V^{-1}[A]_{E,E}V$ und damit
det $[A]_{F,F}$ = (det V^{-1})(det $[A]_{E,E}$)(det V) = det $[A]_{E,E}$.
Es gilt
det (AB) = det A det B,
A ist invertierbar ($\Leftrightarrow$ A ist Isomorphismus $\Leftrightarrow$ A ist Automorphismus $\Leftrightarrow$ det A $\neq$ 0).

2.7 Der Rang

19.04.06

Sei A $\in M_{m,n}(\mathcal{K})$. Eine *Untermatrix der Ordnung s* von A ist eine Matrix, die aus dem Schnitt von s Zeilen und s Spalten von A gebildet wird.
Ein *Minor der Ordnung s* ist die Determinante einer Untermatrix der Ordnung s.
Für s > min(m,n) gibt es keine Untermatrizen der Ordnung s. Für s $\leq$ min(m,n) gibt es genau $\begin{pmatrix} m \\ s \end{pmatrix}\begin{pmatrix} n \\ s \end{pmatrix}$ Untermatrizen der Ordnung s.

Definition 1 *Man sagt, dass $A \in M_{m,n}(\mathcal{K})$ den Rang s hat, wenn A einen von Null verschiedenen Minor der Ordnung s besitzt und alle Minoren von größerer Ordnung als s Null sind. Man bezeichnet den Rang von A mit rg A. Schließlich setzt man rg $\mathcal{O}$ = 0.*

Beispiele:

$$\operatorname{rg}\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} = 2$$

$$\operatorname{rg}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 1$$

$$\operatorname{rg}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = 1$$

Die Zeilen bzw. Spalten einer Matrix aus $M_{m,n}(\mathcal{K})$ sind Elemente von $M_{1,n}(\mathcal{K})$ (bzw. $M_{m,1}(\mathcal{K})$), und damit ist der Begriff der linearen Unabhängigkeit von Zeilen (bzw. Spalten) einer Matrix definiert.

Definition 2 *Man sagt, dass $A \in M_{m,n}(\mathcal{K})$ den Zeilenrang (bzw. Spaltenrang) s hat, wenn A s linear unabhängige Zeilen (bzw. Spalten) besitzt und eine beliebige Anzahl von mehr als s Zeilen (bzw. Spalten) linear abhängig sind.*

Mit anderen Worten: Der Zeilenrang (Spaltenrang) ist die maximale Anzahl von linear unabhängigen Zeilen (Spalten).

	Zeilenrang	Spaltenrang
$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$	2	2
$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	1	1
$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$	1	1

Satz 1 *Für eine beliebige Matrix fallen Rang, Zeilenrang und Spaltenrang stets zusammen.*

Beweis:
Sei A = $(a_{ij}) \in M_{m,n}(\mathcal{K})$ und sei rg A = r.
Es ist leicht zu sehen, dass sich beim Vertauschen von Zeilen oder Spalten der Rang, Zeilenrang, Spaltenrang nicht ändern.
Können daher annehmen, dass

$$\Delta := \begin{vmatrix} a_{11} & \dots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \dots & a_{rr} \end{vmatrix} \neq 0$$

ist. Zeigen, dass der Zeilenrang auch gleich r ist.
Zeigen dazu, dass die Zeilen $z_1, \ldots, z_r$ linear unabhängig sind und dass z_i für i > r eine Linearkombination von $z_1, \ldots, z_r$ ist.
Sei $\alpha_1 z_1 + \ldots + \alpha_r z_r = 0$. Daraus folgt

$$\alpha_1(a_{11}, \ldots, a_{1r}) + \ldots + \alpha_r(a_{r1}, \ldots, a_{rr}) = (0, \ldots, 0),$$

d.h.

$$\begin{array}{lcl} \alpha_1 a_{11} + \quad \ldots \quad + \alpha_r a_{r1} & = & 0 \\ \qquad\qquad\quad \vdots & & \\ \alpha_1 a_{1r} + \quad \ldots \quad + \alpha_r a_{rr} & = & 0 \end{array}$$

d.h.

$$\begin{pmatrix} a_{11} & \ldots & a_{r1} \\ \vdots & & \vdots \\ a_{1r} & \ldots & a_{rr} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_r \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

und wegen $\Delta \neq 0$ liefert die Cramersche Regel $\alpha_1 = \ldots = \alpha_r \neq 0$.

Sei nun i > r. Für $1 \leq j \leq n$ betrachten wir die Determinante

$$\begin{vmatrix} a_{11} & \ldots & a_{1r} & a_{1j} \\ \vdots & & \vdots & \vdots \\ a_{r1} & \ldots a_{rr} & a_{rj} & \\ a_{i1} & \ldots & a_{ir} & a_{ij} \end{vmatrix}.$$

Diese Determinante ist Null: Für $1 \leq j \leq r$ hat sie zwei gleiche Spalten und für $r < j \leq n$ ist sie ein Minor der Ordnung r + 1.
Entwickeln die Determinante nach der letzten Spalte:

$$A_1 a_{1j} + \ldots + A_r a_{rj} + \Delta a_{ij} = 0$$

mit von j unabhängigen Zahlen $A_1, \ldots, A_r$.
Haben also

$$\begin{array}{lcl} A_1 a_{11} + \quad \ldots \quad + A_r a_{r1} + \Delta a_{i1} & = & 0 \\ \qquad\qquad\quad \vdots & & \\ A_1 a_{1n} + \quad \ldots \quad + A_r a_{rn} + \Delta a_{in} & = & 0, \end{array}$$

d.h.

$$A_1(a_{11}, \ldots, a_{1n}) + \ldots + A_r(a_{r1}, \ldots, a_{rn}) + \Delta(a_{i1}, \ldots, a_{in}) = (0, \ldots, 0),$$

$$A_1 z_1 + \ldots + A_r z_r + \Delta z_i = 0,$$

$$z_i = -\frac{1}{\Delta}(A_1 z_1 + \ldots + A_r z_r).$$

Damit ist rg A = Zeilenrang von A bewiesen.
Schließlich ist
Spaltenrang von A = Zeilenrang von A^T = rg A^T = rg A. #

Satz 2 *Sei $A \in M_{m,n}(\mathcal{K})$ und seinen $M \in M_m(\mathcal{K})$ und $N \in M_n(\mathcal{K})$ invertierbare Matrizen. Dann ist rg MAN = rg A.*

Beweis:
Sei rg A = r. Nach Satz 1 hat A dann r linear unabhängige Spalten $s_1, \ldots, s_r$.
Die Spalten Ms_1,...,Ms_r sind Spalten von MA. Die Spaten Ms_1, ..., Ms_r sind linear unabhängig:

$$\alpha_1 M s_1 + \ldots + \alpha_r M s_r = 0$$

$$M(\alpha_1 s_1 + \ldots + \alpha_r s_r) = 0$$

von links mit M $^{-1}$ multiplizieren:

$$\alpha_1 s_1 + \ldots \alpha_r s_r = 0$$

$$\Rightarrow \alpha_1 = \ldots = \alpha_r = 0$$

Somit ist rg MA $\geq$ r = rg A.
Es ergibt sich rg A = rg M^{-1}MA $\geq$ MA, also rg MA = rg A und damit
rg MAN = rg $(MAN)^T$ = rg N^T $(MA)^T$ = rg $(MA)^T$ = rg MA = rg A. #

Definition 3 *Unter einer elementaren Umformung einer Matrix versteht man eine der folgenden drei Transformationen:*
(i) Vertauschen zweier Zeilen (Spalten);
(ii) Multiplikation einer Zeile (Spalte) mit einem von Null verschiedenen Skalar;
(iii) Addition eines skalaren Vielfaches einer Zeile (Spalte) zu einer anderen Zeile (Spalte).

Satz 3 *Elementare Umformungen ändern den Rang einer Matrix nicht.*

Beweis:
Folgt aus Satz 1.
Alternativ kann der Satz aus Satz 2 abgeleitet werden.
Die elementaren Umformungen können nämlich durch Multiplikation der Matrix von links (bei Zeilen) und rechts (bei Spalten) mit den folgenden invertierbaren Matrizen bewirkt werden:

(i)
$$\begin{pmatrix}
1 & & & & & & & \\
& \ddots & & & & & & \\
& & 1 & & & & & \\
& & 0 & \ldots & 1 & & & \\
& & 1 & & & & & \\
& & \vdots & \ddots & \vdots & & & \\
& & & & 1 & & & \\
& & 1 & \ldots & 0 & & & \\
& & & & & 1 & & \\
& & & & & & \ddots & \\
& & & & & & & 1
\end{pmatrix}$$

(ii)
$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

(iii)
$$\begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & 1 & \dots & C & & & \\ & & 1 & & & & & \\ & & \vdots & \ddots & \vdots & & & \\ & & & & 1 & & & \\ & & 0 & \dots & 1 & & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix}$$

#
Damit ergibt sich eine Methode, den Rang einer Matrix ziemlich effektiv zu bestimmen.
Beispiel:

$$rg \begin{pmatrix} 2 & 2 & 0 & 2 \\ 4 & 6 & 4 & 7 \\ 5 & 6 & 2 & 7 \\ 2 & 3 & 2 & 4 \end{pmatrix} = rg \begin{pmatrix} 1 & 1 & 0 & 1 \\ 4 & 6 & 4 & 7 \\ 5 & 6 & 2 & 7 \\ 2 & 3 & 2 & 4 \end{pmatrix} = rg \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 2 & 4 & 3 \\ 0 & 1 & 2 & 2 \\ 0 & 1 & 2 & 2 \end{pmatrix}$$

$$= rg \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 2 & 2 \\ 0 & 2 & 4 & 3 \\ 0 & 1 & 2 & 2 \end{pmatrix} = rg \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = rg \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$= rg \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = rg \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = rg \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \underline{\underline{3}}$$

Satz 4 *Eine Matrix $A \in M_{m,n}(\mathcal{K})$ vom Rang r kann durch elementare Umformung auf die Gestalt*

$$\left(\begin{array}{ccc|c} 1 & & & 0 \\ & \ddots & & 0 \\ & & 1 & 0 \\ \hline & & 0 & 0 \end{array}\right) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

gebracht werden.

21.04.06

Beweis: Man verfahre wie im obigen Beispiel. #

Man sagt, dass A *vollen Rang* hat, wenn r = min(m,n).

Satz 5 *(a) Für eine Matrix $A \in M_{m,n}(\mathcal{K})$ mit rg A = r exisiteren invertierbare Matrizen $M \in M_m(\mathcal{K})$ und $N \in M_n(\mathcal{K})$ mit*

$$MAN = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

(b) Sind X und Y endlichdimensionale lineare Räume und $A \in \mathcal{L}(X,Y)$, so existieren Basen E und F mit

$$[A]_{E,F} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

(c) Zwei Matrizen aus $M_{m,n}(\mathcal{K})$ sind genau dann äquivalent, wenn sie den gleichen Rang haben.

Beweis:
(a) Nach Satz 4 kann A durch elementare Umformungen auf die gewünschte Gestalt gebracht werden und im Beweis von Satz 3 haben wir gesehen, dass elementare Umformungen durch Multiplikation mit invertierbaren Matrizen realisiert werden können.
(b) Folgt aus (a) und Satz 4/2.6.
(c) Ist B = MAN, so golt rg B = rg A aus Satz 2. Umgekehrt sei rg B = rg A. Nach (a) gibt es invertierbare Matrizen M,N,K,L mit

$$MBN = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}, KAL = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Also ist MBN = KAL, d.h. $B = M^{-1}KALN^{-1}$. #

2.8 Unterräume

Dies sind die Analoga von Geraden und Ebenen in allgemeinen linearen Räumen.

Definition 1 *Eine Teilmenge $U \neq \varnothing$ eines linearen Raumes heißt (linearer) Unterraum von X, wenn U mit den Operationen aus X selbst ein linearer Raum ist.*

Offenbar ist U $\subset$ X genau dann ein Unterraum, wenn gilt
(i) x + y $\in$ U $\forall$ x,y $\in$ U
(ii) αx $\in$ U $\forall$ x $\in$ U, $\forall$ $\alpha \in \mathcal{K}$.
Insbesondere ist 0 $\in$ U und mit jedem x enthält U auch -x.

1. Beispiel:
In jedem linearen Raum X gibt es die beiden sogenannten *trivialen Unterräume* {0} und X.

2. Beispiel:

Die Unterräume von $\mathcal{R}$ sind $\{0\}$ und $\mathcal{R}$. Die Unterräume von $\mathcal{R}^2$ sind $\{0\}$, die Geraden durch den Ursprung und ganz $\mathcal{R}^2$. Unterräume des $\mathcal{R}^3$: $\{0\}$, Geraden durch den Ursprung, Ebenen durch den Ursprung und ganz $\mathcal{R}^3$.

3. Beispiel:
Sei $\mathcal{K}$[x] der lineare Raum aller Polynome in X mit Koeffizienten aus $\mathcal{K}$, d.h. die Menge aller Polynome p(x) = $p_0 + p_1$x + ... + $p_k\ x^k$ mit $p_0, \ldots, p_k \in \mathcal{K}$. Für jede Zahl $z_0 \in \mathcal{C}$ ist dann
U_{z_0} = {p(x) $\in \mathcal{K}$[x]: p(z_0) = 0} ein Unterraum von $\mathcal{K}$[x].

Definition 2 *Seien X und Y lineare Räume und A $\in \mathcal{L}$(X,Y). Die Mengen*

ker A := { x $\in$ X: Ax = 0}
Im A := {Ax: x $\in$ X}

heißen Kern (= Nullraum) und Bild (= Bildraum) von A. Statt Im A trifft man auch auf ran A.

4. Beispiel:
ker A $\subset$ X und Im A $\subset$ Y sind stets Unterräume. Insbesondere ist die Menge aller Lösungen des Gleichungssystems Ax = 0 ein Unterraum des $\mathcal{K}^n$. Und die Menge aller rechten Seiten $(b_1, \ldots, b_m)$, für die das Gleichungssystem Ax = b lösbar ist, ist ein Unterraum des $\mathcal{K}^m$.

5. Beispiel:
Sind U_α ($\alpha \in$ A) Unterräume von X, so ist auch $\bigcap_{\alpha \in\ A} U_\alpha$ ein Unterraum von X. Allerdings muss $\bigcup_{\alpha \in\ A} U_\alpha$ kein Unterraum sein.

Definition 3 *Sei X ein linearer Raum und S eine Teilmenge von X. Der Durchschnitt aller Unterräume von X, die S enthalten (X ist ein solcher Unterraum), wird Aufspannung oder lineare Hülle von S genannt und mit span S = lin S = [S] = <S> bezeichnet.*

Die Aufspannung von S ist also der kleinste Unterraum, der S enthält. Es ist leicht zu sehen, dass die Aufspannung von S mit der Menge aller endlichen Linearkombinationen von Elementen aus S (mit Koeffizienten aus $\mathcal{K}$) zusammenfällt, d.h.

span S = $\{\alpha_1 x_1 + \ldots \alpha_k x_k : k \geq 1; \alpha_1, \ldots, \alpha_k \in \mathcal{K};\ x_1, \ldots, x_k \in\ S\}$.

In der Tat, bezeichnen die geschweifte Klammer mit U. Die Menge span S ist ein linearer Raum und muss somit alle Elemente aus U enthalten, d.h. U $\subset$ span S.
Die Menge U ist ein Unterraum, der S enthält, und somit ist U einer der Unterräume U_α ($\alpha \in$ A), die S enthalten. Dies ergibt
span S = $\bigcap_{\alpha \in\ A} U_\alpha \subset U$.

6. Beispiel:
Seien P und Q zwei Punkte in der Ebenen (im $\mathcal{R}^2$). Was ist span {P,Q}?

P = Q = 0 → span {P,Q} = {0}
P = Q ≠ 0 → span {P,Q} = Gerade durch 0 und Punkt P = Q
P ≠ Q, P ≠ 0, Q ≠ 0 → span {P,Q} = $\mathcal{R}^2$ oder Gerade durch P und Q

7. Beispiel:
Wenn U_α ($\alpha \in$ A) Unterräume von X sind, so ist span $\bigcap_{\alpha\in\ A} U_\alpha$ ein Unterraum von X. Dieser Unterraum wird mit $\sum_{\alpha\in\ A} U_\alpha$ bezeichnet.
Es ist leicht zu sehen, dass

$$\sum_{\alpha\in\ A} U_\alpha = \{x_1 + \ldots + x_k; k \geq 1; x_1, \ldots, x_k \in \bigcap_{\alpha\in\ A} U_\alpha\}$$

gilt.
Bei unendlicher Indexmenge A schreibt man

$$\sum_{\alpha=1}^{n} U_\alpha = U_1 + \ldots + U_n.$$

Es ist erneut leicht zu sehen, dass

$$U_1 + \ldots + U_n = \{x_1 + \ldots + x_n : x_1 \in U_1, \ldots, x_n \in U_n\}$$

ist. Für n = 2 ist insbesondere

$$U + V = span(U \cup V) = \{x + y : x \in U, y \in V\}.$$

Satz 1 *Sei X ein endlichdimensionaler linearer Raum und {e_1,…,e_n} eine Basis in X. Seien desweiteren f_1,…,f_m Elemente von X und es gelte*

$$\begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \ldots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \ldots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} =: A \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.$$

Dann ist
dim span $\{f_1, \ldots, f_m\}$ = rg A.

26.04.06

Beweis:
Sei rg A = r. Betrachten die Matrix $A^T = (\alpha_{ij})$. Dann ist rg A^T = r.
O.B.d.A. mögen die ersten r Spalten von A^T linear unabhängig sein. Die restlichen Spalten sind dann Linearkombinationen der ersten.

Behaupten, dass $f_1, \ldots, f_r$ linear unabhängig sind.
Sei dazu

$$\begin{aligned} 0 = \beta_1 f_1 + \ldots + \beta_r f_r &= \beta_1(\alpha_{11}e_1 + \ldots + \alpha_{1n}e_n) + \ldots + \beta_r(\alpha_{r1}e_1 + \ldots + \alpha_{rn}e_n) \\ &= (\beta_1\alpha_{11} + \ldots + \beta_r\alpha_{r1})e_1 + \ldots + (\beta_1\alpha_{1n} + \ldots + \beta_r\alpha_{rn})e_n. \end{aligned}$$

Also folgt

$$\begin{array}{lcl} \beta_1\alpha_{11} + \ \ldots \ + \beta_r\alpha_{r1} & = & 0 \\ \qquad\qquad\ \vdots & & \\ \beta_1\alpha_{1n} + \ \ldots \ + \beta_r\alpha_{rn} & = & 0 \end{array}$$

d.h.

$$\beta_1 \begin{pmatrix} \alpha_{11} \\ \vdots \\ \alpha_{1n} \end{pmatrix} + \ldots + \beta_r \begin{pmatrix} \alpha_{r1} \\ \vdots \\ \alpha_{rn} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

woraus $\beta_1 = \ldots = \beta_r = 0$ folgt
Analog zeigt man, dass f_i mit i > r eine Linearkombination von $f_1, \ldots, f_r$ ist.

Somit ist dim span $\{f_1, \ldots, f_r, f_{r+1}, \ldots, f_m\}$ = dim span $\{f_1, \ldots, f_r\}$ = r, da $f_1, \ldots, f_r$ eine Basis in span $\{f_1, \ldots, f_r\}$ ist. #

Beispiel:
Seien zum Beispiel m Elemente des $\mathcal{R}^n$ gegeben. Schreiben diese als Spalten:

$$\begin{pmatrix} f_{11} \\ \vdots \\ f_{1n} \end{pmatrix}, \ldots, \begin{pmatrix} f_{m1} \\ \vdots \\ f_{mn} \end{pmatrix}.$$

Sei also $\{e_1, \ldots, e_n\}$ die Standardbasis im $\mathcal{R}^n$. Dann ist

$$\begin{pmatrix} f_{i1} \\ \vdots \\ f_{in} \end{pmatrix} = f_{i1} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \ldots + f_{in} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

und Satz 1 liefert somit

$$dim\ span \begin{pmatrix} f_{11} \\ \vdots \\ f_{1n} \end{pmatrix}, \ldots, \begin{pmatrix} f_{m1} \\ \vdots \\ f_{mn} \end{pmatrix} = rg \begin{pmatrix} f_{11} & \ldots & f_{m1} \\ \vdots & & \vdots \\ f_{1n} & \ldots & f_{mn} \end{pmatrix}.$$

Satz 2 *Seien X und Y endlichdimensionale lineare Räume und A ∈ $\mathcal{L}$(X,Y). Seien E und F Basen in X und Y und r := rg* $[A]_{E,F}$. *Dann gilt:*
dim Im A = r,
dim ker A = dim X - r.

Beweis:
Nach Satz 5b/2.7. existieren Basen G und H mit $[A]_{G,H} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.
Dann ist rg $[A]_{E,F}$ = rg $[A]_{G,H}$.
Für x ∈ X haben wir

$$x = x_1 g_1 + \ldots + x_n g_n,$$
$$Ax = y_1 h_1 + \ldots y_m h_m.$$

Dabei ist

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = [A]_{G,H} \qquad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

d.h.
$y_1 = x_1$,
$\vdots$
$y_r = x_r$,
$y_{r+1} = 0$,
$y_m = 0$.

Also ist Im A = {Ax : x ∈ X} = $\{x_1h_1 + \ldots + x_rh_r : x_j \in \mathcal{K}\}$ = span $\{h_1, \ldots, h_r\}$ ⇒ dim Im A = r

ker A = { x ∈ X: Ax = 0} = {x ∈ X: $x_1h_1 + \ldots + x_rh_r = 0$} = {x ∈ X: x = $x_{r+1}g_{r+1} + \ldots + x_ng_n$} = span $\{g_{r+1}, \ldots, g_n\}$
⇒ dim ker A = n - r = dim X - r
#

Satz 3 *Sei X ein linearer Raum und dim X = n. Sind $e_1, \ldots, e_k$ ($k < n$) beliebige linear unabhängige Elemente von X, so existieren Elemente $e_{k+1}, \ldots, e_n$ von X, sodass $\{e_1, \ldots, e_n\}$ eine Basis in X ist.*

Beweis:
Sei X_k = span $\{e_1, \ldots, e_k\}$.
Dann ist $X_k \neq$ X. Wählen $e_{k+1} \in$ X \ X_k. Die Elemente $e_1, \ldots, e_k, e_{k+1}$ sind linear unabhängig:

$$\alpha_1e_1 + \ldots + \alpha_ke_k + \alpha_{k+1}e_{k+1} = 0$$

⇒ $\alpha k + 1 = 0$ (da sonst $e_{k+1} \in X_k$)
⇒ $\alpha_1 = \ldots = \alpha_k = 0$ (da $e_1, \ldots, e_n$ linear unabhängig).
Wiederholen dies für X_{k+1} = span $\{e_1, \ldots, e_{k+1}\}$ usw., bis X_k = X entsteht. #

Satz 4 *Seien U und V Unterräume eines endlichdimensionalen linearen Raumes X. Dann gilt*

dim (U+V) = dim U + dim V - dim (U ∩ V).

Beweis:
Sei dim (U ∩ V) = r, dim U = u, dim V = v.
Wählen Basis $e_1, \ldots, e_r$ in U ∩ V. (1)
Nach Satz 3 existieren $f_1, \ldots f_{u-r}$, sodass
$e_1, \ldots, e_r, f_1, \ldots, f_{u-r}$ Basis in U ist. (2)
Ebenfalls nach Satz 3 existieren $g_1, \ldots, g_{v-r}$, sodass
$e_1, \ldots, e_r, g_1, \ldots, g_{v-r}$ Basis in V ist. (3)
Behauptung folgt, wenn wir bewiesen haben, dass
$e_1, \ldots, e_r, f_1, \ldots, f_{u-r}, g_1, \ldots, g_{v-r}$ eine Basis in U + V ist.
(⇒ dim (U + V) = u + v - r).
Zeigen lineare Unabhängigkeit. Sei

$$\underbrace{\alpha_1e_1 + \ldots + \alpha_re_r}_{e} + \underbrace{\beta_1f_1 + \ldots + \beta_{u-r}f_{u-r}}_{f} + \underbrace{r_1g_1 + \ldots + r_{v-r}g_{v-r}}_{g} = 0.$$

Haben e + f = -g mit e + f ∈ U (wegen (2)) und -g ∈ V (wegen (3)). Also ist g ∈ U ∩ V. Damit ist
$r_1g_1 + \ldots + r_{v-r}g_{v-r} = \delta_1e_1 + \ldots + \delta_re_r$ (wegen (1))
und aus (3) folgt $r_1 = \ldots = r_{v-r} = 0$ (und $\delta_1 = \ldots = \delta_r = 0$).
Somit ist $\alpha_1e_1 + \ldots + \alpha_re_r + \beta_1f_1 + \ldots + \beta_{u-r}f_{u-r} = 0$ und (2) liefert $\alpha_1 = \ldots = \alpha_r = \beta_1 = \ldots = \beta_{u-r} = 0$.

Zeigen noch, dass jedes Element aus U + V eine Linearkombination der Elemente $e_1, \ldots, e_r, f_1, \ldots, f_{u-r}, g_1, \ldots, g_{v-r}$ ist (4).
Ein Element aus U + V ist von der Form x + y mit x ∈ U und y ∈ V. Aus (2) und (3) folgt, dass x + y eine Linearkombination der Elemente (4) ist. #

Haben dim (U ∩ V) = dim U + dim V - dim (U + V) und dies ergibt
→ zwei 3-dimensionale Unterräume eines 4-dimensionalen Raumes schneiden sich mindestens in einem 2-dimensionalen Unterraum
→ zwei 3-dimensionale Unterräume eines 5-dimensionalen Raumes schneiden sich mindestens in einem 1-dimensionalen Unterraum
→ zwei 3-dimensionale Unterräume eines 6-dimensionalen Raumes schneiden sich im 0-dimensionalen Unterraum

→ z.B. U = {(x,y,z,0,0,0) ∈ $\mathcal{R}^6$}, V = {(0,0,0,u,v,w) ∈ $\mathcal{R}^6$}

Definition 4 *Sei X ein linearer Raum und seien U,V Unterräume von X. Man sagt, dass X in die direkte Summe von U und V zerfällt und schreibt X = U ⊞ V = (U + V), wenn gilt: X = U + V, U ∩ V = {0}.*

(Hinweis: Eigentlich kommt ein Kreis um das Plus, jedoch lässt sich dies hiermit nicht darstellen.)

28.04.06

Satz 5 *Es gilt X = U ⊞ V genau dann, wenn sich jedes x ∈ X eindeutig in der Form x = u + v mit u ∈ U und v ∈ V darstellen lässt.*

Beweis:
Sei X = U ⊞ V. Dann ist X = U + V und somit hat jedes x ∈ X eine Darstellung x = u + v mit u ∈ U, v ∈ V. Ist x = u + v = u' + v', so folgt u - u' = v' - v ∈ U ∩ V = {0}, d.h. u - u' = v' - v = 0, d.h. u = u' und v = v' (also ist die Darstellung eindeutig).

Umgekehrt möge sich jedes x ∈ X eindeutig als x = u + v schreiben lassen. Dies liefert zunächst X = U + V. Sei a ≠ 0 und a ∈ U ∩ V. Dann ist x = (u + a) + (v - a) und u + a ∈ U, v - a ∈ V in Widerspruch zur Eindeutigkeit der Darstellung. Also ist U ∩ V = {0}, d.h. X = U ⊞ V. #

Satz 6 *Ist X endlichdimensional, so gilt X = U ⊞ V genau dann, wenn X = U + V und*
dim X = dim U + dim V ist.

Beweis: Folgt aus Satz 4. #

Definition 5 *Sei X ein linearer Raum und U ein Unterraum von X. Jeder Unterraum V von X mit X = U ⊞ V heißt Komplementärraum von U in X.*

8. Beispiel:
$X = \mathcal{R}^2$
U : Gerade durch den Ursprung
Dann ist jede andere Gerade durch den Ursprung ein Komplementärraum,

X = $\mathcal{R}^3$
U : Ebene durch den Ursprung
Jede Gerade V durch den Ursprung, die nicht in U enthalten ist, ist ein Komplementärraum.

Satz 7 *Ist X endlichdimensional, so hat jeder Unterraum von X einen Komplementärraum.*

Beweis:
Sei U ein Unterraum von X. Wählen in U eine Basis $\{e_1, \ldots, e_m\}$ und ergänzen diese zu einer Basis $\{e_1, \ldots, e_m, e_{m+1}, \ldots, e_n\}$ von X (Satz 3). Haben U = span $\{e_1, \ldots, e_m\}$ und setzen V = span $\{e_{m+1}, \ldots, e_n\}$.
Dann ist X = U + V und wegen dim X = n, dim U = m, dim V = n-m ist dim X = dim U + dim V. Satz 6 liefert also X = U $\boxplus$ V. #
Sei X ein linearer Raum und U ein Unterraum. Definieren eine Relation ~ in X über

$$x \tilde{} y \Leftrightarrow x - y \in U.$$

Dies ist eine Äquivalenzrelation in X:
(R) x - x = 0 ∈ U,
(S) x - y ∈ U ⇒ - (x - y) ∈ U ⇒ y - x ∈ U,
(T) x - y ∈ U, y - z ∈ U ⇒ (x - y) + (y - z) ∈ U ⇒ x - z ∈ U.

Die Äquivalenzklassen sind Mengen der Form

$$a + U := \{a + u : u \in U\}.$$

In der Tat, sei M eine Äquivalenzklasse und a ∈ M. Dann ist:
x ∈ M ⇔ x ~ a ⇔ x - a ∈ U ⇔ ∃ u ∈ U: x - a = u ⇔ ∃ u ∈ U: x = a + u ⇔ x ∈ a + U

Definition 6 *Sei X ein linearer Raum und U ein Unterraum von X. Teilmengen von X der Form a + U mit a ∈ X heißen lineare Untermannigfaltigkeiten von X. Die Dimension einer linearen Untermannigfaltigkeit a + U ist definiert durch*

dim (a + U) := dim U

9. Beispiel:
X = $\mathcal{R}^2$
U sei Gerade durch den Ursprung.
Dann gilt: x ~ y ⇔ x und y liegen auf einer Geraden, die parallel zu U ist
Äquivalenzklassen sind also zu U parallele Geraden.

Also: Lineare Untermannigfaltigkeiten im $\mathcal{R}^2$ sind Punkte der Ebene (Dimension 0), Geraden (Dimension 1), Ebene (Dimension 2). Lineare Untermannigfaltigkeiten sind also „verschobene Unterräume“.

Definieren Summe und skalares Vielfaches von linearen Untermannigfaltigkeiten wie folgt:

$$(a + U) + (b + U) := (a + b) + U$$

$$\alpha(a+U) := \alpha a + U$$

Diese Definition ist korrekt:
a_1 + U = a_2 + U, b_1 + U = b_2 + U
$\Leftrightarrow$ a_1 - a_2 $\in$ U, b_1 - b_2 $\in$ U
$\Rightarrow$ (a_1 + b_1) - (a_2 + b_2) $\in$ U
$\Rightarrow$ (a_1 + b_1) + U = (a_2 + b_2) + U,

a_1 + U = a_2 + U
$\Rightarrow$ a_1 - a_2 $\in$ U
$\Rightarrow$ αa_1 - αa_2 $\in$ U
$\Rightarrow$ αa_1 + U = αa_2 + U.

Haben hierbei benutzt, dass gilt a + U = b + U $\Leftrightarrow$ a - b $\in$ U.

Die Menge aller linearen Untermannigfaltigkeiten der Form a + U (mit fixiertem U) bildet bezüglich der oben definierten Addition und Multiplikation mit Skalaren einen linearen Raum. Das Nullelement ist 0 + U = U.

Definition 7 *Sei X ein linearer Raum und U ein Unterraum von X. Der lineare Raum aller linearen Untermannigfaltigkeiten der Form a + U (a $\in$ X) heißt Faktorraum (quotient space) von X nach U und wird mit X/U bezeichnet.*

Satz 8 *Sei X ein linearer Raum, U ein Unterraum von X und V ein Komplementärraum von U in X, d.h.*
X = $\boxplus$ V. Dann ist V isomorph zu X/U.

Beweis:
Betrachten die Abbildung

$$A : V \to X/U, v \mapsto v + U.$$

Injektivität:
Av_1 = Av_2 $\Rightarrow$ v_1 + U = v_2 + U
$\Rightarrow$ v_1 - v_2 $\in$ U.
Haben auch v_1 - v_2 $\in$ V. Also ist v_1 - v_2 $\in$ U $\cap$ V = {0},d.h. v_1 = v_2.

Surjektivität:
Sei a $\in$ X. Müssen zeigen, dass ein v $\in$ V mit Av = v + U = a + U existiert.
Wegen X = U + V gilt a = u + v mit u $\in$ U und v $\in$ V. Damit ist

a + U = {a + w: w $\in$ U} = {u + v + w: w $\in$ U} = {v + z: z $\in$ U} = v + U.

Linearität:

$$A(v_1 + v_2) = Av_1 + Av_2$$
$$(v_1 + v_2) + U = (v_1 + U) + (v_2 + U)$$
$$A(\alpha v) = \alpha(Av)$$
$$\alpha v + U = \alpha(v + U)$$

Somit ist A ein Isomorphismus von V auf X/U. #

X/U ≅ V 03.05.06

Ist X endlichdimensional, so gilt
dim (X/U) = dim X - dim U
Dies folgt aus X/U ≅ V mit U ⊞ V = X und dim X = dim U + dim V - $\underbrace{dim(U \cap V)}_{=0}$.

Satz 9 *Sei X ein linearer Raum und sei Y ein linearer Raum. Für A ∈ $\mathcal{L}$(X,Y) gilt dann*
X/ker A ≅ Im A.

Beweis:
Betrachten die Abbildung
φ: X/ker A → Im A, x + ker A ↦ Ax.

Korrektheit:
x + ker A = y + ker A ⇒ x - y ∈ ker A ⇒ A(x - y) = 0
⇒ Ax - Ay = 0 ⇒ Ax = Ay.

Linearität:
φ((x + ker A) + (y + ker A)) = φ((x + y) + ker A) = A(x + y) = Ax + Ay
= φ(x + ker A) + φ(y + ker A).
Analog zeigt man $\varphi(\alpha$(x + ker A)) = α φ(x + ker A).

Injektivität:
φ(x + ker A) = φ(y + ker A)
⇒ Ax = Ay
⇒ A(x - y) = 0
⇒ x - y ∈ ker A
⇒ x + ker A = y + ker A

Surjektivität:
y = Ax ∈ Im A ⇒ y = φ(x + ker A) ∈ Im φ. #

Definition 8 *Seien X und Y endlichdimensionale lineare Räume und $A \in \mathcal{L}(X,Y)$. Die Zahlen*

$$r(A) := \dim \operatorname{Im} A,$$
$$n(A) := \dim \ker A,$$

heißen Rang und Kerndimension (nullity) von A. Der Faktorraum Y/Im A heißt Cokern von A und wird mit Coker A bezeichnet. Die Zahl

$$d(A) := \dim \operatorname{Coker} A \ (= \dim Y/\operatorname{Im} A)$$

nennt man Defekt von A. Schließlich heißt

$$\operatorname{Ind} A := \dim \ker A - \dim \operatorname{Coker} A = n(A) - d(A)$$

Index von A.

Satz 2 liefert r(a) = rg $[A]_{E,F}$ für ein beliebiges Basenpaar E,F. Dies rechtfertigt den Namen Rang für r(A).
Haben dim ker A = dim X - dim Im A d.h. n(a) = dim X - r(A).

Des Weiteren ist d(A) = dim Y - dim Im A = dim Y - r(A)
und somit Ind A = (dim X - r(A)) - (dim Y - r(A)) = dim X - dim Y.
Alle Operatoren aus $\mathcal{L}$(X,Y) haben also den gleichen Index. Insbesondere ist Ind(A) = 0 für alle A $\in \mathcal{L}$(X).

Fredholmsche Alternative

Ist A $\in \mathcal{L}$(X) mit X endlichdimensional, so ist entweder die Gleichung Ax = y für alle y $\in$ X lösbar oder die Gleichung Ax = 0 hat nichttriviale Lösungen.

Beweis:
Haben dim ker A = dim X - dim Im A
und somit ist entweder dim Im A = dim X, d.h. Im A = X,
oder dim ker A > 0. #
Umformulierung:
Entweder Im A = X oder ker A $\neq \{0\}$ ist äquivalent zu Im A = X $\Leftrightarrow$ ker A = $\{0\}$ oder A surjektiv $\Leftrightarrow$ A injektiv
Fredholmsche Alternative halbiert also das Problem, einen Operator (aus X in X) auf Bijektivität zu untersuchen. Es reicht, Surjektivität oder Injektivität festzustellen. Mit anderen Worten:
A bijektiv $\Leftrightarrow$ A surjektiv $\Leftrightarrow$ A injektiv.

2.9 Der Gauß'sche Algorithmus

10.05.06

Betrachten das lineare System

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

von m Gleichungen mit n Unbekannten.

Kennen bereits Aussagen über die Lösung solcher Systeme (z.B. Cramersche Regel) oder über deren Lösbarkeit (z.B. Fredholmsche Alternative).

Der Gauß'sche Algorithmus ist eine universelle Lösungsmethode.

Schreiben das System formal als

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

Suchen die erste von Null verschiedene Spalte. Dies sei die Spalte s_j. Darin suchen wir ein von Null verschiedenes Element a_{ij} und durch Vertauschen der ersten und der i-ten Zeile bringen wir dieses in die erste Zeile.
Multiplizieren dann die erste Zeile mit $\frac{1}{a_{ij}}$ und so wird das a_{ij} zu 1.
Subtrahieren dann a_{kj} mal die erste Zeile von der k-ten Zeile. So entstehen Nullen unter der geschaffenen 1. Erhalten eine Matrix der Form

$$\left(\begin{array}{ccccccc|c} 0 & \dots & 0 & 1 & * & \dots & * & * \\ 0 & \dots & 0 & 0 & . & . & . & . \\ \vdots & & \vdots & \vdots & . & . & . & . \\ 0 & \dots & 0 & 0 & . & . & . & . \end{array}\right).$$

Wiederholen alles für die gepunktete Matrix rechts unten usw.

Am Ende erhalten wir eine Matrix in sogenannter Zeilenstufenform, z.B.:

$$\left(\begin{array}{ccccccccc|c} 0 & 0 & 1 & * & * & * & * & * & * & c_1 \\ 0 & 0 & 0 & 0 & 1 & * & * & * & * & c_2 \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * & * & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & c_{m-1} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & c_m \end{array}\right).$$

Die letzten l Zeilen seien Null (l = m - rg A). Ist eine der Zahlen $c_{m-l+1}, \dots, c_m$ von Null verschieden, so hat das System keine Lösung.

Ist aber $c_{m-l+1} = \dots = c_m = 0$, so streichen wir die letzten l Zeilen.

Die Anfangseinsen der verbleibenden Zeilen heißen Leitkoeffizienten.

Ziehen geeignete Vielfache der Zeilen von unten nach oben ab und erzeugen so Nullen an Stellen über den Leitkoeffizienten. Im obigen Beispiel würde entstehen:

$$\left(\begin{array}{ccccccccc|c} 0 & 0 & 1 & * & 0 & 0 & * & * & * & \tilde{c_1} \\ 0 & 0 & 0 & 0 & 1 & 0 & * & * & * & \tilde{c_2} \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * & * & \tilde{c_3} \end{array}\right).$$

Bezeichnen mit $j_1, \ldots, j_s$ (s = n - rg A) die Spalten vor dem Strich, in denen keine Leitkoeffizienten stehen. Im Beispiel wären dies 1.,2.,4.,7.,8.,9. Spalte.

Die Unbekannten $x_{j_1}, \ldots, x_{j_s}$ kann man dann frei wählen und die restlichen Unbekannten sind eindeutig bestimmt und ergeben sich von unten nach oben automatisch.

Beispiel:

$$\left(\begin{array}{ccccc|c} 0 & 0 & 1 & 3 & 3 & 2 \\ 1 & 2 & 1 & 4 & 3 & 3 \\ 1 & 2 & 2 & 7 & 6 & 5 \\ 2 & 4 & 1 & 5 & 3 & 4 \end{array}\right) \Rightarrow \left(\begin{array}{ccccc|c} 1 & 2 & 1 & 4 & 3 & 3 \\ 0 & 0 & 1 & 3 & 3 & 2 \\ 1 & 2 & 2 & 7 & 6 & 5 \\ 2 & 4 & 1 & 5 & 3 & 4 \end{array}\right) \Rightarrow \left(\begin{array}{ccccc|c} 1 & 2 & 1 & 4 & 3 & 2 \\ 0 & 0 & 1 & 3 & 3 & 2 \\ 0 & 0 & 1 & 3 & 3 & 2 \\ 0 & 0 & -1 & -3 & -3 & -2 \end{array}\right)$$

$$\Rightarrow \left(\begin{array}{ccccc|c} 1 & 2 & 1 & 4 & 3 & 3 \\ 0 & 0 & 1 & 3 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \Rightarrow \left(\begin{array}{ccccc|c} 1 & 2 & 1 & 4 & 3 & 3 \\ 0 & 0 & 1 & 3 & 3 & 2 \end{array}\right) \Rightarrow \left(\begin{array}{ccccc|c} 1 & 2 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 & 3 & 2 \end{array}\right)$$

$\rightarrow x_2, x_4, x_5$ frei wählen,
$x_1 + 2x_2 + x_4 = 1$
$x_3 + 3x_4 + 3x_5 = 2$

$x_3 = 2 - 3x_4 - 3x_5$
$x_1 = 1 - 2x_2 - x_4.$

$x_2 = u, x_4 = v, x_5 = w$

$x_3 = 2 - 3v - 3w$
$x_1 = 1 - 2u - v$
d.h.

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 - 2u - v \\ u \\ 2 - 3v - 3w \\ v \\ w \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \\ 0 \end{pmatrix} + u \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + v \begin{pmatrix} -1 \\ 0 \\ -3 \\ 1 \\ 0 \end{pmatrix} + w \begin{pmatrix} 0 \\ 0 \\ -3 \\ 0 \\ 1 \end{pmatrix}$$

Lösungsmenge einer beliebigen linearen Gleichung Ax = b ist stets von der Form
x* + ker A, wobei x* irgendeine Lösung ist
(Ax = b $\Leftrightarrow$ x = x* + z mit A(x* + z) = b $\Leftrightarrow$ x = x* + z, Az = 0 $\Leftrightarrow$ x $\in$ x* + ker A).

In unserem Beispiel ist

$$x* = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}, ker\ A\ = span\ \left[\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ -3 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -3 \\ 0 \\ 1 \end{pmatrix} \right].$$

2.10 Eigenwerte

Definition 1 *Sei X ein linearer Raum und A ∈ $\mathcal{L}$(X). Ein Unterraum U ⊂ X heißt invarianter Unterraum von A, wenn A(U) ⊂ U gilt, d.h. wenn gilt u ∈ U ⇒ Au ∈ U.*

Sei X endlichdimensional, U ein invarianter Unterraum von A und V ein Komplementärraum von U, d.h. U ⊞ V = X. [Dann muss V nicht notwendigerweise ein invarianter Unterraum von A sein.]
Wählen Basis $\{e_1, \ldots, e_m\}$ in U und Basis $\{f_1, \ldots, f_k\}$ in V.

Dann ist E = $\{e_1, \ldots, e_m, f_1, \ldots, f_k\}$ eine Basis in X und die Matrixdarstellung von A in der Basis E ist von der Form

$$[A]_{E,E} = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix},$$

wobei A_{11} eine m x m-Matrix ist.

Definition 2 *Sei X ein linearer Raum und A ∈ $\mathcal{L}$(X). Zwei Unterräume U und V von X heißen ein Paar reduzierender (invarianter) Unterräume von A, wenn U und V invariante Unterräume von A sind und U ⊞ V = X ist.*

Ist U, V ein reduzierendes Paar, so liefern obige Basen

$$[A]_{E,E} = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}.$$

Allgemeiner schreibt man
X = $U_1 \boxplus \ldots \boxplus U_m$, wenn $U_1, \ldots, U_m$ Unterräume von X sind und sich jedes x ∈ X eindeutig in der Form x = $u_1 + \ldots + u_m$ mit $u_1 \in U_1$, …, $u_m \in U_m$ schreiben lässt. Man sagt dann, dass X die direkte Summe von U_1, …, U_m ist. (Dafür ist notwendig, aber nicht hinreichend, dass
X = $U_1 + \ldots + U_m$ und $U_i \cap U_j = \{0\}$ für i ≠ j gilt.)

Ist X endlichdimensional und X = $U_1 \boxplus \ldots \boxplus U_m$ und wählt man sich Basen E_i in U_i und setzt E = $E_1 \cup \ldots \cup E_m$, so ergibt sich für A ∈ $\mathcal{L}$(X) die Matrixdarstellung

$$[A]_{E,E} = \begin{pmatrix} A_{11} & & & \\ & A_{22} & & \\ & & \ddots & \\ & & & A_{mm} \end{pmatrix},$$

falls alle U_i invariante Unterräume von A sind. Die Matrixdarstellung ist also blockdiagonal.

Sind alle U_i eindimensional, so ist $[A]_{E,E}$ eine Diagonalmatrix.

Man nennt Operatoren *diagonalisierbar*, wenn es eine Basis gibt, in der die Matrixdarstellung eine Diagonalmatrix ist.

Ist U ein eindimensionaler invarianter Unterraum von A, so gilt Ax = λ x $\forall$ U mit einer gewissen Zahl λ.

Definition 3 *Sei X ein linearer Raum und A $\in \mathcal{L}(X)$. Eine Zahl $\lambda \in \mathcal{K}$ heißt Eigenwert (EW) von A, wenn es ein $x \in \mathcal{K}$ {0} gibt mit $Ax = \lambda x$. Jedes $x \in X$ {0} mit $Ax = \lambda x$ heißt zu λ gehörender Eigenvektor oder Eigenelement (EV). Die Menge {$x \in X$: $Ax = \lambda x$ } heißt der zu λ gehörende Eigenunterraum (EUR). Die Dimension des Eigenunterraums nennt man geometrische Vielfachheit des Eigenwertes λ und bezeichnet diese mit $\gamma(\lambda)$.*

Haben also
λ EW von A $\Leftrightarrow$ Ax - λx = 0 hat eine nichttriviale Lösung
$\Leftrightarrow$ ker (A - λI) $\neq$ {0}
x EV $\Leftrightarrow$ x $\in$ ker (A - λI) \ {0}
EUR = ker (A - λI), $\gamma(\lambda)$ = dim ker (A - λI).

1. Beispiel 12.05.06

A: $\mathcal{R}^2 \to \mathcal{R}^2$
Sei gegeben durch die Matrix $\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$.
A$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$, d.h. λ ist Eigenwert und jedes Element aus $\mathcal{R}^2$ \{0} ist ein Eigenvektor. Also $\gamma(\lambda) = 2$.

Sei A: $\mathcal{R}^3 \to \mathcal{R}^3$ gegeben durch die Matrix $\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$. Dann ist

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \lambda_1 x \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} \lambda x & \\ \lambda y & cr\lambda z \end{pmatrix}$$

Also: λ_1 ist EW, EV ist $\begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix}$,
λ_2 ist EW, EV ist $\begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix}$,
λ_3 ist EW, EV ist $\begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix}$.

2. Beispiel:
A: $\mathcal{R}^2 \to \mathcal{R}^2$ sei Spiegelung an der x-Achse.
x ist EV genau dann, wenn x $\neq$ 0 ist und x auf der x-Achse oder der y-Achse liegt.
Für x auf der x-Achse ist AX = x, d.h. EW ist 1. Für x auf der y-Achse ist Ax = -x, d.h. EW ist -1. Geometrische Vielfachheiten sind immer 1.

3. Beispiel:
A: $\mathcal{R}^2 \to \mathcal{R}^2$ sei Drehung um Winkel $\alpha \notin \{0°,180°\}$
Dann gibt es kein x $\neq$ 0 mit Ax = λx, d.h. A hat keine EW und keine EV.

4. Beispiel:

A: $\mathcal{R}^3 \to \mathcal{R}^3$ sei gegeben durch die Matrix $\frac{1}{9}\begin{pmatrix} 19 & -2 & 4 \\ 4 & 10 & -2 \\ 4 & -8 & 25 \end{pmatrix}$.

Dann gilt

$$A\begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, A\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = 2\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, A\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}.$$

Operator hat also die 3 EW 1,2,3.

Definition 4 *Sei $A \in M_n$ ($\mathcal{K}$). Unter den Eigenwerten, Eigenvektoren, Eigenunterräumen, Geometrischen Vielfachheiten der Matrix A versteht man die entsprechenden Objekte für A als Operator auf dem Spaltenraum $M_{n,1}$ ($\mathcal{K}$).*

Sei A $\in M_n$ ($\mathcal{K}$). Dann ist

$$\det(\mathrm{A} - \lambda\mathrm{I}) = \det\begin{pmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{pmatrix}$$

$= (-1)^n\lambda^n + p_{n-1}\lambda^{n-1} + \ldots + p_1\lambda + p_0$.

Vermerken, dass $p_0 = \det$ A, $p_{n-1} = (-1)^{n-1}(a_{11} + a_{22} + \ldots + a_{nn}) =: (-1)^{n-1}$ *tr A* $=: (-1)^{n-1}$ *spur A* gilt.

Ist A $\in \mathcal{L}$(X) (dim X = n), so wählen wir uns eine Basis E in X und definieren
$\det(\mathrm{A} - \lambda\mathrm{I}) := \det[\mathrm{A} - \lambda\mathrm{I}]_{E,E}$.

Nach Satz 6/2.6. sind alle möglichen Matrixdarstellungen $[\mathrm{A} - \lambda\mathrm{I}]_{E,E}$ ähnlich zueinander, d.h. $\det[\mathrm{A} - \lambda\mathrm{I}]_{E,E}$ hängt nicht von der speziellen Wahl von E ab.
($\det \mathrm{C}\,[\mathrm{A} - \lambda\mathrm{I}]_{E,E}\,\mathrm{C}^{-1} = \det \mathrm{C} \det[\mathrm{A} - \lambda\mathrm{I}]_{E,E} \det \mathrm{C}^{-1} = \det[\mathrm{A} - \lambda\mathrm{I}]_{E,E}$).

Definition 5 *Sei $A \in M_n$ ($\mathcal{K}$) oder $A \in \mathcal{L}(X)$. Das Polynom $\det(A - \lambda I)$ heißt charakteristisches Polynom von A. Die Menge der komplexen Nullstellen des charakteristischen Polynoms, d.h. die $\lambda \in \mathcal{C}$ mit $\det(A - \lambda I) = 0$, wird Spektrum von A genannt und mit sp A bezeichnet.*

Andere Bezeichungen: sp A $= \sigma(\mathrm{A}) = \Lambda(\mathrm{A}) \ldots$

Satz 1 *Sei $A \in M_n$ ($\mathcal{K}$) oder $A \in \mathcal{L}(X)$ über $\mathcal{K}$. Die Menge der Eigenwerte von A ist dann gleich sp A $\cap\, \mathcal{K}$.*

Beweis:
Sei $\lambda \in \mathcal{K}$ ein EW. Dann ist A - λI nicht injektiv und somit nicht invertierbar, d.h. det

(A - λI) = 0. Also gilt Menge der Eigenwerte $\subset \mathcal{K} \cap$ sp A.
Umgekehrt sei $\lambda \in \mathcal{K}$ und det (A - λI) = 0. Dann ist r = rg(A - λI) < n. Nach Satz 2/2.8. ist somit dim ker (A - λI) = n - r > 0, d.h. λ ist EW. #
Haben also erstmals ein Resultat, bei dem der Skalarkörper eine entscheidende Rolle spielt.

5. Beispiel:
Sei A wie im 3. Beispiel. Wöhlen Standardbasis E im $\mathcal{R}^2$. Also ist

$$[A]_{E,E} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix},$$

$$det\ [A-\lambda I]_{E,E} = \begin{vmatrix} \cos\alpha-\lambda & -\sin\alpha \\ \sin\alpha & \cos\alpha-\lambda \end{vmatrix}$$

$$= \lambda^2 - 2\lambda\cos\alpha + \cos^2\alpha + \sin^2\alpha$$

$$= \lambda^2 - 2\lambda\cos\alpha + 1 \stackrel{!}{=} 0$$

$$\lambda_{1,2} = \cos\alpha \pm \sqrt{\cos^2\alpha - 1} = \cos\alpha \pm \sqrt{-\sin^2\alpha} = \cos\alpha \pm i\sin\alpha$$

Für $\alpha \neq \{0°,180°\}$ sind λ_1, λ_2 nicht reell, d.h. A hat keine Eigenwerte als Operator im $\mathcal{R}^2$. Aber A hat die beiden EW λ_1, λ_2 als Operator im $\mathcal{C}^2$.

Definition 6 *Sei $A \in M_n(\mathcal{K})$ oder $A \in \mathcal{L}(X)$. Die algebraische Vielfachheit eines Eigenwerts λ_0 von A ist das größte μ, für das det(A - λI) durch $(\lambda - \lambda_0)^\mu$ teilbar ist.*
Die algebraische Vielfachheit von λ_0 bezeichnen wir mit $\alpha(\lambda_0)$. Ein Eigenwert heißt einfach, wenn die algebraische Vielfachheit 1 ist.

6. Beispiel:
Sei A = $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Haben

$$\begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} = \lambda^2 - 2\lambda + 1 = (\lambda-1)^2.$$

Also sp A = {1}, Menge der EW von A = {1}, $\alpha(1) = 2$.
$\begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ ist EV $\Leftrightarrow \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$
$\Leftrightarrow \begin{pmatrix} x+y \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \Leftrightarrow$ y = 0,
d.h. EUR = $\{\begin{pmatrix} x \\ 0 \end{pmatrix}$: x $\in \mathcal{K}\}$ = span $\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\}$. Also $\gamma(1)$ = dim EUR = 1.

Satz 2 *Es gilt stets $\gamma(\lambda) \leq \alpha(\lambda)$.*

Beweis:
Sei λ_0 ein EW von A mit $\gamma(\lambda_0) = \gamma$, $\alpha(\lambda_0) = \alpha$. Wählen einen Komplementärraum V zu U := ker (A - λ_0I).

Sei $\{e_1, \ldots, e_\gamma\}$ eine Basis in U und $\{e_{\gamma+1}, \ldots, e_n\}$ eine Basis in V. Dann ist E = $\{e_1, \ldots, e_\gamma, e_{\gamma+1}, \ldots, e_n\}$ eine Basis in gesamten Raum. Haben

$$[A]_{E,E} = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$$

mit B = $\begin{pmatrix} \lambda_0 & & 0 \\ & \ddots & \\ 0 & & \lambda_0 \end{pmatrix}$.
Dies folgt aus A $e_j = \lambda_0\, e_i$ für $1 \leq j \leq \gamma$. Es ergibt sich
det (A - λI) = det $[A - \lambda I]_{E,E}$ = det $\begin{pmatrix} B - \lambda I & C \\ 0 & D - \lambda I \end{pmatrix}$
= det (B - λI) * det (D - λI) = det $\begin{pmatrix} \lambda_0 - \lambda & & \\ & \ddots & \\ & & \lambda_0 - \lambda \end{pmatrix}_{\gamma \times \gamma}$ det (D - λI)
= $(\lambda_0 - \lambda)^\gamma$ det (D - λ I).

Sehen also, dass det (A - λ I) durch $(\lambda - \lambda_0)^\gamma$ teilbar ist. Somit ist $\alpha \geq \gamma$. #

17.05.06

Man bestimmt also zunächst alle Nullstellen von det (A - λ I). Damit hat man auch die algebraischen Vielfachheiten. Für $\lambda_0 \in$ sp A $\cap\, \mathcal{K}$ löst man dann die Gleichung (A - λ_0 I)x = 0 (z.B. über Gauß'schen Algorithmus) und ermittelt so alle Eigenvektoren und damit auch die geometrischen Vielfachheiten.

Definition 7 *Ein Operator $A \in \mathcal{L}(X)$ heißt diagonalisierbar, wenn es eine Basis E in X gibt, sodass*

$$[A]_{E,E} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

mit $\lambda_i \in \mathcal{K}$ gilt.

Eine Matrix $A \in M_n(\mathcal{K})$ heißt diagonalisierbar, wenn es eine invertierbare Matrix $C \in M_n(\mathcal{K})$ gibt, sodass

$$C^{-1}AC = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

mit $\lambda_i \in \mathcal{K}$ ist.

Aus Satz 6/2.6. folgt, dass ein Operator genau dann diagonalisierbar ist, wenn irgendeine seiner Matrixdarstellungen $[A]_{F,F}$ diagonalisierbar ist, und in diesem Fall sind alle Matrixdarstellungen $[A]_{F,F}$ diagonalisierbar.

Wenn A $\in \mathcal{L}$(X) diagonalisierbar ist, dann sind die Zahlen λ_i alles Eigenwerte von A (ist E = $\{e_1, \ldots, e_n\}$, so ist A$e_i = \lambda_i e_i$).

Aus $C^{-1}AC = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ folgt $AC = C \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ und für die i-te Spalte s_i von C gilt somit A$s_i = \lambda_i s_i$, d.h. die Spalten von C sind gerade die Eigenvektoren.

Lemma: Seien $\lambda_1, \ldots, \lambda_m$ verschiedene Eigenwerte eines Operators A $\in \mathcal{L}$(X) und seien $E_1, \ldots, E_m$ die entsprechenden Eigenunterräume. Dann gilt

$$E_1 + \ldots + E_m = E_1 \boxplus \ldots \boxplus E_m,$$

d.h. jedes x $\in E_1 + \ldots + E_m$ ist eindeutig in der Form x = $x_1 + \ldots + x_m$ mit $x_i \in E_i$ darstellbar.

Beweis:
Ist für m = 1 klar. Behauptung sei für m - 1 richtig. Zeigen sie für m.
Sei dazu $x_i \in E_i$ und $x_1 + \ldots + x_m = 0$.
Anwendung von A ergibt

$$\lambda_1 x_1 + \ldots + \lambda_{m-1} x_{m-1} + \lambda_m x_m = 0$$

und Multiplikation mit λ_m liefert

$$\lambda_m x_1 + \ldots + \lambda_m x_{m-1} \lambda_m x_m = 0.$$

Subtraktion bringt dann

$$\underbrace{(\lambda_1 - \lambda_m)x_1}_{\in E_1} + \ldots + \underbrace{(\lambda_{m-1} - \lambda_m)x_{m-1}}_{\in E_{m-1}} = 0,$$

und nach Induktionsvoraussetzung ist also (λ_i - λ_m)x_i = 0 für i = 1, ..., m-1. Somit ist x_i = 0 für i = 1, ...,m-1. Es folgt x_m = 0.

Also gilt $x_1 = \ldots = x_m = 0$. #

Folgerung: Zu verschiedenen Eigenwerten gehörende Eigenvektoren sind stets linear unabhängig.

Satz 3 *Ein Operator $A \in \mathcal{L}(X)$ ist genau dann diagonalisierbar, wenn sp $A \subset \mathcal{K}$ und $\gamma(\lambda) = \alpha(\lambda)$ für alle Eigenwerte von λ von A ist.*

Beweis:
Sei sp A $\subset \mathcal{K}$ und $\gamma(\lambda) = \alpha(\lambda)$ $\forall\lambda$. Haben

$$det\ (A - \lambda I) = (\lambda_1 - \lambda)^{\alpha(\lambda_1)} \dots (\lambda_m - \lambda)^{\alpha(\lambda_m)},$$

wobei $\lambda_1, \dots, \lambda_m$ verschieden sind.

Die Zahlen $\lambda_1, \dots, \lambda_m$ sind alles Eigenwerte von A (sp A $\subset \mathcal{K}$ und Satz 1). Seien $E_1, \dots, E_m$ die Eigenunterräume. Nach dem Lemma ist
dim $(E_1 + \dots + E_m)$ = dim $(E_1 \boxplus \dots \boxplus E_m)$ = dim E_1 + … + dim E_m
= $\gamma(\lambda_1)$ + … + $\gamma(\lambda_m)$ = $\alpha(\lambda_1)$ + … + $\alpha(\lambda_m)$ = n = dim X,
woraus X = $E_1 \boxplus \dots \boxplus E_m$ folgt.
Wählen Basen in $E_1, \dots, E_m$. Deren Vereinigung ist dann also eine Basis E = $\{e_1, \dots, e_n\}$ in X.
Alle Elemente e_i sind Eigenvektoren, A$e_i = \mu_i e_i$, d.h. $[A]_{E,E} = \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix}$.

Umgekehrt sei A diagonalisierbar:

$$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Bezeichnen mit $\mu_1, \dots, \mu_m$ die verschiedenen Elemente auf der Diagonalen. Dabei möge μ_i genau q_i mal vorkommen. Haben also

$$det\ (A - \lambda I) = (\lambda_1 - \lambda) \dots (\lambda_n - \lambda) = (\mu_1 - \lambda)^{q_1} \dots (\mu_m - \lambda)^{q_m}.$$

Jedes μ_i ist ein Eigenwert von A.
Es gibt q_i Elemente x aus der Basis E mit Ax = μ_ix.
Bezeichnen mit Q_i die Aufspannung dieser Elemente. Dann ist $Q_i \subset E_i$ = EUR von μ_i. Dies ergibt zunächst q_i = dim $Q_i \leq$ dim $E_i = \gamma(\mu_i)$.

q_i ist die algebraische Vielfachheit von μ_i, d.h. $q_i = \alpha(\mu_i)$. Nach Satz 2 ist $\alpha(\mu_i) \geq \gamma(\mu_i)$.
Zusammen mit der Ungleichung $q_i \leq \gamma(\mu_i)$ ergibt dies $\alpha(\mu_i) = \gamma(\mu_i)$ für alle i. Schließlich ist sp A = $\{\mu_1, \dots, \mu_m\} \subset \mathcal{K}$, da nach Definition 7 die Zahlen $\lambda_1, \dots, \lambda_n$ alle zu $\mathcal{K}$ gehören. #

7. Beispiel:
Sei A = $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.
Wissen, dass sp A = {1} mit $\alpha(1) = 2$ und $\gamma(1) = 1$ (früheres Beispiel) ist.
Diese Matrix lässt sich also weder über $\mathcal{R}$ noch über $\mathcal{C}$ diagonalisieren.

Ist A = $\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$ mit $\alpha \notin \{0°, 180°\}$,
so hat A keine Eigenwerte über $\mathcal{R}$, ist über $\mathcal{R}$ also nicht diagonalisierbar.
Wissen, dass sp A = {cos α + isin α, cos α - isin α}.
Wegen sp A $\not\subset \mathcal{R}$ ist A über $\mathcal{R}$ nicht diagonalisierbar. Aus sp A $\subset \mathcal{C}$, $\alpha(\lambda) = 1$, $\gamma(\lambda) = 1$ folgt, dass A über $\mathcal{C}$ diagonalisierbar ist.

Folgerung: Sind alle Eigenwerte eines Operatos (einer Matrix) einfach, so ist der Operator (die Matrix) über $\mathcal{C}$ diagonalisierbar. Ist überdies $\mathcal{K} = \mathcal{R}$ und sind alle Punkte aus dem Spektrum reell, so ist der Operator (die Matrix) über $\mathcal{R}$ diagonalisierbar.

Beweis:
Voraussetzung impliziert, dass sp A $\subset \mathcal{K}$ gilt und $\alpha(\lambda) = 1\ \forall\lambda$ ist.
Nach Satz 2 ist $\gamma(\lambda) = 1\ \forall\lambda$ und Satz 3 liefert somit die Behauptung. #
Hatten in 2.6. Äquivalenz von Matrizen betrachtet (A ~ B $\Leftrightarrow$ $\exists$ M,N invertirbar mit MAN = B) und im Satz 5/2.7. festgestellt, dass jede Matrix aus M_n ($\mathcal{K}$) zu genau einer Matrix aus der Liste

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} I_1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} I_2 & 0 \\ 0 & 0 \end{pmatrix}, \ldots, \begin{pmatrix} I_{n-1} & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$

äquivalent ist.
Dies impliziert, dass es für jeden Operator A $\in \mathcal{L}$(X) ($\dim_{\mathcal{K}}$ X = n) Basen E,F gibt, sodass $[A]_{E,F}$ eine Matrix obiger Liste ist und dass man auf diese Weise niemals zu zwei verschiedenen Matrizen aus der Liste kommt.

Suchen nun eine solche Liste bezüglich der Ähnlichkeit von Matrizen (A ähnlich zu B $\Leftrightarrow$ $\exists$ C invertierbar mit $C^{-1}AC = B$) bzw. bezüglich der Matrixdarstellung $[A]_{E,E}$ von Operatoren.

Definition 8 *Eine (r x r)-Matrix mit Einträgen aus $\mathcal{K}$ heißt Jordankästchen, wenn sie von der Form*

$$J_r(\lambda) := \begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & & 1 \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$$

mit $\lambda \in \mathcal{K}$ ist.
Haben also

$$J_1(\lambda) = (\lambda), J_2(\lambda) = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, J_3(\lambda) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}, \ldots$$

Eine Matrix mit Einträgen aus $\mathcal{K}$ heißt Jordanmatrix, wenn sie blockdiagonal ist und die Diagonalblöcke alle Jordankästchen mit Einträgen aus $\mathcal{K}$ sind.

24.05.06

Jordanmatrizen sind zum Beispiel

$$\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}, \left(\begin{array}{c|cc} \lambda & 0 & 0 \\ \hline 0 & \mu & 1 \\ 0 & 0 & \mu \end{array}\right), \left(\begin{array}{ccc|c|cc} \lambda & 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & \lambda & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \mu & 1 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{array}\right).$$

Satz 4 *(a) Zwie Jordanmatrizen sind genau dann (über $\mathcal{K}$) ähnlich, wenn sie sich lediglich in der Reihenfolge (Anordnung) der Jordankästchen unterscheiden.*

(b) Jede Matrix $A \in M_n$ ($\mathcal{K}$) mit der Eigenschaft sp $A \subset \mathcal{K}$ ist zu einer bis auf die Anordnung der Jordankästchen eindeutig bestimmten Jordanmatrix ähnlich, der sogenannten Jordanschen Normalform.

(c) Für jeden Operator $A \in \mathcal{L}(X)$ mit der Eigenschaft sp $A \subset \mathcal{K}$ existiert eine Basis E in X, sodass $[A]_{E,E}$ eine bis auf die Reihenfolge eindeutig bestimmte Jordanmatrix ist, die Jordansche Normalform von A.

Beweis: siehe Literatur. #

Jordansche Normalformen von Matrizen aus M_1 ($\mathcal{C}$) sind also (λ). Für Matrizen aus M_2 ($\mathcal{C}$) ist die Liste der JNF wie folgt:

$$\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} (\lambda \succ \mu), \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}, \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

($\lambda \succ \mu \Leftrightarrow$ Re $\lambda >$ Re μ oder Re $\lambda =$ Re μ, Im $\lambda >$ Im μ)

JNF für Matrizen aus $M_3(\mathcal{C})$:

$$\begin{pmatrix} \lambda & & \\ & \mu & \\ & & \nu \end{pmatrix} (\lambda \succ \mu \succ \nu), \begin{pmatrix} \lambda & & \\ & \lambda & \\ & & \mu \end{pmatrix} (\lambda \neq \mu), \begin{pmatrix} \lambda & 1 & \\ 0 & \lambda & \\ & & \mu \end{pmatrix} (\lambda \neq \mu),$$

$$\begin{pmatrix} \lambda & & \\ & \lambda & \\ & & \lambda \end{pmatrix}, \begin{pmatrix} \lambda & 1 & \\ 0 & \lambda & \\ & & \lambda \end{pmatrix}, \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

Die Bestimmung der JNF ist im Allgemeinen nicht leicht. Klar ist, dass die Einträge auf der Hauptdiagonalen der JNF gerade die Eigenwerte sind (fordern stets sp A $\subset \mathcal{K}$) und dort entsprechend ihrer algebraischen Vielfachheit auftreten.
Man kann zeigen, dass die geometrische Vielfachheit eines Eigenwerts gleich der Anzahl der Jordankästchen mit diesem Eigenwert ist.

Im Prinzip kann man die JNF wie folgt bestimmen:
Bezeichnen mit $D_k(\lambda)$ (k=1,. . .,n) den ggT aller Minoren der Ordnung k von A -λI (A $\in M_n(\mathcal{K})$, sp A $\subset \mathcal{K}$). Dies sind Polynome in λ. Insbesondere ist $D_n(lambda) =$ det (A - λI). Man setzt noch $D_0(\lambda)$:=1. Man kann zeigen, dass $D_k(\lambda)$ durch D_{k-1} (λ) teilbar ist. Die Polynome

$$E_k(\lambda) = \frac{D_k(\lambda)}{D_{k-1}(\lambda)} \quad (k = 1, \ldots, n)$$

nennt man Elementarteiler oder *invariante Polynome* von A.

Man kann wieder zeigen, dass $\mathrm{E}_k(\lambda)$ durch $\mathrm{E}_{k-1}(\lambda)$ teilbar ist. Haben also

$$E_n(\lambda) = (\lambda_1 - \lambda)^{k_1}(\lambda_2 - \lambda)^{m_1} \dots$$

$$E_{n-1}(\lambda) = (\lambda_1 - \lambda)^{k_2}(\lambda_2 - \lambda)^{m_2} \dots$$

$$\vdots$$

$$E_1(\lambda) = (\lambda_1 - \lambda)^{k_n}(\lambda_2 - \lambda)^{m_n} \dots$$

mit $k_1 \geq k_2 \geq \dots \geq k_n$, $m_1 \geq m_2 \geq \dots \geq m_n$, ...
Die JNF von A hat dann die Jordankästchen

$$J_{k_1}(\lambda_1), J_{k_2}(\lambda_1), \dots, J_{k_n}(\lambda_1),$$

$$J_{m_1}(\lambda_2), J_{m_2}(\lambda_2), \dots, J_{m_n}(\lambda_2), \dots$$

mit der Vereinbarung, dass $J_0(\lambda)$ nicht gezählt wird (leeres Kästchen).

Beispiel:
Sei A $\in M_{10}(\mathcal{K})$ und

$$E_{10}(\lambda) = (3-\lambda)^3(4-\lambda)^2$$
$$E_9(\lambda) = (3-\lambda)^2(4-\lambda)$$
$$E_8(\lambda) = (4-\lambda)$$
$$E_7(\lambda) = 4-\lambda$$
$$E_6(\lambda) = \dots = E_1(\lambda) = 1.$$

Die JNF ist

$$\begin{pmatrix}
3 & 1 & 0 & & & & & & & \\
0 & 3 & 1 & & & & & & & \\
0 & 0 & 3 & & & & & & & \\
 & & & 3 & 1 & & & & & \\
 & & & 0 & 3 & & & & & \\
 & & & & & 4 & 1 & & & \\
 & & & & & 0 & 4 & & & \\
 & & & & & & & 4 & & \\
 & & & & & & & & 4 & \\
 & & & & & & & & & 4
\end{pmatrix}$$

Folgerung: A mit sp A $\subset \mathcal{K}$ ist diagonalisierbar genau dann, wenn alle Elementarteiler nur einfache Nullstellen haben.

Folgerung: Zwei Matrizen A,B $\in M_n\ (\mathcal{K})$ mit Spektrum in $\mathcal{K}$ (sp A, sp B $\subset \mathcal{K}$) sind genau dann ähnlich, wenn sie die gleichen Elementarteiler haben.

Die Elementarteiler bilden also einen vollständigen Satz von Invarianten für die Ähnlichkeit.

Definition 9 *Ein Polynom*

$$p(\lambda) = p_m\lambda^m + p_{m-1}\lambda^{m-1} + \ldots + p_1\lambda + p_0$$

mit Koeffizienten aus $\mathcal{C}$ heißt annulierendes Polynom für eine Matrix $A \in M_n(\mathcal{C})$, wenn

$$p(A) = p_m A^m + p_{m-1}A^{m-1} + \ldots + p_1 A + p_0 I = \mathcal{O}$$

ist.

Es ist unschwer zu sehen, dass $(\lambda_0 - \lambda)^r$ ein annulierendes Polynom für $J_r(\lambda_0)$ ist. Klar für r = 1:
$J_1(\lambda_0) = \lambda_0$, $(\lambda_0 - \lambda_0)^1 = 0$.

Beweis für r = 2:

$$J_2(\lambda_0) = \begin{pmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{pmatrix}, J_2^2(\lambda_0) = \begin{pmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{pmatrix}\begin{pmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{pmatrix} = \begin{pmatrix} \lambda_0^2 & 2\lambda_0 \\ 0 & \lambda_0^2 \end{pmatrix}$$

$$p(J_2(\lambda_0)) = \lambda_0^2 I - 2\lambda_0 J_2(\lambda_0) + J_2^2(\lambda_0)$$

$$= \begin{pmatrix} \lambda_0^2 & 0 \\ 0 & \lambda_0^2 \end{pmatrix} - 2\lambda_0\begin{pmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{pmatrix} + \begin{pmatrix} \lambda_0^2 & 2\lambda_0 \\ 0 & \lambda_0^2 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Definition 10 *Das annulierende Polynom von minimalem Grad und mit höchstens Koeffizienten 1 ($p_m = 1$) heißt minimales Polynom oder Minimalpolynom von A.*

Satz 5 *(a) Das Minimalpolynom von A ist $\pm E_n(\lambda)$.*

(b) (Satz von Cayley-Hamilton)
Das charakteristische Polynom $D_n(\lambda)$ = det (A - λI) ist ein annulierendes Polynom von A.

(c) Der Grad des Minimalpolynoms ist die Summe der Dimensionen (Ordnungen) der maximalen Jordankästchen für jeden Eigenwert, d.h. ist $k_1 + m_1 + \ldots$

(d) Das Minimalpolynom ist $\pm$ mal das charakteristische Polynom genau dann, wenn es für jeden Eigenwert nur ein Jordankästchen gibt

Beweis: HA
$((\lambda_0 - \lambda)^r$ ist annulierend für $J_r(\lambda_0)$

$\Rightarrow E_n(\lambda) = (\lambda_1 - \lambda)^{k_1}(\lambda_2 - \lambda)^{m_1} \ldots$ ist annulierend für JNF.)

Kommen nun zu Matrizen A $\in M_n(\mathcal{R})$, für die nicht notwendigerweise sp A $\subset \mathcal{R}$ ist.

Definition 11 *Eine reelle r x r-Matrix heißt reelles Jordankästchen, wenn sie von der Form*

$$\begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & 1 \\ & & \ddots & \lambda \end{pmatrix}$$

mit $\lambda \in \mathcal{R}$ oder von der Form

$$\begin{pmatrix} \sigma & \tau & 1 & 0 & & & \\ -\tau & \sigma & 0 & 1 & & & \\ & & \sigma & \tau & \ddots & & \\ & & -\tau & \sigma & & & \\ & & & & \ddots & 1 & 0 \\ & & & & & 0 & 1 \\ & & & & & \sigma\tau & \\ & & & & & -\tau & \sigma \end{pmatrix}$$

mit $\sigma, \tau \in \mathcal{R}$ ist.
(Letztere Jordankästchen gibt es bloß für gerade r.)
Eine reelle Jordanmatrix ist eine Blockdiagonalmatrix, deren Diagonalblöcke reelle Jordankästchen sind.

26.05.06

Satz 6 *(a) Zwei reelle Jordanmatrizen sind genau dann ähnlich, wenn sie sich nur in der Reihenfolge der reellen Jordankästchen unterscheiden.*
(b) Jede Matrix aus $M_n(\mathcal{R})$ ist über $\mathcal{R}$ ähnlich zu einer (bis auf die Reihenfolge der Jordankästchen eindeutig bestimmten) reellen Jordanmatrix, ihrer sogenannten reellen Jordanschen Normalform.
(c) Für jeden Operator $A \in \mathcal{L}(X)$ über einem endlichen linearen Raum X über $\mathcal{R}$ gibt es eine Basis E in X, sodass $[A]_{E,E}$ eine (bis auf die Reihenfolge der Jordankästchen eindeutig bestimmte) reelle Jordanmatrix ist, die sogenannte reelle Jordansche Normalform von A.

Beweis: Literatur. #

JNFn von Matrizen aus $M_4(\mathcal{C})$:

$$\begin{pmatrix} \lambda & & & \\ & \mu & & \\ & & \nu & \\ & & & \sigma \end{pmatrix}, \begin{pmatrix} \lambda & & & \\ & \lambda & & \\ & & \mu & \\ & & & \nu \end{pmatrix}, \begin{pmatrix} \lambda & 1 & & \\ 0 & \lambda & & \\ & & \mu & \\ & & & \nu \end{pmatrix}, \begin{pmatrix} \lambda & & & \\ & \lambda & & \\ & & \mu & \\ & & & \mu \end{pmatrix}, \begin{pmatrix} \lambda & 1 & & \\ 0 & \lambda & & \\ & & \mu & \\ & & & \mu \end{pmatrix},$$

$$\begin{pmatrix} \lambda & 1 & & \\ 0 & \lambda & & \\ & & \mu & 1 \\ & & 0 & \mu \end{pmatrix}, \begin{pmatrix} \lambda & & & \\ & \lambda & & \\ & & \lambda & \\ & & & \mu \end{pmatrix}, \begin{pmatrix} \lambda & 1 & & \\ 0 & \lambda & & \\ & & \lambda & \\ & & & \mu \end{pmatrix}, \begin{pmatrix} \lambda & 1 & 0 & \\ 0 & \lambda & 1 & \\ 0 & 0 & \lambda & \\ & & & \mu \end{pmatrix}, \begin{pmatrix} \lambda & & & \\ & \lambda & & \\ & & \lambda & \\ & & & \lambda \end{pmatrix},$$
$$\begin{pmatrix} \lambda & 1 & & \\ 0 & \lambda & & \\ & & \lambda & \\ & & & \lambda \end{pmatrix}, \begin{pmatrix} \lambda & 1 & & \\ 0 & \lambda & & \\ & & \lambda & 1 \\ & & 0 & \lambda \end{pmatrix}, \begin{pmatrix} \lambda & 1 & 0 & \\ 0 & \lambda & 1 & \\ 0 & 0 & \lambda & \\ & & & \lambda \end{pmatrix}, \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}.$$

JNFn von Matrizen aus $M_4(\mathcal{R})$ sind die Obigen mit reellen λ, μ, ν und σ und die Folgenden:

$$\begin{pmatrix} \sigma & \tau & & \\ -\tau & \sigma & & \\ & & \lambda & \\ & & & \mu \end{pmatrix}, \begin{pmatrix} \sigma & \tau & & \\ -\tau & \sigma & & \\ & & \lambda & \\ & & & \lambda \end{pmatrix}, \begin{pmatrix} \sigma & \tau & & \\ -\tau & \sigma & & \\ & & \lambda & 1 \\ & & 0 & \lambda \end{pmatrix}, \begin{pmatrix} \sigma & \tau & & \\ -\tau & \sigma & & \\ & & \alpha & \beta \\ & & -\beta & \alpha \end{pmatrix},$$
$$\begin{pmatrix} \sigma & \tau & & \\ -\tau & \sigma & & \\ & & \sigma & \tau \\ & & -\tau & \sigma \end{pmatrix}, \begin{pmatrix} \sigma & \tau & 1 & 0 \\ -\tau & \sigma & 0 & 1 \\ & & \sigma & \tau \\ & & -\tau & \sigma \end{pmatrix}.$$

3 Räume mit Skalarprodukt und ihre Operatoren

3.1 Euklidische und unitäre Räume

31.05.06

Mengen: zählen
Lineare Räume: rechnen (Addition, Multiplikation mit Skalaren)
Wollen Geometrie betreiben. Dazu müssen wir Längen und Winkel messen können. Dies führ zu Räumen mit Skalarprodukt.

Definition 1 *Sei X ein linearer Raum über $\mathcal{K}$ ($\in \{\mathcal{R}, \mathcal{C}\}$). Ein Skalarprodukt (= inneres Produkt) auf X ist eine Abbildung von $X \times X$ in $\mathcal{K}$, die jedem geordneten Paar (x,y) $\in X \times X$ einen Skalar (x,y) zuordnet. (= <x,y>, x * y, <x|y>) und folgende Eigenschaften hat:*
(i) (x,y) = $\overline{(y,x)}$ $\forall$ x,y $\in X$ (Symmetrie)
(ii) (αx,y) = α(x,y) $\forall$ $\alpha \in \mathcal{K}$ $\forall$ x,y $\in X$
(iii) (x + y,z) = (x,z) + (y,z) $\forall$ x,y,z $\in X$
(iv) (x,x) $\in [0,\infty)$ $\forall$ x $\in X$ und (x,x) = 0 $\Leftrightarrow$ x = 0 (Positive Definitheit)

Zunächst einige wichtige Bemerkungen:

$\mathcal{K} = \mathcal{R}$	$\mathcal{K} = \mathcal{C}$
(x,y) = (y,x)	(x,y) = $\overline{(x,y)}$
(αx,y) = α(x,y)	(αx,y) = α(x,y)
(x,αy) = α(x,y)	(x,αy) = $\overline{\alpha}$(x,y)
(x + y,z) = (x,z) + (y,z)	(x + y,z) = (x,z) + (y,z)
(x,y + z) = (x,y) + (x,z)	(x,y + z) = (x,y) + (x,z)
Skalarprodukt ist eine bilineare Form	Skalarprodukt ist eine sesquilineare Form

Manchmal (besonders in der Physik) ersetzt man Axiom (ii) durch

$$(x, \alpha y) = \alpha(x, y) \quad \forall \alpha \in \mathcal{K} \quad \forall x, y \in X$$

Dann hat man

$$(\alpha x, y) = \overline{\alpha}(x, y) \quad \forall \alpha \in \mathcal{K} \quad \forall x, y \in X.$$

Definition 2 *Ein linearer Raum über $\mathcal{K}$ mit Skalarprodukt heißt Euklidischer Raum für $\mathcal{K} = \mathcal{R}$ und Unitärer Raum für $\mathcal{K} = \mathcal{C}$.*

1. Beispiel:
X = $\mathcal{R}^3$ (über $\mathcal{R}$) mit (x,y) = |x||y| cos φ ist Euklidischer Raum

2. Beispiel:
X = $\mathcal{K}^n$ mit (x,y) = $((x_1, \ldots, x_n),(y_1, \ldots, y_n)) = x_1\overline{y_1} + \ldots + x_n\overline{y_n}$ ist Raum mit Skalarprodukt
(y,x) = $y_1\overline{x_1} + \ldots + y_n\overline{x_n} = \overline{x_1\overline{y_1} + \ldots + x_n\overline{y_n}} = \overline{(x, y)}$
(ii),(iii) trivial
(x,x) = $x_1\overline{x_1} + \ldots + x_n\overline{x_n} = |x_1|^2 + \ldots + |x_n|^2$

Eine Verallgemeinerung ist wie folgt:
Sei A = $(a_{ij})_{i,j=1}^n \in M_n(\mathcal{K})$. Definieren $(x, y)_A = (x, Ay)$, wobei (-,-) wie oben ist.
Ausgeschrieben:

$$(x,y)_A = \left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right) = \left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} a_{11}y_1 + \ldots + a_{1n}y_n \\ \vdots \\ a_{n1}y_1 + \ldots + a_{nn}y_n \end{pmatrix} \right)$$

$$= x_1(\overline{a_{11}y_1 + \ldots + a_{1n}y_n}) + \ldots + x_n(\overline{a_{n1}y_1 + \ldots + a_{nn}y_n}) = \sum_{i,j=1}^n \overline{a_{ij}} x_i \overline{y_j}.$$

Symmetrie:

$$\overline{(y,x)_A} = \overline{\sum_{i,j=1}^n \overline{a_{ij}} y_i \overline{x_j}} = \sum_{i,j=1}^n a_{ij} x_j \overline{y_i} = \sum_{i,j=1}^n a_{ji} x_i \overline{x_j}$$

Haben also Symmetrie genau dann, wenn $\overline{a_{ij}} = a_{ji}$ ∀ i,j ist, d.h. genau dann, wenn A = A* ist (A* := $\overline{A^T}$) (solche Matrizen heißen *hermitesch*).

$\mathcal{K} = \mathcal{R}$
A = $\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$
A = $\begin{pmatrix} 1 & 2+i \\ 2-i & 3 \end{pmatrix}$ ist keine Matrix aus $M_2(\mathcal{R})$

$\mathcal{K} = \mathcal{C}$
A = $\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$
A = $\begin{pmatrix} 1 & 2+i \\ 2-i & 3 \end{pmatrix}$
A = $\begin{pmatrix} 1 & 2+i \\ 2+i & 3 \end{pmatrix}$ nicht hermitesch

Axiome (ii) und (iii) sind stets erfüllt.
Axiom (iv) ist diffizil. Sei z.B. A = $\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$ und $\mathcal{K} = \mathcal{R}$. Dann ist
$(x, x)_A = x_1^2 + 2x_1x_2 + 2x_2x_1 + 5x_2^2 = x_1^2 + 4x_1x_25x_2^2 = (x_1 + 2x_2)^2 + x_2^2 \geq 0$

und $(x,x)_A = 0 \Leftrightarrow x_1 + 2x_2 = 0,\ x_2 = 0 \Leftrightarrow x_1 = x_2 = 0$.
Sei andererseits A = $\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$. Dann ist
$(x,x)_A = x_1^2 + 4x_1x_2 + 3x_2^2 = (x_1 + 2x_2)^2 - x_2^2 < 0$ für x_1 = -2, $x_2 = 1$.

Hermitesche Matrizen, für die das Axiom (iv) gilt, nennt man *positiv definit*. Es gilt
A = $\begin{pmatrix} a & b \\ b & c \end{pmatrix} \in M_2\ (\mathcal{R})$ ist positiv definit $\Leftrightarrow$ a > 0, $\begin{vmatrix} a & b \\ b & c \end{vmatrix} = ac - b^2 > 0$ (Baby-Sylvester)

3. Beispiel:
X = $C_{\mathcal{K}}$ [a,b]
ist die Menge aller stetigen Abbildungen f: [a,b] $\to \mathcal{K}$. Addition und Multiplikation mit Skalaren punktweise. Skalarprodukt ist

$$(f,g) = \int_a^b f(x)\overline{g(x)}dx.$$

z.B. ist

$$(f,f) = \int_a^b |f(x)|^2 dx \geq 0 \qquad (f,f) = 0 \Leftrightarrow f(x) = 0 \forall x \Leftrightarrow f = 0.$$

Satz 1 *Cauchy-Schwarzsche Ungleichung*
In einem Raum mit Skalarprodukt gilt

$$|(x,y)| \leq ||x||||y|| \forall x, y.$$

Das Gleichheitszeichen gilt genau dann, wenn ein $\alpha \in \mathcal{K}$ mit $x = \alpha y$ oder $y = \alpha x$ existiert.

Hatten dabei folgende Definition vergessen:

Definition 3 *Sei X ein linearer Raum mit Skalarprodukt. Die Norm eines Elements $x \in X$ ist definiert als*

$$||x|| = \sqrt{(x,x)} (\geq 0)$$

und der Abstand zweier Elemente x,y $\in$ X wird erklärt durch

$$d(x,y) = ||x - y||.$$

4. Beispiel:
Für (-,-) wie in Beispiel 1 ist ||x|| = Länge des Vektors x = Abstand des Punktes x vom Ursprung.
($||x|| = \sqrt{(x,x)} = \sqrt{|x||x|\cos\varphi} = \sqrt{x^2 * \cos 0} = |x|$)
d(x,y) ist der übliche Abstand von x und y.

Für (-,-) wie in Beispiel 2 ist

$$||x|| = \sqrt{(x,x)} = \sqrt{|x_1|^2 + \ldots + |x_n|^2}$$

$$d(x,y) = \sqrt{|x_1 - y_1|^2 + \ldots + |x_n - y_n|^2}$$

Kreis ist {x $\in \mathcal{R}^2$: d(x,0) = 1} = {(x_1,x_2) $\in \mathcal{R}^2$: $\sqrt{x_1^2 + x_2^2} = 1$} = {(x_1,x_2) $\in \mathcal{R}^2$: x_1^2 + $x_2^2 = 1$}.

Sei z.B. A = $\begin{pmatrix} \frac{1}{a^2} & 0 \\ 0 & \frac{1}{b^2} \end{pmatrix} \in M_2\ (\mathcal{R})$.
Skalarprodukt im $\mathcal{R}^2$ ist

$$(x,y)_A = \frac{1}{a^2}x_1y_1 + \frac{1}{b^2}x_2y_2$$

$$(x,x)_A = \frac{1}{a^2}x_1^2 + \frac{1}{b^2}x_2^2$$

$$d_A(x,y) = \sqrt{\frac{1}{a^2}(x_1-y_1)^2 + \frac{1}{b^2}(x_2-y_2)^2}$$

Kreis ist $\{x \in \mathcal{R}^2: d_A(x,0) = 1\} = \{(x_1,x_2) \in \mathcal{R}^2: \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1\}$ = Ellipse

Ist (-,-) wie in Beispiel 3, so ist

02.06.06

$$||f|| = \sqrt{\int_a^b |f(x)|^2 dx}$$

$$d(f,g) = \sqrt{\int_a^b |f(x)-g(x)|^2 dx}$$

Fläche zwsischen f und g ist

$$\int_a^b |f(x)-g(x)|dx =: d_1(f,g)$$

Sinnvoller Abstand zwischen f und g ist auch

$$max_{x\in[a,b]}|f(x)-g(x)| =: d_\infty(f,g)$$

Beweis von Satz 1:
Für y = 0 ist die Ungleichung trivial. Sei also y ≠ 0. Setzen $\lambda = \frac{(x,y)}{||y||^2}$. Haben

$$0 \le (x-\lambda y, x-\lambda y) = (x,x) - \overline{\lambda}(x,y) - \lambda(y,x) + \lambda\overline{\lambda}(y,y)$$

$$= (x,x) - \frac{\overline{(x,y)}(x,y)}{||y||^2} - \frac{(x,y)(y,x)}{||y||^2} + \frac{|(x,y)|^2}{||y||^4}(y,y)$$

$$= ||x||^2 - \frac{|(x,y)|^2}{||y||^2} - \frac{|(x,y)|^2}{||y||^2} + \frac{|(x,y)|^2}{||y||^2}$$

$$= ||x||^2 - \frac{|(x,y)|^2}{||y||^2},$$

d.h. $|(x,y)|^2 \le ||x||^2\ ||y||^2$.
Gleichheit gilt genau dann, wenn y = 0 oder x - λ y = 0 ist. #

5. Beispiel:
Cauchy-Schwarzsche Ungleichung konkret

Im 1. Beispiel:

$$||a||||b||\,|\cos\varphi| \le ||a||||b||$$

Im 2. Beispiel:
Ist (x,y) = $x_1\overline{y_1} + \ldots + x_n\overline{y_n}$, so lautet es

$$|x_1\overline{y_1} + \ldots + x_n\overline{y_n} \le \sqrt{|x_1|^2 + \ldots + |x_n|^2}\sqrt{|y_1|^2 + \ldots + |y_n|^2}$$

$$|x_1\overline{y_1} + \ldots + x_n\overline{y_n}|^2 \le (|x_1|^2 + \ldots + |x_n|^2)(|y_1|^2 + \ldots + |y_n|^2)$$

$(\Rightarrow (a_1 + \ldots + a_n)^2 \le (1^2 + \ldots + 1^2)(a_1^2 + \ldots + a_n^2) = n\ (a_1^2 + \ldots + a_n^2 \Leftrightarrow \frac{a_1+\ldots+a_n}{n} \le \sqrt{\frac{a_1^2+\ldots+a_n^2}{n}}))$

3. Beispiel:

$$\left|\int_a^b f(x)\overline{g(x)}dx\right|^2 \leq \left(\int_a^b |f(x)|^2 dx\right)\left(\int_a^b |g(x)|^2 dx\right).$$

Satz 2 *In einem Raum mit Skalarprodukt gilt stets*
(a) $//x// \geq 0 \; \forall \; x \in X$, $//x// = 0 \Leftrightarrow x = 0$,
(b) $//\alpha x// = /\alpha/ \; //x// \; \forall \alpha \in \mathcal{K} \; \forall \; x \in X$,
(c) $//x + y// \leq //x// + //y// \; \forall \; x,y \in X$
(Dreiecksungleichung oder Minkowskische Ungleichung)
(d) $d(x,y) \geq 0 \; \forall \; x,y \in X$, $d(x,y) = 0 \Leftrightarrow x = y$,
(e) $d(x,y) = d(y,x) \; \forall \; x,y \in X$,
(f) $d(x,y) \leq d(x,z) + d(z,y) \; \forall \; x,y,z \in X$
(Dreiecksungleichung)

Bemerkung:
Eigenschaften (a),(b),(c) sind Axiome eines sogenannten *linearen Raumes*. Eigenschaften (d),(e),(f) sind Axiome eines sogenannten *metrischen Raumes*.

Beweis:
(a) klar
(b) $||\alpha x||^2 = \alpha \; \overline{\alpha}(x,x) = |\alpha|^2 ||x||^2$.
(c) $||x + y||^2 = (x + y, x + y) = (x,x) + (x,y) + (y,x) + (y,y) = ||x||^2 + 2 \; \mathrm{Re}(x,y) + ||y||^2 \leq ||x||^2 + 2|(x,y)| + ||y||^2 \leq ||x||^2 + 2 \; ||x|| \; ||y|| = (||x|| + ||y||)^2$
(d),(e) klar
(f) $d(x,y) = ||x - y|| = ||x + z - z - y|| \leq ||x - z|| + ||z - y|| = d(x,z) + d(z,y)$. #

6. Beispiel: Minkowski konkret

$||x + y|| \leq ||x|| + ||y||$,

$$\sqrt{|x_1 + y_1|^2 + \ldots + |x_n + y_n|^2} \leq \sqrt{|x_1|^2 + \ldots + |x_n|^2} + \sqrt{|y_1|^2 + \ldots + |y_n|^2}$$

$$\sqrt{\int_a^b |f(x) + g(x)|^2 dx} \leq \sqrt{\int_a^b |f(x)|^2 dx} + \sqrt{\int_a^b |g(x)|^2 dx}.$$

Definition 4 *Sei X ein linearer Raum über $\mathcal{K}$ und (-,-) ein Skalarprodukt. Für $x,y \in X$ ohne $\{0\}$ und $\mathcal{K} = \mathcal{R}$ ist der Winkel φ zwischen x und y definiert durch*

$$\cos\varphi = \frac{(x,y)}{||x||||y||}, \quad \varphi \in [0,\pi].$$

Für $\mathcal{K} = \mathcal{C}$ wird der Begriff des Winkels nicht definiert.
Sowohl für $\mathcal{K} = \mathcal{R}$ als auch für $\mathcal{K} = \mathcal{C}$ nennt man zwei Elemente $x,y \in X$ orthogonal ($x \perp y$), wenn $(x,y) = 0$ ist.

7. Beispiel:
Zwei Vektoren sind genau dann orthogonal, wenn sie dies im üblichen Sinne sind. Zwei Punkte x,y $\in \mathcal{R}^n$ sind genau dann orthogonal, wenn die Ortsvektoren, die in x und y enden, orthogonal sind.

f,g $\in C_{\mathcal{K}}$ [a,b] sind orthogonal $\Leftrightarrow \int_a^b f(x)\overline{g(x)}dx = 0$.

Satz 3 *Pythagoras*

Sind x und y orthogonal, so gilt

$$||x+y||^2 = ||x||^2 + ||y||^2.$$

Beweis:

$$||x+y||^2 = (x+y, x+y) = ||x||^2+(x,y)+(y,x)+||y||^2 = ||x||^2+\underbrace{(x,y)}_{=0}+\underbrace{\overline{(x,y)}}_{=0}+||y||^2 = ||x||^2+||y||^2.$$

#

3.2 Orthogonalsysteme

Definition 1 *Sei X ein linearer Raum mit Skalarprodukt. Eine Menge $\{x_1, \ldots, x_m\}$ von Elementen aus X ohne {0} heißt*
Orthogonalsystem (OGS), wenn (x_i,x_j) = 0 für $i \neq j$ ist,
Orthonormalsystem (ONS), wenn (x_i,x_j) = δ_{ij} $\forall$ i,j,
Orthogonalbasis (OGB), wenn $\{x_1, \ldots, x_m\}$ gleichzeitig ein OGS und eine Basis in X ist,
Orthonormalbasis (ONB), wenn $\{x_1, \ldots, x_m\}$ gleichzeitig ein ONS und eine Basis in X ist.

Ein einzelnes Element ist OGS genau dann, wenn es von Null verschieden ist.
Die Elemente eines beliebigen OGS sind stets linear unabhängig:

$$\alpha_1 x_1 + \ldots + \alpha_m x_m = 0 \Rightarrow (\alpha_1 x_1 + \ldots \alpha_m x_m, x_j) = 0$$
$$\Rightarrow \alpha_1(x_1, x_j) + \ldots + \alpha_m(x_m, x_j) = 0 \Rightarrow \alpha_j \underbrace{(x_j, x_j)}_{\neq 0} = 0 \Rightarrow \alpha_j = 0.$$

Also: Ein OGS $\{x_1, \ldots, x_m\}$ ist genau dann eine Basis in X, wenn m = dim X gilt.

Ist $\{e_1, \ldots, e_n\}$ eine ONB in X, so gilt
x = (x,e_1)e_1 + ... + (x,e_n) e_n (Fourierreihe)
für jedes x $\in$ X. In der Tat,

$$x = x_1 e_1 + \ldots + x_n e_n$$

$$(x, e_j) = (x_1 e_1 + \ldots + x_n e_n, e_j) = x_1(e_1, e_j) + \ldots x_n(e_n, e_j) = x_j \underbrace{(e_j, e_j)}_{=1} = x_j.$$

Eine ONB im $\mathcal{K}^n$ ist z.B. {(1,0,...,0),...,(0,...,0,1)}.

Gibt es in jedem linearen Raum mit Skalarprodukt eine ONB?
Z.B. bildet {1,x, ..., x^n} eine Basis im Unterraum aller Polynome vom Grad $\leq$ n in $C_{\mathcal{K}}$ [a,b], aber keine ONB. Für endlichdimensionale Räume ist die Antwort ja, wie der folgende Satz zeigt.

Satz 4 *(Gram/Schmidt):*
Sei X ein linearer Raum mit Skalarprodukt und $\{e_1, \ldots, e_n\}$ eine Basis in X. Dann existiert eine ONB $\{f_1, \ldots, f_n\}$ in X mit span $\{e_1, \ldots, e_k\}$ = span $\{f_1, \ldots, f_k\}$ für alle k mit $1 \leq k \leq n$.

Folgerung: In jedem endlichdimensionalen Raum mit Skalarprodukt existiert eine ONB.

07.06.06

Beweis:
Wählen eine beliebige Basis $\{e_1, \ldots, e_n\}$. Dann bildet die durch Satz 1 gegebene ONB $\{f_1, \ldots, f_n\}$ die gewünschte ONB. #

Beweis von Satz 1:
Setzen $f_1 = \frac{e_1}{||e_1||}$.

Nehmen an, dass wir ein ONS $\{f_1, \ldots, f_m\}$ mit span $\{e_1, \ldots, e_k\}$ = span $\{f_1, \ldots, f_k\}$ für $1 \leq k \leq m$ haben.
Setzen

$$g = e_{m+1} - (e_{m+1}, f_1)f_1 + \ldots - (e_{m+1}, f_m)f_m.$$

Dann ist $g \neq 0$, da ansonsten $e_{m+1} \in$ span $\{f_1, \ldots, f_m\}$ = span $\{e_1, \ldots, e_m\}$ wäre.

Setzen $f_{m+1} = \frac{g}{||g||}$.

Dann ist
span $\{f_1, \ldots, f_{m+1}$ = span $\{f_1, \ldots, f_m, e_{m+1}\}$ = span $\{e_1, \ldots, e_m, e_{m+1}\}$
und $||f_{m+1}|| = 1$,

$$(f_{m+1}, f_j) = \frac{1}{||g||}(g, f_j) = \frac{1}{||g||}(e_{m+1}, f_j) - \frac{1}{||g||}\sum_{k=1}^{m}(e_{m+1}, f_k)(f_k, f_j) = \frac{1}{||g||}(e_{m+1}, f_j) - \frac{1}{||g||}(e_{m+1}, f_j) * 1 = 0.$$

$(1 \leq j \leq m)$. Zum Schluss erhält man so ein ONS $\{f_1, \ldots, f_n\}$ mit span $\{f_1, \ldots, f_n\}$ = span $\{e_1, \ldots, e_n\}$ = X, d.h. eine ONB $\{f_1, \ldots, f_n\}$. #

Der Beweis liefert zugleich eine Methode, um eine Basis zu orthogonaliseren („Schmidtsches Orthogonalisierungsverfahren"):

$$f_1 = \frac{e_1}{||e_1||}$$

$$f_2 = \frac{e_2 - (e_2, f_1)f_1}{||e_2 - (e_2, f_1)f_1||}$$

$$f_3 = \frac{e_3 - (e_3, f_1)f_1 - (e_3, f_2)f_2}{||e_3 - (e_3, f_1)f_1 - (e_3, f_2)f_2||}$$

1. Beispiel:

X = $C_{\mathcal{K}}$[-1,1]
e_0(x) = 1, e_1(x) = x, e_2(x) = x^2, ...

Orthogonalisieren die ersten n dieser Funktionen und erhalten so ein ONS in X bzw. eine ONB im Unterraum aller Polynome vom Grad ≤ n - 1.

$$f_0(x) = \frac{e_0(x)}{||e_0(x)||}$$

$$||e_0(x)|| = \left(\int_{-1}^{1} 1 * 1 dx\right)^{\frac{1}{2}} = \sqrt{2}$$

$$\boxed{f_0(x) = \frac{1}{\sqrt{2}}}$$

$$f_1(x) = \frac{e_1(x) - (e_1(x), f_1(x))f_0(x)}{||e_1(x) - (e_1(x), f_1(x))f_0(x)||}$$

$$(e_1(x), f_0(x)) = \int_{-1}^{1} x \frac{1}{\sqrt{2}} dx = 0$$

$$||e_1(x)|| = \left(\int_{-1}^{1} x * x dx\right)^{\frac{1}{2}} = \left(\frac{2}{3}^{\frac{1}{2}}\right)$$

$$\boxed{f_1(x) = \sqrt{\frac{3}{2}} x}$$

$$f_2(x) = \frac{e_2(x) - (e_2(x), f_1(x))f_1(x) - (e_2(x), f_0(x))f_0(x)}{||e_2(x) - (e_2(x), f_1(x))f_1(x) - (e_2(x), f_0(x))f_0(x)||}$$

$$(e_2(x), f_1(x)) = \int_{-1}^{1} x^2 \sqrt{\frac{3}{2}} x dx = 0$$

$$(e_2(x), f_1(x)) = \int_{-1}^{1} x^2 \frac{1}{\sqrt{2}} dx = \frac{1}{\sqrt{2}} \frac{2}{3} = \frac{\sqrt{2}}{3}$$

$$||\dots||^2 = \int_{-1}^{1} \left(x^2 - \frac{\sqrt{2}}{3} \frac{1}{\sqrt{2}}\right)^2 dx = \int_{-1}^{1} \left(x^4 - \frac{2}{3}x^2 + \frac{1}{9}\right) dx = \frac{8}{45}$$

$$\boxed{f_2 = \sqrt{\frac{45}{8}} \left(x^2 - \frac{1}{3}\right)}$$

Man kann zeigen, dass

$$f_n = \sqrt{\frac{2n+1}{2}} \frac{1}{2^n n!} \frac{d^n}{dx^n}(x^2-1)^n$$

gilt.
(n = 2: $f_2(x) = \sqrt{\frac{5}{2}} * \frac{1}{8}[(x^2-1)^2]$
$f_2(x) = \sqrt{\frac{5}{2}}\frac{1}{8}12(x^2-\frac{1}{3}) = \sqrt{\frac{5}{2}}\frac{3}{2}(x^2-\frac{1}{3}) = \sqrt{\frac{45}{8}}(x^2-\frac{1}{3}))$
Die orthogonalen Polynome auf [-1,1] mit (f,g) = $\int_{-1}^{1} f(x)\overline{g(x)}dx$ heißen *Legendresche Polynome.*

Definition 2 *Sei X ein linearer Raum mit Skalarprodukt und U ein Unterraum von X. Die Menge*

$$U^\perp := \{x \in X : (x,u) = 0 \forall u \in U\}$$

heißt orthogonales Komplement von U.

Satz 1 *Sei X ein linearer Raum mit Skalarprodukt und U ein endlichdimensionaler Unterraum von X. Dann ist $U^\perp$ ein Unterraum von X und es gilt $X = U \boxplus U^\perp$.*

Beweis:
Offenbar ist $U^\perp$ ein Unterraum (auch wenn U unendlichdimensional ist). Wählen ONB $\{e_1, \ldots, e_m\}$ in U. Für x ∈ X setzen wir

$$x_1 = (x,e_1)e_1 + \ldots + (x,e_m)e_m,$$

$$x_2 = x - x_1.$$

Dann ist x = x_1 + x_2 mit x_1 ∈ U. Wegen

$$(x_2,e_j) = (x,e_j) - (x_1,e_j) = (x,e_j) - \sum_{k=1}^{m}(x,e_k)(e_k,e_j) = (x,e_j) - (x,e_j) * 1 = 0$$

ist (x_2,u) = 0 ∀ u ∈ U, d.h. $x_2 \in U^\perp$.
Also ist $U + U^\perp$.
Ist y ∈ $U \cap U^\perp$, so ist (y,y) = 0, d.h. y = 0. Also ist $X = U \boxplus U^\perp$. #

Definition 3 *Sei X ein endlichdimensionaler linearer Raum mit Skalarprodukt und U ein Unterraum von X. Nach Satz 2 lässt sich jedes $x \in X$ eindeutig in der Form $x = x_1 + x_2$ mit $x_1 \in U$ und $x_2 \in U^\perp$ schreiben. Die Abbildung, die x in x_1 überführt, wird mit P_U bezeichnet und orthogonaler Projektor von X auf U genannt.*

Es ist leicht zu sehen, dass gilt: $P_U^2 = P_U$, ker $P_U = U^\perp$, Im P_U = U.
Aus Beweis von Satz 2 folgt:

Ist $\{e_1, \ldots, e_m\}$ eine ONB in U, so ist

$$P_U x = (x, e_1)e_1 + \ldots + (x, e_m)e_m.$$

Betrachten noch folgendes Problem: 09.06.06
Gegeben sei ein endlichdimensionaler linearer Raum X mit Skalarprodukt und ein Unterraum U $\subset$ X. Für ein x $\in$ X suchen wir ein u $\in$ U, sodass ||x - u|| minimal wird, d.h. wir suchen die bestmögliche Approximation von x durch Elemente aus U.

Satz 2 *Die bestmögliche Approximation von x in U existiert und ist eindeutig bestimmt. Sie ist gegeben durch P_U x. Mit anderen Worten, es gilt*

$$||x - P_U x|| < ||x - u|| \quad \forall u \in U\{P_U x\}.$$

Bemerkung:
Existenz und Eindeutigkeit sind nicht trivial:
bei offenen Mengen existiert z.B. keine optimale Approximation. Wählt man die Manhattan-Metrik lässt sich die optimale Approximation nicht eindeutig wählen.

Beweis:
Es gilt u - P_Ux $\in$ U für jedes u $\in$ U und x - P_U x = (I - P_U)x $\in$ ker(I - P_U) = $U^\perp$.

Also ist

$$||x - u||^2 = ||\underbrace{x - P_u x}_{\in U^\perp} + \underbrace{P_U x - u}_{\in U}||^2 = ||x - P_U x||^2 + ||P_U x - u||^2 \geq ||x - P_U x||^2$$

mit Gleichheit genau dann, wenn u = P_U x ist. #

2. Beispiel:
Betrachten $C_{\mathcal{R}}$ [0, 2π] mit (f,g) = $\int_0^{2\pi} f(x)\overline{g(x)}dx$.
Setzen $c_0(x) = \frac{1}{\sqrt{2\pi}}$, $c_n(x) = \frac{1}{\sqrt{pi}} \cos nx$ (n $\geq$ 1), $s_n(x) = \frac{1}{\sqrt{\pi}} \sin nx$ (n $\geq$ 1).
Sei U = span $\{c_0, c_1, \ldots, c_n, s_1, \ldots, s_n\}$.
Für f $\in$ $C_{\mathcal{R}}$ [0,2π] suchen wir die beste Approximation in U, d.h. wir suchen $\tilde{a_0}, \tilde{a_1}, \ldots, \tilde{a_n}, \tilde{b_1}, \ldots, \tilde{b_n}$ $\in \mathcal{R}$ mit

$$||f - \tilde{a_0}c_0 - \tilde{a_1}c_1 - \ldots - \tilde{a_n}c_n - \tilde{b_1}c_1 - \ldots - \tilde{b_n}c_n||^2$$

$$= \int_0^{2\pi} |f(x) - \underbrace{a_0 - \sum_{k=1}^{n}(a_k \cos kx + b_k \sin kx)}_{trigonometrisches\ Polynom\ vom\ Grad\ n}|^2 dx \to min.$$

Es ist leicht zu sehen, dass $\{c_0, c_1, \ldots, c_n, s_1, \ldots, s_n\}$ ein ONS in $C_{\mathcal{R}}$ [0,2π] und damit eine ONB in U sind.

$$\int_0^{2\pi} \cos kx \cos jk dx = 0 \quad (k \neq j), \ldots$$

Obiges Minimum wird also genau dann angenommen, wenn gilt:

$$\tilde{a_0} = (f, c_0), \tilde{a_k} = (f, c_k), \tilde{b_k} = (f, s_k)$$

$$a_0 = \frac{1}{\sqrt{2\pi}} * \int_0^{\pi} f(x)\frac{1}{\sqrt{2\pi}}dx = \frac{1}{\sqrt{2\pi}}\int_0^{2\pi} f(x)dx$$

$$a_k = \frac{1}{\sqrt{\pi}}\int_0^{2\pi} f(x)\frac{1}{\sqrt{\pi}}\cos kxdx = \frac{1}{\sqrt{\pi}}\int_0^{2\pi} f(x)\cos kxdx$$

$$b_k = \frac{1}{\sqrt{\pi}}\int_0^{2\pi} f(x)\frac{1}{\sqrt{\pi}}\sin kxdx = \frac{1}{\sqrt{\pi}}\int_0^{2\pi} f(x)\sin kxdx$$

Die Zahlen a_k, b_k heißen *Fourierkoeffizienten* von f und die Reihe

$$a_0 + \sum_{k=1}^{\infty}(a_k \cos kx + b_k \sin kx)$$

heißt *Fourierreihe* von f.

3.3 Orthogonale und unitäre Operatoren und Matrizen

Definition 1 *Seien X, Y lineare Räume mit Skalarprodukt (-,-)$_X$, (-,-)$_Y$. Ein linearer Operator $A \in \mathcal{L}(X,Y)$ heißt orthogonal, wenn $\mathcal{K} = \mathcal{R}$ und (Ax,Ay)$_Y$ = (x,y)$_X$ $\forall$ x,y $\in$ X gilt und unitär, wenn $\mathcal{K} = \mathcal{C}$ und (Ax,Ay)$_Y$ = (x,y)$_X$ $\forall$ x,y $\in$ X ist.*
Ein Operator $A \in \mathcal{L}(X,Y)$ heißt Isometrie, wenn $//Ax//_Y = //x//_X$ $\forall$ x $\in$ X gilt.

Offenbar sind orthogonale und unitäre Operatoren auch Isometrien.

$$||Ax||^2 = (Ax, Ax) = (x, x) = ||x||^2.$$

Erstaunlicherweise gilt auch die Umkehrung, d.h. Isometrien sind automatisch orthogonal ($\mathcal{K} = \mathcal{R}$) bzw. unitär ($\mathcal{K} = \mathcal{C}$). Dies folgt aus

$$(x, y) = \frac{1}{2}(||x + y||^2 - ||x||^2 - ||y||^2) \quad (\mathcal{K} = \mathcal{R}),$$

$$(x, y) = \frac{1}{4}(||x + y||^2 - ||x - y||^2 + i||x + iy||^2 - i||x - iy||^2) \quad (\mathcal{K} = \mathcal{C})$$

(*Polarisierungsidentitäten*)
Isometrien sind immer injektiv:
Ax = Ay $\Rightarrow$ A(x - y) = 0 $\Rightarrow$ ||A(x - y)|| = 0 $\Rightarrow$ ||x - y|| = 0 $\Rightarrow$ x = y.
Isometrien in $\mathcal{L}$(X) = $\mathcal{L}$(X,X) sind stets invertierbar (Fredholmsche Alternative).

Definition 2 *Zwei lineare Räume mit Skalarprodukt heißen isometrisch isomorph, wenn es eine bijektive Isometrie zwischen ihnen gibt.*

Satz 1 *Zwei endlichdimensionale lineare Räume mit Skalarprodukte sind genau dann isometrisch isomorph, wenn sie die gleiche Dimension haben.*

Beweis:
X isometrisch isomorph zu Y $\Rightarrow$ X isomorph zu Y $\Rightarrow$ dim X = dim Y.

Sei umgekehrt dim X = dim Y = n. Wählen ONB $\{e_1, \ldots, e_n\}$ in X und ONB $\{f_1, \ldots, f_n\}$ in Y. Definieren eine Abbildung A: X $\rightarrow$ Y durch

$$A(x_1e_1 + \ldots + x_ne_n) = x_1f_1 + \ldots + x_nf_n.$$

Offenbar ist A linear und bijektiv. Haben

$||A(x_1e_1+\ldots+x_ne_n)||^2 = ||x_1f_1+\ldots+x_nf_n||^2 = ||x_1f_1||^2+\ldots+||x_nf_n||^2 = |x_1|^2||f_1||^2+\ldots+|x_n|^2||f_n||^2$

$= |x_1|^2+\ldots+|x_n|^2 = |x_1|||e_1||^2+\ldots+|x_n|^2||e_n||^2 = ||x_1e_1||^2+\ldots+||x_ne_n||^2 = ||x_1e_1+\ldots+x_ne_n||^2,$

d.h. A ist Isometrie. #

Definition 3 *Eine Matrix $A \in M_n$ ($\mathcal{R}$) heißt orthogonal, wenn die Spalten von A ein ONS im $\mathcal{R}^n$ mit dem üblichen Skalarprodukt bilden, d.h. wenn*

$$a_{1j}a_{1k} + \ldots + a_{nj}a_{nk} = \delta_{jk} \quad \forall j,k$$

gilt. Eine Matrix $A \in M_n$ ($\mathcal{C}$) heißt unitär, wenn die Spalten von A ein ONS im $\mathcal{C}^n$ mit dem üblichen Skalarprodukt bilden, d.h. wenn

$$a_{1j}\overline{a_{1k}} + \ldots + a_{nj}\overline{a_{nk}} = \delta_{jk} \quad \forall j,k$$

gilt.
Für $A = (a_{jk})$ setzt man
$A^T = (a_{kj})$ (transponierte Matrix)
$A^* = (\overline{a_{kj}})$ (adjungierte Matrix).

Satz 2 *Folgende Bedingungen sind äquivalent für $A \in M_n$ ($\mathcal{K}$):*

$\mathcal{K} = \mathcal{R}$	$\mathcal{K} = \mathcal{C}$
(i) A ist orthogonal	*(i) A ist unitär*
(ii) Spalten von A bilden ONS	*(ii) Spalten von A bilden ONS*
(iii) Zeilen von A bilden ONS	*(iii) Zeilen von A bilden ONS*
(iv) $A^T A = I$	*(iv) $A^*A = I$*
(v) $A A^T = I$	*(v) $AA^* = I$*
(vi) $A^T A = AA^T = I$	*(vi) $A^*A = AA^* = I$*
(vii) A invertierbar und $A^{-1} = A^T$	*(vii) A invertierbar und $A^{-1} = A^*$.*

Beweis:
Offenbar sind (ii) und (iv) äquivalent: $(s_j,s_k) = s_j^T\ s_k$.

Und der Rest ist dann klar. #

Folgerung: Die Determinante einer orthogonalen Matrix aus M_n ($\mathcal{R}$) ist stets +1 oder -1, und die Determinante einer unitären Matrix aus M_n ($\mathcal{C}$) ist immer eine Zahl aus $\mathcal{T} := \{z \in \mathcal{C}: |z| = 1\}$. 14.06.06

Beweis: A*A = I ⇒ det A* det A = 1 und det A* = $\overline{det\ A}$. #

Setzen O(n) = {A ∈ M_n ($\mathcal{R}$): A orthogonal},
U(n) = {A ∈ M_n ($\mathcal{C}$): A unitär}.

Dies sind Gruppen bezüglich der Matrizenmultiplikation (Untergruppen von GL(n,$\mathcal{R}$) bzw. GL(n,$\mathcal{C}$)).
((AB)*AB = B*A*AB = B*IB = I)
Haben
O(1) = {-1,1} und U(1) = $\mathcal{T}$.
Bestimmen O(2):
Haben A = $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ∈ O(2) genau dann, wenn gilt
$a^2 + c^2 = 1 \Leftrightarrow a = \cos\alpha, c = \sin\alpha (\alpha \in [0, 2\pi])$
$b^2 + d^2 = 1 \Leftrightarrow b = \cos\beta, d = \sin\beta (\beta \in [0, 2\pi])$
$ab + cd = 0 \Leftrightarrow \cos(\alpha - \beta) = 0 \Leftrightarrow \alpha - \beta \in \{\frac{\pi}{2}, \frac{-\pi}{2}, \frac{3\pi}{2}, \frac{-3\pi}{2}\}$

$\alpha = \beta + \frac{\pi}{2} \quad \cos\alpha = -\sin\beta$
$\alpha = \beta - \frac{\pi}{2} \quad \cos\alpha = \sin\beta$
$\alpha = \beta + \frac{3\pi}{2} \quad \cos\alpha = \sin\beta$
$\alpha = \beta - \frac{3\pi}{2} \quad \cos\alpha = -\sin\beta$

$\alpha - \beta \in \{\frac{\pi}{2}, \frac{-3\pi}{2}\}$:

$$A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$$

ist Achsenspiegelung

$\alpha - \beta \in \{\frac{-\pi}{2}, \frac{3\pi}{2}\}$:

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

ist Drehung um Winkel α.

Also O(2) = {$\begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$ ($\alpha \in [0,2\pi)$), $\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$ ($\alpha \in [0,2\pi)$)}
= Menge der Achsenspiegelungen ∪ Menge der Drehungen.

Man setzt
SO(n) = {A ∈ O(n): det A = 1}
SU(n) = {A ∈ U(n): det A = 1}
Dies sind Untergruppen von O(n) bzw. U(n).
SO(1) = {1}, SU(1) = {1},
SO(2) = {($\begin{matrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{matrix}$) ($\alpha \in [0,2\pi)$)} = Menge der Drehungen.

Satz 3 *Sei X ein endlichdimensionaler linearer Raum mit Skalarprodukt und $A \in \mathcal{L}(X)$. Dann sind folgende Bedingungen äquivalent:*
(i) A ist Isometrie;
(ii) A überführt eine ONB in eine ONB;
(iii) A überführt jede ONB in eine ONB;
(iv) es existiert eine ONB E, sodass $[A]_{E,E}$ orthogonal bzw. unitär ist;
(v) für jede ONB E ist $[A]_{E,E}$ orthogonal bzw. unitär.

Beweis:
(i) ⇒ (iii):
Sei $\{e_1, \ldots, e_n\}$ eine ONB. Dann ist (Ae$_i$,Ae$_j$) = (e_i,e_j) = δ_{ij}. Also ist {Ae$_1$,...,Ae$_n$} eine ONB.
(iii) ⇒ (v):
Sei $\{e_1, \ldots, e_n\}$ =: E eine ONB. Die i-te Spalte von $[A]_{E,E}$ sind die Koeffizienten in

$$Ae_i = x_1 e_1 + \ldots + x_n e_n,$$

die in der j-ten Spalte die Koeffizienten in

$$Ae_j = y_1 e_1 + \ldots + y_n e_n.$$

Haben $x_1\overline{y_1} + \ldots + x_n\overline{y_n} = (x_1 e_1 + \ldots + x_n e_n, y_1 e_1 + \ldots + y_n e_n) = (Ae_i, Ae_j) = \delta_{ij}$, da $\{Ae_1, \ldots, Ae_n\}$ eine ONB ist.
(v) ⇒ (iv) trivial
(iv) ⇒ (ii) wie (iii) ⇒ (v)
(ii) ⇒ (i):
Sei $\{e_1, \ldots, e_n\}$ eine ONB, für die $\{Ae_1, \ldots, Ae_n\}$ ebenfalls eine ONB ist. Für x = $x_1 e_1 + \ldots + x_n e_n$ ist dann

$$||Ax||^2 = ||x_1 Ae_1 + \ldots + x_n Ae_n||^2 = |x_1|^2 + \ldots + |x_n|^2 = ||x_1 e_1 + \ldots + x_n e_n||^2 = ||x||^2.$$

#

Satz 4 *Die EW einer orthogonalen Matrix liegen in {-1,1}, die EW einer unitären Matrix liegen auf $\mathcal{T}$. In beiden Fällen sind zu verschiedenen EW gehörende EV orthogonal.*

Beweis:
Ist λ ein EW in A und u $\neq$ 0 ein EV, so gilt

$$||u|| = ||Au|| = ||\lambda u|| = |\lambda|||u||,$$

d.h. $|\lambda| = 1$.
Ist $\lambda \neq \mu$, Au = λu, Av = μv mit u $\neq$ 0, v $\neq$ 0, so folgt

$$(u, v) = (Au, Av) = (\lambda u, \mu v) = \lambda\overline{\mu}(u, v) \Rightarrow \lambda(u, v) = \lambda\mu\overline{\mu}(u, v) = \lambda(u, v) \Rightarrow (\mu - \lambda)(u, v) = 0 \Rightarrow (u, v) = 0.$$

#

Satz 5 *(a) Sei X ein endlichdimensionaler euklidischer Raum. Eine Abbildung $A \in \mathcal{L}(X)$ ist genau dann orthogonal, wenn eine ONB E existiert, sodass gilt*

$$[A]_{E,E} = \begin{pmatrix} D(\alpha_1) & & & & & & & & \\ & \ddots & & & & & & & \\ & & D(\alpha_m) & & & & & & \\ & & & 1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & 1 & & & \\ & & & & & & -1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & -1 \end{pmatrix},$$

wobei $D(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$ *ist.* ($\alpha \neq 0, \pi$).
(b) Sei X ein endlichdimensionaler unitärer Raum. Ein Operator $A \in \mathcal{L}(X)$ ist genau dann unitär, wenn eine ONB E existiert, sodass

$$[A]_{E,E} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

mit $\lambda_j \in \mathcal{T}$ $\forall$ j gilt.

Beweis:
(a) Ist $[A]_{E,E}$ von der angegebenen Gestalt, so ist $[A]_{E,E}$ orthogonal und damit A $\in$ $\mathcal{L}$(X) orthogonal nach (iv) $\Rightarrow$ (i) aus Satz 3.
Umgekehrt sei A $\in \mathcal{L}$(X) orthogonal. Für dim X = 1 ist Behauptung trivial. Für dim X = 2 ist A eine Drehung oder Achsenspiegelung (nachdem X mit $\mathcal{R}^2$ identifiziert wurde), d.h. es existiert eine ONB E mit $[A]_{E,E}$ = D(α) oder $[A]_{E,E} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.
Sei nun dim X $\geq$ 3. Nehmen an, dass A einen EW λ hat. Dann ist $\lambda = \pm$ 1. Sei u $\neq$ 0 ein EV und U = span {u}. Dann ist U ein invarianter Unterraum für A. Behaupten, dass $U^\perp$ ebenfalls ein invarianter Unterraum für A ist.
Sei also y $\in U^\perp$. Für v $\in$ U ist dann

$$(v, Ay) = (A(\lambda^{-1}v), Ay) = (\lambda^{-1}v, y) = \lambda^{-1}(v, y) = 0,$$

da (v,y) = 0 $\forall$ u $\in$ U. Damit ist U, $U^\perp$ ein reduzierendes Paar invarianter Unterräume mit Au = ($\pm$ 1)u $\forall$ u $\in$ U. Wegen dim $U^\perp$ = dim X - 1 folgt Behauptung durch Induktion.

Nehmen an, dass A keine EW hat. Reelle JNF von A zeigt, dass A einen zweidimensionalen invarianten Unterraum U hat. Die Einschränkung A|U ist orthogonal und somit existiert eine ONB F in U mit $[A|U]_{E,E}$ = D(α) mit $\alpha \neq 0$, $\alpha \neq \pi$.
Zeigen, dass $U^\perp$ ebenfalls invarianter Unterraum von A ist. Sei y $\in U^\perp$ Dann ist für v $\in$ U

$$(v, Ay) = (A(A^{-1}v), Ay) = (A^{-1}v, y) = 0,$$

da A auf U invertierbar ist und somit $A^{-1}v \in U$ gilt.

Also ist $Ay \in U^{\perp}$. Wieder bilden U und $U^{\perp}$ ein reduzierendes Paar invarianter Unterräume und Behauptung ergibt sich so durch Induktion.

(b) Beweis wie (a), sogar einfacher, da A stets EW hat. #

Folgerung: (Satz von Euler)
Jede orthogonale Abbildung A im $\mathcal{R}^3$ mit det A = 1 (d.h. jede lineare Abbildung im $\mathcal{R}^3$ mit einer Matrixdarstellung in einer ONB, die eine Matrix in SO(3) ist) ist eine Achsendrehung.

16.06.06

Kommentar:
Lineare Abbildungen mit det A > 0 sind orientierungserhaltend, d.h. überführen rechte Schuhe in rechte Schuhe; solche mit det A < 0 ändern die Orientierung (überführen rechte Schuhe in linke Schuhe).
Achsendrehung: $E = \{e_1, e_2, e_3\}$ ONB

$$[A]_{E,E} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$$

Jede Achsendrehung ist also aus SO(3).
Insbesondere folgt aus dem Satz von Euler, dass das Produkt von beliebig vielen Achsendrehungen (um verschiedene Achsen) wieder eine Achsendrehung ist.

Beweis:
Nach Satz 5 existiert eine ONB E mit

$$[A]_{E,E} = \begin{pmatrix} D(\alpha) & 0 \\ 0 & \pm 1 \end{pmatrix} \text{ oder } \begin{pmatrix} \pm 1 & & \\ & \pm 1 & \\ & & \pm 1 \end{pmatrix}.$$

Dies gilt für jede Abbildung aus O(3). Ist aber die Abbildung aus SO(3), so muss det $[A]_{E,E} = 1$ sein, d.h. erhalten

$$[A]_{E,E} = \underbrace{\begin{pmatrix} D(\alpha) & 0 \\ 0 & 1 \end{pmatrix}}_{Achsendrehung\ um\ Winkel\ \alpha} \text{ oder } \underbrace{\begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}}_{Drehung\ um\ Winkel\ 0} \text{ oder } \underbrace{\begin{pmatrix} -1 & & \\ & -1 & \\ & & 1 \end{pmatrix}}_{Drehung\ um\ 180^\circ}. \#$$

Folgerung: ebenfalls von Euler
Sei A eine orthogonale Abbildung im $\mathcal{R}^3$ mit det A = 1 (d.h. aus SO(3)) und seien u und v zwei orthogonale Vektoren im $\mathcal{R}^3$. Dann gibt es drei Winkel α,β,γ (die sogenannten Eulerschen Winkel), sodass A das Produkt der folgenden drei Achsendrehungen ist:
Winkel α um u, Winkel β um v, Winkel γ um u.

Beweis:
Nehmen u und v als Einheitsvektoren und ergänzen diese zu einer ONB {u,v,w}.
Achsendrehung B um v bewirkt, dass BAu ∈ span {u,w}.
Achsendrehung C um v bewirkt CBAu = u.
Nach Satz von Euler ist CBA eine Achsendrehung D, und da CBA den Vektor u invariant lässt, muss D Drehung um u sein. Aus CBA = D folgt

$$A = B^{-1}C^{-1}D.$$

#

3.4 Dualer Raum und adjungierter Operator

Definition 1 *Sei X ein linearer Raum. Die Abbildung aus $\mathcal{L}(X,\mathcal{K})$, d.h. die linearen Abbildungen f: $X \to \mathcal{K}$ heißen lineare Funktionale oder Linearformen. Die Menge aller linearen Funktionale nennt man dualen Raum und bezeichnet diesen mit X*, d.h. $X^* = \mathcal{L}(X,\mathcal{K})$.*

Menge X* ist wieder ein linearer Raum. Haben $f \in (\mathcal{K}^n)^*$ genau dann, wenn f von der Form $f(x_1,\ldots,x_n) = f_1x_1 + \ldots + f_nx_n$ ($f_j \in \mathcal{K}$) ist. Die Abbildung
$(\mathcal{K}^n)^* \to \mathcal{K}^n$, $f \mapsto (f_1,\ldots,f_n)$
ist offenbar ein Isomorphismus. Also $(\mathcal{K}^n)^* \cong \mathcal{K}^n$.

Satz 1 *Ist X ein endlichdimensionaler linearer Raum, so ist auch X* endlichdimensional, und es gilt dim X* = dim X.*

Beweis:
Wählen Basis $\{e_1,\ldots,e_n\}$ in X und definieren $f_1,\ldots,f_n \in$ X* durch $f_j(x_1e_1 + \ldots + x_ne_n) = x_j$.
Behaupten, dass $\{f_1,\ldots,f_n\}$ eine Basis in X* ist.
Sei $\sum \alpha_j f_j = 0$. Dann ist

$$0 = (\sum \alpha_f f_j)(e_i) = \alpha_1 f_j(e_1) + \ldots + \alpha_n f_j(e_n) = \alpha_j \quad \forall j.$$

Sei g $\in$ X*. Dann ist g = $\sum$ g(e_j) f_j:

$$g(x) = g(x_1e_1 + \ldots + x_ne_n) = x_1g(e_1) + \ldots + x_ng(e_n)$$

$$= g(e_1)x_1 + \ldots + g(e_n)x_n = g(e_1)f_1(x) + \ldots + g(e_n)f_n(x)$$

$$= (g(e_1)f_1 + \ldots + g(e_n)f_n)(x).$$

#

Definition 2 *Sei X ein endlichdimensionaler linearer Raum und $E = \{e_1,\ldots,e_n\}$ eine Basis in X.*
Dann gibt es genau eine Basis $\{f_1,\ldots,f_n\}$ in X mit $f_i(e_j) = \delta_{ij} \quad \forall i,j$.*
Diese Basis wird mit E bezeichnet und zu E duale Basis genannt.*

Existenz: $f_i(x_1e_1 + \ldots + x_ne_n) = x_j$ (Satz 1)
Eindeutigkeit: $f_i(e_j) = g_i(e_j)$
$\Rightarrow f_i(x_1e_1 + \ldots + x_ne_n) = g_i(x_1e_1 + \ldots + x_ne_n)$

$\Rightarrow f_i = g_i$.

Definition 3 *Seien X und Y lineare Räume und A ∈ $\mathcal{L}$(X,Y). Der adjungierte Operator A* ist der Operator aus $\mathcal{L}$(Y*,X*), der durch A*f = f ∘ A, d.h. durch (A*f)(x) = f(Ax) ∀ x ∈ X. gegeben ist.*

Satz 2 *Seien X und Y endlichdimensionale lineare Räume und A ∈ $\mathcal{L}$(X,Y). Seien E = $\{e_1, \ldots, e_n\}$ und F = $\{f_1, \ldots, f_m\}$ Basen in X und Y und seien E* = $\{e_1^*, \ldots, e_n^*\}$ und F* = $\{f_1^*, \ldots, f_m^*\}$ die dualen Basen in X* und Y*.*
Dann gilt $[A^*]_{F*,E*} = [A]_{E,F}^T$.

Beweis:
Die Zahl $([A^*]_{F*,E*})_{ij}$ ist das α_i aus der Zerlegung

$$A * f_j^* = \alpha_1 e_1^* + \ldots + \alpha_n e_n^*$$

und $([A]_{E,F})_{ji}$ ist das β_j aus

$$Ae_i = \beta_1 f_1 + \ldots + \beta_m f_m.$$

Haben

$$\alpha_i = (\alpha_1 e_1^* + \ldots + \alpha_n e_n^*)(e_i) = (A^* f_j^*)(e_i) = f_j^*(Ae_i) = f_j^*(\beta_1 f_1 + \ldots + \beta_m f_m) = \beta_j.$$

#

Satz 3 *Baby Riesz*
Sei X ein endlichdimensionaler linearer Raum mit Skalarprodukt (-,-). Für jedes f ∈ X ist dann die durch f(x) = (x,f) definierte Abbildung f*: X → $\mathcal{K}$ ein Element von X*. Die Abbildung R: X → X*, f ↦ f* ist eine sesquilineare Bijektion, d.h. eine bijektive Abbildung mit*

$$R(f + g) = Rf + Rg,$$
$$R(\alpha f) = \overline{\alpha} Rf.$$

Beweis:
Es ist klar, dass f* linear ist. Haben
R(f + g) = (f + g)*, Rf = f*, Rg = g*,
(f + g)*(x) = (x,f + g) = (x,f) + (x,g) = f*(x) + g*(x).

R(αf) = (αf)*, Rf = f*,
(αf)*(x) = (x,αf) = $\overline{\alpha}$(x,f) = $\overline{\alpha}$f*(x).

Injektivität:
Rf = Rg ⇒ f* = g* ⇒ (x,f) = (x,g) ∀ x ∈ X ⇒ (x,f - g) = 0 ∀ x ∈ X ⇒ (f - g,f - g) = 0 ⇒ f - g = 0 ⇒ f = g.

Surjektivität:
Sei g ∈ X*. Wählen ONB $\{e_1, \ldots, e_n\}$ in X und setzen

$$f = \overline{g(e_1)}e_1 + \ldots + \overline{g(e_n)}e_n.$$

Dann ist Rf = g:
(Rf)(x) = f*(x) = (x,f) = (x,$\overline{g(e_1)}e_1 + \ldots + \overline{g(e_n)}e_n$) = g($e_1$)(x,$e_1$)+…+g($e_n$)(x,$e_n$);
x = e_i ergibt
(Rf)(e_i) = g(e_i) und damit Rf = g. #

Definition 4 *Sei X ein endlichdimensionaler Raum mit Skalarprodukt. Man identifiziert dann X* mit X über die Rieszsche Abbildung:*

$$R\text{: } X \to X^*,\ f \mapsto f^*$$
$$((Rf)(x) = f^*(x) = (x,f)).$$

Ist $E = \{e_1, \ldots, e_n\}$ eine Basis in X, so existiert genau eine Basis $E^* = \{f_1, \ldots, f_n\}$ in X mit $(e_i,f_j) = \delta_{ij}\ \forall$ i,j, die sogenante duale oder biorthogonale Basis E.*

21.06.06

Ist $A \in \mathcal{L}(X)$, so ist der Rieszsche adjungierte Operator definiert als $R^{-1}A^*R$ ($\in L(X)$), wobei $A^* \in \mathcal{L}(X^*)$ der adjungierte Operator ist.
Man bezeichnet den Rieszschen adjungierten Operator ebenfalls mit A und nennt ihn einfach den adjungierten Operator, wenn keine Verwechslungen möglich sind.*

Satz 4 *Sei X ein endlichdimensionaler linearer Raum mit Skalarprodukt (-,-) und $A \in \mathcal{L}(X)$. Dann gibt es genau einen Operator $B \in \mathcal{L}(X)$ mit*

$$(x, Ay) = (Bx, y) \qquad \forall x, y \in X.$$

Dieser Operator ist der Rieszsche adjungierte Operator von A, d.h. es gilt

$$(x, Ay) = (A * x, y) \qquad \forall x, y \in X.$$

Desweiteren gilt für den Rieszschen adjungierten Operator
$A^{**} = A$, $(A + B)^* = A^* + B^*$, $(\alpha A)^* = \overline{\alpha}A^*$, $(AB)^* = B^* A^*$.

Beweis:
Sei A* der Rieszsche Adjungierte. Dann ist

$$(\underbrace{A^*}_{Riesz} x, y) = (R^{-1} \underbrace{A^*}_{Adj.} Rx, y) = \overline{(y, R^{-1}ARx)}$$

$$\overset{(Rf)(y)=(y,f)}{=} \overline{(A * Rx)(y)} \overset{(A^*f)(y)=f(Ay)}{=} \overline{(Rx)(Ay)} \overset{(Rf)(z)(zf)}{=} \overline{(Ay, x)} = (x, Ay).$$

Seien B,C $\in \mathcal{L}$(X) zwei Operatoren mit (Bx,y) = (Cx,y) $\forall$ x,y $\in$ X.
Setzen D = B - C und haben dann (Dx,y) = 0 $\forall$ x,y $\in$ X.
$\Rightarrow$ (Dx,Dx) = 0 $\forall$ x $\in$ X
$\Rightarrow$ Dx = 0 $\forall$ x $\in$ X
$\Rightarrow$ D = 0 $\Rightarrow$ B = C.
Damit ist die Eindeutigkeit gezeigt.

Schließlich ist

$$(A^{**}x, y) = ((A^*)^*x, y) = (x, A^*y) = \overline{(A^*y, x)} = \overline{(y, Ax)} = (Ax, y) \qquad \forall x, y \in X \Rightarrow A^{**} = A.$$

$$((AB)^*x, y) = (x, ABy) = (A^*x, By) = (B^*A^*x, y) \qquad \forall x, y \in X \Rightarrow (AB)^* = B^*A^*.$$

#

Merken uns also:
Man kann A beliebig umschaufeln, indem man einen * ranmacht.
(Ax,y) = (x,A*y), (x,Ay) = (A*x,y).

Satz 5 *Sei X ein endlichdimensionaler linearer Raum mit Skalarprodukt und A $\in \mathcal{L}(X)$. Ist E eine Basis in X und E* die biorthogonale Basis in X, so gilt*

$$[A^*]_{E*,E*} = [A]^*_{E,E}.$$

Ist insbesondere E eine ONB ($\Leftrightarrow$ E = E), so ist*

$$[A^*]_{E,E} = [A]^*_{E,E}.$$

Beweis:
Sei $([A^*]_{E*,E*})_{i,j} = \alpha$, $([A]_{E,E})_{j,i} = \beta$.
Müssen zeigen: $\alpha = \overline{\beta}$
Sei E = $\{e_1, \ldots, e_n\}$, E* = $\{f_1, \ldots, f_n\}$. Haben dann

$$A^*f_j = \alpha_1 f_1 + \ldots + \alpha_n f_n, \quad \alpha = \alpha_i$$

$$Ae_i = \beta_1 e_1 + \ldots + \beta_n f_n, \quad \beta = \beta_j.$$

Damit ist

$$\overline{\alpha} = \overline{\alpha_i} = (e_i, \alpha_1 f_1 + \ldots + \alpha_n f_n) = (e_i, A^*f_j) = (Ae_i, f_j) = (\beta_1 e_1 + \ldots + \beta_n f_n, f_j) = \beta_j = \beta.$$

#

Satz 6 *Sei X ein endlichdimensionaler linearer Raum mit Skalarprodukt und A $\in \mathcal{L}(X)$. Dann sind folgende Bedingungen äquivalent:*
(i) A ist unitär bzw. orthogonal;
*(ii) A*A = I;*
(iii) AA = I;*
(iv) A invertierbar und $A^{-1} = A^*$.

Beweis:
Folgt aus Satz 5 und Sätzen 2 und 3 aus 3.3.

Direkter Beweis:
(i) $\Rightarrow$ (ii):
(Ax,Ay) = (x,y) $\forall$ x,y $\in$ X $\Rightarrow$ (A*Ax,y) = (x,y) $\forall$ x,y $\in$ X $\Rightarrow$ A*A = I.
(ii) $\Rightarrow$ (iii):
A*A = I $\Rightarrow$ (det A*)(det A) = 1 $\Rightarrow$ det A $\neq$ 0. $\Rightarrow$ A invertierbar.
Multiplizieren A*A = I von links mit A:
Erhalten AA*A = A und multiplizieren von rechts mit A^{-1}: AA* = I.
(iii) $\Rightarrow$ (ii): analog.
(ii) $\wedge$ (iii) $\Rightarrow$ (iv): folgt aus Definition des inversen Operators
(ii) $\Rightarrow$ (i):
A*A = I $\Rightarrow$ (A*Ax,y) = (x,y) $\forall$ x,y $\in$ X $\Rightarrow$ (Ax,Ay) = (x,y) $\forall$ x,y $\in$ X. #

Satz 7 *Sei X ein endlichdimensionaler linearer Raum mit Skalarprodukt und $A \in \mathcal{L}(X)$, $A^* \in \mathcal{L}(X)$ der Rieszsche Adjungierte. Dann gilt*

$$\dim \operatorname{coker} A = \dim \ker A^*,$$
$$\dim \operatorname{coker} A^* = \dim \ker A.$$

Beweis:
Haben coker A = X/Im A und nach Satz 8 aus 2.8. ist X/Im A isomorph zu jedem direkten Komplement von Im A in X. Nach Satz 2 aus 3.2. ist $(\text{Im A})^{\perp}$ ein direktes Komplement von Im A in X. Zeigen $(\text{Im A})^{\perp}$ = ker A*:

y $\in$ $(\text{Im A})^{\perp}$ $\Rightarrow$ (x,y) = 0 $\forall$ x $\in$ Im A $\Leftrightarrow$ (Az,y) = 0 $\forall$ z $\in$ X $\Leftrightarrow$ (z,A*y) = 0 $\forall$ z $\in$ X $\Leftrightarrow$ (A*y, A*y) = 0 $\Leftrightarrow$ A*y = 0 $\Leftrightarrow$ y $\in$ ker A*. #

3.5 Selbstadjungierte Operatoren und Matrizen

Im folgenden sei X stets ein endlichdimensionaler linearer Raum mit Skalarprodukt (-,-). Unter A* verstehen wir stets den Rieszschen Adjungierten.

Defintion 1 *Ein Operator $A \in \mathcal{L}(X)$ heißt selbstadjungiert, wenn $A = A^*$ ist, d.h. wenn gilt*

$$(Ax,y) = (Ax,y) \ \forall \ x,y \in X.$$

Im Falle $\mathcal{K} = \mathcal{R}$ (Euklidischer Raum) nennt man selbstadjungierte Operatoren auch symmetrische Operatoren und für $\mathcal{K} = \mathcal{C}$ (Unitärer Raum) auch hermitesche Operatoren.

Eine Matrix $A \in M_n(\mathcal{K})$ heißt selbstadjungiert, wenn $A = A^*$ ist, d.h. wenn $A_{ij} = A^*_{ji}$ $\forall$ i,j gilt. Für $\mathcal{K} = \mathcal{R}$ spricht man auch von symmetrischen Matrizen, für $\mathcal{K} = \mathcal{C}$ von hermiteschen Matrizen.

23.06.06

Satz 1 *Für $A \in \mathcal{L}(X)$ sind folgende Bedingungen äquivalent:*
(i) A ist selbstadjungiert;
(ii) Es existiert eine ONB E, sodass $[A]_{E,E}$ selbstadjungiert ist;
(iii) $[A]_{E,E}$ ist selbstadjungiert für jede ONB E.

Beweis: Folgt aus Satz 5 aus 3.4. #

Satz 2 *Sei $A \in \mathcal{L}(X)$ selbstadjungiert. Dann ist sp $A \subset \mathcal{R}$. Jeder selbstadjungierte Operator und jede selbstadjungierte Matrix hat insbesondere einen EW und alle EW sind reell. Zu verschiedenen EW gehörende EVn sind orthogonal.*

Beweis:
Wählen eine ONB in X und erhalten eine selbstadjungierte Matrix B = $[A]_{E,E} \in M_n$ ($\mathcal{C}$).
Ein Punkt λ gehört zum Spektrum genau dann, wenn det (B-λI) = 0 ist, d.h. wenn x $\in \mathcal{C}^n \setminus \{0\}$ mit (B-λI)x = oder oder Bx = λx existiert.

Sei (-,-) das übliche Skalarprodukt im $\mathcal{C}^n$.
Dann ist (Bx,x) = $\sum b_{ij} x_i \overline{x_j}$ und

$$(x, Bx) = \overline{(Bx, x)} = \sum \overline{b_{ij}}\,\overline{x_i} x_j = \sum b_{ji} x_j \overline{x_i},$$

d.h. es gilt (Bx,x) = (x,Bx). Somist ist

$$\lambda(x,x) = (\lambda x, x) = (Bx, x) = (x, Bx) = (x, \lambda x) = \overline{\lambda}(x,x)$$

und wegen (x,x) $\neq 0$ folgt $\lambda = \overline{x}$, d.h. $\lambda \in \mathcal{R}$.
Zweite Behauptung folgt aus EW(A) = sp A $\cap$ $\mathcal{K}$ = sp A.
Sei Ax = λx und Ay = μy mit $\lambda \neq \mu$ und λ, $\mu \in \mathcal{R}$. Haben

$$(Ax, y) = (x, Ay) \Rightarrow (\lambda x, y) = (x, \mu y) \Rightarrow \lambda(x,y) = \overline{\mu}(x,y) \Rightarrow \lambda(x,y) = \mu(x,y) \Rightarrow (\lambda - \mu)(x,y) = 0 \Rightarrow (x,y) = 0.$$

#

Satz 3 *Für jeden selbstadjungierten Operator $A \in \mathcal{L}(X)$ existiert eine ONB E, sodass*

$$[A]_{E,E} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

mit $\lambda_j \in \mathcal{R}$ ist (Die $\lambda_1, \ldots, \lambda_n$ sind die EW von A).

Lemma: Sei A $\in \mathcal{L}$(X) selbstadjungiert und U $\subset$ X ein invarianter Unterraum von A. Dann ist auch $U^{\perp}$ ein invarianter Unterraum von A.

Beweis:
Sei y $\in$ U$^\perp$. Haben dann für x $\in$ U folgendes:
(Ay,x) = (y,$\underbrace{Ax}_{\in U}$) = 0.
Also ist Ay ebenfalls $\in$ U$^\perp$. #

Beweis von Satz 3:
Nach Satz 2 hat einen EW λ_1. Sei e_1 ein EV mit $||e_1|| = 1$ und U_1 = span $\{e_1\}$. Dann ist U_1 ein invarianter Unterraum von A, nach dem Lemma also auch $U_1^\perp$. Die Einschränkung von A auf $U_1^\perp$ ist wieder selbstadjungiert. Sei $\lambda_2 \in U_1^\perp$ ein EW und $e_2 \in U_1^\perp$ ein EV mit $||e_2|| = 1$. Setzen U_2 = span $\{e_2\}$. Dann ist U_2 ein invarianter Unterraum von A/$U_1^\perp$, nach dem Lemma also auch $U_2^\perp \subset U_1^\perp$. Fahren wir so fort, so erhalten wir $e_1, \ldots, e_n$ mit A$e_j = \lambda_j e_j$, $||e_j|| = 1$, $(e_i, e_j) = 0$ für i $\neq$ j (da $e_2, \ldots, e_n \in U_1^\perp$, $e_3, \ldots, e_n \in U_2^\perp$, ...). #

Wissen also, dass sich unitäre und selbstadjungierte Operatoren in einer ONB diagonalisieren lassen (stimmt nicht für jeden orthogonalen Operator).

Defintion 2 *Sei X ein unitärer Raum (d.h. $\mathcal{K} = \mathcal{C}$). Ein Operator $A \in \mathcal{L}(X)$ heißt normal, wenn $AA^* = A^*A$ gilt. Eine Matrix $A \in M_n(\mathcal{C})$ heißt normal, wenn $A^*A = AA^*$ ist.*

Selbstadjungierte Operatoren sind normal: AA = AA
Unitäre Operatoren sind normal: A* = A^{-1}; A^{-1} = AA^{-1} = I
Diagonaloperatoren sind normal: DD* = D*D
(D = $\begin{pmatrix} 1+i & 0 \\ 0 & 0 \end{pmatrix}$ ist normal, aber weder selbstadjungiert noch unitär)

Satz 4 *Für einen Operator $A \in \mathcal{L}(X)$ in einem unitären Raum X sind folgende Bedingungen äquivalent:*
(i) A ist normal;
(ii) Es existiert eine ONB E, sodass $[A]_{E,E}$ eine Diagonalmatrix ist.

Beweis:
(ii) $\Rightarrow$ (i): trivial, folgt aus Satz 5 aus 3.3
(i) $\Rightarrow$ (ii):
Sei λ EW von A und E_λ = {x $\in$ X: Ax = λx }.
Behaupten, dass E_λ und $E_\lambda^\perp$ invariante Unterräume von A und A* sind.

A(E_λ) $\subset$ E_λ: x $\in$ E_λ $\Rightarrow$ Ax = λx $\in$ E_λ.

A*(E_λ) $\subset$ E_λ: x $\in$ E_λ $\Rightarrow$ A(A*x) = A*(Ax) = A*(λx) = λ(A*x) $\Rightarrow$ A*x $\in$ E_λ.

A($E_\lambda^\perp$) $\subset$ $E_\lambda^\perp$: y $\in$ $E_\lambda^\perp$ (d.h. (y,x) = 0 $\forall$ x $\in$ E_λ) $\Rightarrow$ (Ay,x) = (y,A*x) = 0 $\forall$ x $\in$ E_λ $\Rightarrow$ Ay $\in$ $E_\lambda^\perp$.

A*($E_\lambda^\perp$) $\subset$ $E_\lambda^\perp$: y $\in$ $E_\lambda^\perp$ $\Rightarrow$ (A*y,x) = (y,Ax) = 0 $\forall$ x $\in$ E_λ $\Rightarrow$ A*y $\in$ $E_\lambda^\perp$.

Wählen ONBn in E_λ und $E_\lambda^\perp$ und bilden deren Vereinigung. Dies ist eine ONB F in X. Haben

$$[A]_{F,F} = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}.$$

Die Einschränkung von A auf E_λ und auf $E_\lambda^\perp$ sind wieder normal und haben die Matrixdarstellung A_1 und A_2 in den obigen Basen. Behauptung folgt so durch vollständige Induktion. #

Jede komplexe Zahl a lässt sich als a = β + iγ mit $\beta, \gamma \in \mathcal{R}$ schreiben. Das Analogon für Operatoren ist wie folgt:
Ist A $\in \mathcal{L}$(X), so ist A = B + iC mit B = $\frac{1}{2}$(A + A*), C = $\frac{1}{2}$(A - A*) und es gilt B = B*, C = C*.
Jede komplexe Zahl a $\neq$ 0 lässt sich ferner eindeutig in der Form a = r γ mit r > 0 und $\gamma \in \mathcal{T}$ schreiben.
Suchen hierfür ein Analogon für Operatoren.

Defintion 3 *Ein Operator A $\in \mathcal{L}(X)$ heißt positiv (bzw. positiv definit), wenn A selbstadjungiert ist und (Ax,x) $\geq$ 0 $\forall$ x $\in$ X ist (bzw. (Ax,x) > 0 $\forall$ x $\in$ X ohne {0} ist).*

A selbstadjungiert $\Rightarrow$ (Ax,x) = (x,Ax) = $\overline{(Ax,x)}$ $\Rightarrow$ (Ax,x) reell

Satz 5 *Für A $\in \mathcal{L}(X)$ sind folgende Bedingungen äquivalent:*

(i) A ist positiv;	*(i) A ist positiv definit;*
*(ii) Es existiert ein Operator C $\in \mathcal{L}(X)$: A = C*C.*	*(ii) Es existiert ein Operator C $\in G\mathcal{L}(X)$: A = C*C.*
(iii) $\exists$ B $\in \mathcal{L}(X)$: B selbstadjungiert, A = B^2;	*(iii) $\exists$ B $\in G\mathcal{L}(X)$: B selbstadjungiert, A = B^2;*
(iv) Alle EW von A sind $\geq$ 0;	*(iv) Alles EW von A sind > 0;*
(v) $\exists$ ONB E: $[A]_{E,E} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ mit $\lambda_j \geq 0$.	*(v) $\exists$ ONB E: $[A]_{E,E} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ mit $\lambda_j > 0$.*

(Mit G$\mathcal{L}$(X) bezeichnen wir die Menge der invertierbaren linearen Operatorn in X.)
Beweis:
(i) $\Rightarrow$ (v): 28.06.06

Nach Satz 3 existiert eine ONB, sodass $[A]_{E,E} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$. Sei E = $\{e_1, \ldots, e_n\}$.

Dann ist A$e_j = \lambda_j e_j$ und somit

$$\lambda_j = \lambda_j(e_j, e_j) = (\lambda_j e_j, e_j) = (Ae_j, e_j) \geq (>)0.$$

(iv) $\Rightarrow$ (v) klar.
(v) $\Rightarrow$ (iii):
Sei B der Operator mit

$$[B]_{E,E} = \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix},$$

d.h. $B(x_1e_1 + \ldots + x_ne_n) = \sqrt{\lambda_1}x_1e_1 + \ldots + \sqrt{\lambda_n}x_ne_n$.
Dann ist B selbstadjungiert und $B^2 = A$.
(iii) $\Rightarrow$ (ii) Man setze C = B.
(ii) $\Rightarrow$ (i):
Haben A* = (C*C)* = C*C** = C*C = A und (Ax,x) = (C*Cx,x) = (Cx,Cx) ≥ 0
Es ist (Cx,Cx) = 0 $\Leftrightarrow$ Cx = 0 $\Leftrightarrow$ x = 0, falls C invertierbar ist. #

Satz 6 *Polardarstellung invertierbarer Operatoren*
Sei $A \in G\mathcal{L}(X)$. Dann existiert eine eindeutig bestimmte Isometrie $U \in G\mathcal{L}(X)$ und ein eindeutig bestimmter positiv definiter Operator $P \in G\mathcal{L}(X)$ mit $A = UP$.

Bemerkung: Satz angewandt auf A* ergibt A* = UP $\Rightarrow$ A = P*U* $\Rightarrow \tilde{P}\tilde{U}$ mit $\tilde{P}$ positiv definit und $\tilde{U}$ Isometrie.

Beweis:
Der Operator A*A ist positiv definit (Satz 5). Nach Satz 5 existiert also ein invertierbarer Operator P mit P = P* und $A^*A = P^2$. Da P invertierbar ist, ist P positiv definit (folgt aus Beweis von Satz 5). Setzen $U = AP^{-1}$. Dann ist A = UP.
Haben
$U^*U = (AP^{-1})^*AP^{-1} = (P^{-1})^*A^*AP^{-1} = P^{-1}PPP^{-1} = I$,
d.h. U ist orthogonal bzw. unitär.
Seien UP = VQ zwei Darstellungen von A. Dann ist
$P^2 = PU^*UP = (UP)^*UP = A^*A = (VQ)^*VQ\ QV^*VQ = Q^2$.
Sei E eine ONB mit $[P]_{E,E} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ und $\lambda_j > 0$ (Satz 5). Dann ist $[P^2]_{E,E} = \begin{pmatrix} \lambda_1^2 & & \\ & \ddots & \\ & & \lambda_n^2 \end{pmatrix} = [Q^2]_{E,E}$, woraus $[Q]_{E,E} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ folgt, weil alle EW von Q positiv sind (Satz 5).
Also ist P = Q.
Es folgt UP = VP und damit U = V. #

Satz 7 *Sei E eine ONB in X. Eine Basis F ist genau dann eine ONB in X, wenn die Übergangsmatrix $U_{E,F}$ orthogonal bzw. unitär ist.*

Beweis:
Sei

$$\begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} = \begin{pmatrix} u_{11} & \ldots & u_{1n} \\ \vdots & & \vdots \\ u_{n1} & \ldots & u_{nn} \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.$$

Dann ist

$$(f_i, f_j) = (u_{i1}e_1 + \ldots + u_{in}e_n, u_{j1}e_1 + \ldots + u_{jn}e_n) = u_{i1}\overline{u_{j1}} + \ldots + u_{in}\overline{u_{jn}},$$

d.h. $(f_i,f_j) = \delta_{ij} \Leftrightarrow$ Zeilen von U bilden ONS. #

Satz 8 *Sei $A \in M_n(\mathcal{K})$ selbstadjungiert. Dann existiert eine orthogonale ($\mathcal{K} = \mathcal{R}$) bzw. unitäre ($\mathcal{C} = \mathcal{R}$) Matrix C mit*

$C^T AC$ *bzw.* $C^*AC = C^{-1} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$,

wobei $\lambda_j \in \mathcal{R}$ ist (die λ_j sind die EW von A).

Beweis:
Fassen A als Operator $\tilde{A}$ auf $\mathcal{K}^n$ mit der Standardbasis S und dem üblichen Skalarprodukt (x,y) = $\sum x_i \overline{y_i}$ auf. Dann ist A = $[\tilde{A}]_{S,S}$.
Nach Satz 3 existiert eine ONB E mit

$$[\tilde{A}]_{E,E} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

wobei $\lambda_j \geq 0$ die EW von A bzw $\tilde{A}$ sind.
Nach Satz 3 aus 2.6. ist

$$\underbrace{[\tilde{A}]_{S,S}}_{A} = \underbrace{(U^T_{E,S})^{-1}}_{C} \quad \underbrace{[\tilde{A}]_{E,E}}_{\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}} \quad \underbrace{U^T_{E,S}}_{C^{-1}}$$

d.h. $C^{-1}AC = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ und C ist nach Satz 7 orthogonal bzw. unitär. #

Defintion 4 *Sei $A \in \mathcal{L}(X)$ oder $A \in M_n(\mathcal{K})$. Die Singulärwerte $s_1, \ldots, s_n$ sind definiert als die (nichtnegativen) Wurzeln aus den Eigenwerten von A*A.*

A*A ist stets positiv (Satz 5), für die EW von A*A gilt also $\lambda_1 \geq 0, \ldots, \lambda_n \geq 0$.
Die Singulärwerte von A sind dann

$$s_j = \sqrt{\lambda_j(A^*A)}.$$

Satz 9 *Singulärwertzerlegung*
Sei $A \in M_n(\mathcal{K})$. Dann existieren orthogonale bzw. unitäre Matrizen U und V mit

$$A = U \begin{pmatrix} s_1 & & \\ & \ddots & \\ & & s_n \end{pmatrix} V,$$

wobei $s_1, \ldots, s_n$ die Singulärwerte von A sind.

Beweis:
Die Matrix A*A ist positiv, nach Satz 8 existiert also eine orthogonale bzw. unitäre Matrix W mit

$$W^{-1}A^*AW = \begin{pmatrix} s_1^2 & & \\ & \ddots & \\ & & s_n^2 \end{pmatrix},$$

d.h.

$$A^*AW = W \begin{pmatrix} s_1^2 & & \\ & \ddots & \\ & & s_n^2 \end{pmatrix}.$$

Sei e_j die j-te Spalte von W. Dann ist E = $\{e_1, \ldots, e_n\}$ eine ONB in $\mathcal{K}^n$. Haben A*Ae_j = $s_j^2 e_j$ $\forall$ j = 1,...,n.
Sei $s_1 \geq \ldots \geq s_r > 0$ und $s_{r+1} = \ldots = s_n = 0$.
Für j = 1,...,r setzen wir $f_j = \frac{1}{s_j} Ae_j$. Dann ist

$$(f_i, f_j) = \frac{1}{s_i s_j}(Ae_i, Ae_j) = \frac{1}{s_i s_j}(A^*Ae_i, e_j) = \frac{1}{s_i s_j}(s_i^2 e_i, e_j) = \delta_{ij}.$$

Also ist $f_1, \ldots, f_j$ ein ONS. Ergänzen dieses zu einer ONB $\{f_1, \ldots, f_r, f_{r+1}, \ldots, f_n\}$ im $\mathcal{K}^n$, die wir F nennen.

$$[A]_{E,F} = \begin{pmatrix} s_1 & & \\ & \ddots & \\ & & s_n \end{pmatrix},$$

denn $1 \leq$ j $\leq$ r $\Rightarrow$ A$e_j = s_j f_j$,
r + 1 $\leq$ j $\leq$ n $\Rightarrow$ A*A$e_j = 0 \Rightarrow$ (A*Ae_j,e_j) = 0 $\Rightarrow$ (Ae_j,Ae_j) = 0 $\Rightarrow$ A$e_j = 0 = 0f_1 + \ldots + 0f_n$.

Die Situation ist also wie folgt:
$\mathcal{K}_s^n \xrightarrow{A} \mathcal{K}_s^n$
$\mathcal{K}_E^n \xrightarrow{[A]_{E,F}} \mathcal{K}_F^n$.

Wieder nach Satz 3 aus 2.6. ist also

$$A = [A]_{S,S} = \underbrace{(U_{F,S}^T)^{-1}}_{U} \quad \underbrace{[A]_{E,F}}_{\begin{pmatrix} s_1 & & \\ & \ddots & \\ & & s_n \end{pmatrix}} \quad \underbrace{U_{E,S}^T}_{V}$$

und U und V sind nach Satz 7 orthogonal bzw. unitär. #

Sei ||-|| eine Norm in X (z.B. ||x|| = $\sqrt{(x,x)}$). Für jeden Operator A $\in \mathcal{L}$(X) ist dann

$$||\text{A}||:= \max_{x \in X ||x|| \leq 1} ||\text{Ax}||$$

endlich. Man nennt ||A|| die mit der Norm ||-|| auf X *assoziierte Operatornorm*.

Man kann zeigen:
||A|| = $\sup_{x \neq 0} \frac{||Ax||}{||x||}$ = kleinste Zahl M mit ||Ax|| $\leq$ M||x|| $\forall$ x $\in$ X.

Bestimmung von ||A|| ist im Allgemeinen schwierig. Ist X = $\mathcal{C}^n$ mit

$$||\mathrm{x}||_p = (|x_1|^p + \ldots + |x_n|^p)^{\frac{1}{p}} \qquad (1 \leq \mathrm{p} \leq \infty)$$
$$||\mathrm{x}||_\infty = \max\,(|x_1|, \ldots, |x_n|),$$

so ist $||\mathrm{A}||_1$ = Spaltensummennorm, $||\mathrm{A}||_\infty$ = Zeilensummennorm, $||\mathrm{A}||_2 = s_{max}(\mathrm{A})$ = max.Singulärwert von A = $\sqrt{\lambda_{max}(A^*A)}$.

Es gilt desweiteren $||\mathrm{A}^{-1}||_2 = \frac{1}{s_{min}(A)}$ = 1/min. Singulärwert von A.

Die Zahl $||\mathrm{A}||_2||\mathrm{A}^{-1}||_2$ nennt man *Konditionszahl* (*Spektrale*) und die Norm $||\mathrm{A}||_2$ nennt man die *Spektralnorm* von A. Haben also

$$||\mathrm{A}||\ ||\mathrm{A}^{-1}|| = \frac{s_{max}(A)}{s_{min}(A)}.$$

30.06.06

Sei A invertierbar und betrachten Ax = y. Sei γ mit Fehler behaftet. Haben A(x + ϵ) = y + δ. Also ist Aϵ = δ. Das kleinste M mit $||\epsilon|| \leq \mathrm{M}||\delta||$ ist $||\mathrm{A}^{-1}||$, d.h. haben $||\epsilon|| \leq ||\mathrm{A}^{-1}||\ ||\delta||$ (und es gibt δ, bei denen $\leq$ zu = wird). Ist $||\mathrm{A}^{-1}||$ groß, so ist das System also schlecht. Man kann ein beliebiges System schlecht machen, indem man es mit einer kleinen Zahl multipliziert, z.B. Ax = y zu 10^{-6}Ax = 10^{-6}y mit der Folge, dass $||(10^{-6}\mathrm{A})^{-1}|| = ||10^6\ \mathrm{A}^{-1}|| = 10^6\ ||\mathrm{A}^{-1}||$ entsteht.
Deshalb nimmt man $||\mathrm{A}||\ ||\mathrm{A}^{-1}||$ ($||\alpha\mathrm{A}||\ ||(\alpha\mathrm{A})^{-1}|| = \alpha||\mathrm{A}||\ ||\alpha^{-1}||\ ||\mathrm{A}^{-1}|| = ||\mathrm{A}||\ ||\mathrm{A}^{-1}||$).
Systeme mit kleinem bzw. großem $||\mathrm{A}||\ ||\mathrm{A}^{-1}||$ heißen gut bzw. schlecht konditioniert. Ist Ax = y schlecht konditioniert, so versucht man ein B zu finden, sodass BAx = By besser konditioniert ist (der Idealfall wäre B = A^{-1}). Dies nennt man Vorkonditionierung.

3.6 Bilinearformen und Räume mit indefinitier Metrik

Im folgenden sei X stets ein endlichdimensionaler linearer Raum.

Definition 1 *Eine Bilinearform auf X ist eine Abbildung g: X × X → $\mathcal{K}$ mit folgenden Eigenschaften:*
g(x + y,z) = g(x,z) + g(y,z)
g(x,y + z) = g(x,y) + g(x,z) $\quad \forall\ x,y,z \in X$
g(αx,y) = αg(x,y)
g(x,αy) = $\overline{\alpha}$g(x,y) $\quad \forall\ \alpha \in \mathcal{K}, \forall\ x,y \in X$

Eigentlich müssten wir (bei $\mathcal{K} = \mathcal{C}$) von sesquilinearen Formen sprechen, bleiben aber bei Bilinearform.

Wählen in X eine Basis E = $\{e_1, \ldots, e_n\}$. Für x = $x_1e_1 + \ldots + x_ne_n \in$ X schreiben wir dann

$$[x]_E = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Setzen $[x]_E^T = (x_1 \ldots x_n)$, $\overline{[x]_E} = \begin{pmatrix} \overline{x_1} \\ \vdots \\ \overline{x_n} \end{pmatrix}$.

Sei A $= (a_{ij})_{i,j=1}^n \in M_n(\mathcal{K})$ eine beliebige Matrix. Dann ist durch g(x,y) $= [x]_E^T$ A $\overline{[y]_E}$ offenbar eine Bilinearform auf X gegeben.

Haben

$$g(x,y) = (x_1 \dots x_n) \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} \overline{y_1} \\ \vdots \\ \overline{y_n} \end{pmatrix} = (x_1 \dots x_n) \begin{pmatrix} a_{11}\overline{y_1} + \dots + a_{1n}\overline{y_n} \\ \vdots \\ a_{n1}\overline{y_1} + \dots + a_{nn}\overline{y_n} \end{pmatrix}$$

$$= x_1(a_{11}\overline{y_1} + \dots + a_{1n}\overline{y_n}) + \dots + x_n(a_{n1}\overline{y_1} + \dots + a_{nn}\overline{y_n}) = \sum_{i,j=1}^n a_{ij} x_i \overline{y_j}$$

Der folgende Satz zeigt, dass damit alle Bilinearformen gegeben sind.

Satz 1 *Sei g eine Bilinearform auf X und $E = \{e_1, \dots, e_n\}$ eine Basis in X. Dann gibt es eine Matrix $A \in M_n(\mathcal{K})$ mit*

$$g(x,y) = [x]_E^T \; A \; \overline{[y]_E} \qquad \forall\, x,y \in X.$$

Diese Matrix A ist eindeutig bestimmt, und zwar ist $A = (g(e_i, e_j))_{i,j=1}^n$.
Man nennt A die Gramsche Matrix von g in der Basis E und bezeichnet A mit $[g]_E$.

Beweis: Haben
$g(x_1e_1 + \dots + x_ne_n, y_1e_1 + \dots + y_ne_n) = \sum_{i,j=1}^n g(x_ie_i, y_je_j) = \sum_{i,j=1}^n g(e_i,e_j)x_i\overline{y_j}$
$= [x]_E^T \begin{pmatrix} g(e_1,e_1) & \dots & g(e_1,e_n) \\ \vdots & & \vdots \\ g(e_n,e_1) & \dots & g(e_n,e_n) \end{pmatrix} \overline{[y]_E}$,
d.h. $(g(e_i,e_j))_{i,j=1}^n$ ist die gewünschte Matrix A.

Ist A irgendeine Matrix mit der angegebenen Eigenschaft, so ist

$$g(e_i,e_j) = (0 \dots 0\ 1\ 0 \dots 0) \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = a_{ij}. \qquad \#$$

Die Übergangsmatrix zwischen zwei Basen ist gegeben durch

$$\begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} = \underbrace{\begin{pmatrix} u_{11} & \dots & u_{1n} \\ \vdots & & \vdots \\ u_{n1} & \dots & u_{nn} \end{pmatrix}}_{U_{E,F}} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix},$$

und wir wissen, dass

$$[x]_F = (U_{E,F}^T)^{-1}[x]_E = U_{F,E}^T[x]_E.$$

Satz 2 *Sind E und F Basen in X, so gilt für jede Bilinearform g auf X die Transformationsformel*

$$[g]_F = U_{E,F}[g]_E U^*_{E,F}$$

Beweis:
Haben

$$g(x,y) = [x]^T_E[g]_E\overline{[y]_E} = ((U^T_{F,E})^{-1}[x]_F)^T[g]_E\overline{(U^T_{F,E})^{-1}[y]_F}$$
$$= [x]^T_F \underbrace{U_{E,F}[g]_E U^*_{E,F}}_{=[g]_F} \overline{[y]_F}$$

#

Gramsche Matrizen transformieren sich also nicht wie Matrixdarstellungen von Operatoren nach der Regel A → C^{-1}AC, sondern nach der Regel A → C*AC.

Durch A ~ B ⇔ ∃ inv. C: A = C*BC ist in $M_n(\mathcal{K})$ eine Äquivalenzrelation gegeben, die *Kongruenz* heißt.
Es ergibt sich also wieder ein Klassifizierungsproblem.

Definition 2 *Seien g und h Bilinearformen auf X bzw. Y. Man nennt (X,g) und (Y,h) isometrisch isomorph, wenn es eine bijektive Abbildung A ∈ $\mathcal{L}$(X,Y) mit*

$$h(Ax,Ay) = g(x,y) \qquad \forall\ x,y \in X.$$

Satz 3 *Sei E eine Basis in X unf F eine Basis in Y. Die Paare (X,G) und (Y,h) sind genau dann isometrisch isomophr, wenn gilt*

(a) dim X = dim Y,
(b) ∃ eine invertierbare Matrix C mit $[h]_F = C^*[g]_E C$.

Beweis:
Sei (X,g) ≅ (Y,h). Dann sind X und Y als lineare Räume isomorph, d.h. dim X = dim Y.
Sei A ein isometrischer Isomorphismus. Dann ist
h (Ax,Ay) = g(x,y)

$$[Ax]^T_F[h]_F\overline{[Ay]_F} = [x]^T_E[g]_E\overline{[y]_E}$$
$$([A]_{E,F}[x]_E)^T[h]_F\overline{[A]_{E,F}[y]_E} = [x]^T_E[g]_E\overline{[y]_E}$$
$$[x]^T_E[A]^T_{E,F}[h]_F\overline{[A]_{E,F}[y]_E} = [x]^T_E[g]_E\overline{[y]_E}$$

Nach Satz 1 ist also

$$\underbrace{[A]^T_{E,F}}_{(C^{-1})^*}[h]_F\ \underbrace{\overline{[A]_{E,F}}}_{=:C^{-1}} = [g]_E,$$

d.h. $[h]_F = C^*[g]_E C$.

Gelten (a) und (b), so ist die Abbildung A $\in \mathcal{L}$(X,Y), die durch $[A]_{E,F} = \overline{C^{-1}}$ gegeben ist, ein isometrischer Isomorphismus. #

Definition 3 *Sei g eine Bilinearform auf X, zwei Elemente x,y $\in$ X heißen orthogonal, wenn g(x,y) = 0 ist. Die Bilinearform heißt symmetrisch, wenn $\mathcal{K} = \mathcal{R}$ und g(x,y) = g(y,x) $\forall$ x,y $\in$ X ist, hermitesch, wenn $\mathcal{K} = \mathcal{C}$ und $g(x,y) = \overline{g(y,x)}$ $\forall$ x,y $\in$ X ist, positiv definit, wenn g symmetrisch bzw. hermitesch ist und $g(x,x) \geq 0 \; \forall \; x \in X$, $g(x,x) = 0 \Leftrightarrow x = 0$ ist, hermitesch, wenn g symmetrisch bzw. hermitesch ist und ein x $\in$ X ohne 0 mit g(x,x) = 0 existiert.*
Ein Raum mit symmetrischer bzw. hermitescher Bilinearform (ein Paar (X,g)) heißt Raum mit Metrik (nicht zu verwechseln mit metrischem Raum).
Positiv definite Bilinearformen sind also Skalarprodukte und Räume mit positiv definiter Metrik sind Euklidische bzw. Unitäre Räume.

05.07.06

Offenbar sind folgende Bedingungen äquivalent:
(i) g ist symmetrisch / hermitesch;
(ii) es existiert eine Basis E in X, sodass $[g]_E$ symmetrisch bzw. hermitesch ist;
(iii) in jeder Basis E von X ist $[g]_E$ symmetrisch bzw. hermitesch

(Beweis: $[g]_E = (g(e_i, e_j))_{i,j=1}^n$)

Definition 4 *Eine Matrix $A \in M_n(\mathcal{K})$ heißt positiv definit, wenn gilt: A symmetrisch bzw. hermitesch und*

$$(x_1 \dots x_n) \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} \overline{x_1} \\ \vdots \\ \overline{x_n} \end{pmatrix} \geq 0$$

für alle $x_1, \dots, x_n \in \mathcal{K}$ und

$$(x_1 \dots x_n) \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} \overline{x_1} \\ \vdots \\ \overline{x_n} \end{pmatrix} = 0$$

genau dann, wenn $(x_1, \dots, x_n)$ = (0,...,0).

Nun erhalten wir also, dass die folgenden Bedingungen äquivalent sind:
(i) g ist positiv definit;
(ii) es existiert eine Basis E, sodass $[g]_E$ positiv definit ist;
(iii) für jede Basis E ist $[g]_E$ positiv definit.

(Beweis: g(x,x) = $[x]_E^T [g]_E \overline{[x]_E}$.)

Beispiele:

$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ positiv definit $(2x_1^2 + x_2^2)$

$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ indefinit $(x_1^2$ - $x_2^2)$

$\begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix}$ indefinit $(4x_1^2)$

Satz 4 *Sei (X,g) ein Raum mit Metrik, d.h. g sei symmetrisch bzw. hermitesch.*
(a) Es existiert eine Basis F in X, sodass gilt

$$[g]_E = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix} = \begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0_r \end{pmatrix}.$$

(b) (X,g) ist isometrisch isomorph zu $(\mathcal{K}^n, k_{pq})$ mit

$$k_{pq}((x_1,\ldots,x_n),(y_1,\ldots,y_n)) = x_1\overline{y_1} + \ldots + x_p\overline{y_p} - x_{p+1}\overline{y_{p+1}} - \ldots - x_{p+q}\overline{y_{p+q}}.$$

(c) Trägheitssatz von Sylvester
Das Tripel (p,q,r) aus (a),(b) ist eindeutig bestimmt.

Beweis:
(a) Wählen irgendeine Basis F. Dann ist $[g]_F$ symmetrisch bzw. hermitesch.
Nach Satz 8/3.5. existiert eine orthogonale bzw. unitäre Matrix C mit

$$C^*[g]_F C = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

mit $\lambda_j \in \mathcal{R}$.
O.B.d.A. sei $\lambda_1,\ldots,\lambda_p > 0$, $\lambda_{p+1},\ldots,\lambda_{p+q} < 0$, $\lambda_{p+q+q} = \ldots = \lambda_n = 0$
(ansonsten C durch CP mit einer geeigneten Permutationsmatrix ersetzen)
Haben

$$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = \underbrace{\begin{pmatrix} \sqrt{|\lambda_1|} & & \\ & \cdots & \\ & & \sqrt{|\lambda_n|} \end{pmatrix}}_{\Lambda^*} \begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0_r \end{pmatrix} \underbrace{\begin{pmatrix} \sqrt{|\lambda_1|} & & \\ & \ddots & \\ & & \sqrt{|\lambda_n|} \end{pmatrix}}_{\Lambda}$$

mit der Vereinbarung, dass $\sqrt{|\lambda_j|} := 1$ für p + q + 1 ≤ j ≤ n.

Die Diagonalmatrix Λ ist invertierbar und es folgt

$$\underbrace{(\Lambda^*)^{-1}C^*}_{(C\Lambda^{-1})^*}[g]_F C\Lambda^{-1} = \begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0_r \end{pmatrix}.$$

Definieren E durch E = $U_{F,E}$ F mit $U_{F,E} = (C\Lambda^{-1})^*$.

Nach Satz 2 ist dann

$$[g]_E = (C\Lambda^{-1})^*[g]_F C\Lambda^{-1} = \begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0_r \end{pmatrix}$$

(b) Folgt aus (a) und Satz 3.

(c) Seien E = $\{e_1, \ldots, e_n\}$ und F = $\{f_1, \ldots, f_n\}$ Basen mit

$$[g]_E = \begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0_r \end{pmatrix}, [g]_F = \begin{pmatrix} I_s & & \\ & -I_t & \\ & & 0_u \end{pmatrix}.$$

Sei p > s. Setzen P = span $\{e_1, \ldots, e_n\}$, T = span $\{f_{s+1}, \ldots, f_n\}$.
Dann ist
n ≥ dim (P + T) = dim P + dim T - dim (P ∩ T) = p + n - s - dim (P ∩ T),

d.h. dim (P ∩ T) ≥ p - s > 0.
Es existiert also ein x ∈ (P ∩ T) \ {0}.
Wegen x ∈ P ist x = $\alpha_1 e_1 + \ldots + \alpha_p e_p$
und damit

$$g(x,x) = [x]_E^T [g]_E \overline{[x]_E} = (\alpha_1, \ldots, \alpha_p, 0, \ldots, 0) \begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0_r \end{pmatrix} \begin{pmatrix} \overline{\alpha_1} \\ \vdots \\ \overline{\alpha_p} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = |\lambda_1|^2 + \ldots + |\alpha_p|^2 > 0$$

Andererseits ist x ∈ T, d.h. x = $\beta_{s+1} f_{s+1} + \ldots + \beta_n f_n$
und somit

$$g(x,x) = [x]_F^T [g]_F \overline{[x]_F} = (0 \ldots 0 \beta_{s+1} \ldots \beta_n) \begin{pmatrix} I_s & & \\ & -I_t & \\ & & 0_u \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \overline{\beta_{s+1}} \\ \vdots \\ \overline{\beta_n} \end{pmatrix} = -|\beta_{s+1}|^2 - \ldots - |\beta_{s+t}|^2 \leq 0.$$

Widerspruch!
Widerspruch zeigt, dass unsere Annahme p > s falsch ist ⇒ p ≤ s.
Analog zeigt man s ≤ p, q ≤ t, t ≤ q. #

Definition 5 *Das durch Satz 4 festgelegte Tripel (p,q,r) nennt man die Signatur der Metrik g.*

Euklidischer Raum ist ($\mathcal{R}^n$,g) mit Signatur (n,0,0).
Unitärer Raum ist ($\mathcal{C}^n$,g) mit Signatur (n,0,0).
Raum ($\mathcal{R}^4$,g) mit Signatur (3,1,0) heißt *Minkowskiraum*.

Folgerung: Zwei endlichdimensionale Räume mit Metrik sind genau dann isometrisch isomorph, wenn ihre Metriken die gleiche Signatur haben.

Beweis: Satz 3, Satz 4, Definition 5. #

Bemerkung.
Manche Autoren betrachten nur *nichtausgeartete Metriken*, d.h. solche, für die die Signatur von der Form (p,q,0) ist. Diese nennen dann p - q die Signatur. Aus σ = p - q und n = p + q lassen sich p und q eindeutig bestimmen.

Satz 5 *Sylvester*

Eine Matrix $A = (a_{ji})_{i,j=1}^n \in M_n\ (\mathcal{K})$ *ist positiv definit genau dann, wenn gilt*

$$a_{11} > 0, det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} > 0, \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0, \ldots, \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} > 0.$$

Beweis:
Notwendigkeit:
Sei A positiv definit. Für $x_1, \ldots, x_r \in \mathcal{K}$ gilt dann, dass

$$(x_1 \ldots x_r 0 \ldots 0) \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} \overline{x_1} \\ \vdots \\ \overline{x_r} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

immer nichtnegativ ist und zu Null nur für $x_1 = \ldots = x_r = 0$ wird.
Obiger Ausdruck ist aber

$$(x_1 \ldots x_r) \begin{pmatrix} a_{11} & \ldots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \ldots & a_{rr} \end{pmatrix} \begin{pmatrix} \overline{x_1} \\ \vdots \\ \overline{x_r} \end{pmatrix}$$

, d.h. $\begin{pmatrix} a_{11} & \ldots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \ldots & a_{rr} \end{pmatrix} =: A_r$ ist positiv definit.
Nach Satz 5/3.5. ist $A_r = C_r^* C_r$ mit einer invertierbaren Matrix C_r (A_r als Operator auf $\mathcal{K}^n$ in der Standardbasis auffassen). Dies ergibt
$\det A_r = \det C_r^* \det C_r = \overline{det\ C_r} \det C_r = |\det C_r|^2 > 0$.

Hinlänglichkeit:
Ist für n = 1 klar. 07.07.06

Sei Behauptung für n - 1 richtig. Zeigen sie für n.
Sei also

$$A = \begin{pmatrix} a_{11} & \dots & a_{n1} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{nn} \end{pmatrix} = \left(\begin{array}{c|c} a_{11} & S^* \\ \hline S & A_{11} \end{array}\right).$$

Setzen Q = $\left(\begin{array}{c|c} 1 & -\frac{1}{a_{11}}S^* \\ \hline 0 & I \end{array}\right)$ und haben dann

$$\begin{aligned} Q^*AQ &= \left(\begin{array}{c|c} 1 & 0 \\ \hline -\frac{1}{a_{11}}S & I \end{array}\right)\left(\begin{array}{c|c} a_{11} & S^* \\ \hline S & A_{11} \end{array}\right)\left(\begin{array}{c|c} 1 & -\frac{1}{a_{11}}S^* \\ \hline 0 & I \end{array}\right) \\ &= \left(\begin{array}{c|c} 1 & 0 \\ \hline -\frac{1}{a_{11}}S & I \end{array}\right)\left(\begin{array}{c|c} a_{11} & 0 \\ \hline S & A_{11} - \frac{1}{a_{11}}SS^* \end{array}\right) \\ &= \left(\begin{array}{c|c} a_{11} & 0 \\ \hline 0 & A_{11} - \frac{1}{a_{11}}SS^* \end{array}\right) =: \begin{pmatrix} a_{11} & 0 \\ 0 & C \end{pmatrix} =: B. \end{aligned}$$

Weiterhin ist

$$\begin{vmatrix} a_{11} & \dots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \dots & a_{rr} \end{vmatrix} (z_2 - \frac{a_{i1}}{a_{11}}z_1) = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1r} \\ 0 & a_{22} - \frac{a_{21}}{a_{11}}a_12 & \dots & \frac{a_{2r}-a_{21}}{a_{11}a_{1r}} \\ \vdots & \vdots & & \vdots \\ 0 & a_{r2} - \frac{a_{r1}}{a_{11}}a_{12} & \dots & a_{rr} - \frac{a_{r1}}{a_{11}}a_{1r} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & * & \dots & * \\ 0 & c_{11} & \dots & c_{1,r-1} \\ \vdots & \vdots & & \vdots \\ 0 & c_{r-1,1} & \dots & c_{r-1,r-1} \end{vmatrix} = a_{11}\begin{vmatrix} c_{11} & \dots & c_{1,r-1} \\ \vdots & & \vdots \\ c_{r-1,1} & \dots & c_{r-1,r-1} \end{vmatrix} > 0,$$

d.h. die Determinanten der s x s- Hauptabschnitte der (n-1)x(n-1)-Matrix C sind alle positiv. Nach Induktionsvoraussetzung ist C positiv definit und somit existiert nach Satz 5/3.5. eine invertierbare Matrix D mit C = D*D. Es ergibt sich

$$Q * AQ = \begin{pmatrix} a_{11} & 0 \\ 0 & D^*D \end{pmatrix} = \begin{pmatrix} \sqrt{a_{11}} & 0 \\ 0 & D^* \end{pmatrix}\begin{pmatrix} \sqrt{a_{11}} & 0 \\ 0 & D \end{pmatrix} (a_{11} > 0)$$

d.h. A = $(Q^*)^{-1}E^*EQ^{-1}$ = $(EQ^{-1})^*(EQ^{-1})$ mit invertierbaren EQ^{-1}.
Wiederum nach Satz 5/3.5. ist also A positiv definit. #

3.7 Quadratische Formen

Eine Kurve m-ter Ordnung im $\mathcal{R}^2$ war gegeben durch eine Gleichung der Form

$$\sum_{k+l\le m, k,l\ge 0} a_{kl}x^k y^l = 0$$

Für m = 1 erhalten wir ax + by + c = 0, d.h. lineare Untermannigfaltigkeiten des $\mathcal{R}^2$, insbesondere Geraden.

Für m = 2 ergeben sich Ellipsen, Hyperbeln, Parabeln und deren Entartungen.

Eine *Hyperfläche* m-ter Ordnung im $\mathcal{R}^n$ ist gegeben durch eine Gleichung der Form

$$\sum_{k_1+\ldots+k_n\leq m, k_1,\ldots,k_n\geq 0} a_{k_1\ldots k_n} x_1^{k_1}\ldots x_n^{k_n} = 0.$$

Für m = 1 haben wir

$$b_1x_1 + \ldots + b_nx_n + c = 0,$$

d.h. erhalten lineare Untermannigfaltigkeiten.

Uns interessiert der Fall m = 2:

$$\sum_{i,j=1}^{n} a_{ij}x_ix_j + \sum_{k=1}^{n} b_kx_k + c = 0$$

Welche Gestalt Hyperflächen zweiter Ordnung haben, ist für viele Probleme von Bedeutung, z.B. im Zusammenhang mit partiellen Differentialgleichungen der Form

$$\sum_{i,j=1}^{n} a_{ij}\frac{\delta^2 u}{\delta x_i x_j} + \sum_{k=1}^{n}\frac{\delta u}{\delta x_k} + cu = 0$$

Beispiel:
Wellengleichung

$$u_{tt} - c^2u_{xx} + \ldots = 0$$
$$t^2 - c^2x^2 + \ldots = 0$$

Hyperbel

Laplacegleichung

$$u_{xx} + u_{yy} + \ldots = 0$$
$$x^2 + y^2 + \ldots = 0$$

Ellipse

Wärmeleitgleichung

$$u_t - a^2u_{xx} + \ldots = 0$$

u(x,t) Temperatur zur Zeit t im Punkt x

$$t - a^2x^2 + \ldots = 0$$

Parabel

Defintion 1 *Sei X ein reeller linearer Raum. Eine quadratische Form auf X ist eine Abbildung q: X → $\mathcal{R}$ mit der Eigenschaft, dass eine Bilinearform g: X × X → $\mathcal{R}$ mit q(x) = g(x,x) ∀ x ∈ X existiert.*

Sei X endlichdimensional und E = $\{e_1, \ldots, e_n\}$ eine Basis in X. Haben dann

$$q(x) = g(x,x) = [x]_E^T[g]_E[x]_E$$

$$(x_1,\ldots,x_n)\begin{pmatrix} g_{11} & \cdots & g_{1n} \\ \vdots & & \vdots \\ g_{n1} & \cdots & g_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i,j=1}^{n} g_{ij}x_ix_j.$$

Satz 1 *Sei q: $X \to \mathcal{R}$ eine quadratische Form. Dann gibt es genau eine symmetrische Bilinearform mit q(x) = g(x,x) $\forall$ x $\in$ X.*

14.07.06

Beweis:
Es sei q(x) = h(x,x) mt einer Bilinearform h. Dann ist

$$g(x,y) = \frac{1}{2}[h(x,y) + h(x,y)]$$

eine symmetrische Bilinearform mit q(x) = g(x,x).
Sind g_1,g_2 symmetrische Bilinearformen mit g_1(x,x) = g_2(x,x) $\forall$ x $\in$ X, so ist l = g_1 - g_2 eine symmetrische Bilinearform mit l(x,x) = 0 $\forall$ x $\in$ X.
Das ergibt

$$l(x,y) = \frac{1}{2}[l(x+y,x+y) - l(x,x) - l(y,y)] = 0 \forall x \in X.$$

#

Sei zum Beispiel

$$q(x,y) = (x\ y)\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

$$= (x\ y)\begin{pmatrix} ax+by \\ cx+dy \end{pmatrix} = ax^2 + bxy + cxy + dy^2.$$

Dann ist

$$q(x,y) = ax^2 + \frac{b+c}{2}xy + \frac{b+c}{2}xy + dy^2 = (x\ y)\begin{pmatrix} a & \frac{b+2}{2} \\ \frac{b+c}{2} & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Satz 2 *Hauptachsentransformation*
(a) Sei X ein n-dimensionaler linearer Raum mit Skalarprodukt und q: $X \to \mathcal{R}$ eine quadratische Form. Dann gibt es eine ONB E in X und reelle Zahlen $\lambda_1,\ldots,\lambda_n$ mit

$$q(x) = [x]_E^T \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} [x]_E,$$

d.h. $q(x_1e_1 + \ldots + x_ne_n) = \lambda_1x_1^2 + \ldots + \lambda_nx_n^2$.

(b) Sei X ein n-dimensionaler linearer Raum und q: $X \to \mathcal{R}$ eine quadratische Form. Dann gibt es eine Basis F in X mit

$$q(x) = [x]_F^T \begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0_r \end{pmatrix} [x]_F,$$

d.h. $q(x_1f_1 + \ldots + x_nf_n) = x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2$.

Beweis:
(a) Nach Satz 1 existiert eine symmetrische Bilinearform mit q(x) = g(x,x).
Nach den Sätzen 7/3.5., 8/3.5. und 2/3.6. existiert eine ONB E mit

$$[g]_E = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

(b) Folgt analog aus Satz 4/3.6. #

Vermerken, dass $\lambda_1, \ldots, \lambda_n$ die Eigenwerte der (einer) Gramschen Matrix von g in der ONB E sind.

1. Beispiel:
Sei M = {(x,y) $\in \mathcal{R}^2$: $ax^2 + 2bxy + cy^2 = 1$}.
Haben $ax^2 + 2bxy + cy^2 = (x\ y) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$.
Seien λ,μ die Eigenwerte von $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$.

Nach Satz 2 existiert eine ONB $\{e_1, e_2\}$ im $\mathcal{R}^2$ mit

$$q(\xi e_1, \eta e_2) = \lambda\xi^2 + \mu\eta^2.$$

Erhalten: M = $\{(\xi, \eta) \in \mathcal{R}^2$: $\lambda\xi^2 + \mu\eta^2 = 1\}$
$\lambda > 0$, $\mu > 0$: Ellipse
$\lambda > 0$, $\mu = 0$: parallele Geraden
$\lambda > 0$, $\mu < 0$: Hyperbel
$\lambda = 0$, $\mu > 0$: parallele Geraden
$\lambda = 0$, $\mu = 0$: $\varnothing$
$\lambda = 0$, $\mu < 0$: $\varnothing$
$\lambda < 0$, $\mu > 0$: Hyperbel
$\lambda < 0$, $\mu = 0$: $\varnothing$
$\lambda < 0$, $\mu < 0$: $\varnothing$

2. Beispiel:
Sei M = {(x,y,z) $\in \mathcal{R}^3$: $ax^2 + by^2 + cz^2 + 2dxy + 2exz - 2fyz = 1$}
Sind λ, μ, ν die EW von

$$\begin{pmatrix} a & d & e \\ d & b & f \\ e & f & c \end{pmatrix},$$

so haben wir in einer ONB E die Darstellung
M = {(ξ,η,ψ) $\in \mathcal{R}^3$: $\lambda\xi^2 + \mu\eta^2 + \nu\psi^2 = 1$}
$\lambda > 0$, $\mu > 0$, $\nu > 0$: Ellipsoid
$\lambda > 0$, $\mu > 0$, $\nu = 0$: elliptischer Zylinder
$\lambda > 0$, $\mu > 0$, $\nu < 0$: einschaliges Hyperboloid
$\lambda > 0$, $\mu < 0$, $\nu = 0$: hyperbolischer Zylinder
$\lambda > 0$, $\mu < 0$, $\nu < 0$: zweischaliges Hyperboloid
usw. (alles weitere sind Entartungen der obigen Fälle)

Betrachten schließlich noch die Gleichung

$$\sum_{i,j=1}^{n} a_{ij} x_i x_j + \sum_{k=1}^{n} b_k x_k + c = 0.$$

Nach Satz 2 existiert eine ONB, in deren Koordinaten dies zu

$$\sum_{i=1}^{n} \lambda_i y_i^2 + \sum_{i=1}^{n} d_i y_i + c = 0$$

wird.
Ist $\lambda_i = 0 \; \forall$ i, so haben wir eine lineare Untermannigfaltigkeite des $\mathcal{R}^n$.
Für $\lambda_i \neq 0$ ist

$$\lambda_i y_i^2 + d_i y_i = \lambda_i \underbrace{(y_i + \frac{d_i}{2\lambda_i})^2}_{w_i} - \frac{d_i^2}{4\lambda_i}$$

und in einem parallel verschobenen Koordinantensystem ergibt sich nach entsprechender Numerierung der Achsen die Gleichung

$$\lambda_1 w_1^2 + \ldots + \lambda_m w_m^2 + d_{m+1} w_{m+1} + \ldots + d_n w_n + f = 0$$

mit m $\leq$ n ($w_i = z_i$ für i $\geq$ m + 1)
Für $d_{m+1} = \ldots = d_n = 0$ nennt man dies eine *Fläche mit Zentrum* und zwar eine *echte* für f $\neq 0$ und eine *konische* für f = 0.

n = 2
$\lambda_1 w_1^2 + f = 0$ parallele Geraden
$\lambda_1 w_1^2 + \lambda_2 w_2^2 + f = 0$ Ellipse, Hyperbel, (0,0) oder zwei sich schneidende Geraden

n = 3
$\lambda_1 w_1^2 + f = 0$ parallele Ebenen
$\lambda_1 w_1^2 + \lambda_2 w_2^2 + f = 0$ elliptische, hyperbolische Zylinder (mit Entartungen)
$\lambda_1 w_1^2 + \lambda_2 w_2^2 + \lambda_3 w_3^2 + f = 0$ Ellipsoide, Hyperboloide (mit Entartungen)
$\lambda_1 w_1^2 + \lambda_2 w_2^2 + \lambda_3 w_3^2 = 0$ konische Flächen

Sei schließlich M:= $d_{m+1}^2 + \ldots + d_n^2 > 0$

Führen neues Koordinantensystem ein durch

$$\begin{pmatrix} z_1 \\ \vdots \\ z_m \\ z_{m+1} \\ z_{m+2} \\ \vdots \\ z_n \end{pmatrix} = \left(\begin{array}{ccc} 1 & & \\ & \ddots & 0 \\ & & 1 \\ \hline 0 & & -\frac{d_{m+1}}{M} - \ldots - \frac{d_n}{M} \\ \hline 0 & & \textit{beliebig} \end{array}\right) \begin{pmatrix} w_1 \\ \vdots \\ w_m \\ w_{m+1} \\ w_{m+2} \\ \vdots \\ w_n \end{pmatrix}.$$

Es entsteht die Gleichung

$$\lambda_1 z_1^2 + \ldots + \lambda_m z_m^2 \underbrace{-M z_{m+1} + f}_{=-M(z_{m+1} - \frac{f}{M})} = 0$$

Setzen noch $\xi_i = z_i$ für i $\neq$ m + 1, $\xi_{m+1} = z_{m+1} - \frac{f}{M}$
und erhalten

$$\lambda_1 \xi_1^2 + \ldots + \lambda_m \xi_m^2 - M \xi_{m+1} = 0.$$

Dies nennt man eine *nichtzentrale Fläche.*

n = 2
$\lambda\xi_1^2 - M\xi_2 = 0$ Parabel

n = 3
$\lambda\xi_1^2 - M\xi_3 = 0$ „Rinne“
$\lambda_1\xi_1^2 + \lambda_2\xi_2^2 - M\xi_3 = 0$
für $\lambda_1 > 0$, $\lambda_2 > 0$ elliptisches Paraboloid
für $\lambda_1 > 0$, $\lambda_2 < 0$ hyperbolisches Paraboloid (Sattelfläche)

4 Algebraische Strukturen und ihre Morphismen

Kennen bisher Gruppen, lineare Räume, Körper (Schiefkörper). Dies sind Mengen mit einer algebraischen Struktur (innere oder äußere Verknüpfung). Im Folgenden beschäftigen wir uns näher mit solchen algebraischen Strukturen.

topologische Strukturen = metrischer Raum d(x,y)
algebraische und topologische Struktur -> echte Mathematik: normierter Raum (linearer Raum mit einer Norm ||x+y|| = ||x|| + ||y||, ||αx|| = |α| ||x||
Differenzieren $\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0}$)

4.1 Gruppen

4.1.1 Grundbegriffe

Ein *Monoid* ist eine Menge M mit einer Abbildung f = M × M → M. Jedem geordneten Paar (a,b) mit a,b ∈ M wird also ein Element f(a,b) ∈ M zugeordnet. Man schreibt a * b, a • b, a + b,... statt f(a,b) und nennt f (oder *, •, +, ...) *innere Verknüpfung* oder *innere Operation* in M.

Seien (M_1,*) und (M_2,∘) zwei Monoide. Eine Abbildung φ: $M_1 \to M_2$ heißt *Homomorphismus*, wenn φ(a * b) = φ(a) ∘ φ(b) ∀ a,b ∈ M gilt. Bijektive Homomorphismen heißen *Isomorphismen*. Man kann zeigen, dass die Inversen von Isomorphismen wieder isomorph sind. Zwei Monoide heißen *isomorph*, wenn es einen Isomorphismus zwischen ihnen gibt. Isomorphie von Monoiden ist eine Äquivalenzrelation (reflexiv, symmetrisch, transitiv).

Ein Monoid (H,•) heißt *Halbgruppe*, wenn die Verknüpfung assoziativ ist, d.h., wenn gilt

$$a \bullet (b \bullet c) = (a \bullet b) \bullet c \;\forall\; a,b,c \in H$$

Können also abcd usw. bilden.

Eine Halbgruppe (G,•) heißt *Gruppe*, wenn folgende Axiome gelten:

$$\exists\, e \in G\colon e \bullet a = a \bullet e = a \;\forall\; a \in G$$
$$\forall\, a \in G\; \exists\, b \in G\colon a \bullet b = b \bullet a = e$$

Das Element e nennt man *neutrales Element* (*Einselement* bei Multiplikation und *Nullelement* bei Addition). Das Element e ist eindeutig bestimmt. Das Element b aus dem letzten Axiom heißt das zu a *inverse Element* und wird mit a^{-1} bezeichnet bei multiplikativer Schreibweise und mit -a bei additiver Schreibweise. Es ist ebenfalls eindeutig bestimmt.

Eine Gruppe heißt *abelsch*, wenn das Kommutativgesetz gilt:

$$a \bullet b = b \bullet a \; \forall \; a,b \in G$$

Beispiel:

Betrachten M_n ($\mathcal{R}$) mit den Verknüpfungen:
φ_1: (A,B) $\mapsto$ AB
φ_2: (A,B) $\mapsto$ AB + BA
φ_3: (A,B) $\mapsto$ AB - BA
Dann ist φ_1 assoziativ, nicht kommutativ
φ_2 nicht assoziativ, kommutativ
φ_3 nicht assoziativ, nicht kommutativ.

Einige Standardgruppen
$\mathcal{Z}$, $\mathcal{Q}$, $\mathcal{R}$ mit + (abelsch)
$\mathcal{Q}_+$ $\mathcal{R}_+$ $\mathcal{Q}\setminus\{0\}$, $\mathcal{R}\setminus\{0\}$, $\mathcal{C}\setminus\{0\}$ mit $\bullet$ (abelsch)
$\mathcal{Z}_n = \mathcal{Z}/\mathrm{n}\mathcal{Z}$ mit + (abelsch), sogennante *Zyklische Gruppe* der Ordnung n, ist Menge {0,1,...,n-1} mit Addition modulo n,
$\mathcal{Z}_p = \mathcal{Z}/\mathrm{p}\mathcal{Z}$ mit $\bullet$ (abelsch) genau dann, wenn p eine Primzahl ist
Kleinsche Vierergruppe {e,a,b,c}

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

(abelsch)

S_n ist Menge aller Permutationen von n Elementen (z.B. der Zahlen 1, ...,n) mit Hintereinanderausführung. Sie heißt *symmetrische Gruppe vom Grad n.*
S_n abelsch $\Leftrightarrow$ n = 1,2,3

$A_n = \{\sigma \in S_n: \operatorname{sgn}\sigma = 1\}$ *alternierende Gruppe vom Grad n*
$M_n(\mathcal{R})$, $M_n(\mathcal{C})$ mit + (abelsch)
$\mathrm{GL}(\mathrm{n},\mathcal{R}) = \{\mathrm{A} \in M_n(\mathcal{R})$: det A $\neq 0\}$ allgemeine lineare Gruppe über $\mathcal{R}$,$\bullet$
$\mathrm{GL}(\mathrm{n},\mathcal{C}) = \{\mathrm{A} \in M_n(\mathcal{C})$: det A $\neq 0$ $\}$ allgemeine lineare Gruppe über $\mathcal{C}$ mit $\bullet$
$\mathrm{SL}(\mathrm{n},\mathcal{R}) = \{\mathrm{A} \in M_n(\mathcal{R})$: det A $= 1\}$ spezielle lineare Gruppe $\mathcal{R}$,$\bullet$
$\mathrm{SL}(\mathrm{n},\mathcal{C}) = \{\mathrm{A} \in M_n(\mathcal{C})$: det A $= 1\}$ spezielle lineare Gruppe $\mathcal{C}$,$\bullet$
$\mathrm{O}(\mathrm{n}) = \{\mathrm{A} \in M_n(\mathcal{R})$: $\mathrm{AA}^T = \mathrm{I}\}$ orthogonale Gruppe Grad n mit $\bullet$
$\mathrm{U}(\mathrm{n}) = \{\mathrm{A} \in M_n(\mathcal{C})$: $\mathrm{AA}^* = \mathrm{I}\}$ unitäre Gruppe Grad n mit $\bullet$
$\mathrm{SO}(\mathrm{n}) = \{\mathrm{A} \in \mathrm{O}(\mathrm{n})$: det A $= 1\}$ spezielle orthogonale Gruppe mit $\bullet$
$\mathrm{SU}(\mathrm{n}) = \{\mathrm{A} \in \mathrm{U}(\mathrm{n})$: det A $= 1\}$ spezielle unitäre Gruppe mit $\bullet$

Zur Erinnerung: SO(3) Menge aller Achsendrehungen (Euler)

Eine Gruppe heißt *endlich*, wenn die Anzahl der Elemente endlich ist. Diese Anzahl wird dann mit |G| bezeichnet und *Ordnung von G* genannt. Haben also z.B.
$|\mathcal{Z}_n = \mathrm{n}$, $|\mathrm{V}_4| = 4$, $|S_n| = \mathrm{n}!$, $|A_n| = \frac{n!}{2}$

Seien G_1 und G_2 Gruppen und f: $G_1 \to G_2$ ein Homomorphismus. Man nennt

$$\text{Im f:= f}(G_1) = \{\text{b} \in G_2\text{: } \exists \text{ a} \in G_1\text{: f(a) = b}\}$$

und

$$\ker \text{f:= f}^{-1}(\{\text{e}\}) = \{\text{a} \in G_1\text{: f(a) = e}\}$$

das *Bild* und den *Kern* von f. Folgende Aussagen lassen sich leicht überprüfen:
(a) f(e) = e [genauer: f(e_1)=e_2;
(b) [f(a)]$^{-1}$ = f(a^{-1}) $\forall$ a $\in G_1$;
(c) Im f ist eine Gruppe mit der Verknüpfung aus G_2;
(d) ker f ist eine Gruppe mit der Verknüpfung aus G_1;
(e) Wenn f: $G_1 \to G_2$ ein Isomorphismus ist, dann ist auch f^{-1}: $G_2 \to G_1$ ein Isomorphismus;
(f) Abbildung f ist genau dann ein Isomorphismus, wenn gilt:

$$\text{Im f} = G_2 \text{ und ker f} = \{\text{e}\}$$

4.1.2 Untergruppen und Normalteiler

Schreiben im Folgenden Gruppen multiplikativ.

Definition 1 *Eine Teilmenge U einer Gruppe G heißt Untergruppe von G, wenn U mit der Verknüpfung aus G selbst eine Gruppe ist.*

Es ist leicht zu sehen, dass folgende Bedingungen äquivalent sind:
(1) U ist Untergruppe von G;
(2) e $\in$ U; a,b $\in$ U $\Rightarrow$ ab $\in$ U; a $\in$ U $\Rightarrow$ a^{-1} $\in$ U;
(3) a,b $\in$ U $\Rightarrow$ ab^{-1} $\in$ U.

1. Beispiel
Jede Gruppe hat die *trivialen Untergruppen* {e} und G.

2. Beispiel
Die Untergruppen von ($\mathcal{Z}$,+) sind (m$\mathcal{Z}$,+) = {mj: j $\in \mathcal{Z}$} mit + mit m $\in$ {0,1,...}.

3. Beispiel
A_n ist Untergruppe von S_n. $S_n \backslash A_n$ hingegen ist keine Untergruppe von S_n.
S_n heißen *Permutationsgruppen*. Satz von Cayley (Satz 1/2.4.) besagt, dass jede endliche Gruppe zu einer Permutationsgruppe isomorph ist.

4. Beispiel
$\text{E} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\text{I} = \begin{pmatrix} i & 0 \\ 0 & -1 \end{pmatrix}$, $\text{J} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $\text{K} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$.

Dann ist Q = {E,-E,I,-I,J,-J,K,-K} eine Untergruppe von SL(2,$\mathcal{C}$), die sogenannte *Quaternionengruppe*.

5. Beispiel

Sei G eine Gruppe. Bezeichnen mit S(G) die Menge aller bijektiven Abbildungen von G auf sich selbst. Dann ist S(G) eine Gruppe bezüglich der Hintereinanderausführung. Sei Aut(G) die Menge aller Isomorphismen von G auf sich selbst (sogenannter *Automorphismus*). Dann ist Aut(G) eine Untergruppe von S(G).
Ist a ∈ G, so wird durch

$$\varphi_a\text{: G} \to \text{G, g} \mapsto \text{aga}^{-1}$$

ein Automorphismus definiert:
$\varphi_a(gh) = agha^{-1} = ag\underbrace{a^{-1}a}_{=e}ha^{-1} = \varphi_a(g)\varphi_a(h)$,
$g = a(a^{-1}ga)a^{-1} = \varphi_a(a^{-1}ga)$,
$varphi_a(g) = e \Rightarrow aga^{-1} = e \Rightarrow g = aea^{-1} = e$.

Solche Automorphismen heißen *innere Automorphismen*.
Die Menge aller inneren Automorphismen ist eine Untergruppe von Aut(G).

6. Beispiel
Ist f: $G_1 \to G_2$ ein Homomorphismus, so sind Im f $\subset G_2$ und ker f $\subset G_1$ Untergruppen.
Die Gruppe aller inneren Automorphismen wird mit Int(G) bezeichnet.

Definition 2 *Sei G eine Gruppe und U eine Untergruppe. Eine Menge der Form aU:={au: u ∈ U} (a ∈ G) heißt Linksnebenklasse und eine Menge der Form Ua:={ua: u ∈ U} (a ∈ G) wird Rechtsnebenklasse genannt.*

7. Beispiel
Sei G = X ein linearer Raum. Die Gruppenoperation sei die Addition in X.
Dann ist jeder Unterraum von X eine Untergruppe von G.
Die Untermannigfaltigkeiten a + U := {a + u: u ∈ U} sind dann gerade die Linksnebenklassen und wegen der Kommutativität der Addition ist a + U = U + a, d.h. die Links- und Rechtsnebenklassen stimmen überein.

8. Beispiel
Sei G = $\mathcal{Z}$ mit + und U = $3\mathcal{Z}$. Haben

$$\text{a + U} = \{\text{a + 3k: k} \in \mathcal{Z}\} = \text{U + a},$$

das heißt $2 + 3\mathcal{Z} = \{\ldots,\text{-}1,2,5,8,\ldots\} = \text{-}1 + 3\mathcal{Z} = \ldots = 5 + 3\mathcal{Z} = \ldots$
$1 + 3\mathcal{Z} = \{\ldots,\text{-}2,1,4,7,\ldots\} = \text{-}2 + 3\mathcal{Z} = 1 + 3\mathcal{Z} = \ldots$
$0 + 3\mathcal{Z} = \{\ldots,\text{-}3,0,3,6,\ldots\} = \text{-}3 + 3\mathcal{Z} = \ldots$
→ es gibt keine anderen Fälle

Erhalten $\mathcal{Z} = 3\mathcal{Z} \cup (1 + 3\mathcal{Z}) \cup (2 + 3\mathcal{Z})$
(disjunkte Zerlegung).

9. Beispiel
Sei G = S_3:

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},$$
$$\sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \sigma_6 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

Dann ist U = $\{\sigma_1,\sigma_2\}$ eine Untergruppe:

	σ_1	σ_2
σ_1	σ_1	σ_2
σ_2	σ_2	σ_1

Linksnebenklassen:
$\sigma_1 U = \{\sigma_1,\sigma_2\}$, $\sigma_2 U = \{\sigma_1,\sigma_2\}$,
$\sigma_3 U = \{\sigma_3,\sigma_3\sigma_2\} = \{\sigma_3,\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}\} = \{\sigma_3,\sigma_4\}$,
$\sigma_4 U = \{\sigma_4,\sigma_4\sigma_1\} = \{\sigma_4,\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}\} = \{\sigma_4,\sigma_3\}$,
$\sigma_5 U = \{\sigma_5,\sigma_5\sigma_2\} = \{\sigma_5,\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}\} = \{\sigma_5,\sigma_6\}$,
$\sigma_6 U = \{\sigma_6,\sigma_6\sigma_1\} = \{\sigma_6,\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}\} = \{\sigma_6,\sigma_5\}$.

Rechtsnebenklassen:
$U\sigma_1 = \{\sigma_1,\sigma_2\}$,
$U\sigma_2 = \{\sigma_2,\sigma_2\sigma_2\} = \{\sigma_1,\sigma_2\}$,
$U\sigma_3 = \{\sigma_3,\sigma_2\sigma_3\} = \{\sigma_3,\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}\} = \{\sigma_3,\sigma_5\}$,
$U\sigma_4 = \{\sigma_4,\sigma_2\sigma_4\} = \{\sigma_4,\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}\} = \{\sigma_4,\sigma_6\}$,
$U\sigma_5 = \{\sigma_5,\sigma_2\sigma_5\} = \{\sigma_5,\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}\} = \{\sigma_5,\sigma_3\}$,
$U\sigma_6 = \{\sigma_6,\sigma_2,\sigma_6\} = \{\sigma_6,\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}\} = \{\sigma_6,\sigma_4\}$.

Menge der Linksnebenklassen $\neq$ Menge der Rechtsnebenklassen

Beide ergeben aber disjunkte Zerlegungen von S_3.

Satz 1 *Sei G eine Gruppe und U eine Untergruppe. Dann sind folgende Aussagen äquivalent:*
(i) $aU = bU$;
(ii) $b \in aU$;
(iii) $a^{-1}b \in U$.
Ebenso sind folgende Aussagen äquivalent:
(i) $Ua = Ub$;
(ii) $b \in Ua$;
(iii) $ab^{-1} \in U$.

Beweis:
Betrachten nur Linksnebenklassen.
(i) $\Rightarrow$ (ii) Haben b = be $\in$ bU = aU.
(ii) $\Rightarrow$ (iii)
Wegen b $\in$ aU ist b = au mit u $\in$ U. Dies ergibt $a^{-1}b = u \in U$.
(iii) $\Rightarrow$ (i)
Sei g $\in$ aU, d.h. g = au mit u $\in$ U. Dann ist $a^{-1}b = v$, d.h. b = av nach Voraussetzung

und somit g = $\underbrace{av}_{b}\underbrace{v^{-1}u}_{\in U}$ = bw mit w = $v^{-1}u \in U$. Also ist g ∈ bU.
Sei nun g ∈ bU, d.h. g = bu mit u ∈ U. Nach Voraussetzung ist b = av mit v ∈ U.
Also ist g = avu ∈ aU. #

Folgerung 1: Sei G eine Gruppe und U eine Untergruppe. Durch a $\overset{l}{\sim}$ b ⇔ aU = bU ist eine Äquivalenrelation auf G gegeben, deren Äquivalenzklassen die Linksnebenklassen von U sind. Folglich bilden die Linksnebenklassen eine disjunkte Zerlegung von G.
Analog ist durch a $\overset{r}{\sim}$ b ⇔ Ua = Ub eine Äquivalenzrelation auf G gegeben, deren Äquivalenzklassen die Rechtsnebenklassen sind. Und die Rechtsnebenklassen bilden somit eine disjunkte Zerlegung von G.

Beweis:
Müssen lediglich zeigen, dass

$$\{b \in G: b \overset{l}{\sim} a\} = aU$$

gilt. Es ist aber b $\overset{l}{\sim}$ a ⇔ b ∈ aU. #
Wie Beispiel 9 zeigt, müssen beide Zerlegungen nicht übereinstimmen.

Bemerkung:
Die Abbildung aU ↦ Ua^{-1} ist eine Bijektion der Menge der Linksnebenklassen auf die Menge der Rechtsnebenklassen.

Beweis:
Die Abbildung ist korrekt definiert:
aU = bU ⇒ $a^{-1}b \in U \Rightarrow a^{-1}(b^{-1})^{-1} \in U \Rightarrow Ua^{-1} = Ub^{-1}$.
Surjektivität trivial: $Ub = U(b^{-1})^{-1}$.
Injektivität: $Ua^{-1} = Ub^{-1} \Rightarrow a^{-1}(b^{-1})^{-1} \in U \Rightarrow a^{-1}b \in U \Rightarrow aU = bU$. #

Folgerung 2: Satz von Lagrange
Ist G eine endliche Grupe und U eine Untergruppe von G, so ist |G| durch |U| teilbar.

Beweis:
Für beliebiges a ∈ G ist die Abbildung

$$f_a: U \to aU, u \mapsto au$$

bijektiv.
Dies bedeutet, dass |aU| = |U| für alle a ∈ G ist. Nach Folgerung 1 ist also
|G| = |U| * Anzahl der Linksnebenklassen, d.h. |G| ist durch |U| teilbar. #

Folgerung 3: Ist p eine Primzahl und G eine Gruppe der Ordnung p, so ist G zu $(\mathcal{Z}_p,+)$ isomorph.

Beweis:
Wählen a ∈ G\{e}. Unter den Elementen a,a^2,a^3,... (a^n:=$\underbrace{a * \ldots * a}_{n\ mal}$)
müssen zwei gleiche sein, d.h. $a^m = a^n$ und damit a^{m-n} = e (m > n).
Es gibt also ein k ≥ 2 mit a^k = e.

O.B.d.A. sei k die kleinste Zahl mit $a^k = e$. Dann ist $U = \{e,a,\ldots,a^{k-1}\}$ eine Untergruppe von G:
$e \in U$;
$a^r \in U$, $a^s \in U \Rightarrow a^r a^s = a^{r+s} = a^{r+s(mod\ k)} \in U$;
$a^r \in U \Rightarrow a^{k-r}a^r = a^r a^{k-r} = a^k = e$.

Nach Folgerung 2 ist $k \geq 2$ ein Teiler von p, d.h. k = p (da p Primzahl).
Also ist $U = G = \{e,a,a^2,\ldots,a^{p-1}\}$ und
$U \to \mathbb{Z}_p$, $a^r \mapsto r$ ist der gewünschte Isomorphismus. #

Satz 2 *Sei G eine Gruppe und U eine Untergruppe. Dann sind folgende Bedingungen äquivalent:*
(i) Die Menge der Linksnebenklassen von U ist gleich der Menge der Rechtsnebenklassen von U;
(ii) $aU = Ua\ \forall\ a \in G$;
(iii) $aUa^{-1} \subset U\ \forall\ a \in G$;
(iv) $aUa^{-1} = U\ \forall\ a \in G$;
(v) $\varphi(U) \subset U\ \forall\ \varphi \in Int(G)$;
(vi) $\varphi(U) = U\ \forall\ \varphi \in Int(G)$.

Beweis:
(i) $\Rightarrow$ (ii)
Sei $a \in G$. Nach (i) existiert ein $b \in G$ mit $aU = Ub$. Haben dann $a = ae \in Ub$, d.h. es existiert ein $u \in U$ mit $a = ub$. Es folgt $ab^{-1} = u \in U$, und Satz 1 ergibt $Ua = Ub$. Also ist $aU = Ua$.
(ii) $\Rightarrow$ (iii)
Sei $u \in U$ und $a \in G$. Müssen zeigen, dass $aua^{-1} \in U$.
Wegen $au \in aU = Ua$ existiert $v \in U$ mit $au = va$. Also ist $aua^{-1} = vaa^{-1} = v \in U$.
(iii) $\Rightarrow$ (iv)
Sei $u \in U$. Dann ist $u = a\underbrace{(a^{-1}ua)}_{\in U}a^{-1}$, d.h. $u \in aUa^{-1}$.
(iv) $\Rightarrow$ (i)
Zeigen, dass sogar $aU = Ua\ \forall\ a \in G$ ist. Sei $x = au \in aU$. Dann ist $x = \underbrace{aua^{-1}}_{\in U}a = va$ mit $v \in U$, d.h. $x \in Ua$.
Ist $x = ua \in Ua$, so ist $x = a\underbrace{a^{-1}ua}_{\in U} = av$ mit $v \in U$, d.h. $x \in aU$.
(iii) $\Rightarrow$ (v) und (iv) $\Rightarrow$ (vi) trivial. #

Definition 3 *Eine Untergruppe U einer Gruppe G heißt Normalteiler, wenn U einer (und damit jeder) der äquivalenten Bedingungen aus Satz 2 genügt.*

10. Beispiel
Jede Gruppe G hat die beiden trivialen Normalteiler {e} und G ($a\{e\} = a = \{e\}a$, $aG = G = Ga$).
In einer abelschen Gruppe sind alle Untergruppen Normalteiler.
Die Untergrupe von S_3 aus Beispiel 9 ist kein Normalteiler.

Sei f: G $\to$ H ein Homomorphismus. Dann ist das Urbilde jedes Normalteilers von

H ein Normalteiler von G, und wenn f surjektiv ist, dann ist auch das Bild jedes Normalteilers von G ein Normalteiler von H.

Beweis:
Sei V $\subset$ H ein Normalteiler. Seien a $\in$ G, u $\in$ f^{-1}(V). Müssen zeigen, dass aua^{-1} $\in$ f^{-1}(V):
Haben f(aua^{-1}) = f(a)$\underbrace{f(u)}_{\in V}$(f(a))$^{-1}$ $\in$ V,
d.h. aua^{-1} $\in$ f^{-1}(V).
Sei U $\subset$ G ein Normalteiler. Wählen b $\in$ H, V $\in$ f(U). Müssen zeigen, dass bvb^{-1} $\in$ f(U).
Haben b = f(a) mit a $\in$ G wegen Surjektivität und v = f(u) mit u $\in$ U.
Also bvb^{-1} = f(a)f(u)(f(a))$^{-1}$ = f($\underbrace{aua^{-1}}_{=v \in V}$) = f(v) mit v $\in$ U. Das heißt bvb^{-1} $\in$ f(U).
#

Insbesondere ist ker f = f^{-1}({e}) stets ein Normalteiler.

Zur Erinnerung:
X linearer Raum, M Unterraum. Der Faktorraum X/M war die Menge aller Untermannigfaltigkeiten der Form x + M (= Menge der Äquivalenzklassen der Äquivalenzrelation x $\sim$ y $\Leftrightarrow$ x - y $\in$ M) mit den Operationen

$$\alpha(x + M) := \alpha x + M;$$
$$(x + M) + (y + M) := (x + y) + M.$$

Wollen das Analogon für Gruppen.
Sei also G eine Gruppe und U eine Untergruppe. Die Linksnebenklassen aU sind die Äquivalenzklassen der Äquivalenzrelation a $\sim$ b $\Leftrightarrow$ a^{-1}b $\in$ U (Satz 1).
Definieren eine Multiplikation in der Menge der Linksnebenklassen wie folgt:

$$(aU)(bU) := (ab)U.$$

Die Definition ist genau dann korrekt, wenn gilt:
$a_1U = a_2U$, $b_1U = b_2U \Rightarrow (a_1b_1)U = (a_2b_2)U$.

Satz 3 *Die obige Definition ist genau dann korrekt, wenn U ein Normalteiler ist.*

Beweis:
Sei U ein Normalteiler. Sei weiter $a_1U = a_2U$, $b_1U = b_2U$. Nach Satz 1 ist dann $a_1^{-1}a_2 \in U$, $b_1^{-1}b_2 \in U$. Müssen zeigen, dass $(a_1b_1)U = (a_2b_2)U$ ist, d.h. (Satz 1) $(a_1b_1)^{-1}(a_2b_2) \in U$ ist.
Haben

$$(a_1b_1)^{-1}(a_2b_2) = b_1^{-1}a_1^{-1}a_2b_2 = \underbrace{b_1^{-1}\underbrace{a_1^{-1}a_2}_{\in U}a_2b_1}_{\in U}\underbrace{b_1^{-1}b_2}_{\in U} \in U.$$

Sei die obige Definition korrekt. Wählen a $\in$ G, u $\in$ U. Müssen zeigen, dass aua^{-1} $\in$ U, d.h. aua^{-1}U = U ist (Satz 1).
Haben auU = {auv: v $\in$ U } = {aw: w $\in$ U} = aU, a^{-1}U = a^{-1}U, und da die Definition korrekt ist, folgt aua^{-1}U = aa^{-1}U = U. #

Definition 4 *Ist G eine Gruppe und U ⊂ G ein Normalteiler, dann wird die Menge der Linksnebenklassen (= Menge aller Rechtsnebenklassen) mit der Operation (au)(bU):=(ab)U bzw. (Ua)(Ub):=U(ab) Faktorgruppe von G nach U genannt und mit G/U bezeichnet.*

Satz 4 *G/U ist eine Gruppe.*

Beweis:
Haben
(aU bU)cU = (ab)U cU = (ab)cU = a(bc)U = (aU)(bcU) = aU(bU cU),
das Einselement ist U = eU,
das zu aU inverse Element ist $a^{-1}U$. #

Der folgende Satz ist in mehrerer Hinsicht von Bedeutung. Er hilft erstens, Faktorgruppen zu „identifizieren“ (mit bekannten Gruppen). Er zeigt zweitens, dass eine Bijektion zwischen den Normalteilern einer Gruppe und den homomorphen Bildern (= Fotografien) der Gruppe besteht.

Satz 5 *„Wundersatz“*
Sind G und H Gruppen und ist f: G → H ein Homomorphismus, so ist ker f ein Normalteiler von G und G/ker f ≅ Im f.
Umgekehrt, ist G eine Gruppe und U ein Normalteiler, so ist die Abbildung π: G → G/U, a ↦ aU ein Homomorphismus mit ker π = U und Im π = G/U.

Beweis:
Haben bereits gezeigt, dass ker f stets ein Normalteiler ist.
Zeigen, dass die Abbildung φ: G/ker f → Im f, a ker f ↦ f(a) ein Isomorphismus ist.
φ korrekt definiert:
a ker f = b ker f ⇒ $a^{-1}b \in$ ker f ⇒ $f(a^{-1}b) = e \Rightarrow f(a^{-1})f(b) = e \Rightarrow f(a) = f(b)$

Homomorphismus:
φ((a ker f)(b ker f)) = φ((ab)ker f) = f(ab) = f(a)f(b) = φ(a ker f)φ(b ker f)

Injektivität:
φ(a ker f) = φ(b ker f) ⇒ f(a) = f(b) ⇒ $f(a^{-1}b) = e \Rightarrow a^{-1}b \in$ ker f ⇒ a ker f = b ker f

Surjektivität: trivial.

π ist Homomorphismus:
$\pi(ab) = (ab)U = (aU)(bU) = \pi(a)\pi(b)$.
ker π: $\pi(a) = e \Leftrightarrow aU = U \Leftrightarrow a \in U$.
Im π: π ist offenbar surjektiv. #

11. Beispiel
In $\mathcal{Z}$ mit + sind die Untergruppen = Normalteiler die Gruppen $n\mathcal{Z}$ (n=0,1,2,...). Haben

$\mathcal{Z}/0\mathcal{Z} = \mathcal{Z}/\{0\} \cong \mathcal{Z}$
$\mathcal{Z}/1\mathcal{Z} = \mathcal{Z}/\mathcal{Z} \cong \{e\} = \{0\}$ Sei n $\geq$ 2. Was ist $\mathcal{Z}/n\mathcal{Z}$? Betrachten die Abbildung

$$f: \mathcal{Z} \to \mathcal{Z}_n,\ a \mapsto a(\text{mod } n)$$

Dies ist ein Homomorphismus mit der Eigenschaft ker f = n$\mathcal{Z}$. Satz 5 liefert also $\mathcal{Z}/n\mathcal{Z}$ $\cong \mathcal{Z}_n$.

12. Beispiel
Sei $\mathcal{R}$ mit Addition. Dann ist $\mathcal{Z}$ eine Untergruppe = Normalteiler. Die Abbildung

$$\text{f: } \mathcal{R} \to \mathcal{T},\ \text{x} \mapsto e^{2\pi i x}$$

ist ein Homomorphismus von $\mathcal{R}$ auf $\mathcal{T}$ mit Multiplikation.
ker f = $\mathcal{Z}$, d.h. erhalten $\mathcal{R}/\mathcal{Z} \cong \mathcal{T}$.
Analog ist $\mathcal{R}^n/\mathcal{Z}^n \cong \mathcal{T}^n$ (= n-dimensionaler Torus).

13. Beispiel
Die Abbildung f: O(n) $\to$ {-1,1}, A $\mapsto$ det A
ist ein surjektiver Homomorphismus. Haben ker f = SO(n) und damit ist SO(n) ein Normalteiler von O(n) und O(n)/SO(n) $\cong$ {-1,1} $\cong \mathcal{Z}_2$.

14. Beispiel
Eine Gruppe heißt *einfach*, wenn sie außer {e} und sich selbst keine Normalteiler hat. Nach Satz 5 ist eine endliche Gruppe G genau dann einfach, wenn es keinen Homomorphismus f: G $\to$ H mit $1 < |\text{H}| <$ G gibt. Man kann zeigen:
- $\mathcal{Z}_n$ ist einfach $\Leftrightarrow$ n Primzahl
- A_5,A_6,A_7, ... sind einfach
- es gibt noch 16 weitere solche „unendlichen Serien“ von einfachen Gruppen

Mathieu einfache Gruppe mit Ordnung 7920, noch vier weitere Gruppen
Janko einfache Gruppe der Ordnung 175.560
... es gibt insgesamt 26 einfache sporadische Gruppen.

4.1.3 Endliche abelsche Gruppen

Kommen nun zum Klassifizierungsproblem für endliche Gruppen. Dieses ist wie folgt: Gegeben sei eine natürliche Zahl n. Wie viele nichtisomorphe Gruppen der Ordnung n gibt es? Man erstelle ein Verzeichnis von Gruppen, sodass jede Gruppe der Ordnung n zu genau einer Gruppe aus dem Verzeichnis isomorph ist.

Kennen den Anfang des Gesamtverzeichnisses:

n	Anzahl	Verzeichnis
1	1	{e}
2	1	$\mathcal{Z}_2$
3	1	$\mathcal{Z}_3$
4	2	$\mathcal{Z}_4$,V_4
5	1	$\mathcal{Z}_5$
6	2	$\mathcal{Z}_6$,S_3
7	1	$\mathcal{Z}_7$

Das Klassifizierungsproblem ist vollständig gelöst für endliche abelsche Gruppen.
Seien G_1,...,G_m Gruppen in additiver Schreibweise. Das *direkte Produkt* $G_1 \times \ldots G_m$ ist die Menge $G_1 \times \ldots G_m$ mit der Addition

$$(g_1, \ldots, g_m) \times (h_1, \ldots, h_m) := (g_1 + h_1, \ldots, g_m + h_m).$$

Dies ist wieder eine Gruppe.

1. Beispiel
$\mathcal{Z}_2 \times \mathcal{Z}_2$
e = (0,0), a = (0,1), b = (1,0), c = (1,1)

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Erhalten also $\mathcal{Z} \times \mathcal{Z} \cong V_4$.

2. Beispiel
$\mathcal{Z}_2 \times \mathcal{Z}_3$
a_0 = (0,0), a_1 = (1,1), a_2 = (0,2), a_3 = (1,0), a_4 = (0,1), a_5 = (1,2)
Es ist leicht zu sehen, dass die Abbildung $\mathcal{Z}_6 \to \mathcal{Z}_2 \times \mathcal{Z}_3$, k $\mapsto a_k$ ein Isomorphismus ist.
Also $\mathcal{Z}_2 \times \mathcal{Z}_3 \cong \mathcal{Z}_6$.

Allgemein: (a,b) = 1 $\Rightarrow \mathcal{Z}_a \times \mathcal{Z}_b \cong \mathcal{Z}_{ab}$
Beweis:
Die Abbildung $\mathcal{Z}_{ab} \to \mathcal{Z}_a \times \mathcal{Z}_b$, k $\mapsto$ (k mod a, k mod b)
ist ein Homomorphismus. Zeigen die Surjektivität (dann folgt automatisch die Injektivität):
Sei 0 *leq* x $\leq$ a - 1, 0 $\leq$ y $\leq$ b - 1,
d.h. (x,y) $\in \mathcal{Z}_a \times \mathcal{Z}_b$. Wegen (a,b) = 1 gibt es r,s $\in \mathcal{Z}$ mit ar + bs = 1. Setzen k = yar + xbs(mod ab), dann ist 0 $\leq$ k $\leq$ ab - 1.
Haben k (mod a) = xbs (mod a) = x(1-ar)(mod a) = x(mod a) = x,
analog zeigt man k(mod b) = y. #

Sei n eine natürliche Zahl. Bezeichnen mit P(n) die Menge aller Zerlegungen (Partitionen) von n in der Form n = $\gamma_1 + \gamma_2 + \ldots + \gamma_m$, $\gamma_1 \geq \gamma_2 \geq \ldots \geq \gamma_m \geq 1$ und mit |P(n)| die Anzahl der Elemente von P(n).
P(1) = {(1)}, |P(1)| = 1,
P(2) = {(2),(1,1)}, |P(2)| = 2,
P(3) = {(3),(2,1),(1,1,1)}, |P(3)| = 3,
P(4) = {(4),(3,1),(2,2),(2,1,1,),(1,1,1,1)}, |P(4)| = 5,
P(5) = {(5),(4,1),(3,2),(3,1,1),(2,2,1),(2,1,1,1),(1,1,1,1,1)}, |P(5)| = 7.

Satz 1 *Hauptsatz für endliche abelsche Gruppen*
Sei n $\leq$ 2 eine natürliche Zahl und n = $p^\alpha q^\beta \ldots$ die Primfaktorenzerlegung. Dann gibt es genau |P(α)| |P(β)| ... nichtisomorphe abelsche Gruppen der Ordnung n.
Jede solche Gruppe ist zu genau einer Gruppe der Form
$\mathcal{Z}_{p^{\alpha_1}} \times \ldots \times \mathcal{Z}_{p^{\alpha_k}} \times \mathcal{Z}_{q^{\beta_1}} \times \ldots \times \mathcal{Z}_{q^{\beta_l}} \times \ldots$
(α_1,...,α_k) $\in$ P(α), (β_1,...,β_l) $\in$ P(β) ... isomorph.

Beweis:
Der Beweis ist elementar, aber länglich. Verweisen daher auf die Literatur. #

3. Beispiel
n = 200 = $2^3 * 5^2$
P(3) = {(3),(2,1),(1,1,1)}, P(2) = {(2),(1,1)}
$\mathcal{Z}_{2^3} \times \mathcal{Z}_{5^2} \cong \mathcal{Z}_{200}$
$\mathcal{Z}_{2^3} \times \mathcal{Z}_5 \times \mathcal{Z}_5 \cong \mathcal{Z}_{2^3} \times \mathcal{Z}_5 \times \mathcal{Z}_5$
$\mathcal{Z}_{2^2} \times \mathcal{Z}_2 \times \mathcal{Z}_{5^2} \cong \mathcal{Z}_4 \times \mathcal{Z}_{50} \cong \mathcal{Z}_2 \times \mathcal{Z}_{100}$
$\mathcal{Z}_{2^2} \times \mathcal{Z}_2 \times \mathcal{Z}_5 \times \mathcal{Z}_5$
$\mathcal{Z}_2 \times \mathcal{Z}_2 \times \mathcal{Z}_2 \times \mathcal{Z}_{5^2}$
$\mathcal{Z}_2 \times \mathcal{Z}_2 \times \mathcal{Z}_2 \times \mathcal{Z}_5 \times \mathcal{Z}_5$

Kommen noch einmal zu nichtabelschen Gruppen.

4. Beispiel: *Diedergruppen*
Betrachten für n $\geq$ 3 ein reguläres n-Eck in der Ebene mit Zentrum im Ursprung.
Bezeichnen mit D_n die Menge aller orthogonalen Abbildungen der Ebene (= Abbildungen aus O(2) = längen- und winkelerhaltene Abbildungen), die das n-Eck in sich überführen.
Diese Abbildungen bilden eine Gruppe bezüglich der Hintereinanderausführung, die sogenannte *Diedergruppe* D_n.

Sei δ_k die Drehung um Winkel $k * \frac{2\pi}{n}$ und seien δ_1,...,δ_n die Spiegelungen an den Symmetrieachsen des n-Ecks. Dann ist

$$D_n = \{e, \delta_1, \ldots, \delta_{n-1}, \sigma_1, \ldots, \sigma_n\}.$$

Es gibt weitere Realisierungen von D_n. Ist z.B. σ irgendeine der obigen Spiegelungen und $\delta = \delta_1$, so ist

$$D_n = \{e, \delta, \ldots, \delta_{n-1}, \sigma, \sigma\delta, \ldots, \sigma\delta_{n-1}\}$$

mit σ^2 = e, δ^n = e, $\sigma\delta\sigma = \delta_{n-1}$.
Eine weitere Interpretation ist wie folgt:
Sei Δ_n eine Doppelpyramide (= Dieder) mit einem regelmäßigen n-Eck als Basis. Dann kann D_n als Gruppe aller Achsendrehungen im $\mathcal{R}^3$ aufgefasst werden, die Δ_n invariant lassen.
Schließlich kann D_n als die Untergruppe von S_n aufgefasst werden, die $\{e, \delta, \ldots, \delta^{n-1}, \sigma, \sigma\delta, \ldots, \sigma\delta^{n-1}\}$ mit

$$\delta = \begin{pmatrix} 1 & 2 & \ldots & n-1 & n \\ 2 & 3 & \ldots & n & 1 \end{pmatrix}, \sigma = \begin{pmatrix} 1 & 2 & \ldots & n-1 & n \\ 1 & n & \ldots & 3 & 2 \end{pmatrix}$$

ist.

Satz 2 *Ist p eine Primzahl, so gibt es bis auf Isomorphie genau zwei Gruppen der Ordnung 2p:*
$\mathcal{Z}_{2p}$ (abelsch) und D_p (nichtabelsch).

Beweis: Literatur. #

Kennen damit alle Gruppen der Ordnung 2 * 3 = 6, 2 * 5 = 10, 2 * 7 = 14, ...

Der Atlas der nichtabelschen Gruppen beginnt wie folgt:

n	Anzahl	Verzeichnis
1	0	
2	0	
3	0	
4	0	
5	0	
6	1	D_3
7	0	
8	2	D_4,Q (Quaternionengruppe)
9	0	
10	1	D_5
11	0	
12	3	D_6,A_4, noch eine
13	0	
14	1	D_7
15	0	
16	9	D_8, noch 8 weitere

1999: Ordnung $2^9 = 512$
insgesamt 10.494.213 Gruppen (einschließlich abelsche)
später: Ordnung $2^{10} = 1.024$
insgesamt 49.487.365.422 Gruppen (einschließlich abelsche)

4.1.4 Symmetriegruppen

Betrachten $\mathcal{R}^n$ mit üblichem Skalarprodukt. Bezeichnen mit B(n) die Gruppe aller Abbildungen B: $\mathcal{R}^n \to \mathcal{R}^n$, die Längen invariant lassen. B(n) heißt *Bewegungsgruppe* des $\mathcal{R}^n$.

Satz 1 *Eine Abbildung B: $\mathcal{R}^n \to \mathcal{R}^n$ gehört genau dann zu B(n), wenn sie von der Form Bx = Ax + T mit A ∈ O(n) (= Gruppe der othogonalen linearen Abbildungen) und T ∈ $\mathcal{R}^n$ ist.*

Beweis:
Setzten T = B(0) und Cx = Bx - T. Zeigen C ∈ O(n). Für x,y ∈ $\mathcal{R}^n$ ist
(Cx,Cy) = ||Cx|| ||Cy|| cos γ = ||x|| ||y|| cos γ = (x,y).
(dabei ist γ der Winkel zwischen x und y)
Also lässt C Skalarprodukte invariant.
Haben damit
(C(x+y)-Cx-Cy,C(x+y)-Cx-Cy)
= (C(x+y),C(x,y)) - (C(x+y),Cx) - (C(x+y),Cy) - (Cx,C(x+y)) + (Cx,Cx) + (Cx,Cy) - (Cy,C(x+y)) + (Cy,Cx) + (Cy,Cy)
= (x+y,x+y) - (x+y,x) - (x+y,y) - (x,x+y) + (x,x) + (x,y) - (y,x+y) + (y,x) + (y,y)
= (x+y-x-y,x+y-x-y) = (0,0) = 0,
das heißt C(x+y) = Cx + Cy ∀ x,y ∈ $\mathcal{R}^n$.
Analog zeigt man C(λx) = λC(x) ∀ λ ∈ $\mathcal{R}$ ∀ x ∈ $\mathcal{R}^n$.
Somit ist C eine lineare Abbildung, die Skalarprodukte invariant lässt, d.h. Cx = Ax mit A ∈ O(n). #

Eine Abbildung der Form $x \mapsto Ax + T$ wird üblicherweise mit (A|T) bezeichnet (sogenanntes Seitz-Symbol).
Haben also

$$B(n) = \{(A|T): A \in O(n), T \in \mathcal{R}^n\}.$$

Man definiert

$$SB(n) = \{(A|T): A \in SO(n), T \in \mathcal{R}^n\}$$

und nennt dies die *eigentliche Bewegungsgruppe* des $\mathcal{R}^n$.
Sei $M \subset \mathcal{R}^n$ eine Menge. Die *Symmetriegruppe* und die *eigentliche Symmetriegruppe* von M sind definiert als

$$B_M(n) = \{B \in B(n): B(M) = M\},$$
$$SB_M(n) = \{B \in SB(n): B(M) = M\}.$$

Für (A|T) nennt man A den *Drehanteil* und T den *Translationsanteil*. Haben
$(A|T) = (B|S) \Leftrightarrow Ax + T = Bx + S \; \forall \, x \in \mathcal{R}^n \Leftrightarrow T = S, A = B$.
Haben auch
$Ax + T = x_0 + A(x - x_0) + \underbrace{T - x_0 + Ax_0}_{=T'}$,
d.h. können (A|T) auch als orthogonale Abbildung mit x_0 als Ursprung (= Drehung um x_0) mit anschließender Translation T' vorstellen. Der Translationsanteil hat sich dabei geändert, nicht aber der Drehanteil.

Nehmen im Folgenden an, dass M beschränkt ist.

Satz 2 *Die Abbildung Φ: $B_M(n) \to O(n)$, $(A|T) \mapsto A$ ist ein injektiver Homomorphismus.*

Beweis:
Haben $(B|S)(A|T)s = (B|S)(Ax + T) = B(Ax + T) + S = BAx + BT + S$,
d.h. $(B|S)(A|T) = (BA|BT + S)$.
Also
$\Phi((B|S)(A|T)) = \Phi((BA|BT+S)) = BA = \Phi((B|S))\Phi((A|T))$,
d.h. Φ ist ein Homomorphismus.
Haben $\ker \Phi = \{(A|T) \in B_M(n): \Phi((A|T)) = I\} = \{(A|T) \in B_M(n): A = I\}$,
d.h. $\ker \Phi$ besteht aus den Abbildungen der Form $x \mapsto x + T$, die M invariant lassen.
Da M beschränkt ist, muss $T = 0$ sein. Also
$\ker \Phi = \{(I|0)\} = \{id\}$. #

Nach Satz 5/4.1.2. ist also $B_M(n)/\ker \Phi \cong \text{Im } \Phi \subset O(n)$, d.h. $B_M(n) \cong \text{Im } \Phi \subset O(n)$, d.h. $B_M(n)$ ist zu einer Untergruppe von O(n) isomorph.
Analog ist $SB_M(n)$ zu einer Untergruppe von SO(n) isomorph.
SO(2): Drehungen
O(2): Drehungen und Achsenspiegelungen
SO(3): Achsendrehungen (Euler)
O(3): Achsendrehung, Ebenenspiegelungen

Satz 3 *Für jede Abbildung aus $B_M(n)$ gibt es ein $x_0 \in \mathcal{R}^n$, sodass die Abbildung von der Form $x \mapsto x_0 + A(x - x_0)$ mit $A \in O(n)$ ist.*
Hierbei ist $n = 2$ oder $n = 3$.

Beweisskizze:
Haben Ax + T = x_0 + A(x - x_0) + T - x_0 + Ax_0
und behaupten also, dass ein x_0 mit (I-A)x_0 = T existiert.
Ist I-A invertierbar, so ist x_0 = (I-A)$^{-1}$T die Lösung. Sei also I-A nicht invertierbar, d.h. sei 1 ein Eigenwert von A.
A $\in$ SO(2) hat keine Eigenwerte.
Sei A $\in$ O(2)\ SO(2). Dann ist A eine Achsenspiegelung. Man erhält, dass T parallel zur y-Achse sein muss.
Also (A|T): $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ -y + a \end{pmatrix}$.
Es gilt (A|T): $\begin{pmatrix} 0 \\ \frac{a}{2} \end{pmatrix} + T = \begin{pmatrix} 0 \\ fraca2 \end{pmatrix}$ und $x_0 = \begin{pmatrix} 0 \\ \frac{a}{2} \end{pmatrix}$ ist die Lösung.
Dies ist alles für n = 2. #
Alternativer Beweis (für alle n):
„Konstruieren“ um x den kleinsten Kreis, der M enthält.
r_M(x) = $\sup_{g \in M}$ ||x-y||.
Dann ist M $\subset$ U(x,r_M(x))
Brauchen x aus einer kompakten Kugel K (mit hinreichend großem Radius) zu betrachten.
x $\mapsto$ r_M(x) ist stetig.
Damit existiert ein $x_0 \in$ K mit $r_m(x_0) = \min_{x \in K} r_M$(x).
Der Punkt x_0 ist Mittelpunkt der kleinsten Kugel, die M enthält. Behaupten f(x_0) = x_0.
Sei f(x_0) $\neq$ x_0. Dann ist M $\subset$ U(x_0,r_M(x_0)). Da f $\in$ B(N) ist, ist M = f(M) $\subset$ U(f(x_0),r_M(x_0)).
Dann ist M die Schnittmenge der Kreise um x_0 und f(x_0), aber diese ist nicht enthalten in einer Kugel vom Radius < r_M(x_0). #

Beispiele:
Buchstabe F: $SB_F(2) = \{e\}$, $B_F(2) = \{e\}$, $SB_F(3) = \{e\}$, $B_F(3) = \{e, \text{Spiegelung}\}$
Buchstabe I: $SB_I(2) \cong \mathcal{Z}_2$, $B_I(2) \cong V_4 \cong \mathcal{Z}_2 \times \mathcal{Z}_2$
$SB_I(3)$ größer, $B_I(3)$ noch größer.

Bezeichnen mit C_n (n $\geq$ 3) die eigentliche Symmetriegruppe des regelmäßigen n-Ecks im $\mathcal{R}^2$. Offenbar ist $C_n \cong \mathcal{Z}^n$. Setzen noch $C_1 = \{id\}$ und C_2 sei eigentliche Symmetriegruppe einer Strecke im $\mathcal{R}^2$ ($C_2 \cong \mathcal{Z}_2$).
Wie oben sei D_n (n $\geq$ 3) die volle Symmetriegruppe des regelmäßigen n-Ecks im $\mathcal{R}^2$. Dies ist die n-te Diedergruppe. Definieren D_2 als volle Symmetriegruppe einer Strecke im $\mathcal{R}^2$ ($D_2 \cong V_4 \cong \mathcal{Z}_2 \times \mathcal{Z}_2$). D_1 wird nicht definiert.

Buchstaben:

M	$SB_M(2)$	$B_M(2)$
R	C_1	C_1
A	C_1	C_2
H	C_2	D_2
X	C_4	D_4
O	SO(2)	O(2)

Satz 4 *Das Verzeichnis aller eigentlichen Symmetriegruppen $SB_M(n)$ von endlicher Ordnung (= aller endlichen Untergruppen von SO(n)) ist bis auf Isomorphie wie folgt:*
n = 2 C_1,C_2,C_3,...
n = 3 C_1,C_2,C_3,...
D_2,D_3,...
T,O,Y

Beweis: Literatur. #

Mit anderen Worten: Jedes SB_M(n) endlicher Ordnung ist zu genau einer Gruppe aus dem Verzeichnis isomorph und jede Gruppe aus dem Verzeichnis ist zu einem SB_M(n) isomorph.

Tetraedergruppe T

Dies ist $SO_T(3)$ für das regelmäßige Tetraeder.
2 * 4 = 8 Drehungen (120°, 240°)
3 Drehungen (180°)
1 identische Abbildung
|T| = 12, T $\cong A_4$

Oktaedergruppe O

ist $SO_W(3)$ für den Würfel.
3 * 3 Drehungen (90°,180°,270°)
2 * 4 Drehungen (120°,240°)
6 Drehungen (180°)
1 identische Abbildung
|O| = 24, O $\cong S_4$

Ikosaedergruppe Y

ist $SO_I(3) = SO_D(3)$ für reguläres Ikosaeder bzw. reguläres Dodekaeder.
|Y| = 60, Y $\cong A_5$.
(H. Weyl: Symmetrie, Birkhäuser 1955)

Satz 5 *Das Verzeichnis aller vollen Symmetriegruppen $B_M(n)$ endlicher Ordnung (= aller endlichen Untergruppen von O(n)) ist bis auf Isomorphie wie folgt:*
n = 2 C_1,C_2,C_3,...
D_2,D_3,...
(Satz von Leonardo da Vinci)
n = 3 C_1,C_2,C_3,...
D_2,D_3,...
T,O,Y,
$\overline{C_1}$,$\overline{C_2}$,$\overline{C_3}$,...
$\overline{D_2}$,$\overline{D_3}$,...
$\overline{T}$,$\overline{O}$,$\overline{Y}$
($\overline{G}$ *= G ∪ ZG, Z ist Zentralspiegelung)*
D_2C_2,D_3C_3,D_4C_4,...
C_2C_1,C_4C_2,C_6C_3,...
D_4D_2,D_6D_3,D_8D_4,...
OT
(gewisse Produktbildung)

7 unendliche Serien
7 einzelne Gruppen

4.1.5 Tapetenmuster und Kristallgruppen

Betrachten nun SB_M(n) und B_M(n) für den Fall, dass M nicht notwendigerweise beschränkt ist.
B(2) Translationen, Drehungen, Achsenspiegelungen, Gleitspiegelungen

Definition: M $\subset \mathcal{R}^2$ heißt Tapetenmuster, wenn
(i) $B_M(2)$ zwei linear unabhängige Translationen enthält,
(ii) $B_M(2)$ eine diskontinuierliche Gruppe ist, d.h. wenn für jedes x $\in \mathcal{R}^2$ die Menge (= Orbit) {g(x): g $\in B_M(2)$} keinen Häufungspunkt im $\mathcal{R}^2$ hat.

Es zeigt sich, dass der Begriff der Isomorphie von Symmetriegruppen zu weit gefasst ist. In Wirklichkeit sieht man zwei Tapetenmuster nur dann als gleich an, wenn der Isomorphismus der Symmetriegruppen eine spezielle Gestalt hat, und zwar, wenn er von der Form

$$\text{f: G} \to \text{H, f(g)} = \text{C}^{-1}\text{gC}$$

mit einer Abbildung C $\in$ B(n) ist (bei uns n = 2).
Man nennt zwei Untergruppen G und H von B(n) *konjugiert*, wenn ein Isomorphismus der obigen Gestalt existiert. Für den Künstler (Kristallographen) sind zwei Muster (Kristalle) genau dann gleich, wenn sie konjugiert sind.
Konjugiertheit ist eine Äquivalenzrelation in der Menge der Untergruppen von B(n). Die Äquivalenzklassen bezüglich Konjugiertheit sind Teilmengen der Äquivalenzklassen bezüglich Isomorphie und im Allgemeinen kleiner (daher gibt es mehr von ihnen).
Ist B_M(n) endlich und B_N(n) ebenfalls, so gilt

$$B_M(\mathrm{n}) \cong B_N(\mathrm{n}) \Leftrightarrow B_M(\mathrm{n}) \text{ konjugiert zu } B_N(\mathrm{n})$$

Es gilt nun Folgendes.

Satz: Bis auf eigentliche Isomorphie gibt es im $\mathcal{R}^2$ genau 5 Tapetenmuster, bis auf eigentliche Konjugiertheit genau 5 Tapetenmuster, bis auf Isomorphie genau 10 Tapetenmuster, bis auf Konjugiertheit genau 17 Tapetenmuster.

Eigentliche Isomorphien sind Isomorphien innerhalb der Gruppen $SB_M(\mathrm{n})$. Eigentliche Konjugiertheit bedeutet, dass C aus SB(n) genommen wird.

„Es gibt bis auf Konjugiertheit genau 17 Tapetenmuster."
Dies bedeutet: Es existieren 17 Tapetenmuster $M_1,\ldots,M_{17}$, sodass für ein beliebiges Tapetenmuster M der Gruppe $B_M(2)$ zu genau einer der Gruppen $B_{M_j}(2)$ konjugiert ist.

Welches der 17 Muster eine Tapete hat, kann man wie folgt bestimmen:
Nehmen ein Blatt Transparentpapier, legen es auf die Tapete und markieren es wie folgt:

———— Symmetrieachse

- - - - - - Achse einer Gleitspiegelung

∘ Drehung um 180°

□ Drehung um 90°

• Drehung um 60°

▲ Drehung um 120°

Nehmen das Transparentpapier dann wieder herunter. Es hat dann eins von genau 17 Mustern.

In der Kristallographie ist die Betrachtungsweise etwas anders. Eine Untergruppe von SB(n) bzw. B(n) heißt *eigentliche Kristallgruppe* bzw. (volle) *Kristallgruppe*, wenn eine kompakte und zusammenhängende Teilmenge $\mathrm{M} \subset \mathcal{R}^n$ (ein sogenannter *Elementarkristall*) mit nichtleerem Inneren ($\dot{M} \neq \varnothing$) existiert, sodass gilt:

$$\text{(i) } \mathcal{R}^n = \bigcup_{g \in G} \mathrm{g(M)}$$
$$\text{(ii) } \mathrm{g}(\dot{M}) \cap \mathrm{h}(\dot{M}) \neq 0 \Rightarrow \mathrm{g(M)} = \mathrm{h(M)}.$$

Bei Kristallgruppen hat man die Begriffe der (eigentlichen) Isomorphie und (eigentlichen) Konjugiertheit.

Satz: Es besteht eine offensichtliche Bijektion zwischen Tapetenmustern und Kristallgruppen (für n = 2).

Satz: Im $\mathcal{R}^3$ gibt es bis auf Isomorphie genau 32 (volle) Kristallgruppen, bis auf Konjugiertheit genau 230 (volle) Kristallgruppen. Im $\mathcal{R}^4$ gibt es bis auf Konjugiertheit genau 4.783 (volle) Kristallgruppen.

4.1.6 Homotopiegruppen

Sei X ein topologischer Raum (z.B. eine Teilmenge des $\mathcal{R}^n$ oder eines beliebigen metrischen Raumes) und $x_0 \in$ X. Ein *Weg* in X ist eine stetige Abbildung f: $S^1 \to$ X mit f(1) = x_0, wobei S^1 der (orientierte) Einheitskreis ist (z.B. {z $\in \mathcal{C}$: |z| = 1}).
Der Weg „startet“ und „endet“ in x_0.

Das *Produkt* zweier Wege f und g ist der Weg fg, der entsteht, wenn man erst g und dann f durchläuft. Präziser: fg ist wie folgt gegeben:
Einheitskreis wird über eine Einschnürung σ in O und U eingeteilt, auf die dann jeweils f und g angewendet werden:

$$(fg)(z) = \{ (\ g\sigma)(z)z \ \in \ O(f\sigma)(z)z \ \in \ U$$

Die Menge W(X,x_0) aller Wege mit diesem Produkt ist eine Halbgruppe.

Zwei Wege f und g heißen *homotop*, wenn sie sich innerhalb der Menge aller Wege stetig ineinander überführen lassen. Exakt: f ist homotop zu g, wenn eine stetige Abbildung H: $S^1 \times$ [0,1] $\to$ X existiert mit

$$\begin{aligned} \text{H(z,0)} &= \text{f(z)} \ \forall \ \text{z} \in S^1 \\ \text{H(z,1)} &= \text{g(z)} \ \forall \ \text{z} \in S^1 \\ \text{H(1,}\mu\text{)} &= x_0 \ \forall \ \mu \in [0,1]. \end{aligned}$$

Schreiben f $\sim$ g für Homotopie. Dies ist eine Äquivalenzrelation in der Menge aller Wege. Bezeichnen mit [f] die Äquivalenzklasse, die f enthält. Leicht zu sehen:

$$f_1 \sim f_2, g_1 \sim g_2 \Rightarrow f_1 g_1 \sim f_2 g_2.$$

Also ist durch [f][g]:=[fg] eine Multiplikation in der Menge der Äquivalenzklassen korrekt definiert. Auf diese Weise wird die Menge der Äquivalenzklassen zu einer Gruppe, der sogenannten *Fundamentalgruppe* oder 1. Homotopiegruppe π_1(X,x_0).

Das Einselement ist die Äquivalenzklasse W_0(X,x_0) aller Wege, die zum „konstanten Weg“ (f(z) = x_0 $\forall$ z $\in S^1$) homotop sind, d.h. die sich in X stetig auf den Punkt x_0 zusammenziehen lassen.

Bemerkung:
Die Menge W_0(X,x_0) ist ein Normalteiler in W(X,x_0):
a $\in$ W(X,x_0), f $\in W_0$(X,x_0) $\Rightarrow$ a^{-1}fa $\in W_0$(X,x_0)
(a^{-1} ist a in umgekehrter Richtung)
Können also die Faktorhalbgruppe W(X,x_0)/W_0(X,x_0) bilden. Diese ist genau π_1(X,x_0).

Wenn π_1(X,x_0) nicht von x_0 abhängt, so schreibt man einfach π_1(X). Dies ist z.B. der Fall, wenn X zusammenhängend ist.

1. Beispiel
Kugeln (Vollkugeln) B^n:
Haben $W_0(B^n,x_0)$ = W(B^n,x_0) und damit $\pi_1(B^n)$ = {e}, was man als $\pi_1(B^n) = 0$ schreibt.

2. Beispiel

Kreisscheibe mit Loch, S^1
Zwei Wege sind genau dann homotop, wenn sie den gleichen Index haben. Es ergibt sich
π_1(Kreis mit Loch) = $\pi_1(S^1) \cong (\mathcal{Z},+)$.
Mit Wundersatz:
Φ: W(S^1,x_0) $\to \mathcal{Z}$, Weg $\mapsto$ Index
ist Homomorphismus. Im $\Phi = \mathcal{Z}$, ker $\Phi = W_0(S^1,x_0)$
W(S^1,x_0)/ker $\Phi \cong$ Im Φ
$\underbrace{W(S^1,x_0)/W_0(S^1,x_0)}_{\pi_1(S^1,x_0)} \cong \mathcal{Z}$.

3. Beispiel
Vollkugel mit Loch, S^2 (= Kugeloberfläche)
π_1 $(S^2) = 0$.

4. Beispiel
Torusoberfläche:
Index in Längsrichtung, Index in Querrichtung ergibt Paar (k,l) mit k,l $\in \mathcal{Z}$
Also: π_1(Torusoberfläche) = $\mathcal{Z} \times \mathcal{Z}$.

Allgemein: $\pi_1(X \times Y) = \pi_1(X) \times \pi_1(Y)$ für X,Y zusammenhängend

Torusoberfläche = $S^1 \times S^1 \Rightarrow \pi_1(S^1 \times S^1) = \pi_1(S^1) \times \pi_1(S^1) = \mathcal{Z} \times \mathcal{Z}$.

Wozu benötigt man Homotopiegruppen?

Satz: Wenn zwei zusammenhängende topologische Räume X und Y homöomorph sind (d.h. wenn eine bijektive Abbildung f: X $\to$ Y mit f und f^{-1} stetig existiert), dann sind π_1(X) und π_1(Y) isomorph.

Topologisches Problem $\to$ Algebraisches Problem (Algebraische Topologie)

Beispiel: X Torus, Y Kugel
$\pi_1(X) = \mathcal{Z} \times \mathcal{Z}$, $\pi_1(Y) = 0$.
Also sind X und Y nicht homöomorph.

Noch ein Beispiel: Lässt sich die Kleeblattschlinge (ohne Zerschneidung) in den Unknoten überführen?

Ein Knoten K ist das Bild einer stetigen und injektiven Abbildung f: $S^1 \to \mathcal{R}^3$, d.h. K = f(S^1). Die Frage ist, ob es einen Homöomorphismus von $\mathcal{R}^3$ auf sich selbst gibt, der einen Knoten K in den Unknoten S^1 überführt. Dies ist äquivalent dazu, dass $\mathcal{R}^3\backslash$K homoömorph zu $\mathcal{R}^3\backslash S^1$ ist. Zeigen, dass $\pi_1(\mathcal{R}^3\backslash$K) und $\pi_1(\mathcal{R}^3\backslash S^1)$ nicht isomorph sind. Daraus folgt, dass $\mathcal{R}^3\backslash$K und $\mathcal{R}^3\backslash S^1$ nicht homöomorph sind.

Gruppe $\pi_1(\mathcal{R}^3\backslash S_1)$ ist leicht zu bestimmen:
Zwei Wege sind genau dann homotop, wenn sie den gleichen Index haben. Also: $\pi_1(\mathcal{R}^3\backslash S^1)$ = $\mathcal{Z}$.

Um $pi_1(\mathcal{R}^3\backslash$K) zu bestimmen, müssen wir etwas ausholen.
Sei A eine endliche Menge, das sogenannte *Alphabet*. Unter einem A-Wort versteht man einen Ausdruck der Form $a_1^{\epsilon_1} a_2^{\epsilon_2} \dots a_n^{\epsilon_n}$ mit $a_j \in$ A, $\epsilon_j \in$ {-1,1}. Das Produkt zweier

Wörter w = $a_1^{\epsilon_1} \dots a_n^{\epsilon_n}$ und v = $b_1^{\delta_1} \dots b_m^{\delta_m}$ ist das Wort wv = $a_1^{\epsilon_1} \dots a_n^{\epsilon_n} b_1^{\delta_1} \dots b_m^{\delta_m}$.
Erhalten so die Halbgruppe W(A) aller A-Wörter.
Sei nun R = {r_1,...,r_k} eine endliche Teilmenge von W(A). Unter einer R-Elementaroperation versteht man das Einfügen oder Streichen (auch am Anfang und am Ende) eines Wortes aus R oder eines Wortes der Form aa^{-1} oder $a^{-1}a$ mit a ∈ A.

(R = {x,xy,t}: $abx^{-1}u^{-1} = xyabx^{-1}u^{-1} = yabx^{-1}u^{-1} = yabxx^{-1}u^{-1} = yabu^{-1}$)

Zwei Wörter aus W(A) heißen R-äquivalent, wenn sie sich durch eine endliche Anzahl von R-Elementaroperationen ineinander überführen lassen. Dies ist eine Äquivalenzrelation in W(A). Bezeichnen mit [w] die Äquivalenzklasse, die w enthält.
Durch [w][v]:=[wv] ist dann eine Multiplikation in der Menge der Äquivalenzklassen korrekt definiert, durch die diese Menge zu einer Gruppe wird. Das Einselement ist die Äquivalenzklasse, die das leere Wort (oder aa^{-1}) enthält. Das zu $ab^{-1}cc$ inverse Wort ist $c^{-1}c^{-1}ba^{-1}$. Diese Gruppe wird mit

$$<\text{A: } r_1 = \text{e}, \dots, r_k = \text{e}>$$

bezeichnet und die von A erzeugte Gruppe mit den definierenden Relationen r_1 = e, ..., r_k = e genannt.

Beispiele:
<a: a^n = e> = $\mathcal{Z}_n$
<a: - > = $\mathcal{Z}$
<a,b: $\underbrace{aba^{-1}b^{-1}}_{ab=ba}$=e, a^2 = e, b^2 = e> = $V_4 = \mathcal{Z}_2 \times \mathcal{Z}_2$
<a,b: ab = ba> = $\mathcal{Z} \times \mathcal{Z}$
<a,b: - > = freie von zwei Elementen erzeugte Gruppe (recht kompliziert)

Sei nun K ⊂ $\mathcal{R}^3$ ein Knoten. Zeichnen diesen mit Unterbrechungen in der Ebene. Orientieren Knoten in eine Richtung. Er wird dann in eine endliche Anzahl orientierter Stränge $a_1 \dots a_n$ zerlegt. Dann ist A = {a_1, ..., a_n}.
Die definierenden Relationen bestimmt man wie folgt:
Malen um jede Unterbrechung einen kleinen Kreis, wählen darauf einen Punkt, der nicht zum Knoten gehört und laufen den Kreis entgegengesetzt zum Uhrzeigersinn in diesem Punkt beginnend ab. Treffen wir dabei auf Strang a_i, so schreiben wir a_i, wenn a_i nach außen zeigt, und a_i^{-1}, wenn a_i nach innen zeigt. Erhalten so ein Wort r_k.

Satz: $\pi_1(\mathcal{R}^3 \text{ /K}) = <\text{A: } r_1 = \text{e}, \dots, r_m = \text{e}>$.

Erhalten für die Kleeblattschlinge
$r_1 = a_2^{-1}a_1a_2a_3^{-1}$
$r_2 = a_2^{-1}a_1^{-1}a_3a_1$
$r_3 = a_2^{-1}a_3^{-1}a_1a_3$

$\pi_1(\mathcal{R}^3 \backslash \text{K}) = <a_1,a_2,a_3:r_1=\text{e},r_2=\text{e},r_3=\text{e}>$

Warum ist <A:r_1=e,r_2=e,r_3=e> $\ncong \mathcal{Z}$?

Betrachten Gruppe D, die aus den drei Spiegelungen a_1,a_2,a_3 eines gleichmäßigen Dreiecks und aus allen Produkten dieser Spiegelungen besteht.
Für D gilt r_1 = e, r_2 = e, r_3 = e.

Die Abbildung φ: <A:r_1=e,r_2=e,r_3=e> $\to$ D, $[b_1^{\epsilon_1} \dots b_n^{\epsilon_n}] \mapsto b_1^{\epsilon_1} \dots b_n^{\epsilon_n}$
ist korrekt definiert und ein surjektiver Homomorphismus. Wäre <A:r_1=e,r_2=e,r_3=e>
isomorph zu $\mathcal{Z}$ und damit abelsch, so wäre auch φ(<A:r_1=e,r_2=e,r_3=e>) = D abelsch,
was aber nicht der Fall ist ($a_1 a_2 \neq a_2 a_1$).

Betrachten noch höhere Homotopiegruppen. Sei
$S^n = \{(x_1, x_2, \dots, x_{n+1}) \in \mathcal{R}^{n+1}\colon x_1^2 + x_2^2 + \dots + x_{n+1}^2 = 1\}$

Betrachten jetzt stetige Abbildungen f: $S^n \to$ X, f(1,0,...,0) = x_0
Produkt fg ist über Einschnürung definiert.

f homotop zu g, wenn eine stetige Abbildung H: $S^n \times [0,1] \to$ X mit

$$H(\sigma,0) = f(\sigma) \; \forall \; \sigma \in S^n$$
$$H(\sigma,1) = g(\sigma) \; \forall \; \sigma \in S^n$$
$$H((1,0,\ldots,0),t) = x_0 \; \forall \; t \in [0,1]$$

existiert.
Menge der Homotopieklassen [f] mit [f][g]:=[fg] ist eine Gruppe, die n-te Homotopiegruppe $\pi_n(X,x_0)$. Ist X zusammenhängend, so hängt $\pi_n(X,x_0)$ nicht von x_0 ab und man schreibt $\pi_n(X)$.

Satz: X homoömorph zu Y $\Rightarrow \pi_n(X) \cong \pi_n(Y) \; \forall \; n \geq 1$.

Man kann zeigen: $\pi_n(X)$ ist für $n \geq 2$ immer abelsch.

π_k	k=1	2	3	4	5	6	7	8	9	10
S^1	$\mathcal{Z}$	0	0	0	0	0	0	0	0	0
S^2	0	$\mathcal{Z}$	$\mathcal{Z}$	$\mathcal{Z}_2$	$\mathcal{Z}_2$	$\mathcal{Z}_{12}$	$\mathcal{Z}_2$	$\mathcal{Z}_2$	$\mathcal{Z}_3$	$\mathcal{Z}_{15}$
S^3	0	0	$\mathcal{Z}$	$\mathcal{Z}_2$	$\mathcal{Z}_2$	$\mathcal{Z}_{12}$	$\mathcal{Z}_2$	$\mathcal{Z}_2$	$\mathcal{Z}_3$	$\mathcal{Z}_{15}$
S^4	0	0	0	$\mathcal{Z}$	$\mathcal{Z}_{24}$	$\mathcal{Z}_{24}$	$\mathcal{Z} \times \mathcal{Z}_{12}$	$\mathcal{Z}_2 \times \mathcal{Z}_2$	$\mathcal{Z}_2 \times \mathcal{Z}_2$	$\mathcal{Z}_{24} \times \mathcal{Z}_3$

Poincarésche Vermutung

M n-dimensionale Mannigfaltigkeit, wenn M ein topologischer Hausdorffraum mit abzählbarer Basis ist und für jeden Punkt $m_0 \in$ M eine offene Umgebung U und ein Homöomorphismus φ von U auf eine offene Teilmenge des $\mathcal{R}^n$ existiert.

Zum Beispiel ist die Kugeloberfläche und auch der Torus eine 2-dimensionale Mannigfaltigkeit
M_1 ist homöomorph zu $M_2 \Leftrightarrow \exists$ f: $M_1 \to M_2$ mit f bijektiv, f stetig, f^{-1} stetig

Poincarésche Vermutung (1901): M 3-dimensionale zusammenhängende, kompakte, orientierbare Mannigfaltigkeit, $\pi_1(M) = \pi_2(M) = 0 \Rightarrow$ M ist homöomorph zu S^3.

Smale 1960: Beweis für $n \geq 7$
Stallings/Zeeman 1961: gilt für n = 5,6
Freedman 1981: gilt für n = 4
Perelman 2005: gilt für n = 3

Kommen nochmal zu Gruppen <A:r_1=e,...,r_k=e>.
Wortproblem: Gegeben sei <A:r_1=e,...,r_k=e> und zwei Wörter w und v aus W(A). Sind w und v äquivalent? Dies ist ein unentscheidbares Problem! Das heißt, es gibt keinen Algorithmus, der dieses Problem in endlicher Zeit lösen kann.

Isomorphieproblem: Gegeben seien <A:r_1=e,...,r_k=e> und <B:s_1=e,...,s_m=e>. Sind beide Gruppen isomorph? Dies ist auch unentscheidbar!
M,N n-dimensionale Mannigfaltigkeiten
$M \cong N \Rightarrow \pi_1(M) \cong \pi_2(N)$

$\pi_1(M)$,$\pi_2(N)$ sind stets Gruppen der Form <A:r_1=e,...,r_k=e>, <B:s_1=e,...,s_m=e>.

Für Knotengruppen ist das Isomorphieproblem entscheidbar.
n = 4: Jede Gruppe <A:r_1=e,...,r_k=e> ist π_1(M) für eine 4-dimensionale kompakte, zusammenhängende, orientierbare Mannigfaltigkeit M.

Daraus folgt: Für n $\geq$ 4 ist das Problem zu entscheiden, ob zwei n-dimensionale kompakte, zusammenhängende, orientierbare Mannigfaltigkeiten homöomorph sind, unentscheidbar. n = 3 ist offen.

4.2 Ringe

4.2.1 Definitionen und Beispiele

Definition 1 *Ein Ring ist eine Menge R mit zwei inneren Operationen, d.h. Abbildungen von R in R, die Addition und Multiplikation heißen,und für die folgende Axiome gelten:*
(i) $(a + b) + c = a + (b + c) \forall\ a,b,c \in R$
(ii) $a + b = b + a \forall\ a,b \in R$
(iii) $\exists\ 0 \in R: a + 0 = a \forall\ a \in R$
(iv) $\forall\ a \in R \exists\ b \in R: a + b = 0$
(v) $(ab)c = a(bc) \forall\ a,b,c \in R$
(vi) $a(b + c) = ab + ac, (a + b)c = ac + bc \forall\ a,b,c \in R$

Bemerkung:
Das Element 0 ist eindeutig bestimmt ($0_1 = 0_1 + 0_2 = 0_2$) und wird *Nullelement* genannt. Das Element aus (iv) ist auch eindeutig bestimmt und wird mit -a bezeichnet. Man schreib a + (-b) := a - b. Leicht zu zeigen:
0 * a = a * 0 = 0 $\forall$ a $\in$ R,
(-a)b = a(-b) = -ab $\forall$ a,b $\in$ R,
(-a)(-b) = ab $\forall$ a,b $\in$ R,
(a - b)c = ac - bc $\forall$ a,b,c $\in$ R,
a(b - c) = ab - ac $\forall$ a,b,c $\in$ R.

Definition 2 *En Ring R heißt kommutativ, wenn*
(vii) $ab = ba \forall\ a,b$ *in* R
gilt und Ring mit Einselement, wenn
(viii) $\exists\ 1 \in R: 1 * a = a * 1 = a \forall\ a \in R$
gilt.

Wenn ein Einselement existiert, dann ist es eindeutig bestimmt ($1_1 = 1_1 1_2 = 1_2$).

1. Beispiel:
Mit den üblichen Operationen sind $\mathcal{Z}$,$\mathcal{Q}$,$\mathcal{R}$, $\mathcal{C}$ kommutative Ringe mit 1, $M_n(\mathcal{Z})$, $M_n(\mathcal{Q})$, $M_n(\mathcal{R})$, $M_n(\mathcal{C})$ Ringe mit 1 (nicht kommutativ für n $\geq$ 2).
$\mathcal{N}$ ist kein Ring, auch $\mathcal{N} \cup \{0\}$ nicht.

2. Beispiel:
$\mathcal{Z}_n$ = {0,1,...,n-1} mit Addition und Multiplikation modulo n ($\geq$ 2) ist kommutativer Ring mit 1.
Setzen im Folgenden stets voraus, dass R $\neq$ {0} ist.

3. Beispiel:
C[a,b] (sind die stetigen Funktionen auf [a,b]) mit üblicher Addition und Multiplikation ist kommutativer Ring mit 1.
C_0[a,b] = {f ∈ C[a,b]: f(a) = f(b) = 0} ist kommutativer Ring ohne 1.
$M_n(C_0$[a,b]) ist Ring, aber nicht kommutativ und ohne 1.

4. Beispiel:
Sei R ein Ring. Dann ist die MEnge aller formalen Polynome $a_0 + a_1x + \ldots + a_nx^n$ mit Koeffizienten aus R ein Ring mit den üblichen Operationen. Dieser Ring wird mit R[x] bezeichnet.
Analog definiert man den Ring R$[x_1, \ldots, x_k]$ aller formalen Polynome in k kommutierenden Variablen x_1,...,x_k.
Ist R kommutativ bzw. mit 1, so hat auch R$[x_1, \ldots, x_k]$ diese Eigenschaften.

Definition 3 *Ein Integritätsbereich ist ein kommutativer Ring R mit Einselement und der Eigenschaft*
a ≠ 0, b ≠ 0 ⇒ ab ≠ 0.
(sogenannte Nullteilerfreiheit).

Bemerkung:
a ≠ 0, b ≠ 0 ⇒ ab ≠ 0 ist äquivalent zu ab = 0 ⇒ a = 0 oder b = 0.

Beispiele:
$\mathcal{Z}$, $\mathcal{Q}$, $\mathcal{R}$, $\mathcal{C}$ sind Integritätsbereiche.
$\mathcal{Z}_n$ ist Integritätsbereich ⇔ n Primzahl.
C[a,b] ist kein Integritätsbereich

Definition 4 *Ein Element a eines Ringes R mit 1 heißt Einheit, wenn ein b ∈ R mit ab = ba = 1 existiert. Die Menge aller Einheiten von R wird mit GR oder R^* bezeichnet.*

Wenn ab = ba = 1 ist, dann ist b eindeutig bestimmt und wird mit a^{-1} bezeichnet: $ab_1 = b_1a = 1$, $ab_2 = b_2a = 1 \Rightarrow b_2ab_1 = b_2 \Rightarrow b_1 = b_2$.

Beispiele:
G$\mathcal{Z}$ = {-1,1},
GC[a,b] = {f ∈ C[a,b]: f(x) ≠ 0 ∀ x ∈ [a,b]},
G$\mathcal{Q}$ = $\mathcal{Q}$\{0},
G$\mathcal{R}$ = $\mathcal{R}$\{0},
G$\mathcal{C}$ = $\mathcal{C}$\{0},
G$M_n(\mathcal{R})$ = GL(n,$\mathcal{R}$),
G$M_n(\mathcal{C})$ = GL(n,$\mathcal{C}$),
G$M_n(\mathcal{Z})$ = {A ∈ $M_n(\mathcal{Z})$: det A ∈ {-1,1}},
G$\mathcal{Z}_n$ = {a ∈ $\mathcal{Z}_n$: ggT(a,n)=1},
(a ∈ $\mathcal{Z}_n$ ⇒ ∃ b: ab = 1 + kn ⇒ ggT(a,n) = 1;
ggT(a,n) = d ⇒ ∃ x,y ∈ $\mathcal{Z}$: ax + bn = d, ist also ggT(a,n) = 1, so existiert x ∈ $\mathcal{Z}$ mit ax + by = 1 ⇒ ax = 1 mod n)

Definition 5 *Ein Ring R mit 1 heißt Schiefkörper, wen GR = R{0} ist. Ein kommutativer Schiefkörper wird Körper genannt.*

Jeder Körper ist ein Integritätsbereich.

Beispiele:
$\mathcal{Q}$,$\mathcal{R}$,$\mathcal{C}$ sind Körper, $\mathcal{Z}$, C[a,b] sind keine (Schief-)Körper,
$\mathcal{Z}_n$ ist Körper $\Leftrightarrow$ n Primzahl.

4.2.2 Ringhomomorphismen und Ideale

Definition 1 *Seien R_1 und R_2 Ringe. Eine Abbildung f: $R_1 \to R_2$ heißt Ringhomomorphismus, wenn gilt:*

$$f(a + b\ ) = f(a) + f(b)\ \forall\ a,b \in R_1$$
$$f(ab) = f(a)\ f(b)\ \forall\ a,b \in R_1$$

Bijektive Ringhomomorphismen heißen Ringisomorphismen. Zwei Ringe R_1 und R_2 heißen isomorph, $R_1 \cong R_2$, wenn es einen Ringisomorphismus von R_1 auf R_2 gibt.

Satz 1 *Ein Ringhomomorphismus f: $R_1 \to R_2$ ist genau dann ein Ringisomorphismus, wenn Im f = R_2 und ker f = {0} ist. Ist f ein Ringisomorphismus, so ist dies auch f^{-1}.*

Beweis: wie für Gruppen. #

Beispiel:
f: $\mathcal{Z} \to \mathcal{Z}_n$, a $\mapsto$ a (mod n) ist ein Ringhomomorphismus.

Sei R ein Ring. Eine Teilmenge U $\subset$ R heißt *Unterring*, wenn U selbst ein Ring ist, d.h. wenn gilt:
0 $\in$ U; a,b $\in$ U $\Rightarrow$ a + b, ab $\in$ U; a $\in$ U $\Rightarrow$ -a $\in$ U.

Sei U $\subset$ R ein Unterring. Können R/U als Faktorgruppe bezüglich der kommutativen Addition bilden. Elemente von R/U sind Nebenklassen der Form a + U mit a $\in$ R. Haben also

$$(a + U) + (b + U) := (a + b) + U.$$

Bei Multiplikation in R/U gibt es die gleichen Schwierigkeiten wie in nichtabelschen Gruppen. Führte damals zu Normalteilern, diesmal zu Idealen.

Definition 2 *Eine Teilmenge I eines Ringes R heißt Ideal, wenn die folgenden beiden Eigenschaften erfüllt sind:*
(i) I ist Untergruppe der additiven Gruppe von R, d.h. a,b $\in I \Rightarrow a + b \in I$; *0* $\in I$; $a \in I \Rightarrow -a \in I$;
(ii) $j \in I$, $a \in R \Rightarrow ja \in I$, $aj \in I$.

1. Beispiel:
Jeder Ring R hat die beiden trivialen Ideale R und {0}.

2. Beispiel:
Sei R = $\mathcal{Z}$. Ein Ideal ist eine Untergruppe von ($\mathcal{Z}$,+) und damit von der Form {0},$\mathcal{Z}$,m$\mathcal{Z}$ = {0,±m, ±2m,...} (m ≥ 2). Dies sind alles Ideale.

3. Beispiel:
Ein Körper R hat bloß die trivialen Ideale. In der Tat, sei I ≠ {0} ein Ideal. Nehmen j ∈ I\{0}. Da j ∈ R\{0} = GR ist, gibt es ein b ∈ R mit jb = 1. Damit ist 1 ∈ I und somit a = a * 1 ∈ I ∀ a ∈ R. Also ist I = R.

4. Beispiel:
Für jede abgeschlossene Teilmenge F ⊂ [a,b] ist
C_F[a,b] := {f ∈ C[a,b]: f|F = 0} = {f ∈ C[a,b]: f(x) = 0 ∀ x ∈ F} ein Ideal von C[a,b].

Satz 2 *Sei R ein Ring und I* $\neq R$ *ein Ideal. Dann ist durch*

$$a \sim b \Leftrightarrow a - b \in I$$

eine Äquivalenzrelation auf R gegeben, die mit den Operationen verträglich ist, d.h.

$$a_1 \sim b_1,\ a_2 \sim b_2 \Leftrightarrow a_1 + a_2 \sim b_1 + b_2,\ a_1a_2 \sim b_1b_2.$$

Die Äquivalenzklassen sind von der Form a + I mit a $\in R$ *und durch*

$$(a + I) + (b + I) := (a + b) + I,$$
$$(a + I)(b + I) := ab + I$$

sind in der Menge der Äquivalenzklassen eine Addition und eine Multiplikation exakt definiert, die die Menge der Äquivalenzklassen zu einem Ring machen.

Beweis:
Zeigen nur, dass ∼ mit • verträglich ist:
$a_1 \sim b_1, a_2 \sim b_2 \Rightarrow a_1 - b_1 \in$ I, $a_2 - b_2 \in$ I
$\Rightarrow a_1 - b_1 = j_1,\ a_2 - b_2 = j_2\ (j_1, j_2 \in$ I)
$\Rightarrow a_1a_2 = (b_1 + j_1)(b_2 + j_2) = b_1b_2 + \underbrace{b_1j_2}_{\in I} + \underbrace{b_2j_1}_{\in I} + \underbrace{j_1j_2}_{\in I}$
$\Rightarrow a_1a_2 - b_1b_2 \in$ I
$\Rightarrow a_1a_2 \sim b_1b_2$. #

Definition 3 *Ist R ein Ring und $I \neq R$ ein Ideal, so heißt die nach Satz 2 mit Addition und Multiplikation versehene Menge der Äquivalenzklassen Faktorring von R nach I und wird mit R/I bezeichnet.*

Hier der Wundersatz:

Satz 3 *Ist f: $R_1 \to R_2$ ein Ringhomomorphismus, so ist ker f ein Ideal von R_1 und $R_1/\ker f \cong$ Im f.*
Ist umgekehrt I ein Ideal eines Ringes R, so ist die Abbildung

$$\pi: R \to R/I,\ a \mapsto a + I$$

ein Ringhomomorphismus mit ker π = I und Im π = R/I.

Beweis: wie für Gruppen. #

5. Beispiel:
Abbildung f: $\mathcal{Z} \to \mathcal{Z}$, a $\mapsto$ a (mod n)
ist Ringhomomorphismus. Haben ker f = n$\mathcal{Z}$, Im f $\mathcal{Z}_n$ und damit $\mathcal{Z}/$n$\mathcal{Z} \cong \mathcal{Z}_n$.

6. Beispiel:
Was ist C[a,b]/C_0[a,b]?
Abbildung f: C[a,b] $\to \mathcal{C}^2$, f $\mapsto$ (f(a),f(b))
Diese Abbildung ist ein Ringhomomorphismus
(Multiplikation in $\mathcal{C}^2$: (x,y)(u,v) = (xu,yv) oder $\mathcal{C}^2 = \{\begin{pmatrix} z & 0 \\ 0 & w \end{pmatrix}$: z,w $\in \mathcal{C}\}$)
ker f = C_0[a,b], Im f = $\mathcal{C}^2$, d.h. C[a,b]/C_0[a,b] $\cong \mathcal{C}^2$.

Allgemeiner: C_F[a,b] = {f $\in$ C[a,b]: f|F = 0}
Abbildung Φ: C[a,b] $\to$ C(F), f $\mapsto$ f|F ist ein Ringhomomorphismus und damit C[a,b]/C_F[a,b] $\cong$ C(F).

4.2.3 Primfaktorenzerlegung in Ringen

In diesem Abschnitt sei R stets ein Integritätsbereich. Die Eins sei 1. Da R ein Integritätsbereich ist, können wir kürzen:
ab = ac (a $\neq$ 0) $\Rightarrow$ b = c
(Beweis: ab = ac $\Rightarrow$ a(b - c) = 0 $\Rightarrow$ b - c = 0)

Definition 1 *Seien a,b $\in$ R/ {0}. Man sagt, dass b durch a teilbar oder a ein Teiler von b ist und schreibt a|b, wenn ein c $\in$ R mit b = ac existiert. Die Elemente a und b heißen assoziiert, a $\sim$ b, wenn eine Einheit ϵ ($\in$ GR) existiert, sodass a = ϵb ist.*

Es gilt: a $\sim$ b $\Leftrightarrow$ a|b und b|a.
Beweis:
a = ϵb $\Rightarrow$ b = ϵ^{-1}a $\Rightarrow$ a|b
und da a = ϵb gilt b|a.
b = ac, a = bd $\Rightarrow$ bd = adc $\Rightarrow$ a = adc $\Rightarrow$ 1 = dc $\Rightarrow$ c,d Einheiten $\Rightarrow$ a $\sim$ b. #

Assoziiertheit ist eine Äquivalenzrelation in R.

Definition 2 *Die trivialen Teiler eines Elements $a \in R/\{0\}$ sind die Einheiten (d.h. die $b \in R/\{0\}$ mit $b \sim 1$) und die zu a assoziierten Elemente (d.h. die $b \in R/\{0\}$mit $b \sim a$). Ein Primelement aus R ist ein Element $\neq 0$, das nur die trivialen Teiler hat.*

Bemerkung:
GR = {b ∈ R\{0}: b ∼ 1}
Beweis: a ∈ GR ⇒ ∃ b ∈ R\{0}: ab = 1. Es gilt b ∈ GR und damit 1 ∼ a.
a ∼ 1 ⇒ ∃ ϵ ∈ GR: a = ϵ * 1 = ϵ ⇒ a ∈ GR. #

1. Beispiel:
Für R = $\mathcal{Z}$ ist Teilbarkeit die übliche. Haben G$\mathcal{Z}$ = {-1,1}, d.h. a ∼ b ⇔ a = b oder a = -b.
Triviale Teiler der ± 1: {-1,1}
Triviale Teiler von a ≠ 0, ≠ ± 1: {-1,1,-a,a}
Primelemente: {± 2,± 3, ± 5, ± 7, ... }

Ist R ein Körper, so gilt GR = R\ {0}. Damit gilt a|b für alle a,b ∈ R\{0}. Also sind auch zwei beliebige Elemente zueinander assoziiert.
Die trivialen Teiler von a ∈ R\{0} sind alle Elemente aus R\{0}. Primelemente gibt es nicht.

Definition 3 *Ein Ring R heißt ZPE-Ring (Zerlegung in Primfaktoren ist eindeutig), wenn sich jedes Element $a \in R/(\{0\} \cup GR)$ eindeutig als Produkt von Primelementen schreiben lässt, d.h. wenn Primelemente $p_1,\ldots,p_n$ existieren, sodass $a = p_1 \ldots p_n$ ist und wenn diese Darstellung im folgenden Sinne eindeutig ist: Sind $a = p_1 \ldots p_n$ und $a = q_1 \ldots q_m$ zwei Faktorisierungen in Primelemente, so ist $m = n$ und nach geeigneter Umnummerierung gilt $p_j \sim q_j$ für alle j.*

Die Faktorisierungen 6 = 2 * 3 = (-2) * (-3)
-14 = (-2) * 7 = 2 * (-7) sieht man jeweils als gleich an.

Wissen aus der Schule, dass $\mathcal{Z}$ ein ZPE-Ring ist (zeigen wir später aber noch).

2. Beispiel:
Betrachten R = $\mathcal{Z}[\sqrt{-5}]$,d.h.
R = {a + bi$\sqrt{5}$: a,b ∈ $\mathcal{Z}$} ⊂ $\mathcal{C}$.
Haben (a + bi$\sqrt{5}$)(c + di$\sqrt{5}$) = $\underbrace{ac - 5bd}_{\in \mathcal{Z}}$ + i$\underbrace{(ad + bc)}_{\in \mathcal{Z}}$ $\sqrt{5}$
Also ist R ein Ring.
Behaupten, dass R ein Integritätsbereich ist, aber kein ZPE-Ring.
Definieren N(α) für α = a + ib$\sqrt{5}$ durch
N(α) = $|\alpha|^2 = \alpha\overline{\alpha} = a^2 + 5b^2$,
dann gilt
N($\alpha\beta$) = $|\alpha\beta|^2 = |\alpha|^2\ |\beta|^2$ = N(α)N(β).
Sei $\alpha\beta = 0$. Dann ist N(α)N(β) = N(0) = 0,
also N(α) = $a^2 + 5b^2 = 0$ oder N(β) = $c^2 + 5d^2 = 0$, d.h. $\alpha = 0$ oder $\beta = 0$.

Haben $\alpha \in$ GR $\Leftrightarrow \exists\ \beta \in$ R: $\alpha\beta = 1 \Rightarrow$ N(α)N(β) = N(1)
$\Leftrightarrow$ N(α)N(β) = 1 $\Rightarrow a^2 + 5b^2 = 1$
$\Rightarrow$ a $\in$ {-1,1}, b = 0 $\Rightarrow \alpha$ = -1 oder α = 1.
Also GR = {-1,1}.

Es gilt 6 = 2 * 3 = (1 + i$\sqrt{5}$)(1 - i$\sqrt{5}$) und wir zeigen, dass 2,3,1 + i$\sqrt{5}$, 1 - i$\sqrt{5}$ Primelemente sind. Da keine dieser zwei Zahlen assoziiert sind, ergibt dies, dass R kein ZPE-Ring ist.
Sei 2 = $\alpha\beta$. Dann ist N(2) = N(α)N(β), d.h. 4 = N(α)N(β). Also
N(α) = 1, N(β) = 4 $\Rightarrow \alpha$ Einheit,
N(α) = 2, N(β) = 2 $\Rightarrow a^2 + 5b^2 = 2$, $c^2 + 5d^2 = 2$ Widerspruch!
N(α) = 4, N(β) = 1 $\Rightarrow \alpha \sim 2$ (da β Einheit).
Also hat 2 nur triviale Teiler, und da 2 keine Einheit ist, muss 2 Primelement sein.
Analog für 3, 1 $\pm$ i$\sqrt{5}$.
N(1$\pm$i$\sqrt{5}$) = $1^2 + 5^2 = 6$ = N(α)N(β).

Ist R ein Integritätsbereich, so ist für jedes a $\in$ R die Menge aR = {ar: r $\in$ R} ein Ideal von R. Solche Ieal heißen *Hauptideale* (werden in vielen Büchern nicht mit aR, sondern mit (a) bezeichnet).

Definition 4 *Ein Integritätsbereich R heißt Hauptidealring, wenn jedes Ideal in R ein Hauptideal ist.*

Ideale in $\mathcal{Z}$ sind n$\mathcal{Z}$ mit n = 0,1,2,3,..., d.h. $\mathcal{Z}$ ist ein Hauptidealring.

Wollen zeigen, dass jeder Hauptidealring ein ZPE-Ring ist.

Lemma: Sei R ein Integritätsbereich und a,b $\in$ R/0}. Dann gilt
(i) a|b $\Leftrightarrow$ bR $\subset$ aR.
(ii) a $\sim$ b $\Leftrightarrow$ bR = aR.

Beweis:
(i): Sei a|b, d.h. b = ac. Ist x $\in$ bR, so ist x = br = acr $\in$ aR.
Ist bR $\subset$ aR, so ist b = b*1 $\in$ bR von der Form b = ar mit r $\in$ R. Das heißt a|b.
(ii) a $\sim$ b $\Leftrightarrow$ a|b und b|a. #

Satz 1 *Sei R ein Hauptidealring, a,b $\in$ R/{0}, p $\in$ R ein Primelement, p|ab. Dann ist p|a oder p|b.*

Beweis:
Sei p $\not|$ a. Die Menge I:={px+ay: x,y $\in$ R} ist ein Ideal in R. Da R ein Hauptidealring ist, existiert ein c $\in$ R mit I = cR.
Haben p = p*1 + a*0 $\in$ I und damit p $\in$ cR, d.h. p = cr mit r $\in$ R. Also teilt c das Primelement p. Daraus folgt c $\sim$ p oder c $\in$ GR.
Sei c $\sim$ p, d.h. c = ϵp mit $\epsilon \in$ GR. Es ist a = p*0+a*1 $\in$ I = cR, d.h. a = cw mit w $\in$ R. Es folgt a = ϵpw, d.h. p|a. Widerspruch!

Also ist c $\in$ GR und damit c $\sim$ 1, nach dem Lemma also cR = R. Somist ist I =

R, d.h. $1 \in I$, d.h. es existieren x,y $\in$ R mit px + ay = 1.
Multiplikation mit b ergibt
pbx + aby = b, (p|ab), d.h. rechte Seite b muss durch p teilbar sein. #

Satz 2 *Jeder Hauptidealring ist ein ZPE-Ring.*

Beweis:
Nehmen an: $a_0 \in$ R\(GR $\cup \{0\}$) ist nicht in Primfaktoren zerlegbar. Dann ist a_0 ein Primelement, d.h. es gibt a_1,$b_1 \in$ R\(GR $\cup \{0\}$) mit $a_0 = a_1 b_1$. Eines der beiden Elemente a_1,b_1 lässt sich nicht in Primfaktoren zerlegen, da wir ansonsten eine Faktorisierung von a_0 hätten.
Sei a_1 dieses Element. Aus dem Lemma folgt, dass a_0R $\not\subset a_1$R. ($a_1|a_0$ und b_1 ist keine Einheit). Erhalten dann $a_1 = a_2 b_2$ usw., d.h. Elemente $a_1, a_2, \ldots$ mit

$$a_0\mathrm{R} \not\subset a_1\mathrm{R} \not\subset a_2\mathrm{R} \not\subset a_2\mathrm{R} \not\subset \ldots$$

Setzen I = $\bigcup_{i_0}^{\infty} a_i$R. Dann ist I ein Ideal von R:
x,y $\in$ I $\Rightarrow$ x $\in a_i$R, y $\in a_j$R $\Rightarrow$ x + y $\in a_{max(i,j)}$R $\subset$ I;
x $\in$ I, r $\in$ R $\Rightarrow$ x = a_iw $\Rightarrow$ xr = $a_i$wr $\in a_i$R $\subset$ I.
Da R ein Hauptidealring ist, existiert ein c $\in$ R mit I = cR. Haben c = c * 1 $\in$ I und somit existiert ein i mit c $\in a_i$R, d.h. c = a_ir mit r $\in$ R. Nach dem Lemma ist cR $\subset$ a_iR, d.h.

$$\mathrm{cR} \subset a_i\mathrm{R} \not\subset a_{i+1}\mathrm{R} \subset \mathrm{I} = \mathrm{cR}.\ \text{Widerspruch!}$$

Damit ist die Existenz der Primfaktorenzerlegung gezeigt.

Seien $p_1, \ldots, p_m$, $q_1, \ldots, q_n$ Primelemente mit $p_1 \ldots p_m = q_1 \ldots q_n$. O.B.d.A. sei m $\leq$ n. Da die linke Seite durch p_1 teilbar ist, muss es auch die rechte Seite ein. Nach Satz 1 existiert ein q_i mit $p_1|q_i$. O.B.d.A. sei $q_i = q_1$.
Also ist $p_1 \sim q_1$, d.h. es existiert eine Einheit ϵ_1 mit $q_1 = \epsilon_1 p_1$:
$p_1 p_2 \ldots p_m = \epsilon_1 p_1 q_2 \ldots q_n \Rightarrow p_2 \ldots p_m = \epsilon_1 q_2 \ldots q_n$.
Fährt man so fort, entsteht zum Schluss

$$1 = \epsilon_1 \ldots \epsilon_n q_{m+1} \ldots q_n.$$

Es folgt, dass $q_{m+1} \ldots q_n$ Einheiten sind. Das ist nicht möglich, d.h. haben m = n und $q_i = \epsilon_i p_i$ $\forall$ i. #

Eine wichtige Klasse von Hauptidealringen sind die sogenannten Euklidischen Ringe.

Definition 5 *Ein Integritätsbereich R heißt Euklidischer Ring, wenn es eine Abbildung N: R/{0} $\rightarrow$ {0,1,2,...} mit folgenden Eigenschaften gibt:*
(i) a,b $\in$ R/{0}, a/b $\Rightarrow$ N(a) $\leq$ N(b),
(ii) $\forall$ a $\in$ R $\forall$ b $\in$ R/{0} $\exists$ c,q $\in$ R mit
a = bc + q und q = 0 oder N(q) < N(b).

Offenbar ist $\mathcal{Z}$ mit N(a) = |a| ein Euklidischer Ring. Hier sind weitere Beispiele:

3. Beispiel:
Die Menge $\mathcal{Z}$[i] = {a + bi: a,b $\in \mathcal{Z}$} (sogenannter Ring der Gaußschen Zahlen) ist ein Euklidischer Ring mit $N(\alpha) = |\alpha|^2$.
Haben $N(\alpha\beta) = |\alpha\beta|^2 = |\alpha|^2\ |\beta|^2 = N(\alpha)N(\beta)$. Daraus folgt (i):
$\alpha|\beta \Rightarrow \beta = \alpha\gamma \Rightarrow N(\beta) = N(\alpha)N(\gamma) \geq N(\alpha)$.

Seien $\alpha \in \mathcal{Z}$[i], $\beta \in \mathcal{Z}$[i]\{0}. Dann ist

$$\frac{\alpha}{\beta} = \frac{a+bi}{c+di} = \frac{(a+bi)(c-di)}{c^2+d^2} = p + qi$$

mit p,q $\in \mathcal{Q}$. Sei $p^* + p^*$i $\in \mathcal{Z}$[i] die (eine der) nächstgelegene(n) Zahl(en) zu p+qi. Dann ist
$\underbrace{|p-p^*|}_{=u} \leq \frac{1}{2}$, $\underbrace{|q-q^*|}_{=v} \leq \frac{1}{2}$.
Haben
$\alpha = \beta$(p+qi) = $\beta(p^*+q^*i) + \beta$(u + vi)
mit β(u+vi) = 0 oder
N(β(u+vi)) = N(β) * N(u+vi) = N(β) $(u^2 + v^2) \leq$N(β)($\frac{1}{4}+\frac{1}{4}$) = $\frac{1}{2}$N(β) $\leq$ N(β).

4. Beispiel:
Ist R ein Integritätsbereich, so ist dies auch R[x].
Ist K ein Körper, so ist K[x] also ein Integritätsbereich. Darüber hinaus ist K ein Euklidischer Ring mit N(a) = 1 $\forall$ a $\in$ K\{0} und K[x] ein Euklidischer Ring mit $N(a_0 + a_1x + \ldots + a_nx^n)$ = n (für $a_n \neq 0$), d.h. N(a) = deg a für a $\in$ K[x].
Für K klar. Eigenschaft (i) für K[x] klar und Eigenschaft (ii) für K[x] über übliche Division von Polynomen mit Rest.

Menge der Primelemente in K = $\varnothing$
Primelemente in K[x] heißen *über K irreduzible Polynome.*
Unter Benutzung des Hauptsatzes der Algebra erhält man:
Über $\mathcal{C}$ irreduzible Polynome sind die linearen Polynome a + bx mit a $\in \mathcal{C}$, b $\in \mathcal{C}$\{0}, über $\mathcal{R}$ irreduzible Polynome sind die linearen Polynome a + bx mit a $\in \mathcal{R}$, b $\in \mathcal{R}$\{0} und die quadratischen Polynome a + bx + cx^2 mit a $\in \mathcal{R}$, b $\in \mathcal{R}$, c $\in \mathcal{R}$\{0} und ohne reelle Nullstellen.

Kennen damit die Primelemente in $\mathcal{C}$[x] und $\mathcal{R}$[x].
Frage nach irreduziblen Polynomen über $\mathcal{Q}$, d.h. Primelemente in $\mathcal{Q}$[x] ist kompliziert.
$x^2 - 2$ ist reduzibel über $\mathcal{R}$, aber irreduzibel über $\mathcal{Q}$.
$x^2 + 1$ ist reduzibel über $\mathcal{C}$, aber irreduzibel über $\mathcal{R}$.
Weil wir dabei sind: $\mathcal{Z}$[x] ist kein Hauptidealring.
Beweis:
Betrachten I = $2\mathcal{Z}$[x] + x$\mathcal{Z}$[x] = {2p(x)+x*q(x):p(x),q(x) $\in \mathcal{Z}$[x]}.
Nehmen an, dies sei ein Hauptideal, I = c$\mathcal{Z}$[i] mit c(x) $\in \mathcal{Z}$[x].
Insbesondere ist
2 = 2*1 + x*0 $\in$ c$\mathcal{Z}$[x] $\Rightarrow$ 2 = c(x)w(x)
x = 2*0 + x*1 $\in$ c$\mathcal{Z}$[x] $\Rightarrow$ x = c(x)v(x)
mit w(x),v(x) $\in \mathcal{Z}$[x].
Also c(x) $\in$ {$\pm$1, $\pm$2} und aus x = c(x)v(x) folgt c(x) = $\pm$1. Dies ergibt I = $\mathcal{Z}$[x]. Aber 1 + x $\notin$ I, sogar 1 $\notin$ I.
Dennoch ist $\mathcal{Z}$[x] ein ZPE-Ring (siehe Literatur).

Allgemeiner: Ist R ein ZPE-Ring, so ist dies auch R[x].

5. Beispiel:
Betrachten $\mathcal{Z}[\sqrt{d}] = \{a+b\sqrt{d}: a,b \in \mathcal{Z}\}$ mit d quadratfrei.
Dann gilt:

$\mathcal{Z}[\sqrt{1}]$	$\mathcal{Z}[\sqrt{2}]$	$\mathcal{Z}[\sqrt{3}]$	$\mathcal{Z}[\sqrt{5}]$	$\mathcal{Z}[\sqrt{6}]$
euklidisch	euklidisch	euklidisch	nicht euklidisch, ZPE	euklidisch
$\mathcal{Z}[\sqrt{-1}]$	$\mathcal{Z}[\sqrt{-2}]$	$\mathcal{Z}[\sqrt{-3}]$	$\mathcal{Z}[\sqrt{-5}]$	$\mathcal{Z}[\sqrt{6}]$
euklidisch	euklidisch	nicht euklidisch, ZPE	nicht euklidisch, kein ZPE	nicht euklidisch, kein ZPE

Satz 3 *Jeder Euklidische Ring ist ein Hauptidealring.*

Beweis:
Sei $I \neq \{0\}$ ein Ideal von R. Dann hat die Menge $\{N(a): a \in I\backslash\{0\}\}$ ein Minimum. Sei $a_0 \in I\backslash\{0\}$ ein Element mit $N(a_0) \leq N(a)\ \forall\ a \in I\backslash\{0\}$. Behaupten, dass $I = a_0R$ ist.
$a_0R \subset I$:
$x \in a_0R \Rightarrow x = a_0r \Rightarrow x \in I$ (da $a_0 \in I$).
$I \subset a_0R$:
$b \in I \Rightarrow \exists\ c,q \in R: b = a_0c + q$ und $q = 0$ oder $N(q) < N(a_0)$.
Für $q = 0$ folgt $b = a_0c \in a_0R$.
Andernfalls ist $q \neq 0$ und $q = \underbrace{b}_{\in I} - \underbrace{a_0c}_{\in I} \in I$, d.h. $N(q) \geq N(a_0)$. Widerspruch! #

4.2.4 Primideale und maximale Ideale

Im folgenden sei R ein kommutativer Ring mit 1.

Definition 1 *Ein Ideal $I \neq R$ heißt Primideal, wenn R/I ein Integritätsbereich ist und maximales Ideal, wenn R/I ein Körper ist.*

Da jeder Körper ein Integritätsbereich ist, ist jedes maximale Ideal ein Primideal.
Offenbar gilt:
I ist Primideal $\Leftrightarrow$ ($ab \in I \Rightarrow a \in I$ oder $b \in I$)
Beweis: R/I ist ein Integritätsbereich $\Leftrightarrow$ $((a + I)(b + I) = I \Rightarrow a + I = I$ oder $b + I = I)$

I ist maximales Ideal $\Leftrightarrow \forall\ a \in R\backslash I\ \exists\ b \in R: ab - 1 \in I$

1. Beispiel:
Betrachten $n\mathcal{Z}$ ($n \geq 2$) in $\mathcal{Z}$. Dann sind folgende Aussagen äquivalent:
(i) $n\mathcal{Z}$ ist Primideal,
(ii) $n\mathcal{Z}$ ist maximales Ideal,
(iii) n ist Primzahl.
$\mathcal{Z}\backslash\{0\} = \mathcal{Z}$ ist Integritätsbereich, aber kein Körper $\rightarrow$ $\{0\}$ ist Primideal, aber kein maximales Ideal

2. Beispiel:
Sei R der Ring der Diagonalmatrizen im $M_2(\mathcal{C})$, d.h.

$$R = \{\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}: a,b \in \mathcal{C}\}.$$

Betrachten die Abbildung φ: R $\to \mathcal{C}$, $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \mapsto$ a.

Dann ist φ ein Homomorphismus mit Im $\varphi = \mathcal{C}$, ker $\varphi = \{\begin{pmatrix} 0 & 0 \\ 0 & b \end{pmatrix}$: b $\in \mathcal{C}\}$.

Nach dem Wundersatz ist R/ker $\varphi \cong \mathcal{C}$ und damit ist ker φ ein maximales Ideal von R.

3. Beispiel:
Sei R = C[a,b] der Ring der komplexwertigen stetigen Funktionen auf [a,b]. Für F $\subset$ [a,b] setzen wir $I_F = \{$f $\in$ C[a,b]: f|F = 0$\}$. Abbildung
φ: C[a,b] $\to$ C($\overline{F}$), f $\mapsto$ f|$\overline{F}$
ist Ringhomomorphismus, mit Im f = C($\overline{F}$) und ker $\varphi = I_F$. Also ist C[a,b]/$I_F \cong$ C($\overline{F}$), d.h. I_F ist Primideal $\Leftrightarrow$ F = {C} eine einelementige Menge ist und I_F ist maximales Ideal $\Leftrightarrow$ F = {c}.

4. Beispiel:
R wie im 2. Beispiel, aber mit $\mathcal{Z}$ statt $\mathcal{C}$. Dann ist R/ker $\varphi \cong \mathcal{Z}$, d.h. ker φ ist Primideal, aber kein maximales Ideal.

Satz 1 *Ein Ideal $I \neq R$ ist maximales Ideal von R genau dann, wenn es kein Ideal J mit $I \subset J \subset R$ und $J \neq I$, $J \neq R$ gibt.*

Beweis:
Sei I maximal und I $\subset$ J $\subset$ R, J $\neq$ I.
Nehmen a $\in$ J\I. Dann ist es möglich, ein b $\in \mathcal{R}$ zu finden, sodass ab - 1 $\in$ I ist, d.h. ab - 1 = i $\in$ I. Dann ist 1 = $\underbrace{ab}_{\in J}$ - $\underbrace{i}_{\in J}$ $\in$ J, woraus J = R folgt.
Umgekehrt möge kein J mit I $\subset$ J $\subset$ R, J $\neq$ I, J $\neq$ R existieren. Sei a $\in$ R \I. Betrachten J = I + aR = {i + ar: i $\in$ I, r $\in$ R}. Dies ist offenbar ein Ideal mit I $\subset$ J $\subset$ R und J $\neq$ I, da a $\in$ J, aber a $\notin$ I. Also ist J = R.
Damit existieren i $\in$ I und b $\in$ R mit i + ab = 1, d.h. ab - 1 = - i $\in$ I. Also ist I maximal.#

Satz 2 *Wenn R ein Hauptidealring ist, dann sind folgende Aussagen für ein Ideal $I \neq R$, $I \neq \{0\}$ äquivalent:*
(i) I ist maximales Ideal,
(ii) I ist Primideal,
(iii) I = pR mit Primelement p.

Beweis:
(i) $\Rightarrow$ (ii) trivial
(ii) $\Rightarrow$ (iii)
Haben I = aR mit a $\in$ R. Ist a kein Primelement, so gibt es eine Primfaktorenzerlegung a = $p_1 p_2 \dots p_m$ mit m $\geq$ 2. Haben $(p_1)(p_2 \dots p_m)$ = a = a * 1 $\in$ aR = I, und da I Primideal ist, folgt $p_1 \in$ I oder $p_2 \dots p_m \in$ I.
Ist $p_1 \in$ I = aR, so ist p_1 = ar und damit p_1|a und a|p_1, d.h. $a_1 \sim p_1$. Widerspruch!
Für $p_2 \dots p_m \in$ I = aR folgt $p_2 \dots p_m$ = ar und somit $p_2 \dots p_m$|a und a|$p_2 \dots p_m$, d.h. a $\sim p_2 \dots p_m$. Damit wäre p_1 eine Einheit. Widerspruch!
(iii) $\Rightarrow$ (i)
Sei J = aR ein Ideal mit pR $\subset$ J $\subset$ R, d.h. pR $\subset$ aR $\subset$ R. Nach dem Lemma aus 4.2.3

bedeutet dies a|p. Dann ist a $\sim$ p oder a $\sim$ 1. Erneut nach dem Lemma ergibt dies aR = pR oder aR = 1R = R. nach Satz 1 ist aR = I maximal. #

4.3 Körper

4.3.1 Der Quotientenkörper

Konstruktion der rationalen Zahlen aus den ganzen Zahlen ging wie folgt:
Betrachten $\mathcal{Z} \times (\mathcal{Z}\backslash\{0\})$ und darin die Äquivalenzrelation

$$(a,b) \sim (c,d) \Leftrightarrow ad = bc$$

Bezeichnen mit $\frac{a}{b}$ die Äquivalenzklasse, die (a,b) enthält. Durch

$$\frac{a}{b} + \frac{c}{d} := \frac{ad+bc}{bd},$$
$$\frac{a}{b} * \frac{c}{d} := \frac{ac}{bd}$$

ist eine Addition und eine Multiplikation in der Menge $\mathcal{Q}$ aller Äquivalenzklassen korrekt definiert, durch die $\mathcal{Q}$ zu einem Körper wird. Durch j:= $\mathcal{Z} \to \mathcal{Q}$, a $\mapsto \frac{a}{1}$ ist ein injektiver Ringhomomorphismus gegeben. Damit kann $\mathcal{Z}$ mit einer Teilmenge von $\mathcal{Q}$ identifiziert werden. $\mathcal{Q}$ ist in dem Sinne minimal, dass jeder Körper, der $\mathcal{Z}$ als Unterring enthält, auch $\mathcal{Q}$ (oder eine Kopie von $\mathcal{Q}$) als Unterkörper hat.

Obiges lässt sich leicht verallgemeinern:

Satz: Sei R ein Integritätsbereich.
(a) Durch (a,b) $\sim$ (c,d) $\Leftrightarrow$ ad = bc ist in R $\times$ (R/{0}) eine Äquivalenzrelation gegeben.
(b) Bezeichnet man die Menge der Äquivalenzklassen mit R/$\sim$ und mit $\frac{a}{b}$ oder a/b die Äquivalenzklasse, die (a,b) enthält, so ist durch

$$\frac{a}{b} + \frac{c}{d} := \frac{ad+bc}{bd}, \quad \frac{a}{b} * \frac{c}{d} := \frac{ac}{bd}$$

eine Addition und Multiplikation in R/$\sim$ korrekt definiert und R/$\sim$ wird zu einem Körper mit diesen Operationen.
(c) Die Abbildung j: R $\to$ R/$\sim$, a $\mapsto \frac{a}{1}$ ist ein injektiver Ringhomomorphismus.
(d) Ist K ein Körper und l: R $\to$ K ein injektiver Ringhomomorphismus, so existiert ein injektiver Ringhomomorphismus i: R/$\sim$ $\to$ K mit l = i $\circ$ j.
Mit anderen Worten: Jeder Körper, der R enthält, enthält auch R/$\sim$. Kurz: R/$\sim$ ist der kleinste Körper, der R enthält.

Beweis: Hausaufgabe. #

Definition: Der Körper R/$\sim$ heißt Quotientenkörper von R.

Beispiel:
Sei k ein Körper. Dann ist die Menge k[x] aller Polynome mit Koeffizientenaus k ein Integritätsbereich.
Dann ist der Quotientenkörper von k[x] die Menge

$$\{\ \tfrac{p(x)}{q(x)}\text{: p(x)} \in \text{k[x]}, \text{q(x)} \in \text{k[x]}\backslash\{0\}\ \}$$

mit den üblichen Identifizierungen und Operationen. Zum Beispiel

$$\frac{x-1}{x+1} = \frac{x(x-1)}{x(x+1)} = \frac{x^2-x}{x^2+x} = \dots$$

Nullstellen des Nennerpolynoms sind dabei ohne Belang!
Man bezeichnet diesen Körper mit k(x) und nennt ihn den *Körper der rationalen Funktionen mit Koeffizienten aus k.*

4.3.2 Einfache Körpererweiterungen

Definition 1 *Eine Teilmenge k eines Körpers K heißt Unterkörper von K, wenn k mit den Operationen aus K selbst ein Körper ist.*
Eine Körpererweiterung ist ein geordnetes Paar (K,k), in dem k ein Unterkörper von K ist; man schreibt dann einfach K ⊃ k. Der Körper K heißt Oberkörper der Körpererweiterung K ⊃ k.
Ein Körper L heißt Zwischenkörper der Körpererweiterung K ⊃ k, wenn K ⊃ L und L ⊃ k Körpererweiterungen sind.

1. Beispiel
$\mathcal{C} \supset \mathcal{R} \supset \mathcal{Q}$
$\mathcal{C}$(x) ⊃ $\mathcal{R}$(x) ⊃ $\mathcal{Q}$(x)

Der Durchschnitt beliebig vieler Unterringe bzw. Unterkörper eines gegebenen Körpers ist wieder ein Unterring bzw. ein Unterkörper.

Definition 2 *Sei K ⊃ k eine Körpererweiterung und a ∈ K. Man setzt*
k[a] = ⋂{R: R ist Unterring von K und R ⊃ k ∪{a} },
k(a) = ⋂{R: R ist Unterkörper von K und R ⊃ k ∪{a} }.
Man nennt k[a] bzw. k(a) den aus k durch Adjunktion von a erzeugten Ring bzw. Körper.
Mit anderen Worten: k[a] ist der kleinste Unterring von K, der k und a enthält, und k(a) ist der kleinste Unterkörper von K, der k und a enthält.

Definition 3 *Eine Körpererweiterung K ⊃ k heißt einfach, wenn es ein a ∈ K mit K = k(a) gibt.*

Ist a ∈ k, so ist offenbar k[a] = k(a) = k.
Interessant ist nur der Fall a ∈ K\k. Offenbar gilt:
k[a] = {p(a): p(x) ∈ k[x]}
k(a) = {p(a)q(a)$^{-1}$: p(x), q(x) ∈ k[x], q(a) ≠ 0}.
Betrachten die Abbildung

$$V_a\text{: k[x]} \to \text{k[a]}, \text{p(x)} \mapsto \text{p(a)}$$

[p(a) = $p_0 + p_1$a + ... + $p_n a^n$, $p_j \in$ k].
Diese Abbildung ist ein surjektiver Ringhomomorphismus. Damit ist ker V_a (=$\{$p(x) $\in$ k[x]: p(a) = 0$\}$) ein Ideal von k[x], und nach dem Wundersatz ist

$$\text{Im } V_a = \text{k[a]} \cong \text{k[x]/ker } V_a$$

Der Körper k(a) hat keine Nullteiler, damit auch k[a] $\subset$ k(a) nicht. Somit ist k[a] ein Integritätsbereich. Also ist ker V_a ein Primideal von k[x]. Da k[x] ein Hauptidealring ist (sogar ein Euklidischer Ring), gibt es nach Satz 2/4.2.4 genau die beiden folgenden Möglichkeiten:
ker $V_a = \{0\}$ oder ker $V_a = \varphi_a$(x)k[x] mit einem Primelement φ_a(x) $\in$ k[x] (= einem über k irreduziblen Polynom φ_a(x) $\in$ k[x]).
Im ersten Fall gibt es kein Polynom p(x) $\in$ k[x]$\backslash\{0\}$ mit p(a) = 0.
Im zweiten Fall sind dies alle Polynome der Form φ_a(x)p(x) mit p(x) $\in$ k[x].

Definition 4 *Sei $K \supset k$ eine Körpererweiterung. Ein Element $a \in K$ heißt transzendent über k, wenn es kein Polynom $p(x) \in k[x]$ mit $p(a) = 0$ gibt.*
Ein Element $a \in K$ heißt algebraisch über k, wenn es ein Polynom $p(x) \in k[x]/\{0\}$ mit $p(a) = 0$ gibt. Im letzten Fall gibt es von allen Polynomen der Form $x^n + p_{n-1}x^{n-1} + \ldots + p_0 \in k[x]/\{0\}$, die a als Nullstelle haben, eines von minimalem Grad, das Minimalpolynom von a.

Satz 1 *Sei $K \supset k$ eine Körpererweiterung und $a \in K$ transzendent über k. Dann gilt:*

$$k[a] \cong k[x],\ k(a) \cong k(x).$$

Beweis:
Haben ker $V_a = \{0\}$ und damit k[a] $\cong$ k[x]/$\{0\}$ $\cong$ k[x].
Nach dem Satz aus 4.3.1. ist k(a) der Quotientenkörper von k[a] und k(x) der Quotientenkörper von k[x].
Isomorphe Ringe haben auch isomorphe Quotientenkörper. Also ist k(a) $\cong$ k(x). #

Satz 2 *Sei $K \supset k$ eine Körpererweiterung und $a \in K$ algebraisch über k.*
(a) Das Minimalpolynom $f_a(x)$ ist über k irreduzibel. Jedes über k irreduzible Polynom mit a als Nullstelle (insbesondere $\varphi_a(x)$) stimmt bis auf den Leitkoeffizienten mit $f_a(x)$ überein.
(b) Es gilt

$$k[a] \cong k(a) \cong k[x]/f_a(x)\ k[x].$$

(c) Sind $a,b \in K$ Nullstellen eines über k irreduziblen Polynoms, so gibt es einen Isomorphismus Φ: $k(a) \to k(b)$ mit $\Phi/k = id$ und $\Phi(a) = b$.

Beweis:
(a) Ist f_a(x) = g(x)h(x) mit deg g < deg f_a, deg h < deg f_a, so folgt g(a)h(a) = 0, und da K keine Nullteiler hat, folgt g(a) = 0 oder h(a) = 0. Widerspruch!

Sei f(x) über k irreduzibel und f(a) = 0. Dann ist f(x) ∈ ker V_a und somit f(x) = φ_a(x)g(x) mit g(x) ∈ k[x]. Da f(x) irreduzibel ist, folgt f(x) ∼ φ_a(x).
Insbesondere ist f_a(x) ∼ φ_a(x). Damit ist f(x) ∼ f_a(x) für jedes irreduzible Polynom f(x) mit a als Nullstelle.
(b) Aus (a) folgt k[a] ≅ k[x]/φ_a(x)k[x] ≅ k[x]/f_a(x)k[x].
Aus (a) folgt auch, dass f_a(x) ∈ k[x] ein Primelement ist. Da k[x] ein Hauptidealring ist, liefert Satz 2/4.2.4., dass f_a(x)k[x] ein maximales Ideal, k[x]/f_a(x)k[x] also ein Körper ist. Damit ist k[a] bereits ein Körper, d.h. k[a] = k(a).
(c) Die Elemente a und b haben nach (a) das gleiche Minimalpolynom f(x). Nach (b) ist also k(a) ≅ k[x]/f(x)k[x] ≅ k(b), und daraus folgt die Behauptung. #

Betrachten noch $\mathcal{R} \supset \mathcal{Q}$. Solche Zahlen wie $\sqrt{2}$, $\sqrt{\sqrt{3}+\sqrt[5]{7}}$ sind algebraisch über $\mathcal{Q}$. Auch die Nullstellen von $x^{17} + 43x^{13} + 49x + 1 = 0$ sind über $\mathcal{Q}$ algebraisch.
Gibt es überhaupt reelle Zahlen, die über $\mathcal{Q}$ transzendent sind?
Menge aller Polynome mit rationalen Koeffizienten ist abzählbar, damit auch die Menge aller über $\mathcal{Q}$ algebraischen Zahlen. Die Menge der über $\mathcal{Q}$ transzendenten Zahlen ist also überabzählbar. Damit ist gezeigt, dass über $\mathcal{Q}$ transzendente Zahlen existieren.
Hermite 1873: e transzendent über $\mathcal{Q}$
Lindemann 1882: π transzendent über $\mathcal{Q}$

4.3.3 Endliche und algebraische Körpererweiterungen

Definition 1 *Eine Körpererweiterung K ⊃ k heißt endlich, wenn K als linearer Raum mit dem Skalarkörper k endlichdimensional ist, d.h. wenn e_1,...,e_n ∈ K existieren, sodass jedes Element aus K als Linearkombination $\alpha_1 e_1 + \ldots + \alpha_n e_n$ mit $\alpha_1, \ldots, \alpha_n \in k$ darstellbar ist. Das minimale n, für das dies möglich ist, d.h. dim_k K, wird mit [K:k] bezeichnet und Grad der Körpererweiterung genannt.*

Satz 1 *Gradsatz*
Ist K ⊃ k eine Körpererweiterung und L ein Zwischenkörper, so gilt

*[K:k] = [K:L] * [L:k].*

Beweisidee:
Ist $\{x_i\}$ eine Basis in L über k und $\{y_i\}$ eine Basis in K über L, so ist $\{x_i\ y_j\}$ eine Basis in K über k. #

Ist [K:k] eine Primzahl, so gibt es also keine echten Zwischenkörper L (d.h. solche mit K ⊃ L ⊃ k).
Wegen [$\mathcal{C}$:$\mathcal{R}$] = 2 gibt es also keinen von $\mathcal{R}$ und $\mathcal{C}$ verschiedenenen Körper zwischen $\mathcal{R}$ und $\mathcal{C}$.
Sei K ⊃ k und a ∈ K algebraisch über k. Dann ist [k(a):k] gleich dem Grad des Minimalpolynoms von a.
(Sei deg f_a(x) = n. Wissen k(a) ≅ k[a] = {p(a):p(x) ∈ k[x]}, deg p(x) ≤ n-1}.)
Eine einfache Körpererweiterung mit einem algebraischen Element ist also endlich.
Ist K ⊃ k und a transzendent über k, so ist [K:k] = ∞. Man kann zeigen, dass dann

$$k(a) \supset k(a^2) \supset k(a^4) \supset \ldots \text{ k ist.}$$

Es existieren also unendlich viele Zwischenkörper.

Definition 2 *Eine Körpererweiterung $K \supset k$ heißt algebraisch, wenn jedes $a \in K$ algebraisch über k ist.*

$\mathcal{R} \supset \mathcal{Q}$ ist keine algebraische Körpererweiterung
$\mathcal{C} \supset \mathcal{R}$ ist eine algebraische Körpererweiterung,
da a + ib Nullstelle von (x-a-ib)(x-a+ib) = $(x-a)^2 + b^2 = x^2-2ax+a^2+b^2$ ist.

Satz 2 *Für eine Körpererweiterung $K \supset k$ sind folgende Aussagen äquivalent:*
(i) $K \supset k$ ist endlich,
(ii) $K \supset k$ ist algebraisch und es existieren endlich viele $a_1, \ldots, a_n \in K$ mit $K = k(a_1)(a_2)\ldots(a_n)$.

Beweis:
(i) ⇒ (ii)
Sei [K:k] = n < ∞. Für beliebiges a ∈ K sind dann die n + 1 Elemente 1,a,...,a^n linear abhängig, d.h. es existiert $p_0, \ldots, p_n \in k$ mit $p_0 + p_1 a + \ldots + p_n a^n = 0$ (und nicht alle p_j sind 0). Also ist a algebraisch über k.
Ist $\{a_1, \ldots, a_n\}$ eine Basis in K, so lässt sich jedes a ∈ K als $q_1 a_1 + \ldots + q_n a_n$ mit q_j ∈ k darstellen, d.h. a ∈ $k(a_1)(a_2)\ldots(a_n)$.
(ii) ⇒ (i)
Die Elemente a_1,...,a_n sind algebraisch über k. Haben
$(k(a_1):k) = m_1 < \infty$
$(k(a_1)(a_2):k(a_1)) = m_2 < \infty$
$\vdots$
$(k(a_1)(a_2)\ldots(a_n):k(a_1)(a_2)\ldots(a_{n-1})) = m_n < \infty$,
nach Satz 1 also $[K:k] = m_1 m_2 \ldots m_n < \infty$. #

Jede endliche Körpererweiterung lässt sich also über endlich viele Adjunktionen von algebraischen Elementen realisieren.

4.3.4 Algebraische Abschließung und Zerfällungskörper

Satz 1 *Für einen Körper K sind folgende Eigenschaften äquivalent:*
(i) Jedes nicht-konstante Polynom p(x) ∈ K[x] hat eine Nullstelle in K
(ii) Jedes nicht-konstante Polynom p(x) ∈ K[x] zerfällt über K in Linearfaktoren, d.h. es existieren $b, a_1, \ldots, a_n \in K$ mit $p(x) = b(x-a_1)\ldots(x-a_n)$.
(iii) Ein Polynom aus K[x] ist genau dann irreduzibel über K, wenn es den Grad 1 hat.
(iv) Ist $L \supset K$ eine algebraische Körpererweiterung, so ist $L = K$.

Beweis:
(i) ⇒ (ii): Polynomdivision.
(ii) ⇒ (iii): Klar.
(iii) ⇒ (iv)

Sei L $\supset$ K eine algebraische Körpererweiterung und a $\in$ L. Dann ist a algebraisch über K. Sei f_a(x) $\in$ K[x] das Minimalpolynom. Dann ist f_a(x) irreduzibel und somit vom Grad 1, d.h. f_a(x) = x - b mit b $\in$ K.
Aus 0 = f_a(a) = a - b folgt a = b $\in$ K.
(iv) $\Rightarrow$ (i)
Sei p(x) $\in$ K[x] ein nicht-konstantes Polynom. O.B.d.A. sei p(x) irreduzibel über K. Nach Satz 2/4.2.4. ist M:= K[x]/p(x)K[x] ein Körper. Identifiziert man q $\in$ K mit q + p(x)K[x], so wird K zu einer Teilmenge von M und wir erhalten eine Körpererweiterung M $\supset$ K. Das Element a = x + p(x)K[x] $\in$ M ist eine Nullstelle von p(x):

$$p(a) = \sum_{j=1}^{m} p_j a^j = \sum_{j=1}^{m} p_j (x+p(x)K[x])^j = \left(\sum_{j=1}^{m} p_j x^j \right) + p(x)K[x] = p(x)+p(x)K[x] = p(x)K[x] = 0.$$

Die Körpererweiterung K(a) $\supset$ K ist also endlich und somit nach Satz 2/4.3.3. algebraisch.
Nach Bedingung (iv) ist K(a) = K, d.h. a $\in$ K. #

Definition 1 *Ein Körper K, der eine und damit alle der Eigenschaften aus Satz 1 hat, heißt algebraisch abgeschlossen.*

$\mathcal{Q}$, $\mathcal{R}$ nicht algebraisch abgeschlossen,
$\mathcal{C}$ ist algebraisch abgeschlossen (Fundamentalsatz der Algebra).

Definition 2 *Sei $K \supset k$ eine Körpererweiterung. Der Körper K heißt algebraische Abschließung von k, wenn gilt*
(a) K ist algebraisch abgeschlossen,
(b) $K \supset k$ ist eine algebraische Körpererweiterung.

$\overline{\mathcal{Q}}$:= Menge aller über $\mathcal{Q}$ algebraischen komplexen Zahlen
Man kann zeigen: $\overline{\mathcal{Q}}$ ist eine algebraische Abschließung von $\mathcal{Q}$.
$\mathcal{R} \supset \mathcal{Q}$ weder (a) noch (b) ist erfüllt, d.h. $\mathcal{R}$ ist keine algebraische Abschließung von $\mathcal{Q}$.
$\mathcal{C} \supset \mathcal{Q}$ (a) gilt, (b) nicht, d.h. $\mathcal{C}$ ist keine algebraische Abschließung von $\mathcal{Q}$.
$\mathcal{C} \supset \mathcal{R}$ (a) und (b) gelten $\rightarrow$ $\mathcal{C}$ ist eine algebraische Abschließung von $\mathcal{R}$.

Satz 2 *Steinitz:*
Jeder Körper k besitzt eine algebraische Abschließung $K \supset k$. Diese ist bis auf Isomorphie eindeutig. Mehr noch, sind K und L algebraische Abschließungen von k, so existiert ein Isomorphismus Φ: $K \rightarrow L$ mit $\Phi/k = id$.

Beweis: Literatur. #

Man bezeichnet die algebraische Abschließung von k mit $\overline{k}$.
Haben also $\overline{\mathcal{Q}} = \overline{\overline{\mathcal{Q}}}$, $\overline{\mathcal{R}} = \mathcal{C}$, $\overline{\mathcal{C}} = \mathcal{C}$.
In $\overline{k}$[x] zerfällt also jedes nicht-konstante Polynom p(x) $\in$ k[x] in Linearfaktoren. Oftmals hat man es aber nur mit einem speziellen Polynom p(x) $\in$ k[x] zu tun.

1. Beispiel:

k = $\mathcal{Q}$, p(x) = $x^2 - 2$
Dann zerfällt p(x) über $\mathcal{Q}[\sqrt{2}] = \mathcal{Q}(\sqrt{2})$ in Linearfaktoren:
$x^2 - 2 = (x - \sqrt{2})(x + \sqrt{2})$.
Offenbar ist $\mathcal{Q}(\sqrt{2})$ der kleinste Körper, über dem $x^2 - 2$ in Linearfaktoren zerfällt.

Definition 3 *Ein Körper K heißt Zerfällungskörper eines nicht-konstanten Polynoms p(x) ∈ k[x], wenn K ⊃ k eine Körpererweiterung ist und dabei gilt:*
(a) p(x) zerfällt über K in Linearfaktoren, d.h. es existieren b,a_1,...,a_n ∈ K mit p(x) = b(x-a_1)...(x-a_n);
(b) es gibt keinen echten Zwischenkörper L der Körpererweiterung K ⊃ k, über dem p(x) in Linearfaktoren zerfällt.

2. Beispiel:
Zerfällungskörper von $x^2 - 2 \in \mathcal{Q}$[x] ist $\mathcal{Q}[\sqrt{2}]$:
$\mathcal{Q}[\sqrt{2}] \supset \mathcal{Q}$, (a) gilt, wegen $[\mathcal{Q}[\sqrt{2}]:\mathcal{Q}] = 2$ gilt auch (b).

Zerfällungskörper von $x^2 + 1 \in \mathcal{R}$[x] ist $\mathcal{C}$:
$\mathcal{C} \supset \mathcal{R}$, (a) gilt $x^2 + 1$ = (x-i)(x+i), $[\mathcal{C}:\mathcal{R}] = 2 \rightarrow$ (b) gilt

Zerfällungskörper von $x^2 + 1 \in \mathcal{Q}$[x] ist $\mathcal{Q}$[i]:
$\mathcal{Q}$[i] = {a+bi:a,b ∈ $\mathcal{Q}$}, $\mathcal{Q}$[i] ⊃ $\mathcal{Q}$, (a) gilt: (x-i)(x+i), [$\mathcal{Q}$[i]:$\mathcal{Q}$] = 2 → (b) gilt.

Satz 3 *Sei p(x) ∈ k[x] ein nicht-konstantes Polynom.*
(a) Es existiert ein Zerfällungskörper von p(x). Dieser ist bis auf Isomorphie eindeutig. Sind K und L Zerfällungskörper von p(x), so existiert überdies ein Isomorphismus ψ: K → L mit ψ/k = id und der Eigenschaft, dass ψ die Nullstellen von p(x) in K bijektiv auf die Nullstellen von p(x) in L abbildet.
(b) Ist M ⊃ k eine Körpererweiterung, über der p(x) in die Linearfaktoren x-a_1,...,x-a_n zerfällt. Dann ist k(a_1)(a_2)...(a_n) ein Zerfällungskörper von p(x) ∈ k[x].
(c) Ist K ein Zerfällungskörper von p(x) ∈ k[x], so ist K ⊃ k eine endliche und damit algebraische Körpererweiterung.

Beweis: Literatur. #

3. Beispiel:
k = $\mathcal{Z}_2$ = {0,1}, p(x) = $x^4 - x \in \mathcal{Z}_2$[x]
Vorüberlegung: $x^4 - x = x(x^3 - 1) = x(x - 1)(x^2 + x + 1)$
Bezeichnen mit M die Menge {a + bϵ: a,b ∈ $\mathcal{Z}_2$} mit $\epsilon^2 + \epsilon + 1 = 0$ als Rechenregel.
(a + bϵ) + (c + dϵ) = (a + c) + (b + d)ϵ,
(a + bϵ)(c + dϵ) = ac + adϵ + bcϵ + bd $\underbrace{\epsilon^2}_{= -\epsilon - 1}$ = ac - bd + (ad + bc - bd)ϵ
Oder so: M := {(a,b):a,b ∈ $\mathcal{Z}_2$} mit
(a,b) + (c,d) = (a+c,b+d),
(a,b)(c,d) = (ac-bd,ad+bc-bd)
Identifizieren a ∈ k mit a + 0ϵ = (a,0). Haben dann M ⊃ k.
M ist ein Körper:
1 * 1 = 1,
ϵ(ϵ + 1) = $\epsilon^2 + \epsilon$ = - ϵ - 1 + ϵ = -1 = 1.

Haben

$$x(x-1) \quad \underbrace{(x-\epsilon)(x+\epsilon+1)}_{x^2+\epsilon x+x-\epsilon x-\epsilon^2-\epsilon=x^2+x+1} \quad = x^4 - x$$

Zerfällungskörper ist also
$\mathcal{Z}_2(0)(1)(\epsilon)(\text{-}\epsilon\text{-}1) = \mathcal{Z}_2(\epsilon)(\text{-}\epsilon\text{-}1) = \text{M}(\text{-}\epsilon\text{-}1) = \text{M}$.

Definition 4 *Sei k ein Körper und p(x) ∈ k[x] ein nicht-konstantes Polynom. Dann zerfällt p(x) über dem Zerfällungskörper in Linearfaktoren x - a_1, ..., x - a_n.*
Das Polynom p(x) heißt separabel über k, wenn $a_1,\ldots,a_n$ paarweise verschieden sind.

Definition 5 *Sei K ⊃ k eine Körpererweiterung. Ein Element a ∈ K heißt separabel über k, wenn a über k algebraisch ist und das Minimalpolynom separabel über k ist.*

Man kann zeigen, dass es in $\mathcal{Z}_p$(x) nichtseparable Minimalpolynome gibt.

Definition 6 *Sei k ein Körper und $p(x) = p_n x^n + \ldots + p_1 x + p_0 \in k[x]$.*
Die Ableitung von p(x) ist definiert durch $p'(x) = np_n x^{n-1} + (n-1)p_{n-1}x^{n-2} + \ldots + p_1 \in k[x]$, wobei $la := \underbrace{a + \ldots + a}_{l}$ ist.

Satz 4 *Sei k ein Körper und p(x) ∈ k[x] ein nicht-konstantes Polynom. Dann sind folgende Aussagen äquivalent:*
(i) p(x) ist separabel über k,
(ii) p(x) und p'(x) haben keinen nicht-konstanten gemeinsamen Teiler in k[x].

Beweis: Hausaufgabe. #

Insbesondere kann der Euklidische Algorithmus verwendet werden, um festzustellen, ob ein Polynom separabel ist.

Satz 5 *Satz von Abel über das primitive Element*
Sei K ⊃ k eine endliche Körpererweiterung. Nach Satz 2/4.3.3. ist dann $K = k(a_1)(a_2)\ldots(a_n)$ mit über k algebraischen Elementen $a_1, \ldots, a_n$. Wenn $a_1, \ldots, a_n$ über k separabel sind, dann gibt es ein über k algebraisches Element a ∈ K (ein sogenanntes primitives Element) mit K = k(a).

Beweis: Literatur. #
Beispiel:
$\mathcal{Q}(\sqrt{2})(\sqrt{3}) = \mathcal{Q}(\sqrt{2}+\sqrt{3})$.

4.3.5 Endliche Körper (Galois-Felder)

Definition 1 *Sei K ein Körper und φ die Abbildung φ: $\mathcal{Z} \to K$, $n \mapsto n * 1$. Dann ist φ ein Ringhomomorphismus und somit ist ker $\varphi = q\mathcal{Z}$ mit $q \in \{0,1,2,\ldots\}$. Die Zahl q heißt Charakteristik des Körpers K und wird mit $\chi(K)$ bezeichnet.*

χ(K) = 0: ker $\varphi = \{0\} \Leftrightarrow$ n * 1 $\neq 0 \; \forall$ n $\neq 0$
χ(K) = 1: ker $\varphi = \mathcal{Z} \Leftrightarrow$ n * 1 $= 0 \; \forall$ n $\in \mathcal{Z}$ bedeutet insbesondere 1 = 0, ist nicht möglich, da wir immer fordern, dass K mindestens 2 Elemente hat.
χ(K) = q: ker $\varphi = q\mathcal{Z} \Leftrightarrow$ q * 1 = 0, l * 1 $\neq 0$ für l = {1,2,...,q-1}.

Haben
$\chi(\mathcal{Q}) = \chi(\mathcal{R}) = \chi(\mathcal{C}) = 0$
$\chi(\mathcal{Z}_p)$ = p (p ist Primzahl)
$\chi(\mathcal{Z}_p(x))$ = p, obwohl $|\mathcal{Z}_p(x)| = \infty$ (p ist Primzahl)

Satz 1 *Die Chrarakteristik eines Körpers ist immer 0 oder eine Primzahl p.*

Beweis:
Sei χ(K) = mn mit m $\geq$ 2, n $\geq$ 2. Aus (mn)*1=0 folgt (m*1)(n*1)=0, und da es in einem Körper keine Nullteiler gibt, folgt m*1=0 oder n*1=0. Widerspruch wegen m < mn, n < mn. #

Definition 2 *Der Durchschnitt aller Unterkörper eines Körpers K ist wieder ein Unterkörper von K. Dieser Unterkörper heißt Primkörper von K.*

Satz 2 *Sei K ein Körper und P sein Primkörer. Dann gilt*
(a) $\chi(K) = 0 \Leftrightarrow P \cong \mathcal{Q}$,
(b) $\chi(K) = p \Leftrightarrow P \cong \mathcal{Z}_p$.

Beweis:
$\Leftarrow$: klar, da stets 1 $\in$ P.
(a) $\Rightarrow$
Die Abbildung φ: $\mathcal{Z} \to$ P, n $\mapsto$ n*1 ist ein injektiver Ringhomomorphismus (n*1 = m*1 $\Rightarrow$ (n-m)*1=0 $\Rightarrow$ m=n). Damit gilt $\mathcal{Z} \cong \varphi(\mathcal{Z}) \subset$ P, d.h. P enthält eine Kopie des Ringes $\mathcal{Z}$. Da P ein Körper ist, muss P auch eine Kopie des Quotientenkörpers $\mathcal{Q}$ von $\mathcal{Z}$ enthalten. Da P der Durchschnitt aller Unterkörper von K ist, ist P in der Kopie von $\mathcal{Q}$ enthalten, also ist P gleich dieser Kopie von $\mathcal{Q}$, d.h. P $\cong \mathcal{Q}$.
(b) $\Rightarrow$
Die Abbildung φ: $\mathcal{Z}_p \to$ P, n $\mapsto$ n*1 ist ein injektiver Ringhomomorphismus. Damit ist eine Kopie von $\mathcal{Z}_p$ din P enthalten. Da $\mathcal{Z}_p$ selbst schon ein Körper ist, folgt wie oben P $\cong \mathcal{Z}_p$. #

Satz 3 *Sei K ein endlicher Körper, d.h. ein Körper mit einer endlichen Anzahl |K| von Elementen.*
(a) Es gilt: |K| = p^n mit einer Primzahl p und n ∈ {1,2,3,...}.
(b) Der Zerfällungskörper von $x^{p^n} - x \in \mathcal{Z}_p[x]$ hat genau p^n Elemente.
(c) Ist P der Primkörper von K, so ist K ein Zerfällungskörper von $x^{p^n} - x \in P[x]$.
(d) Je zwei Körper mit p^n Elementen sind zueinander isomorph.

Beweis:
(a) Sei P der Primkörper von K. Da K eine endliche Menge ist, ist n = [K:P] (= Dimension von K als linearer Raum über P) endlich und damit ist K als linearer Raum isomorph zu P^n. Der Primkörper ist isomorph zu $\mathcal{Z}_p$ mit p Primzahl nach Sätzen 1 und 2, d.h. K ≅ $\mathcal{Z}_p^n$. Also ist |K| = $|\mathcal{Z}_p^n| = p^n$.
(b) Sei K der Zerfällungskörper von $x^{p^n} - x \in \mathcal{Z}_p^n$[x].
(Bemerkung: $x^{p^n} - x = x^{p^n} + (p-1)x \in \mathcal{Z}_p$[x]).
Sei zunächst n = 1. Die multiplikative Gruppe G$\mathcal{Z}_p$:= $\mathcal{Z}_p\backslash\{0\}$ hat die Ordnung p - 1. Damit ist $a^{p-1} = 1 \; \forall$ a ∈ G$\mathcal{Z}_p$, d.h. a^p - a = 0 ∀ a ∈ $\mathcal{Z}_p$.
Damit ist $\mathcal{Z}_p$ ein Zerfällungskörper von $x^p - x \in \mathcal{Z}_p$[x]. Wegen $|\mathcal{Z}_p|$ = p ist die Behauptung bewiesen.
Sei nun n > 1.
Haben K ⊃ $\mathcal{Z}_p$. Bezeichnen mit NST die Nullstellen von $x^{p^n} - x$ in K. Oben wurde gezeigt, dass a^p = a ∀ a ∈ $\mathcal{Z}_p$ gilt. Es folgt
$(a^p)^p = a^p$ = a, d.h. a^{p^2} = a
$(a^{p^2})^p = a^p$ = a, d.h. a^{p^3} = a usw.,
also $a^{p^n} - a = 0 \; \forall$ a ∈ $\mathcal{Z}_p$.
Somit ist K ⊃ NST ⊃ $\mathcal{Z}_p$.
Zeigen, dass NST ein Körper ist und dass die Anzahl der Nullstellen genau p^n ist. Da K der Zerfällungskörper ist, muss dann K = NST oder NST = $\mathcal{Z}_p$ sein. Letzteres geht wegen p^n > p nicht. Also ist K = NST und somit |K| = |NST| = p^n.
Sei a^{p^n} = a, b^{p^n} = b. Dann ist

$$(a+b)^{p^n} = a^{p^n} + \binom{p^n}{1} a^{p^n-1}b + \binom{p^n}{2} a^{p^n-2}b^2 + \ldots + b^{p^n}$$

p|$\binom{p^n}{k}$ für 1 ≤ k ≤ p-1 (Beweis: Olympiadenaufgabe)
= $a^{p^n} + b^{p^n}$ = a + b,
analog für a - b und

$$(ab)^{p^n} = a^{p^n} b^{p^n} = ab,$$
$$(ab^{-1})^{p^n} = a^{p^n} (b^{p^n})^{-1} = ab^{-1}.$$

Damit ist gezeigt, dass NST ein Körper ist.
Beweisen noch |NST| = p^n. Offenbar ist NST ≤ p^n. Zeigen also noch, dass p(x) separabel über $\mathcal{Z}_p$ ist, d.h. alle Nullstellen verschieden sind. Haben
p'(x) = $p^n x^{p^n-1}$ -1 = -1
Nach Satz 2/4.3.4. ist also p(x) separabel über $\mathcal{Z}_p$. Somit ist |NST| = p^n.
(c) Haben K ⊃ P. Die multiplikative Gruppe GK = K\{0} hat p^n - 1 Elemente (nach (a)). Also ist $a^{p^n-1} = 1 \; \forall$ a ∈ GK, d.h. $a^{p^n} - a = 0 \; \forall$ a ∈ K. Jedes Element aus K ist also Nullstelle von $x^{p^n} - x \in$ P[x], d.h. das Polynom zerfällt über K in Linearfaktoren.

Sei K $\supset$ L $\supset$ P mit Körper L, über dem $x^{p^n} - x$ in Linearfaktoren zerfällt. Dann ist |L| $\leq p^n$. Da p(x) mindestens p^n = |K| Nullstellen hat, ist |L| $\geq p^n$. Es folgt |L| = |K| = p^n. Also ist K = L.
(d) Ein Körper mit p^n Elementen hat die Charakteristik p (z.b. weil |K| = [K:P] * |P| ist und damit |P| ein Teiler von p^n also von p sein muss und sich somit nach Satz 2 |P| = p = χ(K) ergibt).
Der Primkörper ist nach Satz 2 also isomorph zu $\mathcal{Z}_p$. Aus (c) folgt, dass endliche Körper mit isomorphen Primkörpern selbst isomorph sind. #

Definition 3 *Der bis auf Isomorphie eindeutig bestimmte Körper mit p^n Elementen wird mit GF(p^n) oder $\mathcal{F}_{p^n}$ bezeichnet.*

Beispiel:
GF(2^2) = GF(4). Die multiplikative Gruppe GF(4)\{0} hat genau 3 Elemente, d.h. ist von der Form{1,a,a^2}. Also ist GF(4) = {0,1,a,a^2}. Multiplikation ist damit klar.
Addition:

	0	1	a	a^2
0	0	1	a	a^2
1	1	0	a^2	a
a	a	a^2	0	1
a^2	a^2	a	1	0

1 + 1 = 0, da χ = 2 - damit ist klar, dass es sich um die Kleinsche Vierergruppe handelt
Nach Satz 3(b) ist GF(4) der Zerfällungskörper von $x^4 - x \in \mathcal{Z}_2$[x]. Diesen hatten wir schon mal konstruiert: $\mathcal{Z}_2[\epsilon]$ mit $\epsilon^2 + \epsilon + 1 = 0$.
Dies liefert eine alternative Beschreibung von GF(4). Haben GF(4) = {0,1,ϵ,ϵ+1}. Multiplikation:

*	0	1	ϵ	ϵ+1
0	0	0	0	0
1	0	1	ϵ	ϵ+1
ϵ	0	ϵ	ϵ+1	1
ϵ+1	0	ϵ+1	1	ϵ

ϵ^2 = -ϵ-1 = ϵ+1
ϵ(1+ϵ) = ϵ + ϵ^2 = -1 = 1
(1+ϵ)^2 = 1+2ϵ + ϵ^2 = 1 + ϵ^2 = -ϵ = ϵ

Additon:

+	0	1	ϵ	ϵ+1
0	0	1	ϵ	ϵ+1
1	1	0	ϵ+1	ϵ
ϵ	ϵ	ϵ+1	0	1
ϵ+1	ϵ+1	ϵ	1	0

(a+bϵ)+(c+dϵ) = a + b(mod 2) + (b+d(mod2))ϵ

Benötigen über $\mathcal{Z}_p$ irreduzible Polynome.

Beispiel:
Polynome vom Grad 2 über $\mathcal{Z}_2$:
x^2+ax+b a,b ∈ {0,1}
x^2
x^2+x
x^2+1
x^2+x+1
(x+c)(x+d) c,d ∈ {0,1}
x^2
x(x+1) = x^2+x
(x+1)x = x^2+x
(x+1)(x+1) = x^2 + 2x + 1 = x^2 + 1
⇒ x^2+x+1 und nur x^2+x+1 ist ein irreduzibles Polynom über $\mathcal{Z}_2$ vom Grad 2.

Satz 4 *Sei p Primzahl und n ≥ 1.*
(a) Es gibt Polynome über $\mathcal{Z}_p$[x] vom Grad n, die über $\mathcal{Z}_p$ irreduzibel sind.
(b) Ist f(x) ∈ $\mathcal{Z}_p$[x] irreduzibel über $\mathcal{Z}_p$ und vom Grad n, so ist GF(p^n) ≅ $\mathcal{Z}_p$[x]/f(x)$\mathcal{Z}_p$[x].

Beweis:
(a) Wir wissen aus Satz 3(b), dass GF(p^n) der Zerfällungskörper von x^{p^n}-x ∈ $\mathcal{Z}_p$[x] ist. Nach Satz 3(b)/4.3.4. ist also GF(p^n) = $\mathcal{Z}_{p^n}(a_1)\ldots(a_{p^n})$, wobei $a_1, \ldots, a_{p^n}$ die Nullstellen von x^{p^n}-x sind. Nach Satz 4/4.3.4. ist $x^{p^n} - x$ separabel über $\mathcal{Z}_p$.
Damit sind die Minimalpolynome der Nullstellen $a_1, \ldots, a_n$ (als Teiler von $x^{p^n} - x$) auch separabel.
Nach Satz 5/4.3.4. existiert also ein $\alpha \in$ GF(p^n) mit GF(p^n) = $\mathcal{Z}_p(\alpha)$ = $\mathcal{Z}_p$[α]. Das Minimalpolynom von α ist irreduzibel über $\mathcal{Z}_p$. Der Grad ist [$\mathcal{Z}_p(\alpha)$:$\mathcal{Z}_p$] = [GF(p^n):$\mathcal{Z}_p$] = n.
(b) Satz 2/4.2.4. liefert, dass f(x)$\mathcal{Z}_p$[x] ein maximales Ideal in $\mathcal{Z}_p$[x] ist. Also ist $\mathcal{Z}_p$[x]/f(x)$\mathcal{Z}_p$[x] ein Körper. Die Elemente des Körpers sind die Restklassen
$\varphi_0 + \varphi_1 x + \ldots + \varphi_{n-1}x^{n-1} + f(x)\mathcal{Z}_p$[x]
mit $\varphi_0, \varphi_1, \ldots, \varphi_{n-1} \in \mathcal{Z}_p$.
(f(x) = x^n + $f_{n-1}x^{n-1}$ + … + f_0
⇒ x^n = -$f_{n-1}x^{n-1}$ - … - f_0
⇒ x^{n+1} = x(-$f_{n-1}x^{n-1}$ - … - f_0)).
Die Anzahl dieser Restklassen ist p^n. #

Bemerkung:
Man kann zeigen, dass die über $\mathcal{Z}_p$ irreduziblen Polynome vom Grad n immer Teiler von $x^{p^n} - x \in \mathcal{Z}_p$[x] sind.

Nochmals zu GF(4):
Wir braucehn das Polynom vom Grad 2, das über $\mathcal{Z}_2$ irreduzibel ist. Haben

$$x^4 - x = x(x^3 - 1) = x(x-1)(x^2 + x + 1)$$

Also ist $x^2 + x + 1$ irreduzibel über $\mathcal{Z}_2$. Damit ist GF(4) = $\mathcal{Z}_2$[x]/$(x^2 + x + 1)\mathcal{Z}_2$[x].
Elemente von GF(4) sind also {a+bx:a,b∈ $\mathcal{Z}_2$}, mit denen nach der Regel $x^2 + x + 1$ = 0 gerechnet wird.

Beispiel GF(16)

Brauchen über $\mathcal{Z}_2$ irreduzible Polynome vom Grad 4. Haben

$$x^{16} - x = x(x^{15} - 1) = x((x^3)^5 - 1) = x(x^3 - 1)(x^{12} + x^9 + x^6 + x^3 + 1)$$
$$= x \underbrace{(x^3 - 1)}_{(x-1)(x^2+x+1)} (x^4 + x^3 + x^2 + x + 1) \underbrace{(x^8 - x^7 + x^5 - x^4 + x^3 - x + 1)}_{f_8(x)}.$$

Alle 5 Faktoren sind irreduzibel über $\mathcal{Z}$. Das Polynom f_8(x) ist aber reduzibel über $\mathcal{Z}_2$:

$$f_8(x) = (x^4 + x^3 + 1)(x^4 + x + 1).$$

Es zeigt sich, dass die Polynome 4. Grades alle über $\mathcal{Z}_2$ irreduzibel sind. Also haben wir:
GF(16) = $\mathcal{Z}_2$[x]/$(x^4 + x^3 + x^2 + x + 1)\mathcal{Z}_2$[x],
d.h. GF(16) = $\{a+bx+cx^2+dx^3$:a,b,c,d $\in \mathcal{Z}_2\}$ mit der Rechenregel $x^4+x^3+x^2+x+1$ = 0 und außerdem
GF(16) = $\mathcal{Z}_2$[x]/$(x^4 + x^3 + 1)\mathcal{Z}_2$[x],
GF(16) = $\mathcal{Z}_2$[x]/$(x^4 + x + 1)\mathcal{Z}_2$[x].

4.3.6 Codierungstheorie

Bei der Übertragung von Nachrichten (Wörtern) können durch „Rauschen“ Fehler entstehen. In der Codierungstheorie geht es darum, diesen Effekt weitestgeehnd zu eliminieren, d.h. Übertragungsfehler zu entdecken und vielleicht sogar zu berichtigen.
Die grundlegende Herangehensweise ist wie folgt: An das *zu übertragende Wort* $c_1, \ldots, c_k$ wird ein *Kontrollwort* $c_{k+1}, \ldots, c_n$ angehängt und dann $c_1, \ldots, c_k, c_{k+1}, \ldots, c_n$ übertragen.
Aus dem empfangenen Wort $\tilde{c_1}, \ldots, \tilde{c_k}, \tilde{c_{k+1}}, \ldots, \tilde{c_n}$ muss das ursprüngliche Wort $c_1, \ldots, c_k$ irgendwie wiederhergestellt werden bzw. man begnügt sich damit, festzustellen, dass bei der Übertragung Fehler aufgetreten sind.

Beispiel: ISDN
Dies sind 10-stellige Zahlen (Striche werden ignoriert), bei denen die letzte Ziffer a_{10} eine nach der Regel

$$a_{10} = a_1 + 2a_2 + 3a_3 + \ldots + 9a_9 \ mod\ 11$$

gebildete Prüfziffer ist (10 = x).
Prüfziffer erkennt einen Fehler: wird statt a_i die Ziffer b_i eingetippt, so ist die Differenz aus richtiger Prüfziffer und Prüfziffer der eingetippten Nummer gerade i$(a_i - b_i)$ mod $11 \neq 0$ (mod 11).
Prüfziffer erkennt auch Verdrehungen zweier Zahlen:
$ia_i + ja_j - ia_j - ja_i$ mod 11 = $i(a_i - a_j) - j(a_i - a_j)$ mod 11 = (i-j)$(a_i - a_j)$ mod 11 $\neq 0$ mod 11.

Betrachten im Folgenden *binäre Codes*, d.h. alle Wörter sind Folgen von Nullen und Einsen. Wörter sind also Elemente von $\mathcal{Z}_2^m = \mathcal{Z}_2 \times \ldots \times \mathcal{Z}_2$.
Ein *binärer (k,n)-Code* besteht aus zwei Abbildungen
f: $\mathcal{Z}_2^k \to \mathcal{Z}_2^n$ (*Codierung*)
g: $\mathcal{Z}_2^n \to \mathcal{Z}_2^k$ (*Decodierung*)
Die Menge C = f($\mathcal{Z}_2^k$) $\subset \mathcal{Z}_2^n$ heißt *Menge der Codewörter.*

1. Beispiel:

Code mit Fehleranzeige
Betrachten Wörter der Länge k.

$$\text{f: } \mathcal{Z}_2^k \to \mathcal{Z}_2^{k+1},\ c_1, \ldots, c_k \mapsto c_1 \ldots c_k(c_1 + \ldots + c_k)$$

Dann ist C = $\{a_1 \ldots a_{k+1} \in \mathcal{Z}_2^{k+1}\text{: } a_1 + \ldots + a_{k+1} = 0\}$.
Empfangen wir $\tilde{a_1} \ldots \tilde{a_{k+1}}$ mit $\tilde{a_1} + \ldots + \tilde{a_{k+1}} = 0$, so liegt kein Fehler oder genau 2 Fehler oder genau 4 Fehler ... vor.
Decodierung g gibt man hier nicht an. Ist $\tilde{a_1} + \ldots + \tilde{a_{k+1}} = 0$, so decodiert man zu $\tilde{a_1} + \ldots + \tilde{a_k}$, ist $\tilde{a_1} + \ldots + \tilde{a_{k+1}} = 1$, so lässt man sich das Wort wiederholen.

2. Beispiel: Code mit Fehlerberichtigung
Betrachten f: $\mathcal{Z}_2^2 \to \mathcal{Z}_2^6$, $c_1c_2 \to c_1c_2c_1c_2c_1c_2$.
Decodieren wie folgt
g: $\mathcal{Z}_2^6 \to \mathcal{Z}_2^2$, $a_1b_1a_2b_2a_3b_3 \mapsto d_1d_2$
mit d_1 das in a_1, a_2, a_3 am häufigsten auftretende Element und d_2 das in b_1, b_2, b_3 am häufigsten vorkommende Element.
Z.B. 01 $\xrightarrow{f}$ 010101 $\overset{Uebertragung}{\longrightarrow}$ 000101 $\xrightarrow{g}$ 01
Kein Fehler: $d_1d_2 = c_1c_2$
Genau ein Fehler: $d_1d_2 = c_1c_2$
Zwei Fehler: $d_1d_2 \neq c_1c_2$ ist möglich.

Ist die Wahrscheinlichkeit für einen Fehler pro Wort 10%, so ist das empfangene Wort nur zu 90 % richtig. Die Wahrscheinlichkeit für genau 2 Fehler pro Wort ist 10 % * 10 % = 1 %.
Im 1. Beispiel wird genau ein Fehler angezeigt und somit haben wir am Ende mit einer Wahrscheinlichkeit von 99 % ein richtiges Wort. Im 2. Beispiel haben wir ebenfalls mit 99 % ein richtiges Wort.

Beschränken uns auf *matrizielle* binäre (k,n)-Codes. Bei diesen ist

$$f : (c_1 \ldots c_k) \mapsto (c_1 \ldots c_k) \begin{pmatrix} g_{11} & \cdots & g_{1n} \\ \vdots & & \vdots \\ g_{k1} & \cdots & g_{kn} \end{pmatrix},$$

f = cG mit G $\in M_{k,n}(\mathcal{Z}_2)$.
Es wird gefordert, dass G k über $\mathcal{Z}_2$ linear unabhängige Spalten hat $\Leftrightarrow$ G hat k über $\mathcal{Z}_2$ linear unabhängige Zeilen. Diese Forderung garantiert $c_1\text{G} = c_2\text{G} \Rightarrow c_1 = c_2$.
(cG = 0 $\Rightarrow$ c = 0 muss gezeigt werden:

$$(c_1 \ldots c_k) \begin{pmatrix} g_{11} & \cdots & g_{1n} \\ \vdots & & \vdots \\ g_{k1} & \cdots & g_{kn} \end{pmatrix} = c_1(g_{11} + \ldots + g_{1n}) + \ldots + c_k(g_{k1} + \ldots + g_{kn})$$

=(0 ... 0) geht aber nur für $c_1 = \ldots = c_k = 0$.)

Zur Decodierung:
Das empfangene Wort sei $\tilde{a}$. Man sucht das zu $\tilde{a}$ nächstgelegene Codewort, bzw., falls es davon mehrere gibt, ein vorher fixiertes. Dieses Codewort ist von der Form dG, und man definiert

$$\text{g: } \mathcal{Z}_2^n \to \mathcal{Z}_2^k,\ \tilde{a} \mapsto \text{d}.$$

Die Codes aus Beispiel 1 und 2 sind matriziell.

Beispiel 1: $c_1 \dots c_k \mapsto (c_1 \dots c_k) \begin{pmatrix} 1 & 0 & & 0 & 1 \\ 0 & 1 & & 0 & 1 \\ \vdots & 0 & \ddots & \vdots & \vdots \\ 0 & 0 & & 1 & 1 \end{pmatrix}$,

Beispiel 2: $c_1c_2 \mapsto (c_1c_2)\begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$.

Bei matriziellen Codes ist die Abbildung f injektiv und daher C = f($\mathcal{Z}_2^k$) eine zu $\mathcal{Z}_2^k$ isomorphe Untergruppe von $\mathcal{Z}_2^n$.
Für ein Codewort x ∈ C ist das *Gewicht* w(x) definiert als die Anzahl der Einsen in x, d.h. $w(x_1 \dots x_n) = x_1 + \dots + x_n$.
Also w(011101) = 4.
Der *Hamming-Abstand* d(x,y) zweier Codewörter x,y ∈ C ist definiert als d(x,y) = w(x-y) = w(x+y). Dies entspricht der Anzahl der Stellen, in denen sich x und y unterscheiden. Also z.B. d(011011,000101) = 4.
Es ist leicht zu sehen, dass d eine Metrik auf C ist.
Der *Codewörterabstand* ist schließlich definiert als $d_C = \min_{x,y\in\ C, x\neq y}$ d(x,y) = $\min_{x\in C, x\neq 0}$ w(x).
Beispiel 1: $d_C = 2$, Beispiel 2 $d_C = 3$.

Satz 1 *Ein matrizieller binärer (k,n)-Code hat folgende Eigenschaften:*
max. t Fehler werden angezeigt ⇔ $d_C \geq t + 1$,
max. t Fehler werden berichtigt ⇔ $d_C \geq 2t + 1$.

Beweis:
Zeigen, dass für $d_C \geq 2t + 1$ max. t Fehler korrigiert werden:
$c \xrightarrow{f} a \xrightarrow{U} \tilde{a} \xrightarrow{naechstes\ Wort} b \xrightarrow{dG=b} d$
Sei c das zu übertragende Wort, a = cG, $\tilde{a}$ das empfangene Wort, b das zu $\tilde{a}$ nächstgelegene Codewort, d die Lösung von dG = b. Wissen, dass $d_C \geq 2t + 1$ ist. Müssen zeigen, dass d = c ist, falls genau s ≤ t Fehler unterlaufen sind. Es reicht, a = b zu zeigen. Haben
d(a, $\tilde{a}$) = s, da genau s Fehler auftreten.
Da b das zu $\tilde{a}$ nächstgelegene Codewort ist, gilt d($\tilde{a}$,b) ≤ d(a,$\tilde{a}$) = s.
Also folgt d(a,b) ≤ d(a,$\tilde{a}$) + d($\tilde{a}$,b) ≤ s + s ≤ 2t, d.h. a = b. #

3. Beispiel: *Hamming-Code* 1940er Jahre
Für jede natürliche Zahl r ≥ 2 existiert ein matrizieller binärer ($2^r - 1 - r$,2^r-1)-Code mit $d_C \geq 3$.

r = 2: (1,3)-Code
f: c ↦ ccc = c(111)

r = 3: (4,7)-Code
Ein solcher Code ist zum Beispiel $(c_1c_2c_3c_4)\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$

In der Praxis r = 8: (247,255)-Code. 6 Bit für einen Buchstaben. Mit 234 Bit kann man Wörter mit 39 Buchstaben darstellen. Mit 246 Bit kann man sogar Wörter mit 41 Buchstaben darstellen.
Der Hamming-Code ist optimal in folgendem Sinn: Für einen matriziellen binären (k,2^n-1)-Code mit $d_C \geq 3$ gilt stets k $\leq 2^n - 1 - r$.
Suche nach Codes mit akzeptabler Übertragungsgeschwindigkeit und $d_C \geq 5$ gestaltete sich schwierig und nahm über 15 Jahre in Anspruch. Lösung wurde erst 1960 gefunden. Durch

$$c_0 c_1 \dots c_{m-1} \mapsto c_0 + c_1 x + \dots + c_{m-1} x^{m-1}$$

ist ein Isomorphismus des linearen Raumes $\mathcal{Z}_2^m$ auf den linearen Raum $\mathcal{Z}_2$[x]/$(x^m - 1)\mathcal{Z}_2$[x] gegeben. Genauer: c $\mapsto$ c(x) + $(x^m - 1)\mathcal{Z}_2$[x]
Eine Abbildung f: $\mathcal{Z}_2^k \to \mathcal{Z}_2^n$ ist also auch gegeben über
$\tilde{f}$: $\mathcal{Z}_2$[x]/$(x^k - 1)\mathcal{Z}_2$[x] $\to$ $\mathcal{Z}_2$[x]/$(x^n - 1)\mathcal{Z}_2$[x].
Ist φ(x) $\in \mathcal{Z}_2$[x] ein Polynom vom Grad n-k, so ist durch
c(x) + $(x^k - 1)\mathcal{Z}_2$[x] $\to$ c(x)φ(x) + $(x^n - 1)\mathcal{Z}_2$[x] eine Abbildung $\tilde{f}$ wie oben definiert.
Solche Codes heißen *polynomiale Codes*. Jeder polynomiale Code ist ein matrizieller Code.

Beispiel: k = 2, n = 5
c(x) = c_0 + c_1x, φ(x) = φ_0 + φ_1x + $\varphi_2$$x^2$ + φ_3 x^3
Bilden

$$c(x)\varphi(x) = c_0\varphi_0 + c_0\varphi_1 x + c_0\varphi_2 x^2 + c_0\varphi_3 x^3 + c_1\varphi_0 x + c_1\varphi_1 x^2 + c_1\varphi_2 x^3 + c_1\varphi_3 x^4$$

$$(c_0 c_1) \begin{pmatrix} \varphi_0 & \varphi_1 & \varphi_2 & \varphi_3 & 0 \\ 0 & \varphi_0 & \varphi_1 & \varphi_2 & \varphi_3 \end{pmatrix}$$

Allgemein:

$$(c_0 \dots c_{k-1}) \begin{pmatrix} \varphi_0 & \dots \varphi_{n-k} & & \\ & \ddots & & \ddots \\ & & \varphi_0 & \dots \quad \varphi_{n-k} \end{pmatrix}$$

Definition: Ein Element $\alpha \in$ GF(2^m) heißt superprimitiv, wenn gilt
GF(2^m) = {0,1,α,α^2, …,α^{2^m-2}}.

Satz 2 *In GF(2^m) existieren stets superprimitive Elemente.*

Beweis: Literatur. #

Man kann zeigen, dass $\alpha \in$ GF(2^m) genau dann superprimitiv ist, wenn das Minimalpolynom von α den Grad m hat, ein Teiler von $x^{2^m-1} - 1$, aber kein Teiler von $x^l - 1$ für l $< 2^m - 1$ ist.

4. Beispiel:
Betrachten GF(16) = GF(2^4). Das Minimalpolynom eines superprimitiven $\alpha \in$ GF(2^4) muss ein irreduzibles Polynom vom Grad 4 sein, und am Ende von 4.3.5. hatten wir gesehen, dass es genau die folgnden drei Polynome mit dieser Eigenschaft gibt:

$$x^4 + x^3 + x^2 + x + x + 1,$$

$$x^4 + x^3 + 1,$$
$$x^4 + x + 1.$$

Haben

$$x^5 - 1 = (x - 1)(x^4 + x^3 + x^2 + x + 1),$$

d.h. erstes Polynom liefer keine superprimitiven Elemente ($\alpha^5 = 1$).
Man kann zeigen, dass die Nullstellen von $x^4 + x^3 + 1$ und $x^4 + x + 1$ superprimitiv sind.

Satz 3 *Bose/Chaudhure/Hocquenghem 1960*
Sei $d \geq 3$ eine ungerade Zahl, $r \geq 2$ eine natürliche Zahl, $2^r \geq d + 1$. Dann existiert ein polynomialer binärer $(2^r - 1 - \frac{d-1}{2}r, 2^r - 1)$-Code mit $d_C \geq d$.

Beweisansatz:
Sei $\alpha \in \mathrm{GF}(2^r)$ superprimitiv. Bezeichnen mit $\phi_j(\mathrm{x}) \in \mathcal{Z}_2[\mathrm{x}]$ das Minimalpolynom von α^j für j = 1, ..., α - 1:
α $\phi_1(\mathrm{x})$
α^2 $\phi_2(\mathrm{x})$
$\vdots$ $\vdots$
$\alpha^{\alpha-1}$ $\phi_{\alpha-1}(\mathrm{x})$

Bezeichnen mit $\phi(\mathrm{x}) \in \mathcal{Z}_2[\mathrm{x}]$ das kleinste gemeinsame Vielfache von $\phi_1(\mathrm{x}) \ldots \phi_{\alpha-1}(\mathrm{x})$. Man kann zeigen, dass $\phi(\mathrm{x})$ maximal den Grad $\frac{d-1}{2}r$ hat. Der polynomiale Code mit $\phi(\mathrm{x})$ ist dann der gewünschte. #

Nochmal zum 4. Beispiel
Sei $\alpha \in \mathrm{GF}(2^4)$ eine Nullstelle von x^4+x^3+1. Haben dann folgende Minimalpolynome (muss gezeigt werden):
$\alpha: \ x^4+x^3+1$
$\alpha^2: \ x^4+x^3+1$
$\alpha^3: \ x^4+x^3+x^2+x+1$
$\alpha^4: \ x^4+x^3+1$

Wählen r = 4, d = 5. Man kann zeigen, dass das kgV gleich
$\phi(\mathrm{x}) = (x4^+x^3+1)(x^4+x^3+x^2+x+1) = x^8+x^7+x^6+x^5+x^4+x^7+x^6+x^5+x^4+x^3+x^4+x^3+x^2+x+1 = x^8+x^4+x^2+x+1$ ist.
Also ist uns ein binärer (2^4-1-$\frac{5-1}{2}$4,2^4-1) = (7,15)-Code mit $d_C \geq 5$ gegeben durch
$(c_0+c_1x+\ldots+c_6x^6)(x^8+x^4+x^2+x+1)$ oder

$$(c_0c_1\ldots c_6)\begin{pmatrix} 1&1&1&0&1&0&0&0&1&0&0&0&0&0&0\\ 0&1&1&1&0&1&0&0&0&1&0&0&0&0&0\\ 0&0&1&1&1&0&1&0&0&0&1&0&0&0&0\\ 0&0&0&1&1&1&0&1&0&0&0&1&0&0&0\\ 0&0&0&0&1&1&1&0&1&0&0&0&1&0&0\\ 0&0&0&0&0&1&1&1&0&1&0&0&0&1&0\\ 0&0&0&0&0&0&1&1&1&0&1&0&0&0&1 \end{pmatrix}$$

In der Praxis rechnet man in Europa mit r = 8 und d = 7.
$(2^8-1-\frac{7-1}{2}8, 2^8-1) = (231,255)$. Berichtigt 3 Fehler und erkennt 6 Fehler.

4.3.7 Konstruktionen mit Zirkel und Lineal

Betrachten folgende Aufgaben:
(a) Gegeben seien drei Kreise in der Ebene. Man konstruiere einen Kreis, der jeden dieser Kreise berührt.
(b) Gegeben sei ein Kreis in der Ebene. Man konstruiere ein Quadrat mit gleichem Flächeninhalt (Quadratur des Kreises).
(c) Gegeben sei ein Kreis in der Ebene. Man beschreibe diesem Kreis ein regelmäßiges n-Eck ein.
Wissen:
(a) geht mit Zirkel und Lineal,
(b) geht nicht mit Zirkel und Lineal,
(c) geht mit Zirkel und Lineal für gewisse n, nicht aber für alle n.

Um zu zeigen, dass etwas geht, gibt man die Konstruktion an. Uns geht es im Folgenden darum, wie man beweist, dass etwas nicht geht.

Was bedeutet Konstruktion mit Zirkel und Lineal?
Uns ist eine Menge M in $\mathcal{C}$ gegeben, die Menge der Ausgangsdaten. O.B.d.A. sei {0,1} $\subset$ M.
Bezeichnen mit pVq die Gerade durch die Punkte p und q und mit K(p,r) den Kreis vom Radius r mit Mittelpunkt p. Erlaubt sind folgende *Elementarkonstruktionen*:

Typ I
Seien p_1,p_2,q_1,q_2 mit $p_1 \neq p_2$, $q_1 \neq q_2$ gegeben. Dann kann $(p_1 \text{ V } p_2) \cap (q_1 \text{ V } q_2)$ konstruiert werden.

Typ II
Seien p_1,p_2,q_1,q_2 mit $p_1 \neq p_2$ und $q_1 \neq q_2$ gegeben. Dann kann $(p_1 \text{ V } p_2) \cap \text{K(p,}|q_1 - q_2|)$ konstruiert werden.

Typ III
Sei p_1,p_2,q_1,q_2,p,q mit $p_1 \neq p_2$, $q_1 \neq q_2$, p $\neq$ q gegeben. Dann kann $\text{K(p,}|p_1 - p_2|) \cap \text{K(q,}|q_1 - q_2|)$ konstruiert werden.

Bezeichnen mit kon M die Menge aller Punkte in $\mathcal{C}$, die sich aus M durch endlich viele Elementarkonstruktionen erhalten lassen. Desweiteren sei $\mathcal{Q}$(M) der kleinste Unterkörper von $\mathcal{C}$, der $\mathcal{Q}$ und M enthält, d.h. $\mathcal{Q}$(M) ist der Durchschnitt aller Körper, die sowohl $\mathcal{Q}$ als auch M enthalten.

Satz 1 *kon M ist ein Zwischenkörper der Körpererweiterung $\mathcal{C} \supset \mathcal{Q}$(M). Ist b $\in \mathcal{C}$ und $b^2 \in$ kon M, so ist auch b $\in$ kon M.*

Beweis:
Mit Zirkel und Lineal geht: Fällen des Lotes, Errichten der Senkrechten, Parallele.
Zeigen, dass kon M ein Körper ist:
a,b $\in$ kon M $\Rightarrow$ a + b $\in$ kon M
a $\in$ kon M $\Rightarrow$ -a $\in$ kon M
a,b $\in$ kon M $\Rightarrow$ ab $\in$ kon M
a $\in$ kon M, a $\neq 0 \Rightarrow a^{-1} \in$ kon M
Damit ist gezeigt, dass kon M ein Unterkörper von $\mathcal{C}$ ist.
Zeigen, dass kon M $\supset \mathcal{Q} \cup$ M enthält. Da kon M ein Körper ist, folgt daraus kon M $\supset$ $\mathcal{Q}$(M).
kon M $\supset$ M ist trivial. Wegen {0,1} $\in$ kon M und der Tatsache, dass kon M ein Körper ist, folgt kon M $\supset \mathcal{Q}$.
Konstruktion der Quadratwurzel: Sei a $\in$ kon M. Winkelhalbierende lässt sich mit Zirkel konstruieren. Nach dem Höhensatz im rechtwinkligen Dreieck gilt: h^2 = pq = 1* |a| $\Rightarrow$ h = $\sqrt{|a|}$. #

Satz 2 *Für $z \in \mathcal{C}$ sind folgende Aussagen äquivalent:*
(i) z $\in$ kon M,
(ii) $\exists$ eine endliche Kette von Zwischenkörpern $\mathcal{Q}$(M) = $L_0 \subset L_1 \subset \ldots \subset L_m \subset \mathcal{C}$ mit $[L_n:L_{n-1}] = 2 \; \forall$ n = 1, ..., m und $z \in L_m$.

Beweis:

(i) $\Rightarrow$ (ii).
Sei z $\in$ kon M. Dann entsteht durch sukzessive Hinzunahme von Punkten zu $L_0 := \mathcal{Q}(M)$. Sei α ein Punkt, der zu L_0 hinzugenommen wird. Entsteht α über Typ I, so ergibt sich α als Lösung von zwei linearen Gleichungen mit zwei Unbekannten und Daten aus L_0 (α = x+iy ist konstruierbar $\Rightarrow$ x,y sind konstruierbar). Daraus folgt $\alpha \in L_0$.
Ergibt sich α über Typ II oder Typ III, so ist α Lösung einer quadratischen Gleichung mit Koeffizienten aus L_0.
Zerfällt das quadratische Polynom über L_0 in Linearfaktoren, so ist $\alpha \in L_0$.
Zerfällt das quadratische Polynom nicht über L_0, so ist es das Minimalpolynom von α über L_0. Setzen $L_1 := L_0(\alpha)$. Dann ist $[L_1:L_0]$ = Grad des Minimalpolynoms von α = 2. Fährt man so fort, nun mit L_1 statt mit L_0 beginnend, so ergibt sich die Behauptung.
(ii) $\Rightarrow$ (i).
Sei $\alpha \in L_n \backslash L_{n-1}$. Müssen zeigen, dass sich α aus L_{n-1} über endlich viele Elementarkonstruktionen konstruieren lässt.
$\llbracket (L_n:L_{n-1}) \rrbracket = 2 < \infty$.
Nach Satz 2/4.3.3. ist α algebraisch über L_{n-1} Sei d der Grad des Minimalpolynoms. Wegen $\alpha \notin L_{n-1}$ ist d ≥ 2. Haben $L_{n-1} \subset L_{n-1}(\alpha) = L_n$. Der Gradsatz liefert
$\llbracket L_n:L_{n-1} \rrbracket = [L_n:L_{n-1}(\alpha)] * [L_{n-1}(\alpha):L_{n-1}]$
$2 = \circ * d$
$\Rightarrow \circ = 1$, d = 2.
Also ist α Nullstelle eines quadratischen Polynoms $x^2 + ax + b = 0$ mit a,b $\in L_{n-1}$. Haben

$$x^2 + ax + b = (x + \frac{a}{2})^2 + b - \frac{a^2}{4} = 0$$

$\sqrt{\frac{a^2}{4} - b}$ lässt sich konstruieren (Satz 1), damit $\alpha + \frac{1}{2}$ und damit auch α. #

Folgerung: Ist $L_0 := \mathcal{Q}(M)$ und z $\in$ kon M, so ist z algebraisch über L_0 und es gilt: $[L_0(z):L_0]$ ist eine Zweierpotenz.

Beweis:
Nach Satz 2 ist z $\in L_m$ mit Zwischenkörper $L_0 \subset L_1 \subset \ldots \subset L_m$ mit $[L_n:L_{n-1}] = 2^m$. Also ist $L_0 \subset L_m$ eine algebraische Körpererweiterung (Satz 2/4.3.3.) und damit ist z algebraisch über L_0. Haben $L_0 \subset L_0(z) \subset L_m$ und nach dem Gradsatz ist

$$\underbrace{[L_m : L_0]}_{2^m} = [L_m : L_0(z)] * [L_0(z) : L_0],$$

d.h. $[L_0(z):L_0]$ ist eine Zweierpotenz (als Teiler von 2^m). #

Unmöglichkeit der Quadratur des Kreises
Gegeben ist M = {0,1}. Gesucht ist die Seitenlänge eines Quadrates mit der gleichen Fläche wie der Kreis mit Mittelpunkt 0 und Radius 1.
Ist $\sqrt{\pi} \in$ kon M?
In diesem Fall ist $L_0 = \mathcal{Q}(M) = \mathcal{Q}$, und da $\sqrt{\pi}$ transzendet über $\mathcal{Q}$ ist (Lindemann), folgt $\sqrt{\pi} \notin$ kon M.

Unmöglichkeit der Würfelverdopplung (Delisches Problem)
Gegeben sei ein Würfel mit der Kantenlänge 1. Gesucht ist die Kantenlänge eines Würfels mit doppeltem Volumen. Also: M = {0,1}. Die Frage ist, ob $\sqrt[3]{2} \in$ kon M.
Wieder ist $L_0 = \mathcal{Q}$. $\sqrt[3]{2}$ ist Lösung von $x^3 - 2 = 0$ und damit algebraisch über $\mathcal{Q}$. Haben $[\mathcal{Q}(\sqrt[3]{2}:\mathcal{Q}] = 3$ und dies ist keine Zweierpotenz.

Unmöglichkeit der Winkeldreiteilung

Spezielle Winkel (z.B. 180°) lassen sich mit Zirkel und Lineal dreiteilen. Im Allgemeinen geht das aber nicht. Zeigen, dass sich z.B. 60° nicht dreiteilen lässt.
Wegen {0,1} ⊂ M ist Konstruktion von $e^{i\alpha}$ äquivalent zur Konstruktion von $\cos\alpha$.
Haben cos 60° = $\frac{1}{2}$.
Gegeben ist also M = {0,$\frac{1}{2}$,1}, gesucht ist cos 20°. Haben $L_0 = \mathcal{Q}(\mathrm{M}) = \mathcal{Q}$.
Betrachten L_0(cos 20°). Es gilt

$$\cos 3\alpha + i\sin 3\alpha = (\cos\alpha + i\sin\alpha)^3 = \cos^3\alpha - 3\cos\alpha\sin^2\alpha + i(\ldots)$$
$$\rightarrow \cos 3\alpha = \cos^3\alpha - 3\cos\alpha(1-\cos^2\alpha) = 4\cos^3\alpha - 3\cos\alpha$$

Setzen wir x = cos 20°, so ist also $\frac{1}{2} = 4x^3 - 3x$ oder $8x^3 - 6x - 1 = 0$.
Wäre dieses Polynom reduzibel über $\mathcal{Q} = L_0$, so hätte es eine rationale Nullstelle. Für (p,q) = 1 ist aber
$8(\frac{p}{q})^3 - 6(\frac{p}{q}) - 1 = 0 \Leftrightarrow 8p^3 - 6pq^2 - q^3 = 0$ Widerspruch zu (p,q) = 1.
Also ist $8x^3 - 6x - 1$ das Minimalpolynom von x = cos 20° und damit [L_0(x):L_0] = 3 und dies ist keine Zweierpotenz.

Konstruktion von regelmäßigen n-Ecken
Gegeben sei ein Kreis. In diesen soll ein regelmäßiges n-Eck einbeschrieben werden. Also M = {0,1} und die Frage ist, ob $e^{2\pi i/n} \in$ kon M ist.
Wissen, dass dies für n = 2,3,4,5,6 geht, nicht aber z.B. für n = 9 (da 20° nicht konstruiert werden kann).
Zur ultimativen Lösung des Problems brauchen wir noch zwei Begriffe. Für eine natürliche Zahl n bezeichnet man mit φ(n) die Anzahl der zu n teilerfremden Zahlen aus {1,2,...,n-1}.
Eine natürliche Zahl heißt *Fermatsche Primzahl*, wenn sie eine Primzahl von der Form $2^{2^m}+1$ ist.
Haben $F_0 = 2^{2^0}+1 = 3$, $F_1 = 2^{2^1}+1 = 5$, $F_2 = 2^{2^2}+1 = 17$, $F_3 = 2^{2^3}+1 = 257$, $F_4 = 2^{2^4}+1 = 65.537$,
F_6 ist aber durch 641 teilbar (Euler).
→ heutige Vermutung: Fermatsche Primzahlen nur von F_0 bis F_4.

Satz 3 *Gauß 1796*
Sei n ≥ 3 eine natürliche Zahl. Dann sind folgende Aussagen äquivalent:
(i) das regelmäßige n-Eck kann mit Zirkel und Lineal konstruiert werden.
(ii) $\varphi(n)$ ist eine Zweierpotenz.
(iii) $n = 2^m p_1 \ldots p_r$, wobei $m \in$ {0,1,2,...} ist und $p_1, \ldots, p_r$ verschiedene Fermatsche Primzahlen sind.

Teilbeweis:
(i) ⇒ (ii).
Haben $L_0 = \mathcal{Q}(\mathrm{M}) = \mathcal{Q}$. Man kann zeigen, dass [$\mathcal{Q}(e^{2\pi i/n}$:$\mathcal{Q}$] = φ(n) gilt.
Behauptung folgt damit aus der Folgerung.
(ii) ⇒ (iii).
Ist n = $p_1^{l_1} \ldots p_m^{l_m}$ die Primfaktorenzerlegung, so ist
$\varphi(\mathrm{n}) = p_1^{l_1-1} \ldots p_m^{l_m-1}(p_1-1)\ldots(p_m-1)$
(Beweis: Man zeige zunächst φ(mn)=φ(m)φ(n) für (m,n) = 1 und zähle dann aus, um φ(p')=$p^{l-1}(p-1)$ zu zeigen.)
Für $p_j \geq 3$ ist $p_j^{l_j-1}(p_j-1)$ genau dann eine Zweierpotenz, wenn $l_j = 1$ und $p_j = 2^k+1$ ist.
Ist k = uv mit ungeradem u, so ist $2^k+1 = (2^v)^u + 1$ durch 2^v+1 teilbar. Also ist

$2^k + 1$ genau dann eine Primzahl, wenn es eine Fermatsche Primzahl ist (k=2^m).
(iii) ⇒ (i).
Ist der schwierigste Teil des Beweises. Er erfordert Mittel aus der Galoistheorie. #

Explizite Konstruktion:
n = 17 Gauß 1796
n = 257 Richelot 1832
n = 65.537 Hermes 1889

Ist (r,s) = 1, so existieren ganze Zahlen a,b mit ar + bs = 1. Es folgt $\frac{a}{s} + \frac{b}{r} = \frac{1}{rs}$ und somit

$$a\frac{2\pi}{s} + b\frac{2\pi}{s} = \frac{2\pi}{rs}.$$

Mit r-Eck und s-Eck ist also auch das rs-Eck konstruierbar.

4.3.8 Hauptsatz der Galoistheorie

Évariste Galois (1811 - 1832)

Definition 1 *Sei K ein Körper. Die Menge aller Automorphismen von K (d.h. aller Isomorphismen von K auf sich selbst) bildet eine Gruppe bezüglich der Hintereinanderausführung, die die Automorphismengruppe von K genannt und mit Aut(K) bezeichnet wird.*
Sei K ⊃ k eine Körpererweiterung. Dann ist
Aut(K,k):={φ ∈ Aut(K):φ|k = id} = {φ ∈ Aut(K): φ(a) = a ∀ a ∈ k}
eine Untergruppe von Aut(K), die man die Galoisgruppe der Körpererweiterung K ⊃ k nennt.
Ist k ein Körper und f(x) ∈ k[x]/k und ist K ⊃ k der Zerfällungskörper von f(x), so setzt man
Gal(f,k):=Aut(K,k) und nennt dies die Galoisgruppe des Polynoms f über k.

Bemerkung:
Sei f(x) ∈ k[x]\k und K ⊃ k der Zerfällungskörper. Bezeichnen mit N ⊂ K die Menge der Nullstellen von f(x) und setzen n = |N|.
Man kann zeigen, dass φ(N) = N für stets φ ∈ Gal(f,k) gilt und dass die Abbildung Gal(f,k) → S_n, $\varphi \mapsto \varphi|N$ injektiv ist. Man kann daher (und dies wird auch oft gemacht) Gal(f,k) mit einer Untergruppe von S_n identifizieren.

Hausaufgabe: Sei K ein Körper und P der Primkörper. Man zeige, dass Aut(K,P) = Aut(K) ist.

Definition 2 *Sei K ein Körper und G eine Untergruppe von Aut(K). Man nennt dann Fix(K,G):={a ∈ K:φ(a) = a ∀ φ ∈ G} den Fixkörper von G in K.*

Definition 3 *Eine Körpererweiterung K ⊃ k heißt Galoiserweiterung, wenn eine endliche Untergruppe g ⊂ Aut(K) existiert, sodass k = Fix(K,g) ist.*

1. Beispiel:

$\mathcal{Q}(\sqrt[3]{2}) \supset \mathcal{Q}$ ist keine Galoiserweiterung.
$\mathcal{Q}(\sqrt[3]{2}) = \{a + b\sqrt[3]{2} + c(\sqrt[3]{2})^2 : a,b,c \in \mathcal{Q}\}$
Bestimmen zunächst $\mathrm{Aut}(\mathcal{Q}(\sqrt[3]{2}))$. Sei $\varphi \in \mathrm{Aut}(\mathcal{Q}(\sqrt[3]{2}))$. Nach Hausaufgabe ist $\varphi|\mathcal{Q} = \mathrm{id}$. Also ist $2 = \varphi(2) = \varphi(\sqrt[3]{2})^3 = (\varphi(\sqrt[3]{2}))^3$, und einziges $\alpha \in \mathcal{R}$ mit $\alpha^3 = 2$ ist $\alpha = \sqrt[3]{2}$.
Somit ist $\varphi(\sqrt[3]{2}) = \sqrt[3]{2}$ und damit $\varphi = \mathrm{id}$.
Haben also gezeigt, dass $\mathrm{Aut}(\mathcal{Q}(\sqrt[3]{2})) = \{\mathrm{id}\}$. Einzige Untergruppe ist $g = \{\mathrm{id}\}$ und
$\mathrm{Fix}(\mathcal{Q}(\sqrt[3]{2}),\{\mathrm{id}\}) = \{x \in \mathcal{Q}(\sqrt[3]{2}) : \mathrm{id}(x) = x\} = \mathcal{Q}(\sqrt[3]{2}) \neq \mathcal{Q}$
$\rightarrow$ ist keine Galoiserweiterung. #

2. Beispiel:
$\mathcal{Q}(i) \supset \mathcal{Q}$ ist Galoiserweiterung:
Haben nach HA: $-1 = \varphi(-1) = \varphi(i^2) = (\varphi(i))^2 \; \forall \; \varphi \in \mathrm{Aut}(\mathcal{Q}(i))$. Es folgt $\varphi(i) = i$ oder $\varphi(i) = -i$. Abbildungen
$\varphi_1(a+bi) = a+bi$, d.h. $\varphi_1(z) = z$,
$\varphi_2(a+bi) = a-bi$, d.h. $\varphi_2(z) = \overline{z}$
sind wirklich Automorphismen. Somit ist $\mathrm{Aut}(\mathcal{Q}(i)) = \{\varphi_1,\varphi_2\}$.
Nehmen $g = \{\varphi_1,\varphi_2\}$. Dann ist
$\mathrm{Fix}(\mathcal{Q}(i),g) = \{z \in \mathcal{Q}(i) : \varphi_1(z) = z,\ \varphi_2(z) = \overline{z}\} = \{z \in \mathcal{Q}(i) : \overline{z} = z\} = \mathcal{Q}$. #

Formulieren den Hauptsatz der Galoistheorie in drei einzelnen Sätzen. Zum Beweis verweisen wir auf die Literatur (z.B. Fischer/Sacher „Einführung in die Algebra“).

Satz 1 *Sei $K \supset k$ eine Galoiserweiterung, d.h. $k = Fix(K,g)$ mit $g \subset Aut(K)$ endlich. Dann gilt:*
(a) $g = Aut(K,k)$;
(b) Ist $\mathcal{K}$ die Menge aller Zwischenkörper $K \supset L \supset k$ und $\mathcal{G}$ die Menge aller Gruppen $\{e\} \subset G \subset g$ (d.h. die Menge aller endlichen Untergruppen von g), so sind die beiden Abbildungen
$Aut(K,\bullet): \mathcal{K} \to \mathcal{G},\ L \mapsto Aut(K,L)$,
$Fix(K,\bullet): \mathcal{G} \to \mathcal{K},\ G \mapsto Fix(K,G)$
bijektiv und zueinander invers, d.h.
$Aut(K,Fix(K,G)) = G$,
$Fix(K,Aut(K,L)) = L$.

Sei $K \supset k$ eine Galoiserweiterung. Betrachten eine Körperkette
$K \supset L_1 \supset L_2 \supset k$.
Anwendung von $\mathrm{Aut}(K,\bullet)$ liefert dann eine Kette von Gruppen:
$\underbrace{Aut(K,K)}_{\{e\}} \subset \mathrm{Aut}(K,L_1) \subset \mathrm{Aut}(K,k) = g$
Anwendung von $\mathrm{Fix}(K,\bullet)$ und Satz 1 liefern
$\mathrm{Fix}(K,\{e\}) \supset \mathrm{Fix}(K,\mathrm{Aut}(K,L_1)) \supset \mathrm{Fix}(K,\mathrm{Aut}(K,L_2)) \supset \mathrm{Fix}(K,g)$
$K \supset L_1 \supset L_2 \supset k$.

Das heißt, wir erhalten die ursprüngliche Körperkette zurück. Starten wir mit einer Kette von endlichen Untergruppen von Aut(K),
$\{e\} \subset G_1 \subset G_2 \subset g = \mathrm{Aut}(K,k)$
und wenden wir $\mathrm{Fix}(K,\bullet)$ darauf an, so entsteht die Körperkette
$\mathrm{Fix}(K,\{e\}) \supset \mathrm{Fix}(K,G_1) \supset \mathrm{Fix}(K,G_2) \supset \mathrm{Fix}(K,g)$.
Wenden wir darauf $\mathrm{Aut}(K,\bullet)$ an, so ergibt sich

Aut(K,K) $\subset$ Aut(K,Fix(K,G_1)) $\subset$ Aut(K,Fix(K,G_2)) $\subset$ Aut(K,k)
{e} $\subset G_1 \subset G_2 \subset$ g,
d.h. erhalten die ursprüngliche Gruppenkette zurück.

Fazit: Haben eine Konstruktion Aut, die aus Körperketten Gruppenketten macht und eine Konstruktion Fix, die Gruppenketten in Körperketten verwandelt. Dabei gilt Fix $\circ$ Aut = id, Aut $\circ$ Fix = id.

Wie hängen nun Eigenschaften der Körperkette mit Eigenschaften der Gruppenkette zusammen?

Satz 2 *Sei K $\supset$ k eine Galoiserweiterung, d.h. k = Fix(K,g) mit g $\subset$ Aut(K) endlich. Dann gilt:*
(a) K $\supset$ k ist eine endliche Körpererweiterung;
(b) Ist K $\supset$ L $\supset$ k ein Zwischenkörper, so ist
(K:L) = |Aut(K,L)|/|Aut(K,K)| = |Aut(K,L)|,
(L:k) = |Aut(K,k)|/|Aut(K,L)|,
wobei |G| die Anzahl der Elemente einer Gruppe G ist (d.h. die Ordnung von G).

Sei K $\supset$ L $\supset$ k wie in Satz 2. Haben dann
[K:L] = m und [L:k] = n und
Aut(K,K) $\subset$ Aut(K,L) $\subset$ Aut(K,k)
{e} $\subset$ Aut(K,L) $\subset$ g
Ordnung = 1 Ordnung = p Ordnung = q.
Satz 2 liefert m = $\frac{p}{1}$, n = $\frac{q}{p}$.

Bemerkung:
Die Ordnung einer Untergruppe ist stets Teiler der Ordnung der Gruppe.

Satz 2 ist nur für einen Zwischenkörper anwendbar, für mehrere Zwischenkörper haben wir den folgenden Satz:

Satz 3 *Sei K $\supset$ k eine Galoiserweiterung, d.h. k = Fix(K,g) mit g $\subset$ Aut(K) endlich. Sei L ein Zwischenkörper, K $\supset$ L $\supset$ k. Dann gilt:*
(a) K $\supset$ L ist eine Galoiserweiterung.
(b) L $\supset$ k ist eine Galoiserweiterung genau dann, wenn Aut(K,L) ein Normalteiler von Aut(K,k) ist;
(c) Wenn L $\supset$ k eine Galoiserweiterung ist, dann ist Aut(L,k) $\cong$ Aut(K,k)/Aut(K,L).

Zur Erläuterung:
Sei K $\supset$ k wie in Satz 3. Betrachten die Körperkette
K $\supset L_1 \supset L_2 \supset$ k mit [K:L_1] = m, [L_1:L_2] = n, [L_1:k] = l,
Wenden Aut(K,$\bullet$) an:
Aut(K,K) $\supset$ Aut(K,L_1) $\supset$ Aut(K,L_2) $\supset$ Aut(K,k)
{e} $\subset G_1 \subset G_2 \subset$ g
Ordnung 1 Ordnung p Ordnung q Ordnung r.
Satz 3 liefert, dass K $\supset L_1$ eine Galoiserweiterung ist, nach Satz 2 ist also m = p. Nach

Satz 3 ist K $\supset L_2$ eine Galoiserweiterung und nach Satz 2 somit
[K:L_2] = q,
[K:L_2] = [K:L_1]*[L_1:L_2] = m * n
$\Rightarrow$ q = pn $\Rightarrow$ n = $\frac{q}{p}$.
Analog:
mnl = [K:k] = r (Satz 2)
p * $\frac{q}{p}$ * l = r $\Rightarrow$ l = $\frac{r}{q}$.

Haben $L_1 \supset L_2$ ist Galoiserweiterung $\Leftrightarrow$ G_1 ist Normalteiler von G_2.
(Satz 3b auf K $\supset L_1 \supset L_2$ anwenden)
$L_2 \supset$ k ist Galoiserweiterung $\Leftrightarrow$ G_2 ist Normalteiler von g.

Ist in der Körperkette jeder Körper eine Galoiserweiterung des folgenden, so ist in der Kette der Galoisgruppen jede ein Normalteiler der folgenden. Umgekehrt, ist in der Gruppenkette jede Gruppe Normalteiler der folgenden, so ist in der Körperkette jeder Körper Galoiserweiterung des folgenden.

Soviel zu Teil b von Satz 3. Nun zu Teil c:
Haben K $\supset$ L $\supset$ k = Fix(K,g) ,
{e} $\subset$ Aut(K,L) $\subset$ g = Aut(K,k).
Sei L $\supset$ k eine Galoiserweiterung. Anwendung von Aut(L,$\bullet$) ergibt
{e} = Aut(L,L) $\subset$ Aut(L,k)
und Satz 3c liefert
Aut(L,k) $\cong$ g/Aut(K,L).

Allgemeiner: Sind alle $\supset$ Galoiserweiterungen, so ist
K $\supset L_1 \supset L_2 \supset$ k
{e} $\subset$ Aut(K,L_1) $\subset$ Aut(K,L_2) $\subset$ Aut(K,k)
{e} $\subset$ Aut(L_1,L_2) $\subset$ Aut(L_1,k)
{e} $\subset$ Aut(L_2,k)

Hier ist eine erste Anwendung des Hauptsatzes der Galoistheorie:

Satz 4 *Gauß 1796*
Ist $p = 2^{2^n} + 1$ eine Primzahl, so lässt sich das reguläre p-Eck mit Zirkel und Lineal konstruieren.

Bemerkung:
Dies ist der noch unbewiesene Teil von Satz 3/4.3.7.

Beweis:
Haben M = {0,1} und wollen $\epsilon := e^{2\pi i/p} \in$ kon M zeigen. Wissen (Satz 2/4.3.7.), dass dies genau gilt, wenn es eine Kette von Körpern $\mathcal{Q}$(M) = $\mathcal{Q} = L_0 \subset L_1 \subset \ldots \subset L_m$ gibt, sodass [L_{i+1}:L_i] = 2 (i=0,...,n-1) und $\epsilon \in L_m$ ist.
Zeigen zunächst, dass $\mathcal{Q}(\epsilon) \supset \mathcal{Q}$ eine Galoiserweiterung ist.
Haben Aut($\mathcal{Q}(\epsilon)$) = Aut($\mathcal{Q}(\epsilon)$,$\mathcal{Q}$) (HA).
Ist also $\varphi \in$ Aut($\mathcal{Q}(\epsilon)$), so ist $1 = \varphi(1) = \varphi(\epsilon^p) = (\varphi(\epsilon))^p$,
d.h. $\varphi(\epsilon) = \epsilon^l$ mit l = 1,...,p-1 (l = 0 scheidet aus, da dann Im $\varphi = \mathcal{Q}$ wäre).
Haben $\epsilon^p = 1$ und damit $\epsilon^p - 1 = (\epsilon - 1)(1 + \epsilon + \ldots + \epsilon^{p-1}) = 0$.
Also ist $1+\epsilon+\ldots+\epsilon^{p-1} = 0$. Man kann zeigen, dass $1+x+\ldots+x^{p-1}$ über $\mathcal{Q}$ irreduzibel ist, und zwar für jede Primzahl p (Eisensteinsches Kriterium und Substitutionsmetho-

de). Somit ist $1 + x + \ldots + x^{p-1}$ das Minimalpolynom von ϵ und es gilt

$$\mathcal{Q}(\epsilon\ )\ \mathcal{Q}[\epsilon] = \{a_0 + a_1\epsilon + \ldots + a_{p-2}\epsilon^{p-2}\text{: } a_j \in \mathcal{Q}\}.$$

Daraus folgt, dass wirklich jedes $\epsilon \mapsto \epsilon^l$ ein Isomorphismus ist. Also Aut($\mathcal{Q}(\epsilon)$) = $\{\varphi_1, \ldots, \varphi_{p-1}\}$ =:g mit $\varphi_l(\epsilon) = \epsilon^l$.
Bestimmen nun Fix($\mathcal{Q}(\epsilon)$,g).
Sei $\lambda = a_0 + a_1\epsilon + \ldots + a_{p-2}\epsilon^{p-2} \in$ Fix($\mathcal{Q}(\epsilon)$,g).
Setzen a(x) := $a_0 + a_1x + \ldots + a_{p-2}x^{p-2} \in \mathcal{Q}$[x]. Dann ist
$\lambda = \varphi_l(x) = a(\epsilon^l)$ für l=1,...,p-1.

Das Polynom a(x)-λ hat den Grad p - 2 und die p - 1 Nullstellen ϵ^l(l=1,...,p-1). Also ist a(x)-$\lambda = a_0 - \lambda + a_1x + \ldots + a_{p-2}x^{p-2}$ das Nullpolynom.

Insbesondere ist $\lambda = a_0 \in \mathcal{Q}$. Damit ist gezeigt, dass $\mathcal{Q}$ = Fix($\mathcal{Q}(\epsilon)$,g) und somit $\mathcal{Q}(\epsilon) \supset \mathcal{Q}$ eine Galoiserweiterung ist.
Haben g = $\{\varphi_1, \ldots, \varphi_{p-1}\} \cong \mathcal{Z}_{p-1} = \mathcal{Z}_{2^{2^n}}$. Haben damit folgende Gruppenkette:
$\{e\} = \mathcal{Z}_1 \subset \mathcal{Z}_2 \subset \mathcal{Z}_4 \subset \mathcal{Z}_8 \subset \ldots \subset \mathcal{Z}_{2^{2^n}} \cong$ g.
Wenden Fix an und erhalten die Körperkette
Fix($\mathcal{Q}(\epsilon)$,{e}) $\supset \ldots \supset$ Fix($\mathcal{Q}(\epsilon)$,$\mathcal{Z}_{2^k}$) $\supset$ Fix($\mathcal{Q}(\epsilon$,$\mathcal{Z}_{2^{k+1}}$) $\supset \ldots \supset$ Fix($\mathcal{Q}(\epsilon)$,g)
$\mathcal{Q}(\epsilon)) \supset \ldots \supset L_k \supset L_{k+1} \supset \ldots \supset \mathcal{Q}$.
Nach unserem Satz 2 ist
$[L_k\text{':}L_{k+1}\text{']} = |\mathcal{Z}_{2^{k+1}}|/|\mathcal{Z}_{2^k}| = \frac{2^{k+1}}{2^k} = 2$ und $\epsilon \in \mathcal{Q}(\epsilon)$. #

4.3.9 Auflösung von algebraischen Gleichungen über Radikale

Eine algebraische Gleichung ist eine Gleichung der Form

$$a_nx^n + \ldots + a_1x + a_0 = 0\ (\text{n} \geq 1,\ a_n \neq 0)$$

mit $a_0, a_1, \ldots, a_n \in$ k.
Für n ≤ 4 lässt sich eine solche Gleichung stets über Formeln lösen, die nur die vier Grundrechenoperationen und Wurzelziehen enthalten.
n = 2:
$ax^2 + bx + c = 0$
$x^2 + px + q = 0$
$x_{1/2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}$
n = 3:
$ax^3 + bx^2 + cx + d = 0$ 1515 Tartaglia, 1545 Cardaro
$x = y - \frac{b}{3a}$
$y^3 + 3py + 2q = 0$
$\{y_1, y_2, y_3\} = \{\text{u+v},\ \epsilon u + \epsilon^2 v,\ \epsilon^2 u + \epsilon v\}$
mit $\epsilon = e^{2\pi i/3}$ und u = $\sqrt[3]{-q + \sqrt{q^2 + p^3}}$, v = $\sqrt[3]{-q - \sqrt{q^2 - p^3}}$.
n = 4: 1540 Ferrari

Ziel dieses Kapitals ist zu zeigen, dass es für n ≥ 5 keine solchen Formeln gibt.

Definition 1 *Eine Körpererweiterung $K \supset k$ heißt Radikalerweiterung, wenn es eine Zwischenkörperkette*

$$K = L_m \supset L_{m-1} \supset \ldots \supset L_0 = k$$

mit $L_{i+1} = L_i(b_i)$, $b_i \in L_{i+1}$, $b_i^{n_i} \in L_i$ (i=0,...,m-1, $n_i \in \{2,3,4,\ldots\}$) gibt.

Definition 2 *Ein Polynom f(x) ∈ k[x]/k heißt über k durch Radikale lösbar, wenn es eine Radikalerweiterung $K \supset k$ gibt, sodass f(x) über K in Linearfaktoren zerfällt.*
Anders gesagt: f(x) ∈ k[x]/k ist über k durch Radikale lösbar, wenn es eine Radikalerweiterung $K \supset k$ mit $K \supset L \supset k$ gibt, wobei L der Zerfällungskörper von f(x) ist.

Definition 3 *Eine Gruppe G heißt auflösbar, wenn es eine Gruppenkette*

$$G = G_0 \supset G_1 \supset \ldots \supset G_m = \{e\}$$

gibt, sodass G_{i+1} Normalteiler von G_i und G_i/G_{i+1} abelsch ist (i=0,...,m-1).

Abelsche Gruppen sind stets auflösbar: G ⊃ {e}.
S_3 ist auflösbar: $S_3 \supset A_3 \supset$ {e}.

Satz 1 *Sei k ein Körper der Charakteristik 0. Für ein Polynom f(x) ∈ k[x]/k sind dann folgende Aussagen äquivalent:*
(i) f(x) ist über k durch Radikale lösbar,
(ii) Gal(f,k) ist auflösbar.

Zur Erinnerung (Definition 1/4.3.8.): Gal(f,k) = Aut(L,K), wobei L der Zerfällungskörper von f(x) ist.

Beschränken uns darauf, den Beweis von (i) ⇒ (ii) zu skizzieren. Brauchen dazu einige Vorbereitungen.

Definition 4 *Sei G eine Gruppe. Ein Element der Form $aba^{-1}b^{-1}$ heißt Kommutator. Die kleinste Untergruppe von G, die alle Komutatoren enthält (= der Durchschnitt aller Untermengen von G, die die Menge aller Kommutatoren enthalten), heißt Kommutatorgruppe von G und wird mit K(G) bezeichnet.*

Offenbar gilt: K(G)={e} ⇔ G abelsch ($aba^{-1}b^{-1} = e \Leftrightarrow ab = ba$)
Es ist leicht zu sehen, dass K(G) = $\{[a_1,b_1]\ldots[a_n,b_n]: a_1,\ldots,a_n,b_1,\ldots,b_n \in G\}$ gilt, hier ist [a,b]:=$aba^{-1}b^{-1}$.

Lemma 1:
(a) Kommutatorgruppe ist stets ein Normalteiler von G.
(b) Sei N ein Normalteiler von G. Dann gilt G/N abelsch $\Leftrightarrow$ K(G) $\subset$ N.

Beweis:
(a) Sei x $\in$ G. Müssen zeigen, dass $xK(G)x^{-1} \subset K(G)$ ist. Sei also c = $[a_1,b_1]\ldots[a_n,b_n]$ $\in$ K(G). Dann ist

$$xcx^{-1} = x[a_1,b_1]x^{-1}x[a_2,b_2]x^{-1}\ldots x[a_n,b_n]x^{-1} = [xa_1x^{-1}, xb_1x^{-1}]\ldots[xa_nx^{-1}, xb_nx^{-1}] \in K(G).$$

$(x[a,b]x^{-1} = xaba^{-1}b^{-1} = xax^{-1}xbx^{-1}xa^{-1}x^{-1}xb^{-1}x^{-1} = (xax^{-1})(xbx^{-1})(xax^{-1})(xbx^{-1})^{-1})$.
(b) Sei G/N abelsch. Für a,b $\in$ G ist dann

$$aba^{-1}b^{-1}N = aNbNa^{-1}Nb^{-1}N = aNa^{-1}NbNb^{-1}N = aa^{-1}bb^{-1}N = N,$$

d.h. $aba^{-1}b^{-1} \in$ N.
Umgekehrt sei K(G) $\subset$ N. Für a,b $\in$ G ist dann

$$aNbNa^{-1}Nb^{-1}N = aba^{-1}b^{-1}N \subset NN = N,$$

das heißt

$$aNbNa^{-1}Nb^{-1}NbNaN = bNaN \Rightarrow aNbN = bNaN,$$

d.h. G/N ist abelsch. #

Das Bilden von Kommutatorgruppen kann iteriert werden:
$K^0(G) = G$,
$K^1(G) = K(G)$,
$K^2(G) = K(K(G))$,
$\vdots$
$K^{i+1}(G) = K(K^i(G))$

Satz 2 *Eine Gruppe G ist genau dann auflösbar, wenn es ein m mit* $K^m(G) = \{e\}$ *gibt.*

Beweis:
Sei $K^m(G) = \{e\}$. Haben dann $G = K^0(G) \supset K^1(G) \supset \ldots \supset K^m(G) = \{e\}$,
nach Lemma 1(a) ist $K^{i+1}(G) = K(K^i(G))$ ein Normalteiler von $K^i(G)$ und wegen $K^{i+1}(G) \supset K(K^i(G))$ ist nach Lemma 1(b) ist die Gruppe $K^i(G)/K^{i+1}(G)$ abelsch. Also ist G auflösbar.
Umgekehrt sei $G = G_0 \supset G_1 \supset \ldots \supset G_m = \{e\}$ mit G_{i+1} Normalteiler von G_i und G_i/G_{i+1} abelsch. Behaupten, dass $K^i(G) \subset G_i$ gilt; dies liefert $K^m(G) = \{e\}$. Für i = 0 ist $K^0(G) = G$ und $G_0 = G$. Sie die Behauptung für i wahr, zeigen sie für i+1.
Haben $K^{i+1}(G) = K(K^i(G)) \subset K(G_i)$.
Da G_{i+1} Normalteiler von G_i und G_i/G_{i+1} abelsch ist, liefert Lemma 1(b), dass $G_{i+1} \supset K(G_i)$ ist. Also ist $K^{i+1}(G) \subset G_{i+1}$. #

Lemma 2: Untergruppen und homomorphe Bilder von auflösbaren Gruppen sind wieder auflösbar.

Beweis:
Sei G auflösbar und H $\subset$ G eine Untergruppe. Dann ist $K^m(H) \subset K^m(G)$ und Satz 2 liefert, dass H auflösbar ist.

Ist G auflösbar und φ: G $\to$ H ein surjektiver Homomorphismus, so haben wir

$$\varphi(aba^{-1}b^{-1}) = \varphi(a)\varphi(b)\varphi(a)^{-1}\varphi(b)^{-1}$$

und dies ergibt φ(K(G)) = K(H). Iteration liefert $\varphi(K^m(G)) = K^m(H)$. Nach Satz 2 ist also H auflösbar. #

Skizze des Beweises von (i) $\Rightarrow$ (ii) aus Satz 1.
Sei L der Zerfällungskörper von f(x) $\in$ k[x]\k und K $\supset$ k eine Radikalerweiterung mit K $\supset$ L $\supset$ k. Haben dann
K = $L_m \supset L_{m-1} \supset \ldots \supset L_0$ = k mit $L_{i+1} = L_i(b_i)$, $b_i \in L_{i+1}$, $b_i^{p_i} \in L_i$, wobei p_i eine Primzahl ist (ansonsten Kette verfeinern), und $b_i^k \notin L_i$ für 1$\leq$k$\leq$ p_i-1.
Man kann zeigen, dass dann $x^{p_i} - b_i^{p_i}$ das Minimalpolynom von b_i über L_i ist.
Man kann desweiteren o.B.d.A. annehmen, dass die Erweiterungen K $\supset L_i$, $L_{i+1} \supset L_i$, L $\supset$ k alle Galoiserweiterungen sind (nicht trivial).
Wenden Aut an und erhalten die Gruppenkette
Aut(K,K) $\subset \ldots \subset$ Aut(K,L_{i+1}) $\subset$ Aut(K,L_i) $\subset \ldots \subset$ Aut(K,k).
Betrachten K $\supset L_{i+1} \supset L_i$
{e} $\subset$ Aut(K,L_{i+1}) $\subset$ Aut(K,L_i)
Da $L_{i+1} \supset L_i$ eine Galoiserweiterung ist, ist nach Teil 3 des Hauptsatzes Aut(K,L_{i+1}) ein Normalteiler von Aut(K,L_i) und nach Teil 2 des Hauptsatzes folgt
|Aut(K,L_i)/Aut(K,L_{i+1})| = |Aut(K,L_i)|/|Aut(K,L_{i+1})| = $[L_{i+1}:L_i] = [L_i(b_i):L_i] = p_i$.
Jede Gruppe mit p_i Elementen ist zu $\mathcal{Z}_{p_i}$ isomorph und damit abelsch. Haben also gezeigt, dass Aut(K,k) auflösbar ist.
Haben schließlich K $\supset$ L $\supset$ k
Aut(K,K) $\supset$ Aut(K,L) $\supset$ Aut(K,k).
Nach Teil 3 des Hauptsatzes ist Aut(K,L) ein Normalteiler von Aut(K,k) und
Aut(L,k) $\cong$ Aut(K,k)/Aut(K,L).
Damit ist Aut(L,k) isomorph zum Bild des Homomorphismus π:Aut(K,k) $\to$ Aut(K,k)/Aut(K,L), a $\mapsto$ a Aut(K,L).
Lemma 2 ergibt also, dass Gal(f,k) = Aut(L,k) auflösbar ist. #

Hatten fürher vermerkt, dass man Gal(f,k) mit einer Untergruppe von S_n identifizieren kann.
S_1,S_2 sind abelsch und damit auflösbar. Haben $S_3 \supset A_3 \supset$ {e}, $|S_3/A_3| = 6/3 = 2$, $|A_3| = 3$ abelsch.
(K(S_3) = A_3, K(A_3) = {e})
$S_4 \supset A_4 \supset \mathcal{Z}_2 \times \mathcal{Z}_2 \supset$ {e}, $|S_4/A_4| = 24/12 = 2$, $|A_4/\mathcal{Z}_2 \times \mathcal{Z}_2| = 12/4 = 3$, $\mathcal{Z}_2 \times \mathcal{Z}_2$ abelsch
(K(S_4) = A_4, K(A_4) $\cong \mathcal{Z}_2 \times \mathcal{Z}_2$, K($\mathcal{Z}_2 \times \mathcal{Z}_2$) = {e})
Somit ist S_3 und S_4 auflösbar.

Satz 3 *Sei k ein Körper mit der Charakteristik 0. Dann lässt sich jedes Polynom f(x) $\in$ k[x]/k vom Grad $\leq$ 4 über k durch Radikale lösen.*

Beweis:
Haben Gal(f,k) $\subset S_n$ mit n $\leq$ 4. Nach Lemma 2 ist Gal(f,k) auflösbar und Satz 1 liefert dann die Behauptung. #

Satz 4 *Für n $\geq$ 5 ist S_n nicht auflösbar.*

Beweis:
Wegen $S_5 \subset S_n$ reicht es nach Lemma 2 zu zeigen, dass S_5 nicht auflösbar ist. Nehmen an, dass S_5 auflösbar ist:
$S_5 = G_0 \supset G_1 \supset \ldots \supset G_m = \{e\}$ mit G_{i+1} Normalteiler von G_i und G_i/G_{i+1} abelsch.
Ein Dreierzyklus ist eine Permutation der Form (abc). Zeigen folgendes: Ist G eine Untergruppe von S_5 und N ein Normalteiler von G (G/N abelsch) und enthält G alle Dreierzyklen, dann enthält auch N alle Dreierzyklen. Dies liefert offenbar die Behauptung.
Sei (abc) $\in$ G. Setzen x = (abd), y = (ace). Dann ist

$$xyx^{-1}y^{-1}N = xNyNx^{-1}Ny^{-1}N = xNx^{-1}NyNy^{-1}N = xx^{-1}yy^{-1}N = N,$$

d.h. $xyx^{-1}y^{-1} \in$ N. Schließlich ist

$$xyx^{-1}y^{-1} = (abd)(ace)(dba)(eca) = (e)(cab)(d) = (cab) = (abc).$$

Also ist (abc) $\in$ N. Enthält G also alle Dreierzyklen, so tut dies auch N. #

Um zu zeigen, dass sich Gleichungen vom Grad n $\geq$ 5 im Allgemeinen nicht durch Radikale lösen lassen, braucht man nur Polynome f(x) $\in$ k[x]\k mit Gal(f,k) = S_n zu finden.

Satz 5 *Es gilt: Gal($x^5 - x - 1$,$\mathcal{Q}$) = S_5.*

Beweis: Literatur. #

Insbesondere lässt sich $x^n(x^5 - x - 1) = 0$ (Grad n + 5) nicht über $\mathcal{Q}$ durch Radikale lösen.
Beweisen einen schwächeren Satz.

Definition 5 *Sei f(x) = $x^n + a_{n-1}x^{n-1} + \ldots + a_0 \in \mathcal{C}[x]$. Dann wird $\mathcal{Q}(a_0)\ldots(a_{n-1})$ der Koeffizientenkörper von f(x) genannt.*

Satz 6 *Abel 1824*
Es gibt Polynome 5. Grades, die über ihrem Koeffizientenkörper nicht durch Radikale lösbar sind.

Beweis:
Sei $y_1 \in \mathcal{R}$ transzendent über $\mathcal{Q}$. Dann ist $\mathcal{Q}(y_1)$ abzählbar. Somit gibt es ein $y_2 \in \mathcal{R}$, das über $\mathcal{Q}(y_1)$ transzendent ist. Dann sei $y_3 \in \mathcal{R}$ transzendent über $\mathcal{Q}(y_1)(y_2)$ usw. Schließlich sei $y_5 \in \mathcal{R}$ transzendent über $\mathcal{Q}(y_1)(y_2)(y_3)(y_4)$. Setzen
$\sigma_1 = y_1 + \ldots + y_5$, $\sigma_2 = y_1y_2 + y_1y_3 + \ldots + y_4y_5$, $\sigma_3 = y_1y_2y_3 + y_1y_2y_3 + \ldots + y_3y_4y_5$, $\sigma_4 = y_1y_2y_3y_4 + \ldots$, $\sigma_5 = y_1y_2y_3y_4y_5$.
Sei f(x) = $(x-y_1)\ldots(x-y_5) = x^5 - \sigma_1x^4 + \sigma_2x^3 - \sigma_3x^2 + \sigma_4x - \sigma_5$. Für dieses Polynom ist Z = $\mathcal{Q}(y_1)\ldots(y_5)$ der Zerfällungskörper und K = $\mathcal{Q}(\sigma_1)\ldots(\sigma_5)$ der Koeffizientenkörper.
Zeigen, dass Gal(f,K) = Aut(Z,K) nicht auflösbar ist.
Sei $\pi \in S_5$. Betrachten die Abbildung φ_π: Z $\to$ Z, r($y_1 \ldots y_5$) $\mapsto$ r($y_{\pi(1)}, \ldots, y_{\pi(5)}$).
Es ist leicht zu sehen, dass $\varphi_\pi \in$ Aut(Z) ist und es ist klar, dass $\varphi_\pi|$K = id ist. Somit

ist $\varphi_\pi \in$ Aut(Z,K). Zeigen, dass $\varphi_\pi \neq \varphi_\tau$ für $\pi \neq \tau$ ist. Dies ergibt, dass Aut(Z,K) eine Kopie von S_5 enthält und somit nicht auflösbar sein kann.
Sei $\varphi_\pi = \varphi_\tau$. Dann ist

$$r(y_\pi(1) \dots y_{\pi(5)}) = r(y_{\tau(1)} \dots y_{\tau(5)})$$

für alle r $\in \mathcal{Q}(x_1)\dots(x_5)$ und insbesondere für

$$r(x_1, \dots, x_5) = x_1 + x_2^2 + x_3^3 + x_4^4 + x_5^5.$$

Das heißt,

$$y_{\pi(1)} + y_{\pi(2)}^2 + \dots + y_{\pi(5)}^5 = y_{\tau(1)} + y_{\tau(2)}^2 + \dots + y_{\tau(5)}^5.$$

Links und rechts in dieser Gleichung steht y_5 genau einmal. Erhalten ein Polynom mit Koeffizienten aus $\mathcal{Q}(y_1)\dots(y_4)$ mit y_5 als Nullstelle. Da y_5 transzendent über $\mathcal{Q}(y_1)\dots(y_4)$ ist, muss dieses Polynom das Nullpolynom sein, d.h. muss y_5 auf beiden Seiten mit der gleichen Potenz vorkommen. Ist diese Potenz j, so folgt π(j) = τ(j). Streichen y_5 auf beiden Seiten und wiederholen diese Überlegung mit y_4,y_3,y_2 und y_1. Es ergibt sich π(k) = τ(k) $\forall$ k $\in$ {1,2,3,4,5}. Damit ist $\pi = \tau$. #

5 Tensoren

Definition 1 *Seien $V_1, \dots, V_r$ lineare Räume über $\mathcal{R}$. Bezeichnen mit $\mathcal{B}(V_1, \dots, V_R)$ die Menge aller multilinearen Abbildungen f: $V_1 \times \dots \times V_r \to \mathcal{R}$, d.h. aller Abbildungen f: $V_1 \times \dots \times V_r \to \mathcal{R}$, die linear sind, wenn man beliebige r-1 Variablen fixiert. Solche Abbildungen heißen Multilinearformen.*

$\mathcal{B}(\mathcal{R}^n)$: $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \sum_{i=1}^n a_i x_i$

$\mathcal{B}(\mathcal{R}^n,\mathcal{R}^m)$: $\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \right) \mapsto \sum_{i=1}^n \sum_{j=1}^m a_{ij} x_i y_j$

$\mathcal{B}(\mathcal{R}^n,\mathcal{R}^m,\mathcal{R}^k)$: (x,y,z) $\mapsto \sum_{i=1}^n \sum_{j=1}^m \sum_{l=1}^k a_{ijl} x_i y_j z_l$
Offenbar ist $\mathcal{B}(V_1, \dots, V_r)$ ein linearer Raum.

Definition 2 *Sei V ein linearer Raum über $\mathcal{R}$. Die Elemente aus*

$$T_p^q(V) := \mathcal{B}(\underbrace{V, \dots, V}_{p}, \underbrace{V^*, \dots, V^*}_{q})$$

heißen Tensoren der Stufe p+q mit p kovarianten und q kontravarianten Variablen (oder Indizes) oder kurz Tensoren vom Typ(p,q).

T_1(V) = T_1^0(V) = $\mathcal{B}$(V) = V^*, d.h. Tensoren vom Typ (1,0) sind Linearformen (= lineare Funktionale)
T_2(V) = T_2^0(V) = $\mathcal{B}$(V,V), d.h. Tensoren vom Typ (2,0) sind Bilinearformen (z.B. Skalarprodukte)

Wichtige Übereinkunft in der Tensorrechnung:
V^{**} = V, denn es gibt einen kanonischen Isomorphismus ϵ: V $\rightarrow$ V**, $(\epsilon(\mathrm{x}))(\varphi) = \varphi(\mathrm{x})$.
$T^1(\mathrm{V}) = T^1_0(\mathrm{V}) = \mathcal{B}(V^*) = V^{**}$ = V, d.h. Tensoren vom Typ (0,1) sind Vektoren.
$T^2(\mathrm{V}) = T^2_0(\mathrm{V}) = \mathcal{B}(V^*,V^*)$, d.h. Tensoren vom Typ (0,2) sind Bilinearformen auf den Linearformen.
Weitere wichtige Übereinkunft:
$\mathcal{B}(\mathrm{U},V^*) = \mathcal{L}(\mathrm{U},\mathrm{V}) = \mathcal{L}(V^*,U^*)$
Kanonische Isomorphismen:
δ: $\mathcal{L}(\mathrm{U},\mathrm{V}) \rightarrow \mathcal{B}(\mathrm{U},V^*)$, $(\delta\mathrm{A})(\mathrm{x},\varphi) = \varphi(\mathrm{Ax})$
γ: $\mathcal{L}(\mathrm{U},\mathrm{V}) \rightarrow \mathcal{L}(\mathrm{V}^*,\mathrm{U}^*)$, $((\gamma\mathrm{A}),(\varphi))(\mathrm{x}) = \varphi(\mathrm{Ax})$.
$T^1_1(\mathrm{V}) = \mathcal{B}(\mathrm{V},\mathrm{V}^*) = \mathcal{L}(\mathrm{V},\mathrm{V})$, d.h. Tensoren vom Typ (1,1) sind lineare Abbildungen.

Definition 3 *Sei V ein linearer Raum über $\mathcal{R}$, $x_1,\ldots,x_q \in V$, $\xi_1, \ldots, \xi_q \in V^*$. Das Tensorprodukt* $\xi_1 \boxplus \ldots \boxplus \xi_p \boxplus x_1 \boxplus \ldots \boxplus x_q$ *ist das durch*

$$(\xi_1 \boxplus \ldots \boxplus \xi_p \boxplus x_1 \boxplus \ldots \boxplus x_q)(y_1,\ldots,y_p,\eta_1,\ldots,\eta_q) = \xi_1(y_1)\ldots\xi_p(y_p)\eta_1(x_1)\ldots n_q(x_q)$$

definierte Element aus $T^q_p(V)$.

Beispiele:
$(\mathrm{x} \boxplus \mathrm{y})(\xi,\eta) = \xi(\mathrm{x})\eta(\mathrm{y})\ T^2(\mathrm{V})$
$(\xi \boxplus \eta)(\mathrm{x},\mathrm{y}) = \xi(\mathrm{x})\eta(\mathrm{y})\ T_2(\mathrm{V})$
$(\xi \boxplus \mathrm{x})(\mathrm{y},\eta) = \xi(\mathrm{y})\eta(\mathrm{x})\ T_1^1(\mathrm{V})$
$\xi = (1\ 0)$, $\mathrm{x} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ $\mathrm{V} = \mathcal{R}^2$
allgemeiner $\xi = (\xi_1,\xi_2)$, $\mathrm{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ Dann ist $\xi \boxplus \mathrm{x}$ auf $\mathcal{R}^2$ durch die Matrix $(\xi_1\mathrm{x},\xi_2\mathrm{x})$
$= \left(\xi_1 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \xi_2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right)$
$= \begin{pmatrix} \xi_1 x_1 & \xi_2 x_1 \\ \xi_1 x_2 & \xi_2 x_2 \end{pmatrix}$ gegeben (muss bewiesen werden).
$(1\ 0) \boxplus \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$,
$(1\ 0) \boxplus \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$,
$(0\ 1) \boxplus \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$,
$(0\ 1) \boxplus \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

Satz 1 *Sei V ein endlichdimensionaler linearer Raum über $\mathcal{R}$, sei $\{e_1,\dots,e_n\}$ eine Basis in V und $\{e^1,\dots,e^n\}$ die duale Basis in V^* ($e^i(e_j) = \delta^i_j$). Dann gilt:*

$$dim\ T^q_p(V) = (dim\ V)^{p+q}$$

und eine Basis in $T^q_p(V)$ ist gegeben durch die Elemente

$$e^{i_1} \boxplus \dots \boxplus e^{i_p} \boxplus e_{j_1} \boxplus \dots \boxplus e_{j_p}\ (1 \le i_1,\dots,e_p,j_1,\dots,j_q \le dim\ V).$$

Definition 4 *Sind E, E^* wie oben, so nennt man*

$$T^{j_1\dots j_q}_{i_1\dots i_p} := T(e_{i_1},\dots,e_{i_p},e^{j_1},\dots,e^{j_q})$$

die Koordinaten eines Tensors $T \in T^q_p(V)$. Man schreibt
$T = (T^{j_1\dots j_q}_{i_1\dots i_p})$ oder $[T]_E = (T^{j_1\dots j_q}_{i_1\dots i_p})$.

Es gilt:
$\mathrm{T} = \sum T^{j_1\dots j_q}_{i_1\dots i_p}(e^{i_1} \boxplus \dots \boxplus e^{i_p} \boxplus e_{j_1} \boxplus \dots \boxplus e_{j_q})$ $(\mathrm{T} = \sum(\mathrm{T},e_j)e_j)$

Koordinatenwechsel
$\mathrm{E} = \{e_1,\dots,e_n\}$, $\mathrm{E}^* = \{e^1,\dots,e^n\}$
$\tilde{E} = \{\tilde{e_1},\dots,\tilde{e_n}\}$, $\tilde{E}^* = \{\tilde{e^1},\dots,\tilde{e^n}\}$
seien Basen. Basiswechsel:

$$\begin{pmatrix} \tilde{e_1} & & \\ \vdots & & \tilde{e_n} \end{pmatrix} = \underbrace{\begin{pmatrix} a_1^1 & \dots & a_1^n \\ \vdots & & \vdots \\ a_n^1 & \dots & a_n^n \end{pmatrix}}_{U} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}, \begin{pmatrix} \tilde{e^1} \\ \vdots \\ \tilde{e^n} \end{pmatrix} = \underbrace{\begin{pmatrix} b_1^1 & \dots & b_n^1 \\ \vdots & & \vdots \\ b_1^n & \dots & b_n^n \end{pmatrix}}_{(U^{-1})^T} \begin{pmatrix} e^1 \\ \vdots \\ e^n \end{pmatrix}$$

$\tilde{e_i} = \sum_j a_i^j e_j = a_i^j e_j$, $\tilde{e^i} = \sum_j b_j^i e^j = b_j^i e^j$

Satz 2 *Die Koordinaten eines Tensors $T \in T_p^q(V)$ transformieren sich nach der Regel*

$$\tilde{T^{j_1 \dots j_q}_{i_1 \dots i_p}} = a_{i_1}^{k_1} \dots a_{i_p}^{k_p} b_{l_1}^{j_1} \dots b_{l_q}^{j_q} T^{l_1 \dots l_q}_{k_1 \dots k_p}.$$

Man nennt Indizes *kovariant*, wenn sie sich wie Basen transformieren lassen, und *kontravariant*, wenn sie sich wie duale Basen transfomieren. Ein Tensor aus T_p^q(V) hat also genau p kovariante und q kontravariante Indizes.
Tensor in der Physik: Ein Objekt, das in jeder Basis durch n^{p+q} Zahlen gegeben ist und bei dem sich diese Zahlen wie in Satz 2 beim Basiswechsel ändern.

Beispiele:
1. f: $\mathcal{R}^n \to \mathcal{R}$
grad f(a) = f'(a) $\in \mathcal{L}(\mathcal{R}^n,\mathcal{R}) = (\mathcal{R}^n)^* = \mathcal{B}(\mathcal{R}^n) = T_1(\mathcal{R}^n)$
Vektoren $\mathcal{R}^n = (\mathcal{R}^n)^{**} = \mathcal{B}((\mathcal{R}^n)^*) = T^1(\mathcal{R}^n)$
Kräfte = Gradienten sind also Covektoren.

2. T_1^1(V) = $\mathcal{B}$(V,V*) = $\mathcal{L}$(V,V)
$\tilde{T_i^j} = a_i^k b_l^j T_k^l = a_i^k T_k^l b_l^j$
$[\mathrm{T}]_{\tilde{E}} = \mathrm{U}\,[\mathrm{T}]_E \mathrm{U}^{-1}$

3. T_2(V) = $\mathcal{B}$(V,V) enthalten Skalarprodukte
$\tilde{T_{ij}} = a_i^k a_j^l T_{kl} = a_i^k T_{kl} a_j^l$
$[\mathrm{T}]_{\tilde{E}} = \mathrm{U}\,[\mathrm{T}]_E \mathrm{U}^T$

Rechnen mit Tensoren
R + S, αR klar
Äußeres Produkt (Tensorprodukt):
R $\in T_p^q$(V), S $\in T_u^v$(V)
R $\boxplus$ S $\in T_{p+u}^{q+v}$(V)
$(\mathrm{R} \boxplus \mathrm{S})^{j_1 \dots j_q l_1 \dots l_v}_{i_1 \dots i_p k_1 \dots k_u} = R^{j_1 \dots j_q}_{i_1 \dots i_p}\ S^{l_1 \dots l_v}_{k_1 \dots k_u}$

A, B $\in T_1^1$(V) A $\boxplus$ B $\in T_2^2$(V)
A $\boxplus$ B = $(a_{ij}\mathrm{B})_{i,j=1}^n$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \boxplus \begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} a \begin{pmatrix} x & y \\ z & w \end{pmatrix} & b \begin{pmatrix} x & y \\ z & w \end{pmatrix} \\ c \begin{pmatrix} x & y \\ z & w \end{pmatrix} & d \begin{pmatrix} x & y \\ z & w \end{pmatrix} \end{pmatrix}$$

$$= \begin{pmatrix} ax & ay & bx & by \\ az & aw & bz & bw \\ cx & cy & dx & dy \\ cz & cw & dz & dw \end{pmatrix}$$

Verjüngung
R $\in T_p^q$(V), 1≤k≤p, 1≤l≤q
Verjüngung V_k^lR ist definiert über

$$(V_k^l R)_{i_1 \ldots i_{k-1} i_{k+1} \ldots i_p}^{j_1 \ldots j_{l-1} j_{l+1} \ldots j_q} = R_{i_1 \ldots i_{k-1} m i_{k+1} \ldots i_p}^{j_1 \ldots j_{l-1} m j_{l+1} \ldots j_q}.$$

Ergibt V_k^lR $\in T_{p-1}^{q-1}$(V). Man setzt T_0^0(V) = $\mathcal{R}$.

Beispiel:
A $\in T_1^1$(V), V_1^1A = $A_m^m = A_1^1 + \ldots + A_n^n$ = spur A

Inneres Produkt
R $\in T_p^q$(V), S $\in T_u^v$(V), 1≤k≤p+u, 1≤l≤q+v
Das (k,l)-innere Produkt ist dann definiert als $(R,S)_k^l = V_k^l(R \boxplus S)$.
Beispiel:
A $\in T_1^1$(V), B $\in T_1^1$(V)
$(A,B)_1^1 = V_1^1(A \boxplus B) = (a_i^i b_k^l)$ = (spur A)B
$(A,B)_2^2 = V_2^2(A \boxplus B) = (a_i^j b_k^k)$ = (spur B)A
$(A,B)_1^2 = V_1^2(A \boxplus B) = (a_i^j b_k^i) = (b_k^i a_i^j)$ = BA
$(A,B)_2^1 = V_2^1(A \boxplus B) = (a_i^j b_j^l)$ = AB

Krümmung einer Fläche
In der Ebene: Verschieben Tangentialvektor parallel längs einer geschlossenen Kurve.
Allgemein: Tangentialebene in x sei T_xM = V.
Wählen $\xi, \eta \in$ V und konstruieren daraus ein Parallelogramm. $\xi \in$ V wählen und längs des Parallelogramms verschieben. Es entsteht $\xi + \Delta\xi \in$ V. Man kann zeigen:
$\Delta\xi = \epsilon R_x(\xi, \eta)\xi + o(\epsilon)$ für $\epsilon \to 0$,
wobei R_x: V × V → $\mathcal{L}$(V,V) bilinear ist.
Es gilt $\mathcal{B}$(V,V;$\mathcal{L}$(V,V)) ≅ $\mathcal{B}$(V,V,V,V*) = T_3^1(V).
$(\gamma A)(x,y,z,\varphi) = \varphi(A(x,y),z)$.
$R_x \in \mathcal{B}$(V,V,V,V*) heißt *Krümmungstensor in x.*
R_{ijl}^k(x) Krümmungstensor
$R_{ij}(x) = R_{ijm}^m(x)$ Ricci-Tensor
In V = T_xM ist ein Skalarprodukt gegeben, g $\in \mathcal{B}$(V,V) = T_2(V).
g = (g_{ij})
Inverse Matrix (g^{ij})
$g^{ij}(x)R_{ij}(x)$ = R = *Krümmungsskalar*
$V_1^1\ V_2^2$ ($g^{-1} \boxplus$ Ricci)

Einsteinsche Gleichungen
$R_{ij} - \frac{1}{2}Rg_{ij} = \chi T_{ij}$.