Lineare Algebra

Wolfgang Soergel

31. August 2009

Inhaltsverzeichnis

A Grundlagen **7**

I Allgemeine Grundlagen **9**
1 Einstimmung . . . 10
1.1 Vollständige Induktion und binomische Formel . . . 10
1.2 Fibonacci-Folge und Vektorraumbegriff . . . 17
2 Naive Mengenlehre und Kombinatorik . . . 24
2.1 Mengen . . . 24
2.2 Abbildungen . . . 33
2.3 Logische Symbole und Konventionen . . . 42
3 Algebraische Grundbegriffe . . . 44
3.1 Mengen mit Verknüpfung . . . 44
3.2 Gruppen . . . 49
3.3 Körper . . . 55

II Geschichtliches und Philosophisches **61**
1 Zum Ursprung des Wortes Mathematik . . . 62
2 Was ist Mathematik? . . . 62
3 Zum Wesen der Mathematik . . . 63
4 Herkunft einiger Symbole . . . 64

B Algebra **67**

III Lineare Algebra **69**
1 Gleichungssysteme und Vektorräume . . . 71
1.1 Lösen linearer Gleichungssysteme . . . 71
1.2 Ergänzungen zur Mengenlehre . . . 77
1.3 Vektorräume und Untervektorräume . . . 79
1.4 Lineare Unabhängigkeit und Basen . . . 84
1.5 Lineare Abbildungen . . . 92
1.6 Dimensionsformel . . . 98

1.7 Lineare Abbildungen und Matrizen . . . 101
1.8 Dualräume und transponierte Abbildungen . . . 113
1.9 Affine Räume . . . 119
2 Ringe, Polynome, Determinanten . . . 129
2.1 Ringe . . . 129
2.2 Untergruppen der ganzen Zahlen . . . 133
2.3 Polynome . . . 141
2.4 Äquivalenzrelationen . . . 148
2.5 Quotientenkörper . . . 149
2.6 Das Signum einer Permutation . . . 151
2.7 Die Determinante . . . 154
2.8 Eigenwerte und Eigenvektoren . . . 165
3 Euklidische Vektorräume . . . 170
3.1 Modellierung des Anschauungsraums . . . 170
3.2 Geometrie in euklidischen Vektorräumen . . . 173
3.3 Orthogonale und unitäre Abbildungen . . . 180
3.4 Isometrien euklidischer affiner Räume . . . 189
3.5 Winkel, Orientierung, Kreuzprodukt . . . 191
3.6 Spektralsatz und Hauptachsentransformationen . . . 201
4 Bilinearformen . . . 209
4.1 Fundamentalmatrix . . . 209
4.2 Definitheitseigenschaften . . . 211
4.3 Klassifikation symmetrischer Bilinearformen . . . 214
4.4 Alternierende Bilinearformen . . . 219
5 Jordan'sche Normalform . . . 221
5.1 Motivation durch Differentialgleichungen . . . 221
5.2 Summen und Produkte von Vektorräumen . . . 222
5.3 Hauptraumzerlegung . . . 224
5.4 Jordan-Zerlegung . . . 229
5.5 Jordan'sche Normalform . . . 232
6 Algebra und Symmetrie . . . 239
6.1 Gruppenwirkungen . . . 239
6.2 Restklassen . . . 245
6.3 Bahnformel . . . 247
6.4 Normalteiler . . . 249
6.5 Zyklische Gruppen . . . 252
6.6 Endlich erzeugte abelsche Gruppen . . . 256
6.7 Konjugationsklassen . . . 263
6.8 Endliche Untergruppen der Drehgruppe . . . 264
6.9 Skalarprodukte zu Drehgruppen . . . 278
7 Universelle Konstruktionen . . . 285

7.1 Quotientenvektorräume 285
7.2 Tensorprodukte von Vektorräumen 289
7.3 Kanonische Injektionen bei Tensorprodukten 298
7.4 Alternierende Tensoren und Determinante 300
7.5 Das kanonische Skalarprodukt 306

IV Typische Prüfungsfragen **311**
1 Lineare Algebra 312
2 Algebra 313
3 Analysis 314

Literaturverzeichnis **317**

Index **319**

Teil A

Grundlagen

Kapitel I

Allgemeine Grundlagen

In diesem ersten Kapitel habe ich Notationen und Begriffsbildungen zusammengefaßt, von denen ich mir vorstelle, daß sie zu Beginn des Studiums in enger Abstimmung zwischen den beiden Grundvorlesungen erklärt werden könnten.

Inhalt

1 Einstimmung 10

1.1 Vollständige Induktion und binomische Formel . . 10

1.2 Fibonacci-Folge und Vektorraumbegriff 17

2 Naive Mengenlehre und Kombinatorik 24

2.1 Mengen 24

2.2 Abbildungen 33

2.3 Logische Symbole und Konventionen 42

3 Algebraische Grundbegriffe 44

3.1 Mengen mit Verknüpfung 44

3.2 Gruppen 49

3.3 Körper 55

1 Einstimmung

1.1 Vollständige Induktion und binomische Formel

Satz 1.1.1. *Für jede natürliche Zahl* $n \geq 1$ *gilt* $1 + 2 + \ldots + n = \frac{n(n+1)}{2}$.

Beweis. Bei diesem Beweis sollen Sie gleichzeitig das Beweisprinzip der **vollständigen Induktion** lernen. Wir bezeichnen mit $A(n)$ die Aussage, daß die Formel im Satz für ein gegebenes n gilt, und zeigen

Induktionsbasis: Die Aussage $A(1)$ ist richtig. In der Tat gilt die Formel $1 = \frac{1(1+1)}{2}$.

Induktionsschritt: Aus $A(n)$ folgt $A(n+1)$. In der Tat, unter der Annahme, daß unsere Formel für ein gegebenes n gilt, der sogenannten **Induktionsannahme** oder **Induktionsvoraussetzung**, rechnen wir

$$\begin{aligned} 1 + 2 + \ldots + n + (n+1) &= \tfrac{n(n+1)}{2} + \tfrac{2(n+1)}{2} \\ &= \tfrac{(n+2)(n+1)}{2} \\ &= \tfrac{(n+1)((n+1)+1)}{2} \end{aligned}$$

und folgern so, daß die Formel auch für $n+1$ gilt.

Es ist damit klar, daß unsere Aussage $A(n)$ richtig ist alias daß unsere Formel gilt für alle $n = 1, 2, 3, \ldots$. □

1.1.2. Dieser Beweis stützt sich auf unser intuitives Verständnis der natürlichen Zahlen. Man kann das Konzept der natürlichen Zahlen auch formal einführen und so die natürlichen Zahlen in gewisser Weise "besser" verstehen. Das mögen Sie in der Logik lernen. Das Wort "Induktion" meint eigentlich "Hervorrufen", so wie etwa das Betrachten einer Wurst die Ausschüttung von Spucke induziert alias uns den Mund wässrig macht. Im Zussammenhang der vollständigen Induktion ist es dahingehend zu verstehen, daß die Richtigkeit unserer Aussage $A(0)$ die Richtigkeit von $A(1)$ induziert, die Richtigkeit von $A(1)$ hinwiederum die Richtigkeit von $A(2)$, die Richtigkeit von $A(2)$ die Richtigkeit von $A(3)$, und immer so weiter.

1.1.3. Es herrscht keine Einigkeit in der Frage, ob man die Null eine natürliche Zahl nennen soll. In diesem Text ist stets die Null mit gemeint, wenn von natürlichen Zahlen die Rede ist. Wollen wir die Null dennoch ausschließen, so sprechen wir wie oben von einer "natürlichen Zahl $n \geq 1$".

Bemerkung 1.1.4. Ich will kurz begründen, warum es mir natürlich scheint, auch die Null eine natürliche Zahl zu nennen: Hat bildlich gesprochen jedes Kind einer Klasse einen Korb mit Äpfeln vor sich und soll seine Äpfel zählen, so kann es ja durchaus vorkommen, daß in seinem Korb gar kein Apfel liegt, weil es zum Beispiel alle seine Äpfel bereits gegessen hat. In der Begrifflichkeit der Mengenlehre ausgedrückt, die wir in 2.1 einführen werden, muß man die leere Menge endlich nennen, damit jede Teilmenge einer endlichen Menge wieder endlich ist. Will man dann erreichen, daß die Kardinalität jeder endlichen Menge eine natürliche Zahl ist, so darf man die Null nicht aus den natürlichen Zahlen herauslassen.

Bemerkung 1.1.5. Man kann sich den Satz anschaulich klar machen als eine Formel für die Fläche eines Querschnitts für eine Treppe der Länge n mit Stufenabstand und Stufenhöhe eins. In der Tat bedeckt ein derartiger Querschnitt ja offensichtlich ein halbes Quadrat der Kantenlänge n nebst n halben Quadraten der Kantenlänge 1. Ein weiterer Beweis geht so:

$$\begin{aligned} 1+2+\ldots+n &= \begin{array}{cccc} 1/2 & + & 2/2 & +\ldots \\ +\,n/2 & +(n-1)/2 & +\ldots & +\,1/2 \end{array} \\ &= \begin{array}{ccccc} \frac{n+1}{2} & + & \frac{n+1}{2} & +\ldots & + \frac{n+1}{2} \end{array} \\ &= n(n+1)/2 \end{aligned}$$

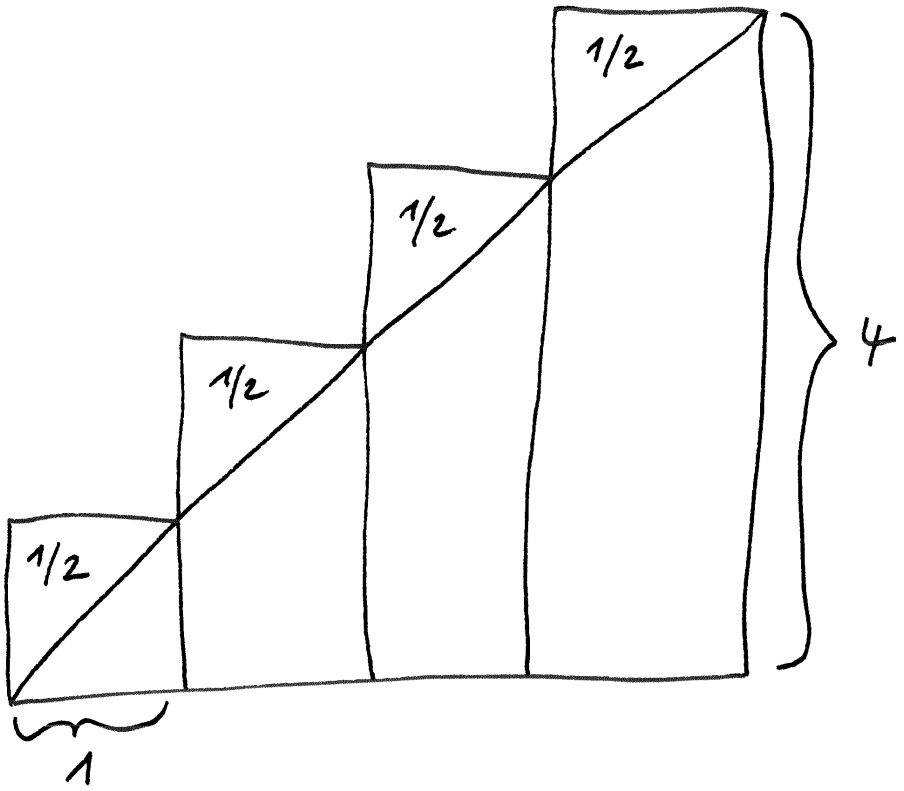
Ich will diesen Beweis benutzen, um eine neue Notation einzuführen.

Definition 1.1.6. Gegeben $a_1, a_2, \ldots, a_n$ schreiben wir

$$\sum_{i=1}^{n} a_i := a_1 + a_2 + \ldots + a_n$$

Das Symbol $\sum$ ist ein großes griechisches S und steht für "Summe". Das Symbol $:=$ deutet an, daß die Bedeutung der Symbole auf der doppelpunktbehafteten Seite des Gleichheitszeichens durch den Ausdruck auf der anderen Seite unseres Gleichheitszeichens definiert ist. In diesem und ähnlichen Zusammenhängen heißen $a_1, \ldots, a_n$ die **Summanden** und i der **Laufindex**, weil er eben etwa in unserem Fall von 1 bis n läuft und anzeigt alias "indiziert", welcher Summand gemeint ist.

1.1.7. Das Wort "Definition" kommt aus dem Lateinischen und bedeutet "Abgrenzung". In Definitionen versuchen wir, die Bedeutungen von Symbolen und Begriffen so klar wie möglich festzulegen. Sie werden merken, daß man in der Mathematik die Angwohnheit hat, in Definitionen Worte der Umgangssprache wie Menge, Gruppe, Körper, Unterkörper, Abbildung etc. "umzuwidmen"



Die Gesamtfläche dieses Treppenquerschnitts ist offensichtlich
$4^2/2 + 4/2 = 4 \cdot 5/2$

und ihnen ganz spezielle und nur noch entfernt mit der umgangssprachlichen Bedeutung verwandte Bedeutungen zu geben. In mathematischen Texten sind dann durchgehend diese umgewidmeten Bedeutungen gemeint, in dieser Weise baut die Mathematik also wirklich ihre eigene Sprache auf. Allerdings wird die Grammatik dann doch noch von den uns geläufigen Sprachen übernommen.

Beispiel 1.1.8. In der $\sum$-Notation liest sich der in 1.1.5 gegebene Beweis so:

$$\begin{aligned}
\textstyle\sum_{i=1}^{n} i &= \textstyle\sum_{i=1}^{n} \frac{i}{2} + \sum_{i=1}^{n} \frac{i}{2} \\
&\quad \text{und nach Indexwechsel } i = n+1-k \text{ hinten} \\
&= \textstyle\sum_{i=1}^{n} \frac{i}{2} + \sum_{k=1}^{n} \frac{n+1-k}{2} \\
&\quad \text{dann mache } k \text{ zu } i \text{ in der zweiten Summe} \\
&= \textstyle\sum_{i=1}^{n} \frac{i}{2} + \sum_{i=1}^{n} \frac{n+1-i}{2} \\
&\quad \text{und nach Zusammenfassen beider Summen} \\
&= \textstyle\sum_{i=1}^{n} \frac{n+1}{2} \\
&\quad \text{ergibt sich offensichtlich} \\
&= n\left(\tfrac{n+1}{2}\right)
\end{aligned}$$

Definition 1.1.9. In einer ähnlichen Bedeutung wie $\sum$ verwendet man auch das Symbol $\prod$, ein großes griechisches P, für "Produkt" und schreibt

$$\prod_{i=1}^{n} a_i := a_1 a_2 \dots a_n$$

Die $a_1, \dots, a_n$ heißen in diesem und ähnlichen Zusammenhängen die **Faktoren** des Produkts.

Definition 1.1.10. Für jede natürliche Zahl $n \geq 1$ definieren wir die Zahl $n!$ (sprich: n **Fakultät**) durch die Formel

$$n! := 1 \cdot 2 \cdot \ldots \cdot n = \prod_{i=1}^{n} i$$

Wir treffen zusätzlich die Vereinbarung $0! := 1$ und haben also $0! = 1$, $1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$ und so weiter.

1.1.11. Wir werden in Zukunft noch öfter Produkte mit überhaupt keinem Faktor zu betrachten haben und vereinbaren deshalb gleich hier schon, daß Produkten, bei denen die obere Grenze des Laufindex um eins kleiner ist als seine untere Grenze, der Wert 1 zugewiesen werden soll. Ebenso vereinbaren wir auch, daß Summen, bei denen die obere Grenze des Laufindex um Eins kleiner ist als seine untere Grenze, der Wert 0 zugewiesen werden soll.

Satz 1.1.12 (**Bedeutung der Fakultät**). *Es gibt genau $n!$ Möglichkeiten, n voneinander verschiedene Objekte in eine Reihenfolge zu bringen.*

Beispiel 1.1.13. Es gibt genau $3! = 6$ Möglichkeiten, die drei Buchstaben a, b und c in eine Reihenfolge zu bringen, nämlich

$$\begin{array}{lll} abc & bac & cab \\ acb & bca & cba \end{array}$$

In gewisser Weise stimmt unser Satz sogar für $n = 0$: In der Terminologie, die wir in **??** einführen werden, gibt es in der Tat genau eine Anordnung der leeren Menge.

Beweis. Hat man n voneinander verschiedene Objekte, so hat man n Möglichkeiten, ein Erstes auszusuchen, dann $(n-1)$ Möglichkeiten, ein Zweites auszusuchen und so weiter, bis schließlich nur noch eine Möglichkeit bleibt, ein Letztes auszusuchen. Insgesamt haben wir also in der Tat wie behauptet $n!$ mögliche Reihenfolgen. □

Definition 1.1.14. Wir definieren für beliebiges n und jede natürliche Zahl k die **Binomialkoeffizienten** $\binom{n}{k}$ (sprich: n **über** k) durch die Regeln

$$\binom{n}{k} := \prod_{j=0}^{k-1} \frac{n-j}{k-j} = \frac{n(n-1)\ldots \quad (n-k+1)}{k(k-1)\ldots \qquad 1} \text{ für } k \geq 1 \text{ und } \binom{n}{0} := 1.$$

Der Sonderfall $k = 0$ wird im Übrigen auch durch unsere allgemeine Formel gedeckt, wenn wir unsere Konvention 1.1.11 beherzigen. Im Lichte des folgenden Satzes schlage ich vor, die Binomialkoeffizienten $\binom{n}{k}$ statt "n über k" inhaltsreicher "k aus n" zu sprechen.

1.1.15. Die Bezeichnung als Binomialkoeffizienten leitet sich von dem Auftreten dieser Zahlen als Koeffizienten in der "binomischen Formel" 3.3.4 ab.

Satz 1.1.16 (**Bedeutung der Binomialkoeffizienten**). *Gegeben natürliche Zahlen n und k gibt es genau $\binom{n}{k}$ Möglichkeiten, aus n voneinander verschiedenen Objekten k Objekte auszuwählen.*

Beispiel 1.1.17. Es gibt genau $\binom{4}{2} = \frac{4\cdot 3}{2\cdot 1} = 6$ Möglichkeiten, aus den vier Buchstaben a, b, c, d zwei auszuwählen, nämlich

$$\begin{array}{lll} a,b & b,c & c,d \\ a,c & b,d & \\ a,d & & \end{array}$$

Beweis. Wir haben n Möglichkeiten, ein erstes Objekt auszuwählen, dann $n-1$ Möglichkeiten, ein zweites Objekt auszuwählen, und so weiter, also insgesamt $n(n-1)\dots(n-k+1)$ Möglichkeiten, k Objekte *der Reihe nach* auszuwählen. Auf die Reihenfolge, in der wir ausgewählt haben, kommt es uns aber gar nicht an, jeweils genau $k!$ von unseren $n(n-1)\dots(n-k+1)$ Möglichkeiten führen also nach 1.1.12 zur Auswahl derselben k Objekte. Man bemerke, daß unser Satz auch im Extremfall $k=0$ noch stimmt, wenn wir ihn geeignet interpretieren: In der Terminologie, die wir gleich einführen werden, besitzt in der Tat jede Menge genau eine nullelementige Teilmenge, nämlich die leere Menge. □

1.1.18. Offensichtlich gilt für alle natürlichen Zahlen n mit $n \geq k$ die Formel

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \binom{n}{n-k}$$

Das folgt einerseits sofort aus der formalen Definition und ist andererseits auch klar nach der oben erklärten Bedeutung der Binomialkoeffizienten: Wenn wir aus n Objekten k Objekte auswählen, so bleiben $n-k$ Objekte übrig. Es gibt demnach gleichviele Möglichkeiten, k Objekte auszuwählen, wie es Möglichkeiten gibt, $n-k$ Objekte auszuwählen. Wir haben weiter $\binom{n}{n} = \binom{n}{0} = 1$ für jede natürliche Zahl $n \geq 0$ sowie $\binom{n}{1} = \binom{n}{n-1} = n$ für jede natürliche Zahl $n \geq 1$.

Definition 1.1.19. Wie in der Schule setzen wir $a^k := \prod_{i=1}^{k} a$, in Worten ist also gemeint "das Produkt von k-mal dem Faktor a", und verstehen im Lichte von 1.1.11 insbesondere $a^0 = 1$.

Satz 1.1.20. *Für jede natürliche Zahl n gilt die* **binomische Formel**

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

Bemerkung 1.1.21. Man beachte, wie wichtig unsere Konvention $a^0 = 1$ und insbesondere auch $0^0 = 1$ für die Gültigkeit dieser Formel ist.

Bemerkung 1.1.22. Die Bezeichung "binomische Formel" leitet sich ab von der Vorsilbe "bi" für Zwei, wie etwa in englisch "bicycle" für "Zweirad" alias "Fahrrad", und dem lateinischen Wort "nomen" für "Namen". Die beiden Namen meinen hier a und b. Mehr dazu wird in **??** erklärt.

Erster Beweis. Beim Ausmultiplizieren erhalten wir so oft $a^k b^{n-k}$, wie es Möglichkeiten gibt, aus unseren n Faktoren $(a+b)$ die k Faktoren auszusuchen, "in denen wir beim Ausmultiplizieren das b nehmen". Dieses Argument werden wir in 2.1.17 noch besser formulieren. □

Zweiter Beweis. Dieser Beweis ist eine ausgezeichnete Übung im Umgang mit unseren Symbolen und mit der vollständigen Induktion. Er scheint mir jedoch auch in einer für Beweise durch vollständige Induktion typischen Weise undurchsichtig. Zunächst prüfen wir für beliebiges n und jede natürliche Zahl $k \geq 1$ die Formel

$$\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$$

durch explizites Nachrechnen. Dann geben wir unserer Formel im Satz den Namen $A(n)$ und prüfen die Formel $A(0)$ und zur Sicherheit auch noch $A(1)$ durch Hinsehen. Schließlich gilt es, den Induktionsschritt durchzuführen als da heißt $A(n+1)$ aus $A(n)$ zu folgern. Dazu rechnen wir

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)(a+b)^n \\
&\quad \text{und mit der Induktionsvoraussetzung} \\
&= (a+b)\textstyle\sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \\
&\quad \text{und durch Ausmultiplizieren} \\
&= \textstyle\sum_{k=0}^{n} \binom{n}{k} a^{k+1} b^{n-k} + \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k+1} \\
&\quad \text{und Indexwechsel } k = i-1 \text{ in der ersten Summe} \\
&= \textstyle\sum_{i=1}^{n+1} \binom{n}{i-1} a^i b^{n-i+1} + \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k+1} \\
&\quad \text{dann mit } k \text{ statt } i \text{ und Absondern von Summanden} \\
&= a^{n+1}b^0 + \textstyle\sum_{k=1}^{n} \binom{n}{k-1} a^k b^{n-k+1} + \\
&\qquad + \textstyle\sum_{k=1}^{n} \binom{n}{k} a^k b^{n-k+1} + a^0 b^{n+1} \\
&\quad \text{und nach Zusammenfassen der mittleren Summen} \\
&= a^{n+1}b^0 + \textstyle\sum_{k=1}^{n} \binom{n+1}{k} a^k b^{n-k+1} + a^0 b^{n+1} \\
&\quad \text{und Einbeziehen der abgesonderten Summanden} \\
&= \textstyle\sum_{k=0}^{n+1} \binom{n+1}{k} a^k b^{n+1-k}
\end{aligned}$$

und folgern so tatsächlich $A(n+1)$ aus $A(n)$. □

Bemerkung 1.1.23. Die Formel $\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$ für $k \geq 1$ kann man zur effektiven Berechnung der Binomialkoeffizienten mit dem sogenannten **Pascal'schen Dreieck** benutzen: Im Schema

$$\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 &
\end{array}$$

seien die Einsen an den Rändern vorgegeben und eine Zahl in der Mitte berechne sich als die Summe ihrer beiden oberen "Nachbarn". Dann stehen in der $(n+1)$-ten Zeile der Reihe nach die Binomialkoeffizienten $\binom{n}{0} = 1, \binom{n}{1} = n, \ldots$ bis $\binom{n}{n-1} = n, \binom{n}{n} = 1$. Wir haben also zum Beispiel

$$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$$

Übung 1.1.24. Man finde und beweise eine Formel für $\sum_{i=1}^{n} i^2$. Hinweis: Man suche zunächst eine Formel für $\sum_{i=1}^{n} i^3 - (i-1)^3$ und beachte $i^3 - (i-1)^3 = 3i^2 - 3i + 1$.

1.2 Fibonacci-Folge und Vektorraumbegriff

Beispiel 1.2.1. Die **Fibonacci-Folge**

$$0, 1, 1, 2, 3, 5, 8, 13, 21, \ldots$$

entsteht, indem man mit $f_0 = 0$ und $f_1 = 1$ beginnt und dann jedes weitere Folgenglied als die Summe seiner beiden Vorgänger bildet. Wir suchen nun für die Glieder dieser Folge eine geschlossene Darstellung. Dazu vereinbaren wir, daß wir Folgen $x_0, x_1, x_2, \ldots$ mit der Eigenschaft $x_n = x_{n-1} + x_{n-2}$ für $n = 2, 3, 4, \ldots$ **Folgen vom Fibonacci-Typ** nennen wollen. Kennen wir die beiden ersten Glieder einer Folge vom Fibonacci-Typ, so liegt natürlich bereits die gesamte Folge fest. Nun bemerken wir, daß für jede Folge $x_0, x_1, x_2, \ldots$ vom Fibonacci-Typ und jedes α auch die Folge $\alpha x_0, \alpha x_1, \alpha x_2, \ldots$ vom Fibonacci-Typ ist, und daß für jede weitere Folge $y_0, y_1, y_2, \ldots$ vom Fibonacci-Typ auch die gliedweise Summe $(x_0 + y_0), (x_1 + y_1), (x_2 + y_2), \ldots$ eine Folge vom Fibonacci-Typ ist. Der Trick ist dann, danach zu fragen, für welche β die Folge $x_i = \beta^i$ vom Fibonacci-Typ ist. Das ist ja offensichtlich genau dann der Fall, wenn gilt $\beta^2 = \beta + 1$, als da heißt für $\beta_\pm = \frac{1}{2}(1 \pm \sqrt{5})$. Für beliebige c, d ist mithin die Folge

$$x_i = c\beta_+^i + d\beta_-^i$$

vom Fibonacci-Typ, und wenn wir c und d bestimmen mit $x_0 = 0$ und $x_1 = 1$, so ergibt sich eine explizite Darstellung unserer Fibonacci-Folge. Wir suchen also c und d mit

$$\begin{array}{rcl} 0 & = & c + d \\ 1 & = & c\left(\frac{1}{2}(1+\sqrt{5})\right) + d\left(\frac{1}{2}(1-\sqrt{5})\right) \end{array}$$

und folgern leicht $c = -d$ und $1 = c\sqrt{5}$ alias $c = 1/\sqrt{5} = -d$. Damit ergibt sich schließlich für unsere ursprüngliche Fibonacci-Folge die explizite

Darstellung

$$f_i = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^i - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^i$$

Es wäre rückblickend natürlich ein Leichtes gewesen, diese Formel einfach zu "raten" um sie dann mit vollständiger Induktion zu beweisen. Diese Art mathematischer Zaubertricks halte ich jedoch für unehrenhaft. Ich werde deshalb stets nach Kräften versuchen, das Tricksen zu vermeiden, auch wenn die Beweise dadurch manchmal etwas länger werden sollten. Eine Möglichkeit, auch den letzten verbleibenden Trick aus den vorhergehenden Überlegungen zu eliminieren, zeigt **??**. Die bei unserer Lösung auftretende reelle Zahl $\frac{1}{2}(1+\sqrt{5})$ ist im Übrigen auch bekannt als "goldener Schnitt" aus Gründen, die im nebenstehenden Bild diskutiert werden. In **??** dürfen Sie dann zur Übung zeigen, daß der Quotient zweier aufeinanderfolgender Fibonacci-Zahlen gegen den goldenen Schnitt strebt, daß also genauer und in Formeln für unsere Fibonacci-Folge $f_0, f_1, f_2, \ldots$ von oben gilt

$$\lim_{i\to\infty} \frac{f_{i+1}}{f_i} = \frac{1+\sqrt{5}}{2}$$

Übung 1.2.2. Kann man für jede Folge $x_0, x_1, \ldots$ vom Fibonacci-Typ Zahlen c, d finden mit $x_i = c\beta_+^i + d\beta_-^i$ für alle i? Finden Sie eine geschlossene Darstellung für die Glieder der Folge, die mit $0, 0, 1$ beginnt und dem Bildungsgesetz $x_n = 2x_{n-1} + x_{n-2} - 2x_{n-3}$ gehorcht.

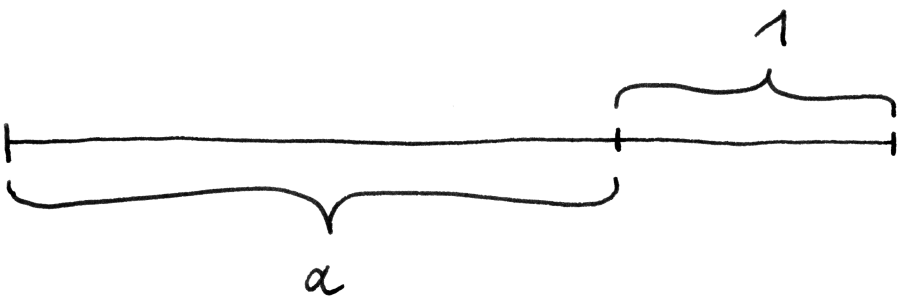
Beispiel 1.2.3. Wir betrachten ein "homogenes lineares" Gleichungssystem alias ein Gleichungssystem der Gestalt

$$\begin{array}{rcl} \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1m}x_m &=& 0 \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \ldots + \alpha_{2m}x_m &=& 0 \\ \vdots & & \\ \alpha_{n1}x_1 + \alpha_{n2}x_2 + \ldots + \alpha_{nm}x_m &=& 0 \end{array}$$

Wie man zu vorgegebenen $\alpha_{i,j}$ für $1 \le i \le n$ und $1 \le j \le m$ die Menge L aller Lösungen $(x_1, \ldots, x_m)$ ermittelt, sollen sie später in dieser Vorlesung lernen. Zwei Dinge aber sind a priori klar:

1. Sind $(x_1, \ldots, x_m)$ und $(x'_1, \ldots, x'_m)$ Lösungen, so ist auch ihre komponentenweise Summe $(x_1 + x'_1, \ldots, x_m + x'_m)$ eine Lösung;

2. Ist $(x_1, \ldots, x_m)$ eine Lösung und α eine reelle Zahl, so ist auch das komponentenweise Produkt $(\alpha x_1, \ldots, \alpha x_m)$ eine Lösung.



Der **goldene Schnitt** ist das Verhältnis, in dem eine Strecke geteilt werden muß, damit das Verhältnis vom größeren zum kleineren Stück gleich dem Verhältnis des Ganzen zum größeren Stück ist, also die positive Lösung der Gleichung $a/1 = (1+a)/a$ alias $a^2 - a - 1 = 0$, also $a = (1+\sqrt{5})/2$.

Beispiel 1.2.4. Wir betrachten die Menge aller Funktionen $f : \mathbb{R} \to \mathbb{R}$, die zweimal differenzierbar sind und der Differentialgleichung

$$f'' = -f$$

genügen. Lösungen sind zum Beispiel die Funktionen $\sin, \cos$, die Nullfunktion oder auch die Funktionen $f(x) = \sin(x + a)$ für konstantes a. Wie man die Menge L aller Lösungen beschreiben kann, sollen Sie nicht hier lernen. Zwei Dinge aber sind a priori klar:

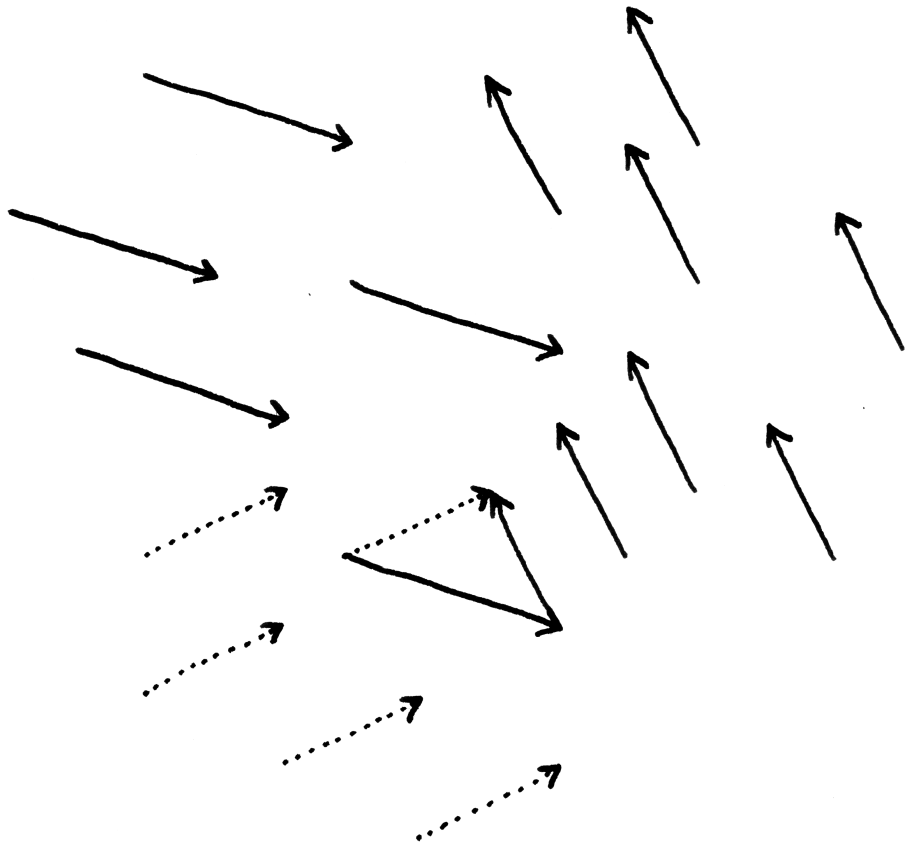
1. Mit f und g ist auch die Funktion $f + g$ eine Lösung;
2. Ist f eine Lösung und α eine reelle Zahl, so ist auch αf eine Lösung.

Beispiel 1.2.5. Wir betrachten die Gesamtheit aller Parallelverschiebungen der Tafelebene. Graphisch stellen wir solch eine Parallelverschiebung dar durch einen Pfeil von irgendeinem Punkt zu seinem Bild unter der Verschiebung. Im nebenstehenden Bild stellen etwa alle gepunktelten Pfeile dieselbe Parallelverschiebung dar. Was für ein Ding diese Gesamtheit P aller Parallelverschiebungen eigentlich ist, scheint mir recht undurchsichtig, aber einiges ist a priori klar:

1. Sind p und q Parallelverschiebungen, so ist auch ihre "Hintereinanderausführung" $p \circ q$, sprich "p nach q" eine Parallelverschiebung.
2. Ist α eine reelle Zahl und p eine Parallelverschiebung, so können wir eine neue Parallelverschiebung αp bilden, das "α-fache von p". Bei negativen Vielfachen vereinbaren wir hierzu, daß eine entsprechende Verschiebung in die Gegenrichtung gemeint ist.
3. Führen wir eine neue Notation ein und schreiben für die Hintereinanderausführung $p \dot{+} q := p \circ q$, so gelten für beliebige Parallelverschiebungen p, q, r der Tafelebene und beliebige reelle Zahlen α, β die Formeln

$$\begin{array}{rcl} (p \dot{+} q) \dot{+} r & = & p \dot{+} (q \dot{+} r) \\ p \dot{+} q & = & q \dot{+} p \\ \alpha(\beta p) & = & (\alpha\beta)p \\ (\alpha + \beta)p & = & (\alpha p) \dot{+} (\beta p) \\ \alpha(p \dot{+} q) & = & (\alpha p) \dot{+} (\alpha q) \end{array}$$

Will man sich die Gesamtheit aller Parallelverschiebungen der Tafelebene anschaulich machen, so tut man im Übrigen gut daran, einen Punkt als "Ursprung" auszuzeichnen und jede Parallelverschiebung mit dem Punkt der Tafelebene zu identifizieren, auf den unsere Parallelverschiebung diesen Ursprung abbildet.



Die Hintereinanderausführung der beiden Parallelverschiebungen der Tafel- oder hier vielmehr der Papierebene, die durch die durchgezogenen Pfeile dargestellt werden, wird die durch die gepunktelten Feile dargestellt.

Beispiel 1.2.6. Analoges gilt für die Gesamtheit der Parallelverschiebung des Raums und auch für die Gesamtheit aller Verschiebungen einer Geraden und, mit noch mehr Mut, für die Gesamtheit aller Zeitspannen.

Bemerkung 1.2.7. Die Formeln unserer kleinen Formelsammlung von 1.2.5.3 gelten ganz genauso auch für die Lösungsmenge unserer Differentialgleichung $f'' = -f$, wenn wir $f \dot{+} g := f + g$ verstehen, für die Lösungsmenge unseres linearen Gleichungssystems, wenn wir

$$(x_1, \ldots, x_m) \dot{+} (x'_1, \ldots, x'_m) := (x_1 + x'_1, \ldots, x_m + x'_m)$$

als "komponentenweise Addition" verstehen, und für die Menge aller Folgen vom Fibonacci-Typ, wenn wir ähnlich die Summe $\dot{+}$ zweier Folgen erklären. Ein wesentliches Ziel der folgenden Vorlesungen über lineare Algebra ist es, einen abstrakten Formalismus aufzubauen, dem sich alle diese Beispiele unterordnen. Dadurch soll zweierlei erreicht werden:

1. Unser abstrakter Formalismus soll uns dazu verhelfen, die uns als Augentieren und Nachkommen von Ästehüpfern angeborene räumliche Anschauung nutzbar zu machen zum Verständnis der bis jetzt gegebenen Beispiele und der vielen weiteren Beispiele von Vektorräumen, denen Sie im Verlauf Ihres Studiums noch begegnen werden. So werden sie etwa lernen, daß man sich die Menge aller Folgen vom Fibonacci-Typ durchaus als Ebene vorstellen darf und die Menge aller Folgen mit vorgegebenem Folgenglied an einer vorgegebenen Stelle als eine Gerade in dieser Ebene. Suchen wir also alle Folgen vom Fibonacci-Typ mit zwei vorgegebenen Folgengliedern, so werden wir im allgemeinen genau eine derartige Lösung finden, da sich eben zwei Geraden in der Ebene im allgemeinen in genau einem Punkt schneiden. In diesem Licht betrachtet soll der abstrakte Formalismus uns also helfen, a priori unanschauliche Fragestellungen der Anschauung zugänglich zu machen. Ich denke, diese Nähe zur Anschauung ist auch der Grund dafür, daß die lineare Algebra meist an den Anfang des Studiums gestellt wird: Von der Schwierigkeit des Formalismus her gesehen gehört sie nämlich keineswegs zu den einfachsten Gebieten der Mathematik, hier würde ich eher an Gruppentheorie oder Graphentheorie oder dergleichen denken.

2. Unser abstrakter Formalismus soll so unmißverständlich sein und seine Spielregeln so klar, daß Sie in die Lage versetzt werden, alles nachzuvollziehen und mir im Prinzip und vermutlich auch in der Realität Fehler nachzuweisen. Schwammige Begriffe wie "Tafelebene" oder "Parallelverschiebung des Raums" haben in einem solchen Formalismus keinen Platz mehr. In diesem Licht betrachtet verfolgen wir mit dem Aufbau des abstrakten Formalismus auch das Ziel einer großen Vereinfachung durch die Reduktion auf die Be-

schreibung einiger weniger Aspekte der uns umgebenden in ihrer Komplexität kaum präzise faßbaren Wirklichkeit.

Die lineare Algebra hat in meinen Augen drei wesentliche Aspekte: Einen **geometrischen Aspekt**, wie ihn das Beispiel 1.2.5 der Gesamtheit aller Parallelverschiebungen illustriert; einen **algorithmischen Aspekt**, unter den ich das Beispiel 1.2.3 der Lösungsmenge eines linearen Gleichungssystems und insbesondere explizite Verfahren zur Bestimmung dieser Lösungsmenge einordnen würde; und einen **abstrakt-algebraischen Aspekt**, eine Art gedankliches Skelett, das Algorithmik und Geometrie verbindet und Brücken zu vielen weiteren Anwendungen schafft, die man dann auch als das Fleisch auf diesem Gerippe ansehen mag. Ich will im weiteren versuchen, diese drei Aspekte zu einer Einheit zu fügen, so daß Sie in die Lage versetzt werden, eine Vielzahl von Problemen mit den verbundenen Kräften Ihrer räumlichen Anschauung, Ihrer algorithmischen Rechenfähigkeiten und Ihres abstrakt-logischen Denkens anzugehen. Als Motivation für den weiteren Fortgang der Vorlesungen über lineare Algebra beschreibe ich nun das "Rückgrat unseres Skeletts" und formuliere ohne Rücksicht auf noch unbekannte Begriffe und Notationen die abstrakte Definition eines reellen Vektorraums.

Definition 1.2.8. Ein **reeller Vektorraum** ist ein Tripel bestehend aus den folgenden drei Dingen:

1. Einer Menge V;
2. Einer Verknüpfung $V \times V \to V$, $(v, w) \mapsto v \dot{+} w$, die V zu einer abelschen Gruppe macht;
3. Einer Abbildung $\mathbb{R} \times V \to V$, $(\alpha, v) \mapsto \alpha v$,

derart, daß für alle $\alpha, \beta \in \mathbb{R}$ und alle $v, w \in V$ gilt:

$$\begin{array}{rcl} \alpha(\beta v) & = & (\alpha\beta)v \\ (\alpha+\beta)v & = & (\alpha v) \dot{+} (\beta v) \\ \alpha(v \dot{+} w) & = & (\alpha v) \dot{+} (\alpha w) \\ 1v & = & v \end{array}$$

Hier ist nun viel zu klären: Was ist eine Menge? Eine Verknüpfung? Eine abelsche Gruppe? Eine Abbildung? Was bedeuten die Symbole $\times$, $\to$, $\mapsto$, $\in$, $\mathbb{R}$? Wir beginnen in der nächsten Vorlesung mit der Klärung dieser Begriffe und Notationen.

2 Naive Mengenlehre und Kombinatorik

2.1 Mengen

2.1.1. Beim Arbeiten mit reellen Zahlen oder räumlichen Gebilden reicht auf der Schule ein intuitives Verständnis meist aus, und wenn die Intuition in die Irre führt, ist ein Lehrer zur Stelle. Wenn Sie jedoch selbst unterrichten oder etwas beweisen wollen, reicht dieses intuitive Verständnis nicht mehr aus. *Im folgenden werden deshalb zunächst der Begriff der reellen Zahlen und der Begriff des Raums zurückgeführt auf Grundbegriffe der Mengenlehre, den Begriff der rationalen Zahlen und elementare Logik.* Bei der Arbeit mit diesen Begriffen führt uns die Intuition nicht so leicht in die Irre, wir geben uns deshalb mit einem intuitiven Verständnis zufrieden und verweisen jeden, der es noch genauer wissen will, auf eine Vorlesung über Logik. Wir beginnen mit etwas naiver Mengenlehre, wie sie von Georg Cantor in den Jahren 1874-1897 begründet wurde, und von der der berühmte Mathematiker David Hilbert einmal sagte: "Aus dem Paradies, das Cantor uns geschaffen, soll uns niemand vertreiben können". Natürlich gab es auch vor der Mengenlehre schon hoch entwickelte Mathematik, bei Carl Friedrich Gauß Tod 1855 gab es diese Theorie noch gar nicht und Fourier fand seine "Fourierentwicklung" sogar bereits zu Beginn des 19.-ten Jahrhunderts. Er behauptete auch gleich in seiner "Théorie analytique de la chaleur", daß sich jede beliebige (periodische) Funktion durch eine Fouriereihe darstellen lasse, aber diese Behauptung stieß bei anderen berühmten Mathematikern seiner Zeit auf Ablehnung und es entstand darüber ein heftiger Disput. Erst im "Paradies der Mengenlehre" konnten die Fourier's Behauptung zugrundeliegenden Begriffe soweit geklärt werden, daß dieser Disput nun endgültig beigelegt ist. Ähnlich verhält es sich auch mit vielen anderen Fragestellungen. Da die Mengenlehre darüber hinaus auch vom didaktischen Standpunkt aus eine äußerst klare und durchsichtige Darstellung mathematischer Sachverhalte ermöglicht, hat sie sich als Grundlage der Mathematik und der Ausbildung von Mathematikern an Universitäten sehr schnell durchgesetzt und ist nun weltweit ein wesentlicher Teil des "Alphabets der Sprache der Mathematiker".

2.1.2. Im Wortlaut der ersten Zeilen des Artikels "Beiträge zur Begründung der transfiniten Mengenlehre (Erster Aufsatz)" von Georg Cantor, erschienen im Jahre 1895, hört sich die Definition einer Menge so an:

> Unter einer **Menge** verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objecten m unserer Anschauung oder unseres Denkens (welche die **Elemente** von M genannt werden) zu einem Ganzen.

Verbinden wir mit einer Menge eine geometrische Vorstellung, so nennen wir ihre Elemente auch **Punkte** und die Menge selbst einen **Raum**. Ein derartiges Herumgerede ist natürlich keine formale Definition und birgt auch verschiedene Fallstricke, vergleiche 2.1.19, aber das Ziel dieser Vorlesung ist auch nicht eine formale Begründung der Mengenlehre, wie Sie sie später in der Logik kennenlernen können. Sie sollen vielmehr die Bedeutung dieser Worte intuitiv erfassen wie ein Kleinkind, das Sprechen lernt: Indem sie mir und anderen Mathematikern zuhören, wie wir mit diesen Worten sinnvolle Sätze bilden, uns nachahmen, und beobachten, welchen Effekt Sie damit hervorrufen. Unter anderem dazu sind die Übungsgruppen da.

Beispiele 2.1.3. Endliche Mengen gibt man oft durch eine vollständige Liste ihrer Elemente in geschweiften Klammern an, zum Beispiel in der Form $X = \{x_1, x_2, \ldots, x_n\}$. Die Elemente dürfen mehrfach genannt werden und es kommt nicht auf die Reihenfolge an, in der sie genannt werden. So haben wir also $\{1, 1, 2\} = \{2, 1\}$. Die Aussage "x ist Element von X" wird mit $x \in X$ abgekürzt, ihre Verneinung "x ist nicht Element von X" mit $x \notin X$. Es gibt auch die sogenannte **leere Menge** $\emptyset = \{\ \}$, die gar kein Element enthält. Andere Beispiele sind die Menge der **natürlichen Zahlen** $\mathbb{N} = \{0, 1, 2, \ldots\}$, die Menge der **ganzen Zahlen** $\mathbb{Z} = \{0, 1, -1, 2, -2, \ldots\}$ und die Menge der **rationalen Zahlen** $\mathbb{Q} = \{p/q \mid p, q \in \mathbb{Z}, q \neq 0\}$. Deren Name kommt von lateinisch "ratio" für "Verhältnis". Man beachte, daß wir auch hier Elemente mehrfach genannt haben, es gilt ja $p/q = p'/q'$ genau dann, wenn $pq' = p'q$. Auf deutsch bezeichnet man die rationalen Zahlen manchmal auch als **Bruchzahlen**.

Bemerkung 2.1.4. In Texten, in deren Konventionen die Null keine natürliche Zahl ist, verwendet man meist die abweichenden Notationen $\mathbb{N}$ für die Menge $\{1, 2, \ldots\}$ und $\mathbb{N}_0$ für die Menge $\{0, 1, 2, \ldots\}$. Die in diesem Text verwendete Notation $\mathbb{N} = \{0, 1, 2, \ldots\}$ stimmt mit der internationalen Norm ISO 31-11 überein.

Definition 2.1.5. Eine Menge Y heißt **Teilmenge** einer Menge X genau dann, wenn jedes Element von Y auch ein Element von X ist. Man schreibt dafür $Y \subset X$ oder $X \supset Y$. Zum Beispiel gilt stets $\emptyset \subset X$ und $\{x\} \subset X$ ist gleichbedeutend zu $x \in X$. Zwei Teilmengen einer gegebenen Menge, die kein gemeinsames Element haben, heißen **disjunkt**.

Bemerkung 2.1.6. Diese Notation weicht ab von der internationalen Norm ISO 31-11, die statt unserem $\subset$ das Symbol $\subseteq$ vorschlägt. In den Konventionen von ISO 31-11 hat das Symbol $\subset$ abweichend die Bedeutung einer **echten**, d.h. von der ganzen Menge verschiedenen Teilmenge, für die wir die Bezeichnung $\subsetneq$ verwenden werden. Meine Motivation für diese Abweichung

ist, daß das Symbol für beliebige Teilmengen sehr häufig und das für echte Teilmengen nur sehr selten vorkommt.

Definition 2.1.7. Wir vereinbaren, daß wir auch die leere Menge endlich nennen wollen, damit jede Teilmenge einer endlichen Menge auch wieder endlich ist. Die Zahl der Elemente einer endlichen Menge X nennen wir ihre **Kardinalität** oder **Mächtigkeit** und notieren sie $|X|$ oder $\mathrm{card}(X)$. In der Literatur findet man auch die Notation $\sharp X$. Ist X unendlich, so schreiben wir kurz $|X| = \infty$ und ignorieren in unserer Notation, daß auch unendliche Mengen "verschieden groß" sein können, für ein Beispiel siehe **??** und für eine genauere Diskussion des Begriffs der Kardinalität **??**. Für endliche Mengen X ist demnach ihre Kardinalität stets eine natürliche Zahl $|X| \in \mathbb{N}$ und $|X| = 0$ ist gleichbedeutend zu $X = \emptyset$.

Definition 2.1.8. Oft bildet man neue Mengen als Teilmengen bestehender Mengen und schreibt $Y = \{x \in X \mid x \text{ hat eine gewisse Eigenschaft}\}$. Zum Beispiel gilt $\mathbb{N} = \{a \in \mathbb{Z} \mid a \geq 0\}$ oder $\{0, 1\} = \{a \in \mathbb{N} \mid a^2 = a\}$. Eine Variante dieser Notation soll hier nur mit zwei Beispielen erklärt werden: $\{2a \mid a \in \mathbb{N}\}$ bezeichnet die Menge aller geraden natürlichen Zahlen, $\{ab \mid a, b \in \mathbb{N}, a \geq 2, b \geq 2\}$ die Menge aller natürlichen Zahlen, die nicht prim und auch nicht Null oder Eins sind.

Definition 2.1.9. Es ist auch erlaubt, die "Menge aller Teilmengen" einer gegebenen Menge X zu bilden. Sie heißt die **Potenzmenge** von X und wird mit $\mathcal{P}(X)$ bezeichnet.

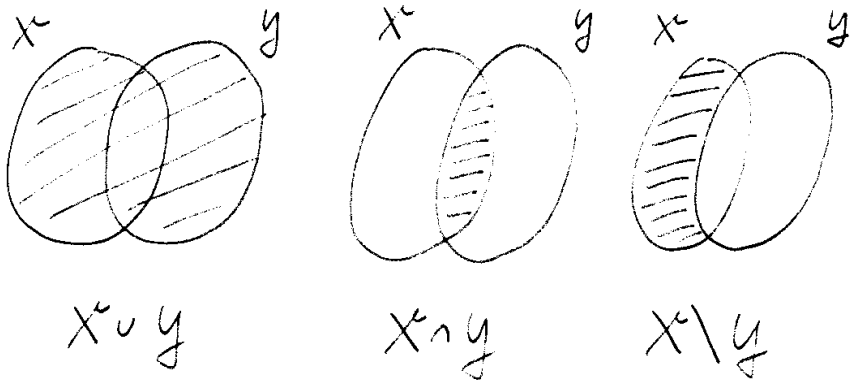
2.1.10. Ist X eine endliche Menge, so ist auch ihre Potenzmenge endlich und es gilt $|\mathcal{P}(X)| = 2^{|X|}$. Für $X = \{1, 2\}$ besteht zum Beispiel $\mathcal{P}(X)$ aus vier Elementen, genauer gilt $\mathcal{P}(X) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$.

Definition 2.1.11. Gegeben zwei Mengen X, Y können wir auf verschiedene Arten neue Mengen bilden:

1. Die **Vereinigung** $X \cup Y := \{z \mid z \in X \text{ oder } z \in Y\}$, zum Beispiel ist $\{1, 2\} \cup \{2, 3\} = \{1, 2, 3\}$.

2. Den **Schnitt** $X \cap Y := \{z \mid z \in X \text{ und } z \in Y\}$, zum Beispiel ist $\{1, 2\} \cap \{2, 3\} = \{2\}$. Zwei Mengen sind also disjunkt genau dann, wenn ihr Schnitt die leere Menge ist.

3. Die **Differenz** $X \backslash Y := \{z \in X \mid z \notin Y\}$, zum Beispiel haben wir $\{1, 2\} \backslash \{2, 3\} = \{1\}$. Man schreibt statt $X \backslash Y$ auch $X - Y$. Ist Y eine Teilmenge von X, so heißt $X \backslash Y$ das **Komplement** von Y in X.



Eine gute Anschauung für die ersten drei Operationen liefern die sogenannten **van-de-Ven-Diagramme** wie sie die obenstehenden Bilder zeigen. Sie sind allerdings nicht zu genau zu hinterfragen, denn ob die Punkte auf einem Blatt Papier im Sinne von Cantor "bestimmte wohlunterschiedene Objekte unserer Anschauung" sind, scheint mir sehr fraglich. Wenn man jedoch jedes der schraffierten Gebiete im Bild auffasst als die Menge aller darin liegenden Kreuzungspunkte auf einem dazugedachten Millimeterpapier und keine dieser Kreuzungspunkte auf den Begrenzungslinien liegen, so können sie wohl schon als eine Menge im Cantor'schen Sinne angesehen werden.

4. Das **kartesische Produkt** $X \times Y := \{(x,y) \mid x \in X,\ y \in Y\}$, als da heißt die Menge aller geordneten Paare. Es gilt also $(x,y) = (x',y')$ genau dann, wenn gilt $x = x'$ und $y = y'$. Zum Beispiel haben wir $\{1,2\} \times \{1,2,3\} = \{(1,1),(1,2),(1,3),(2,1),(2,2),(2,3)\}$. Oft benutzt man für das kartesische Produkt $X \times X$ einer Menge X mit sich selbst die Abkürzung $X \times X = X^2$.

2.1.12. Wir werden in unserer naiven Mengenlehre die ersten drei Operationen nur auf Teilmengen einer gemeinsamen Obermenge anwenden, die uns in der einen oder anderen Weise bereits zur Verfügung steht. Die Potenzmenge und das kartesische Produkt dahingegen benutzen wir, um darüber hinaus neue Mengen zu erschaffen. Diese Konstruktionen erlauben es, im Rahmen der Mengenlehre so etwas wie Abstraktionen zu bilden: Wenn wir uns etwa die Menge T aller an mindestens einem Tag der Weltgeschichte lebenden oder gelebt habenden Tiere als eine Menge im Cantor'schen Sinne denken, so würden wir Konzepte wie "männlich" oder "Hund" oder "Fleischfresser" formal als Teilmengen dieser Menge definieren, d.h. als Elemente von $\mathcal{P}(T)$, und das Konzept "ist Kind von" als eine Teilmenge des kartesischen Produkts dieser Menge T mit sich selbst, also als ein Element von $\mathcal{P}(T \times T)$.

2.1.13. Für das Rechnen mit Mengen überlegt man sich die folgenden Regeln:

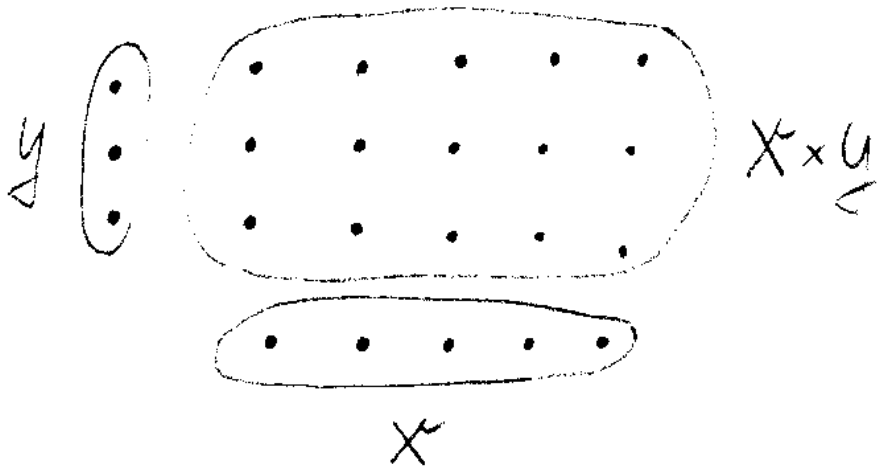
$$\begin{aligned}
X \cap (Y \cap Z) &= (X \cap Y) \cap Z \\
X \cup (Y \cup Z) &= (X \cup Y) \cup Z \\
X \cup (Y \cap Z) &= (X \cup Y) \cap (X \cup Z) \\
X \cap (Y \cup Z) &= (X \cap Y) \cup (X \cap Z) \\
X \backslash (Y \cup Z) &= (X \backslash Y) \cap (X \backslash Z) \\
X \backslash (Y \cap Z) &= (X \backslash Y) \cup (X \backslash Z) \\
X \backslash (X \backslash Y) &= X \cap Y
\end{aligned}$$

Eine gute Anschauung für diese Regeln liefern die van-de-Ven-Diagramme, wie sie die nebenstehenden Bilder zeigen.

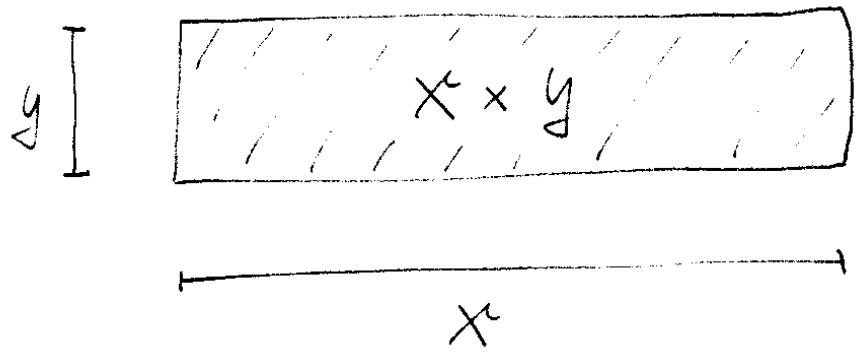
Übung 2.1.14. Sind X und Y endliche Mengen, so gilt für die Kardinalitäten $|X \times Y| = |X| \cdot |Y|$ und $|X \cup Y| = |X \backslash Y| + |X \cap Y| + |Y \backslash X|$.

Satz 2.1.15 (Bedeutung der Binomialkoeffizienten). *Gegeben $n, k \in \mathbb{N}$ gibt der Binomialkoeffizient $\binom{n}{k}$ die Zahl der k-elementigen Teilmengen einer n-elementigen Menge an, in Formeln*

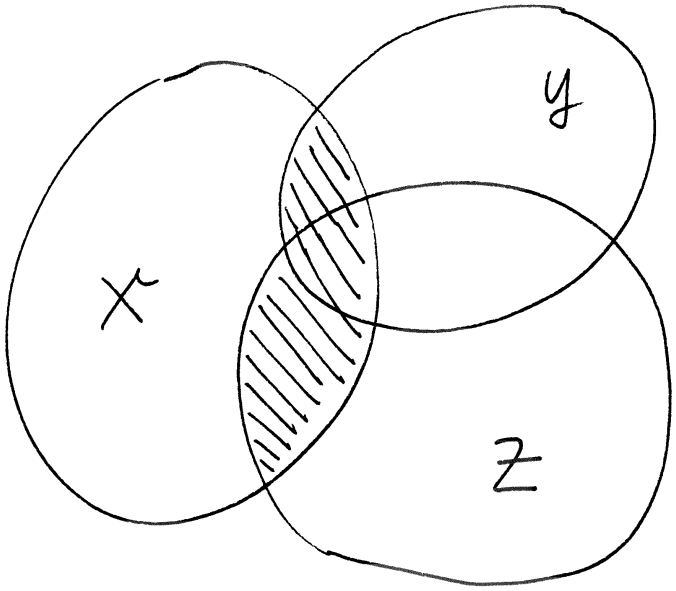
$$|X| = n \quad \textit{impliziert} \quad |\{Y \subset X \mid |Y| = k\}| = \tbinom{n}{k}$$



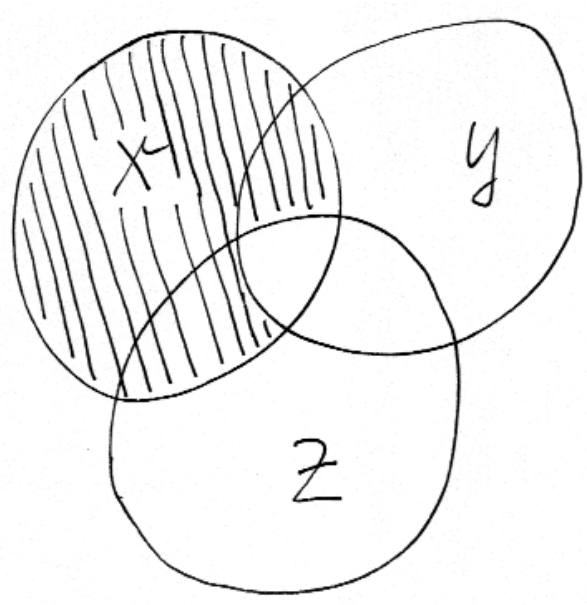
Anschauliche Darstellung des Produkts einer Menge mit fünf und einer Menge mit drei Elementen. Hier wird ein Paar (x, y) dargestellt durch einen fetten Punkt, der über x und neben y liegt.



Dies Bild muß anders interpretiert werden als das Vorherige. Die Mengen X und Y sind nun zu verstehen als die Mengen der Punkte der vertikalen und horizontalen Geradensegmente und ein Punkt des Quadrats meint das Element $(x, y) \in X \times Y$, das in derselben Höhe wie $y \in Y$ senkrecht über $x \in X$ liegt.



$$X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$$



$$X \setminus (Y \cap Z) = (X \setminus Y) \cup (X \setminus Z)$$

Beweis. Vollständige Induktion über n. Für $n = 0$ gilt die Aussage, denn eine nullelementige Menge hat genau eine k-elementige Teilmenge falls $k = 0$ und keine k-elementige Teilmenge falls $k \geq 1$. Nehmen wir nun an, die Aussage sei für ein n schon bewiesen. Eine $(n+1)$-elementige Menge X schreiben wir als $X = M \cup \{x\}$, wo M eine n-elementige Menge ist und $x \notin M$. Ist $k = 0$, so gibt es genau eine k-elementige Teilmenge von $M \cup \{x\}$, nämlich die leere Menge. Ist $k \geq 1$, so gibt es in $M \cup \{x\}$ nach Induktionsannahme genau $\binom{n}{k}$ k-elementige Teilmengen, die x nicht enthalten. Die k-elementigen Teilmengen dahingegen, die x enthalten, ergeben sich durch Hinzunehmen von x aus den $(k-1)$-elementigen Teilmengen von M, von denen es gerade $\binom{n}{k-1}$ gibt. Insgesamt hat $M \cup \{x\}$ damit also genau $\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}$ k-elementige Teilmengen. □

Bemerkung 2.1.16. Wieder scheint mir dieser Beweis in der für vollständige Induktion typischen Weise undurchsichtig. Ich ziehe deshalb den in 1.1.16 gegebenen weniger formellen Beweis vor. Man kann auch diesen Beweis formalisieren und verstehen als Spezialfall der sogenannten "Bahnformel" III.6.3.2, vergleiche III.6.3.3.

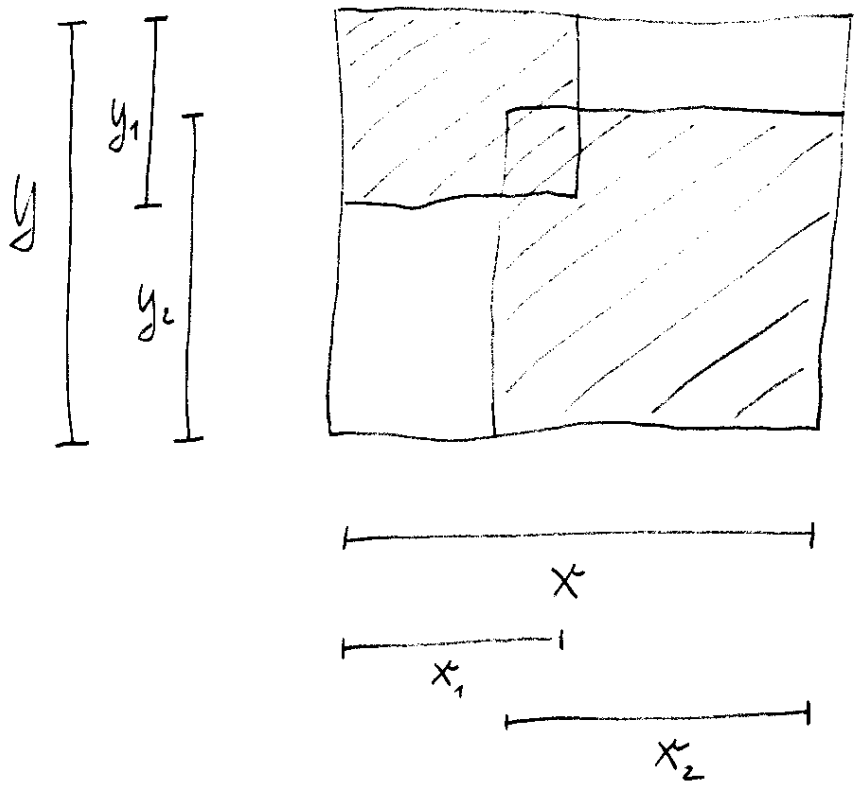
Bemerkung 2.1.17. Wir geben nun die versprochene präzise Formulierung unseres ersten Beweises der binomischen Formel 1.1.20. Wir rechnen dazu

$$(a+b)^n = \sum_{Y \subset \{1,2,\ldots,n\}} a^{|Y|} b^{n-|Y|}$$

wo die rechte Seite in Verallgemeinerung der in Abschnitt 1.1 eingeführten Notation bedeuten soll, daß wir für jede Teilmenge Y von $\{1, 2, \ldots, n\}$ den angegebenen Ausdruck $a^{|Y|}b^{n-|Y|}$ nehmen und alle diese Ausdrücke aufsummieren. Dann fassen wir gleiche Summanden zusammen und erhalten mit 2.1.15 die binomische Formel.

Übung 2.1.18. Es gilt $\sum_k \binom{n}{k} = 2^n$.

2.1.19. Ich will nicht verschweigen, daß der in diesem Abschnitt dargestellte naive Zugang zur Mengenlehre durchaus begriffliche Schwierigkeiten mit sich bringt: Zum Beispiel darf die Gesamtheit $\mathcal{M}$ aller Mengen nicht als Menge angesehen werden, da wir sonst die "Menge aller Mengen, die sich nicht selbst als Element enthalten", gegeben durch die formelhafte Beschreibung $\mathcal{N} = \{A \in \mathcal{M} \mid A \notin A\}$, bilden könnten. Für diese Menge kann aber weder $\mathcal{N} \in \mathcal{N}$ noch $\mathcal{N} \notin \mathcal{N}$ gelten ... Diese Art von Schwierigkeiten kann erst ein formalerer Zugang klären und auflösen, bei dem man unsere naiven Vorstellungen durch Ketten von Zeichen aus einem wohlbestimmten endlichen Alphabet ersetzt und unsere Vorstellung von Wahrheit durch die Verifizierbarkeit vermittels rein algebraischer "erlaubter Manipulationen"



Aus $X = X_1 \cup X_2$ und $Y = Y_1 \cup Y_2$ folgt noch lange nicht
$X \times Y = (X_1 \times Y_1) \cup (X_2 \times Y_2)$

solcher Zeichenketten, die in "Axiomen" festgelegt werden. Diese Verifikationen kann man dann durchaus auch einer Rechenmaschine überlassen, so daß wirklich auf "objektivem" Wege entschieden werden kann, ob ein "Beweis" für die "Richtigkeit" einer unserer Zeichenketten in einem vorgegebenen axiomatischen Rahmen stichhaltig ist. Allerdings kann in derartigen Systemen von einer Zeichenkette algorithmisch nur entschieden werden, ob sie eine "sinnvolle Aussage" ist, nicht aber, ob sie "bewiesen" werden kann. Noch viel stärker zeigt der Unvollständigkeitssatz von Gödel, daß es in einem derartigen axiomatischen Rahmen, sobald er reichhaltig genug ist für eine Beschreibung des Rechnens mit natürlichen Zahlen, stets sinnvolle Aussagen gibt derart, daß entweder sowohl die Aussage als auch ihre Verneinung oder aber weder die Aussage noch ihre Verneinung bewiesen werden können. Mit diesen und ähnlichen Fragestellungen beschäftigt sich die Logik in ihren Grenzbereichen zur Informatik.

2.1.20. Um mich nicht dem Vorwurf auszusetzen, während des Spiels die Spielregeln ändern zu wollen, sei bereits hier erwähnt, daß in III.1.2 noch weitere wichtige Konstruktionen der Mengenlehre eingeführt werden, und daß wir in **??** einige weniger offensichtliche Folgerungen erläutern, die meines Erachtens bereits an den Rand dessen gehen, was man in unserem informellen Rahmen als Argumentation noch vertreten kann.

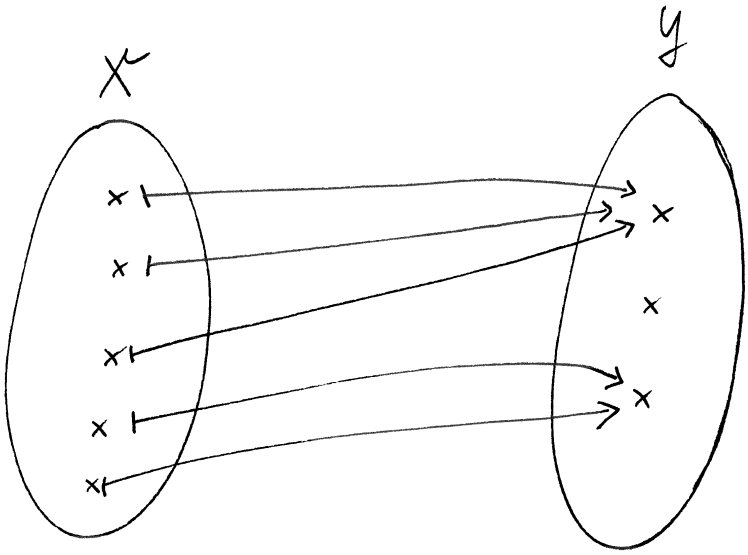
2.2 Abbildungen

Definition 2.2.1. Seien X, Y Mengen. Eine **Abbildung** $f : X \to Y$ ist eine Zuordnung, die jedem Element $x \in X$ genau ein Element $f(x) \in Y$ zuordnet, das **Bild** von x unter f, auch genannt der **Wert** von f an der Stelle x.

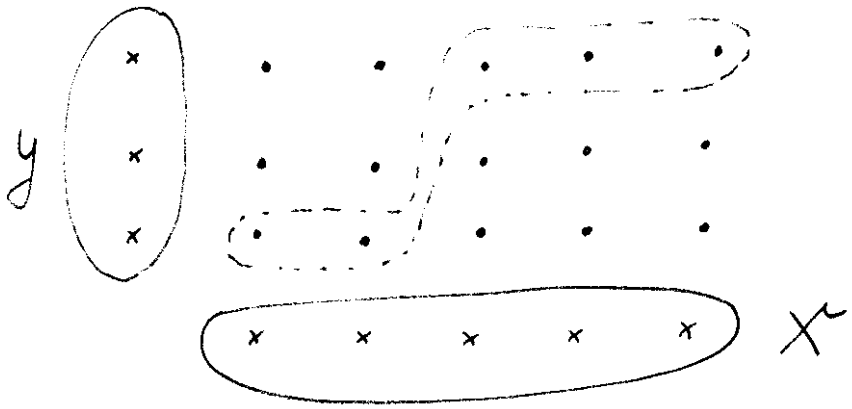
2.2.2. Wem das zu vage ist, der mag die alternative Definition vorziehen, nach der eine **Abbildung** $f : X \to Y$ eine Teilmenge $f \subset X \times Y$ ist derart, daß es für jedes $x \in X$ genau ein $y \in Y$ gibt mit $(x, y) \in f$. Dies eindeutig bestimmte y schreiben wir dann $f(x)$ und sind auf einem etwas formaleren Weg wieder am selben Punkt angelangt. In unseren Konventionen nennen wir besagte Teilmenge den **Graphen von** f und notieren sie mit dem Symbol Γ (sprich: Gamma), einem großen griechischen G, und schreiben also

$$\Gamma(f) := \{(x, f(x)) \mid x \in X\} \subset X \times Y$$

Definition 2.2.3. Ist $f : X \to Y$ eine Abbildung, so nennen wir X ihren **Definitionsbereich** und Y ihren **Wertebereich**. Zwei Abbildungen nennen wir gleich genau dann, wenn sie denselben Definitionsbereich X, denselben Wertebereich Y und dieselbe Abbildungsvorschrift $f \subset X \times Y$ haben. Die



Eine Abbildung einer Menge mit fünf in eine mit drei Elementen



Der Graph der oben angegebenen Abbildung, wobei das X oben mit dem X hier identifiziert wurde durch "Umkippen nach Rechts"

Menge aller Abbildungen von X nach Y bezeichnen wir mit $\mathrm{Ens}(X,Y)$ nach der französischen Übersetzung **ensemble** des deutschen Begriffs "Menge". Üblich ist auch die Notation Y^X.

Bemerkung 2.2.4. Noch gebräuchlicher ist die Bezeichnung $\mathrm{Abb}(X,Y)$ für die Menge aller Abbildungen von X nach Y. Ich will jedoch sehr viel später die "Kategorie aller Mengen" mit Ens bezeichnen und für je zwei Objekte X,Y einer Kategorie $\mathcal{C}$ die Menge aller "Morphismen" von X nach Y mit $\mathcal{C}(X,Y)$, und das motiviert dann auch erst eigentlich die hier gewählte Bezeichnung für Mengen von Abbildungen.

2.2.5. Wir notieren Abbildungen oft in der Form

$$\begin{array}{rccc} f: & X & \to & Y \\ & x & \mapsto & f(x) \end{array}$$

und in verschiedenen Verkürzungen dieser Notation. Zum Beispiel sprechen wir von "einer Abbildung $X \to Y$" oder "der Abbildung $n \mapsto n^3$ von der Menge der natürlichen Zahlen in sich selber". Wir benutzen unsere zwei Arten von Pfeilen auch im allgemeinen in derselben Weise.

Beispiel 2.2.6. Für jede Menge X haben wir die **identische Abbildung** oder **Identität**

$$\begin{array}{rccc} \mathrm{id} = \mathrm{id}_X: & X & \to & X \\ & x & \mapsto & x \end{array}$$

Ein konkreteres Beispiel für eine Abbildung ist das Quadrieren

$$\begin{array}{rccc} q: & \mathbb{Z} & \to & \mathbb{Z} \\ & n & \mapsto & n^2 \end{array}$$

Beispiel 2.2.7. Gegeben zwei Mengen X,Y erklärt man die sogenannten **Projektionsabbildungen** oder **Projektionen** $\mathrm{pr}_X : X \times Y \to X$ bzw. $\mathrm{pr}_Y : X \times Y \to Y$ durch die Vorschrift $(x,y) \mapsto x$ bzw. $(x,y) \mapsto y$.

Definition 2.2.8. Ist $f : X \to Y$ eine Abbildung und $A \subset X$ eine Teilmenge, so definieren wir ihr **Bild** oder genauer ihre **Bildmenge** $f(A)$, eine Teilmenge von Y, durch

$$f(A) := \{f(x) \mid x \in A\}$$

Eine Abbildung, deren Bild aus höchstens einem Element besteht, nennen wir eine **konstante Abbildung**. Eine Abbildung, deren Bild aus genau einem Element besteht, nennen wir eine **einwertige Abbildung**. In anderen Worten ist eine einwertige Abbildung also eine konstante Abbildung mit nichtleerem Definitionsbereich.

Beispiel 2.2.9. Für unsere Abbildung $q : \mathbb{Z} \to \mathbb{Z}$, $x \mapsto x^2$ von eben gilt

$$q(\mathbb{Z}) = \{a^2 \mid a \in \mathbb{Z}\} \subset \mathbb{N}$$

2.2.10. Gegeben ein festes $c \in Y$ schreiben wir oft auch kurz c für die konstante Abbildung $X \to Y$, $x \mapsto c$ für alle $x \in X$ in der Hoffnung, daß aus dem Kontext klar wird, ob die Abbildung $c : X \to Y$ oder das Element $c \in Y$ gemeint sind.

Definition 2.2.11. Ist $f : X \to Y$ eine Abbildung und $B \subset Y$ eine Teilmenge, so definieren wir ihr **Urbild** $f^{-1}(B)$, eine Teilmenge von X, durch

$$f^{-1}(B) := \{x \in X \mid f(x) \in B\}$$

Besteht B nur aus einem Element, so schreiben wir auch $f^{-1}(x)$ statt $f^{-1}(\{x\})$ und nennen diese Menge die **Faser** von f über x. Die Abbildung q aus 2.2.9 hat etwa die Fasern $q^{-1}(1) = \{1, -1\}$ und $q^{-1}(-1) = \emptyset$.

Definition 2.2.12. Sind schließlich drei Mengen X, Y, Z gegeben und Abbildungen $f : X \to Y$ und $g : Y \to Z$, so definieren wir eine Abbildung $g \circ f : X \to Z$, die **Verknüpfung** der Abbildungen f und g, durch die Vorschrift

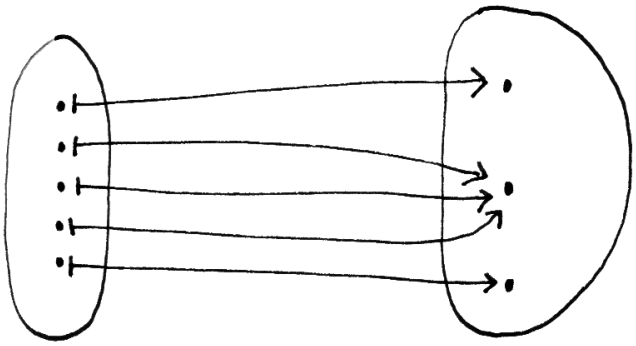
$$\begin{array}{rccc} g \circ f : & X & \to & Z \\ & x & \mapsto & g(f(x)) \end{array}$$

2.2.13. Die Notation $g \circ f$, sprich "g nach f" für "erst f, dann g" ist gewöhnungsbedürftig, erklärt sich aber durch die Formel $(g \circ f)(x) = g(f(x))$. Betrachten wir zum Beispiel zusätzlich zum Quadrieren $q : \mathbb{Z} \to \mathbb{Z}$ die Abbildung $t : \mathbb{Z} \to \mathbb{Z}$, $x \mapsto x+1$, so gilt $(q \circ t)(x) = (x+1)^2$ aber $(t \circ q)(x) = x^2+1$. Natürlich gilt $(g \circ f)(A) = g(f(A))$ für jede Teilmenge $A \subset X$ und umgekehrt auch $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$ für jede Teilmenge $C \subset Z$.

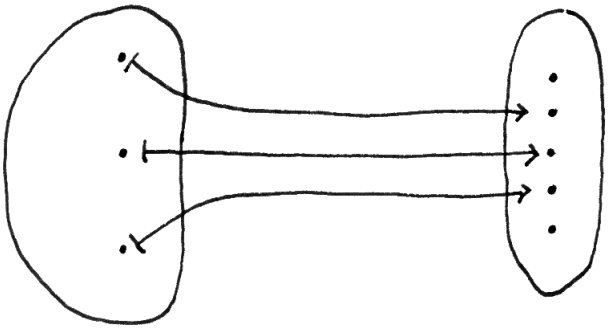
Definition 2.2.14. Sei $f : X \to Y$ eine Abbildung.

1. f heißt **injektiv** oder eine **Injektion** genau dann, wenn aus $x \neq x'$ folgt $f(x) \neq f(x')$. Gleichbedeutend ist die Forderung, daß es für jedes $y \in Y$ höchstens ein $x \in X$ gibt mit $f(x) = y$. Injektionen schreibt man oft $\hookrightarrow$.

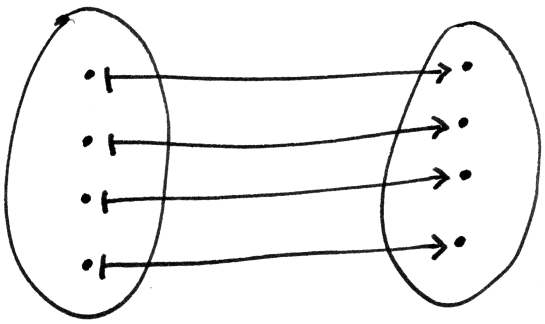
2. f heißt **surjektiv** oder eine **Surjektion** genau dann, wenn es für jedes $y \in Y$ mindestens ein $x \in X$ gibt mit $f(x) = y$. Surjektionen schreibt man manchmal $\twoheadrightarrow$.



Eine Surjektion



Eine Injektion



Eine Bijektion

3. f heißt **bijektiv** oder eine **Bijektion** genau dann, wenn f injektiv und surjektiv ist. Gleichbedeutend ist die Forderung, daß es für jedes $y \in Y$ genau ein $x \in X$ gibt mit $f(x) = y$. Bijektionen schreibt man oft $\xrightarrow{\sim}$.

2.2.15. Ist $X \subset Y$ eine Teilmenge, so ist die **Einbettung** oder **Inklusion** $i : X \to Y$, $x \mapsto x$ stets injektiv. Ist $g : Y \to Z$ eine Abbildung und $X \subset Y$ eine Teilmenge, so nennen wir die Verknüpfung $g \circ i$ von g mit der Inklusion auch die **Einschränkung** von g auf X und notieren sie $g \circ i = g|X = g|_X : X \to Z$. Oft bezeichnen wir eine Einschränkung aber auch einfach mit demselben Buchstaben g in der Hoffnung, daß der Leser aus dem Kontext erschließen kann, welche Abbildung genau gemeint ist.

2.2.16. Ist $f : X \to Y$ eine Abbildung, so ist die Abbildung $f : X \to f(X)$, $x \mapsto f(x)$ stets surjektiv. Der Leser möge entschuldigen, daß wir hier zwei verschiedene Abbildungen mit demselben Symbol f bezeichnet haben. Das wird noch öfter vorkommen.

2.2.17. Gegeben eine Menge X kann eine Abbildung $f : X \to \mathcal{P}(X)$ nie surjektiv sein. In der Tat, betrachten wir $A = \{x \in X \mid x \notin f(x)\}$, so kann es kein $y \in X$ geben mit $f(y) = A$, denn für solch ein y hätten wir entweder $y \in A$ oder $y \notin A$, und aus $y \in A$ alias $y \in f(y)$ folgte $y \notin A$, wohingegen aus $y \notin A$ alias $y \notin f(y)$ folgte $y \in A$.

Beispiele 2.2.18. Unsere Abbildung $q : \mathbb{Z} \to \mathbb{Z}$, $n \mapsto n^2$ ist weder injektiv noch surjektiv. Die Identität $\mathrm{id} : X \to X$ ist stets bijektiv. Sind X und Y endliche Mengen, so gibt es genau dann eine Bijektion von X nach Y, wenn X und Y dieselbe Kardinalität haben, in Formeln $|X| = |Y|$.

Satz 2.2.19. *Seien $f, f_1 : X \to Y$ und $g, g_1 : Y \to Z$ Abbildungen.*

1. *Ist $g \circ f$ surjektiv, so ist g surjektiv.*
2. *Ist $g \circ f$ injektiv, so ist f injektiv.*
3. *Sind g und f injektiv, so auch $g \circ f$.*
4. *Sind g und f surjektiv, so auch $g \circ f$.*
5. *Ist f surjektiv, so folgt aus $g \circ f = g_1 \circ f$ schon $g = g_1$.*
6. *Ist g injektiv, so folgt aus $g \circ f = g \circ f_1$ schon $f = f_1$.*

Beweis. Übung. □

2.2.20. Ist $f : X \to Y$ eine bijektive Abbildung, so ist ofenichtlich die Menge $\{(f(x), x) \in Y \times X \mid x \in X\}$ im Sinne von 2.2.2 eine Abbildung oder, vielleicht klarer, der Graph einer Abbildung $Y \to X$. Diese Abbildung in die Gegenrichtung heißt die **Umkehrabbildung** oder **Umkehrfunktion** auch die **inverse Abbildung** zu f und wird mit $f^{-1} : Y \to X$ bezeichnet. Offensichtlich ist mit f auch f^{-1} eine Bijektion.

Beispiel 2.2.21. Die Umkehrabbildung unserer Bijektion $t : \mathbb{Z} \to \mathbb{Z}, x \mapsto x+1$ ist die Abbildung $\mathbb{Z} \to \mathbb{Z}, x \mapsto x - 1$.

Übung 2.2.22. Gegeben eine Bijektion $f : X \to Y$ ist $g = f^{-1}$ die einzige Abbildung $g : Y \to X$ mit $f \circ g = \mathrm{id}_Y$. Ebenso ist auch $h = f^{-1}$ die einzige Abbildung $h : Y \to X$ mit $h \circ f = \mathrm{id}_X$.

2.2.23. Gegeben drei Mengen X, Y, Z haben wir eine offensichtliche Bijektion $\mathrm{Ens}(X \times Y, Z) \xrightarrow{\sim} \mathrm{Ens}(X, \mathrm{Ens}(Y, Z))$. Etwas vage formuliert ist also eine Abbildung $X \times Y \to Z$ dasselbe wie eine Abbildung, die jedem $x \in X$ eine Abbildung $Y \to Z$ zuordnet, und symmetrisch natürlich auch dasselbe wie eine Abbildung, die jedem $y \in Y$ eine Abbildung $X \to Z$ zuordnet. In der exponentiellen Notation liest sich das ganz suggestiv als kanonische Bijektion $Z^{(X\times Y)} \xrightarrow{\sim} (Z^X)^Y$.

Satz 2.2.24 (**Bedeutung der Fakultät**). *Sind X und Y zwei Mengen mit je n Elementen, so gibt es genau $n!$ bijektive Abbildungen $f : X \to Y$.*

Beweis. Sei $X = \{x_1, \ldots, x_n\}$. Wir haben n Möglichkeiten, ein Bild für x_1 auszusuchen, dann noch $(n - 1)$ Möglichkeiten, ein Bild für x_2 auszusuchen, und so weiter, bis schließlich nur noch 1 Element von Y als mögliches Bild von x_n in Frage kommt. Insgesamt gibt es also $n(n - 1) \cdots 1 = n!$ Möglichkeiten für f. Da wir $0! = 1$ gesetzt hatten, stimmt unser Satz auch für $n = 0$. □

Übung 2.2.25. Seien X, Y endliche Mengen. So gibt es genau $|Y|^{|X|}$ Abbildungen von X nach Y, und unter diesen Abbildungen sind genau $|Y|(|Y| - 1)(|Y| - 2) \ldots (|Y| - |X| + 1)$ Injektionen.

Übung 2.2.26. Sei X eine Menge mit n Elementen, und seien $\alpha_1, \ldots, \alpha_r \in \mathbb{N}$ gegeben mit $n = \alpha_1 + \ldots + \alpha_r$. Man zeige: Es gibt genau $n!/(\alpha_1! \cdots \alpha_r!)$ Abbildungen $f : X \to \{1, \ldots, r\}$, die jedes i genau α_i-mal als Wert annehmen, in Formeln

$$\frac{n!}{\alpha_1! \cdots \alpha_r!} = \mathrm{card}\{f \mid |f^{-1}(i)| = \alpha_i \text{ für } i = 1, \ldots r\}$$

2.2.27. Manche Autoren bezeichnen diese Zahlen auch als **Multinomialkoeffizienten** und verwenden die Notation

$$\frac{n!}{\alpha_1! \cdots \alpha_r!} = \binom{n}{\alpha_1; \ldots; \alpha_r}$$

Mich überzeugt diese Notation nicht, da sie im Gegensatz zu unserer Notation für die Binomialkoeffizienten recht eigentlich nichts kürzer macht.

Übung 2.2.28. Man zeige die Formel

$$(x_1 + \ldots + x_r)^n = \sum_{\alpha_1+\ldots+\alpha_r=n} \frac{n!}{\alpha_1! \cdots \alpha_r!} x_1^{\alpha_1} \cdots x_r^{\alpha_r}$$

Hier ist zu verstehen, daß wir für alle $\alpha_1, \ldots, \alpha_r \in \mathbb{N}$ mit $\alpha_1 + \ldots + \alpha_r = n$ den angegebenen Ausdruck nehmen und alle diese Ausdrücke aufsummieren.

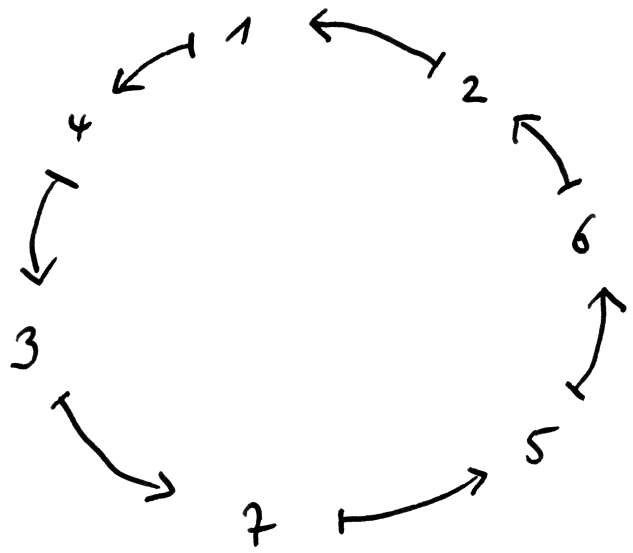
Übung 2.2.29. Eine **zyklische Anordnung** einer endlichen Menge M ist eine Abbildung $z : M \to M$ derart, daß wir durch mehrmaliges Anwenden von z auf ein beliebiges Element $x \in M$ jedes Element $y \in M$ erhalten können. Man zeige, daß es auf einer n-elementigen Menge mit $n \geq 1$ genau $(n-1)!$ zyklische Anordnungen gibt. Die Terminologie "zyklische Anordnung" scheint mir nicht besonders glücklich, da unsere Struktur nun beim besten Willen keine Ordnung im Sinne von **??** ist. Andererseits ist das Angeben einer Anordnung auf einer endlichen Menge M schon auch etwas ähnliches.

Übung 2.2.30. Sei X eine Menge mit $n \geq 1$ Elementen und sei m eine natürliche Zahl. Man zeige, daß es genau $\binom{n+m-1}{n-1}$ Abbildungen $f : X \to \mathbb{N}$ gibt mit $\sum_{x\in X} f(x) = m$. Hinweis: Man denke sich $X = \{1, 2, \ldots, n\}$ und veranschauliche sich dann f als eine Folge auf $f(1)$ Punkten gefolgt von einem Strich gefolgt von $f(2)$ Punkten gefolgt von einem Strich und so weiter, insgesamt also eine Folge aus $n + m - 1$ Symbolen, davon m Punkten und $n - 1$ Strichen.

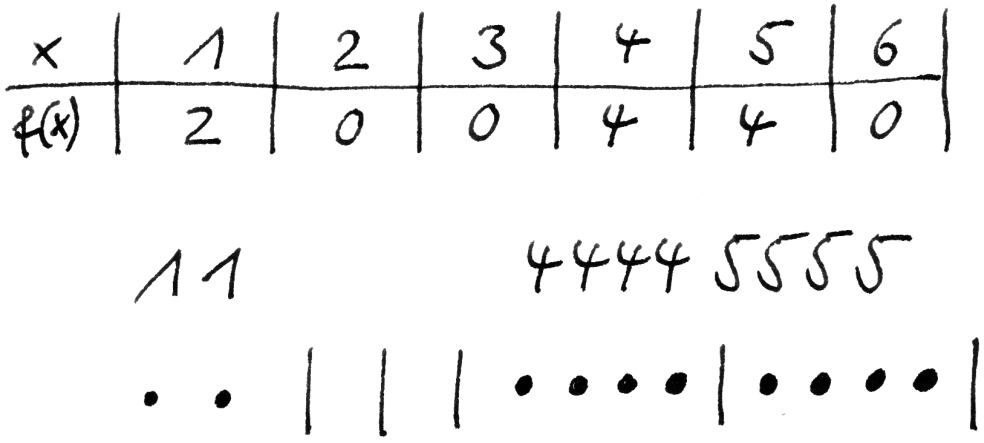
2.2.31. Gegeben eine Menge X mag man sich eine Abbildung $X \to \mathbb{N}$ veranschaulichen als eine "Menge von Elementen von X, in der jedes Element mit einer wohlbestimmten Vielfachheit vorkommt". Aufgrund dieser Vorstellung nennt man eine Abbildung $X \to \mathbb{N}$ auch eine **Multimenge** von Elementen von X. Wir notieren die Gesamtheit aller derartigen Multimengen mit

$$\mathcal{M}(X) = \mathrm{Ens}(X, \mathbb{N})$$

in vager Analogie zur Potenzmenge $\mathcal{P}(X)$ von X, die ja in kanonischer Bijektion zu $\mathrm{Ens}(X, \{0, 1\})$ steht. Unter der Kardinalität einer Multimenge verstehen wir die Summe über alle Werte der entsprechenden Abbildung, aufgefaßt als ein Element von $\mathbb{N} \amalg \{\infty\}$. Ich notiere eine Multimenge mit normalen Mengenklammern, so wäre etwa $\{5, 5, 5, 7, 7, 1\}$ die hoffentlich offensichtliche Multimenge von natürlichen Zahlen der Kardinalität 6, und der Leser muß aus dem Kontext erschließen, wann eine Multimenge und wann eine normale Menge gemeint ist.



Versuch der graphischen Darstellung einer zyklischen Anordnung auf der Menge $\{1, 2, \ldots, 7\}$. Die Pfeile $\mapsto$ sollen jeweils den Effekt der Abbildung z veranschaulichen.



Eine Abbildung $f : \{1, 2, \ldots, n\} \to \mathbb{N}$ im Fall $n = 6$ mit Wertesumme $m = 10$ und die Veranschaulichung nach der Vorschrift aus Übung 2.2.30 als Folge bestehend aus m Punkten und $n - 1$ Strichen.

2.3 Logische Symbole und Konventionen

2.3.1. In der Mathematik meint **oder** immer, daß auch beides erlaubt ist. Wir haben diese Konvention schon benutzt bei der Definition der Vereinigung wenn wir schreiben $X \cup Y = \{z \mid z \in X \text{ oder } z \in Y\}$, zum Beispiel haben wir $\{1,2\} \cup \{2,3\} = \{1,2,3\}$.

2.3.2. Sagt man der Mathematik, es gebe ein Objekt mit diesen und jenen Eigenschaften, so ist stets gemeint, daß es *mindestens ein* derartiges Objekt geben soll. Hätten wir diese Sprachregelung rechtzeitig vereinbart, so hätten wir zum Beispiel das Wörtchen "mindestens" in Teil 2 von 2.2.14 bereits weglassen können. Sagt ihnen also ein Mathematiker, er habe einen Bruder, so kann es auch durchaus sein, daß er noch weitere Brüder hat! Will man in der Mathematik Existenz und Eindeutigkeit gleichzeitig ausdrücken, so sagt man, es gebe **genau ein** Objekt mit diesen und jenen Eigenschaften. Sagt ihnen also ein Mathematiker, er habe genau einen Bruder, so können sie sicher sein, daß er nicht noch weitere Brüder hat.

2.3.3. Die folgenden Abkürzungen erweisen sich als bequem und werden recht häufig verwendet:

$\forall$	für alle (ein umgedrehtes A wie "alle")
$\exists$	es gibt (ein umgedrehtes E wie "existiert")
$\exists!$	es gibt genau ein
$\ldots \Rightarrow \cdots$	aus $\ldots$ folgt $\cdots$
$\ldots \Leftarrow \cdots$	$\ldots$ folgt aus $\cdots$
$\ldots \Leftrightarrow \cdots$	$\ldots$ ist gleichbedeutend zu $\cdots$

Ist zum Beispiel $f : X \to Y$ eine Abbildung, so können wir unsere Definitionen injektiv, surjektiv, und bijektiv etwas formaler so schreiben:

$$\begin{array}{lcl} f \text{ injektiv} & \Leftrightarrow & ((f(x) = f(z)) \Rightarrow (x = z)) \\ f \text{ surjektiv} & \Leftrightarrow & \forall y \in Y \exists x \in X \text{ mit } f(x) = y \\ f \text{ bijektiv} & \Leftrightarrow & \forall y \in Y \exists! x \in X \text{ mit } f(x) = y \end{array}$$

2.3.4. Bei den "für alle" und "es gibt" kommt es in der Mathematik sehr auf die Reihenfolge an, viel mehr als in der weniger präzisen Umgangssprache. Man betrachte zum Beispiel die beiden folgenden Aussagen:

"Für alle $n \in \mathbb{N}$ gibt es $m \in \mathbb{N}$ so daß gilt $m \geq n$"

"Es gibt $m \in \mathbb{N}$ so daß für alle $n \in \mathbb{N}$ gilt $m \geq n$"

Offensichtlich ist die erste richtig, die zweite aber falsch. Weiter mache man sich klar, daß die "für alle" und "es gibt" bei Verneinung vertauscht werden. Äquivalent sind zum Beispiel die beiden folgenden Aussagen

"Es gibt kein $n \in \mathbb{N}$ mit $n^2 = 2$"

"Für alle $n \in \mathbb{N}$ gilt nicht $n^2 = 2$"

2.3.5. Wollen wir zeigen, daß aus einer Aussage A eine andere Aussage B folgt, so können wir ebensogut zeigen: Gilt B nicht, so gilt auch A nicht. In formelhafter Schreibweise haben wir also

$$(A \Rightarrow B) \Leftrightarrow ((\text{nicht } B) \Rightarrow (\text{nicht } A))$$

Wollen wir zum Beispiel zeigen ($g \circ f$ surjektiv) $\Rightarrow$ (g surjektiv), so reicht es, wenn wir uns überlegen: Ist g nicht surjektiv, so ist $g \circ f$ erst recht nicht surjektiv.

3 Algebraische Grundbegriffe

Auf der Schule versteht man unter einer "reellen Zahl" meist einen unendlichen Dezimalbruch, wobei man noch aufpassen muß, daß durchaus verschiedene unendliche Dezimalbrüche dieselbe reelle Zahl darstellen können, zum Beispiel gilt in den reellen Zahlen ja

$$0,99999\ldots = 1,00000\ldots$$

Diese reellen Zahlen werden dann addiert, subtrahiert, multipliziert und dividiert ohne tiefes Nachdenken darüber, wie man denn zum Beispiel mit den eventuell unendlich vielen Überträgen bei der Addition und Subtraktion umgehen soll, und warum dann Formeln wie $(a+b)-c = a+(b-c)$ wirklich gelten, zum Beispiel für $a = b = c = 0,999\ldots$. Dieses tiefe Nachdenken wollen wir im Folgenden vom Rest der Vorlesung abkoppeln und müssen dazu sehr präzise formulieren, welche Eigenschaften für die Addition, Multiplikation und Anordnung in "unseren" reellen Zahlen gelten sollen: In der Terminologie, die in den folgenden Abschnitten eingeführt wird, werden wir die reellen Zahlen charakterisieren als einen angeordneten Körper, in dem jede nichtleere Teilmenge mit einer unteren Schranke sogar eine größte untere Schranke besitzt. Von dieser Charakterisierung ausgehend erklären wir dann, welche reelle Zahl ein gegebener unendlicher Dezimalbruch darstellt, und errichten das Gebäude der Analysis. In demselben Begriffsgebäude modellieren wir auch den Anschauungsraum, vergleiche 1.2.8 oder besser **??**. Um diese Charakterisierungen und Modellierungen verständlich zu machen, führen wir zunächst einige grundlegende algebraische Konzepte ein, die Ihnen im weiteren Studium der Mathematik noch oft begegnen werden.

3.1 Mengen mit Verknüpfung

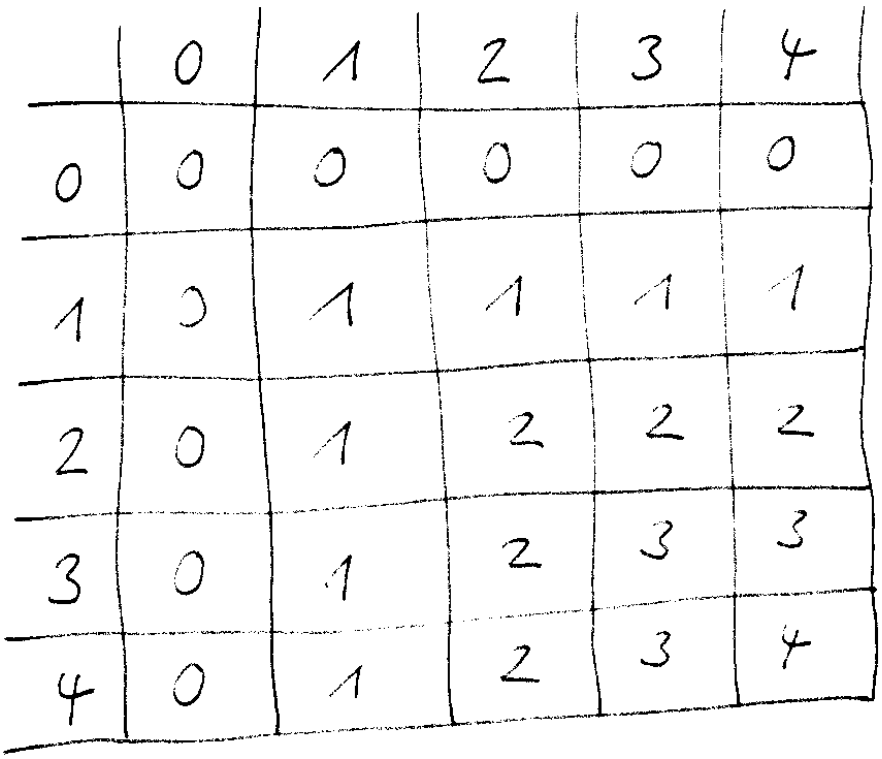
Definition 3.1.1. Eine **Verknüpfung $\top$ auf einer Menge** A ist eine Abbildung

$$\begin{array}{rccc} \top: & A\times A & \to & A \\ & (a,b) & \mapsto & a\top b \end{array}$$

die also jedem geordneten Paar (a,b) mit $a,b \in A$ ein weiteres Element $(a\top b) \in A$ zuordnet.

Beispiele 3.1.2. 1. Die Addition von ganzen Zahlen ist eine Verknüpfung

$$\begin{array}{rccc} +: & \mathbb{Z}\times\mathbb{Z} & \to & \mathbb{Z} \\ & (a,b) & \mapsto & a+b \end{array}$$



	0	1	2	3	4
0	0	0	0	0	0
1	0	1	1	1	1
2	0	1	2	2	2
3	0	1	2	3	3
4	0	1	2	3	4

Man kann Verknüpfungen auf endlichen Mengen darstellen durch ihre **Verknüpfungstabelle**. Hier habe ich etwa die Verknüpfungstabelle der Verknüpfung min auf der Menge $\{0, 1, 2, 3, 4\}$ angegeben. Natürlich muß man sich dazu einigen, ob im Kästchen aus Spalte a und Zeile b nun $a \top b$ oder vielmehr $b \top a$ stehen soll, aber bei einer kommutativen Verknüpfung wie min kommt es zum Glück nicht darauf an.

2. Die Multiplikation von ganzen Zahlen ist eine Verknüpfung

$$\begin{array}{rccc} \cdot : & \mathbb{Z} \times \mathbb{Z} & \to & \mathbb{Z} \\ & (a,b) & \mapsto & a \cdot b \end{array}$$

3. Die Zuordnung min, die jedem Paar von natürlichen Zahlen die kleinere zuordnet (wenn sie verschieden sind, man setzt sonst $\min(a,a) = a$), ist eine Verknüpfung

$$\begin{array}{rccc} \min : & \mathbb{N} \times \mathbb{N} & \to & \mathbb{N} \\ & (a,b) & \mapsto & \min(a,b) \end{array}$$

4. Sei X eine Menge. Die Verknüpfung von Abbildungen liefert eine Verknüpfung auf der Menge $\mathrm{Ens}(X,X)$ aller Abbildungen von X in sich selber

$$\begin{array}{rccc} \circ : & \mathrm{Ens}(X,X) \times \mathrm{Ens}(X,X) & \to & \mathrm{Ens}(X,X) \\ & (f,g) & \mapsto & f \circ g \end{array}$$

   Oft kürzen wir auch $\mathrm{Ens}(X,X) = \mathrm{Ens}(X)$ ab.

5. Die Subtraktion von ganzen Zahlen ist eine Verknüpfung

$$\begin{array}{rccc} - : & \mathbb{Z} \times \mathbb{Z} & \to & \mathbb{Z} \\ & (a,b) & \mapsto & a - b \end{array}$$

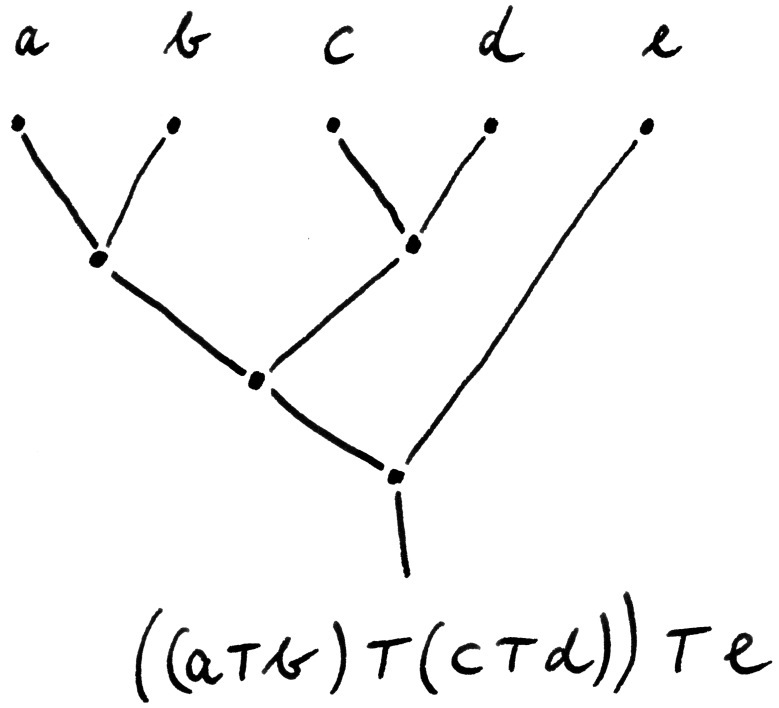
6. Jede Verknüpfung $\top$ auf einer Menge induziert eine Verknüpfung auf ihrer Potenzmenge vermittels der Vorschrift

$$U \top V = \{u \top v \mid u \in U,\ v \in V\}$$

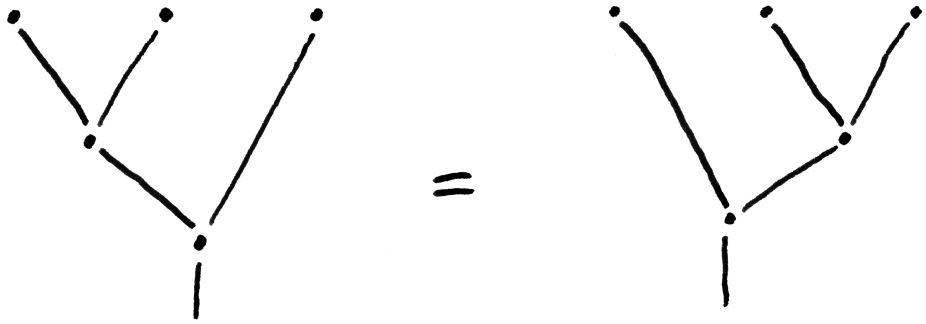
7. Gegeben Mengen mit Verknüpfung $(A, \top)$ und $(B, \perp)$ erhalten wir eine Verknüpfung auf ihrem Produkt $A \times B$ vermittels der Vorschrift $((a,b),(a',b')) \mapsto ((a \top a'),(b \perp b'))$. Sie heißt die **komponentenweise Verknüpfung**.

Definition 3.1.3. Eine Verknüpfung $\top$ auf einer Menge A heißt **assoziativ** genau dann, wenn gilt $a \top (b \top c) = (a \top b) \top c \quad \forall a,b,c \in A$. Sie heißt **kommutativ** genau dann, wenn gilt $a \top b = b \top a \quad \forall a,b \in A$.

Beispiele 3.1.4. Von unseren Beispielen sind die ersten drei assoziativ und kommutativ, das vierte ist assoziativ aber nicht kommutativ falls X mehr als ein Element hat, das fünfte ist weder assoziativ noch kommutativ.



Mögliche “Klammerungen” mag man sich graphisch wie oben angedeutet veranschaulichen. Die Assoziativität bedeutet dann graphisch so etwas wie



wobei das Gleichheitszeichen nur meint, daß beide Klammerungen stets dasselbe liefern, wenn wir oben drei Elemente unserer Menge mit Verknüpfung einfüllen...

Bemerkung 3.1.5. Ist eine Verknüpfung assoziativ, so liefern Ausdrücke der Form $a_1 \top a_2 \ldots \top a_n$ wohlbestimmte Elemente von A, das Resultat ist genauer unabhängig davon, wie wir die Klammern setzen. Um diese Erkenntnis zu formalisieren, vereinbaren wir für so einen Ausdruck die Interpretation

$$a_1 \top a_2 \ldots \top a_n = a_1 \top (a_2 \top (\ldots (a_{n-1} \top a_n) \ldots))$$

und zeigen

Lemma 3.1.6. *Gegeben $(A, \top)$ eine Menge mit einer assoziativen Verknüpfung und $a_1, \ldots, a_n, b_1, \ldots, b_m \in A$ gilt*

$$(a_1 \top \ldots \top a_n) \top (b_1 \top \ldots \top b_m) = a_1 \top \ldots \top a_n \top b_1 \top \ldots \top b_m$$

Beweis. Wir folgern aus dem Assoziativgesetz $(a_1 \top \ldots \top a_n) \top (b_1 \top \ldots \top b_m) = a_1 \top ((a_2 \top \ldots \top a_n) \top (b_1 \top \ldots \top b_m))$ und sind fertig mit vollständiger Induktion über n. □

Bemerkung 3.1.7. Das Wort Lemma, im Plural Lemmata, kommt wohl von griechisch $\lambda\alpha\mu\beta\alpha\nu\epsilon\iota\nu$ "nehmen" und bezeichnet in der Mathematik kleinere Resultate oder auch Zwischenschritte von größeren Beweisen, denen der Autor außerhalb ihres engeren Kontexts keine große Bedeutung zumißt.

3.1.8. Die Zahl der Möglichkeiten, einen Ausdruck in $n+1$ Faktoren so zu verklammern, daß in jedem Schritt nur die Verknüpfung von je zwei Elementen zu berechnen ist, heißt die **n-te Catalan-Zahl** und wird C_n notiert. Die ersten Catalan-Zahlen sind $C_0 = C_1 = 1$, $C_2 = 2$, $C_3 = 5$. Im allgemeinen zeigen wir in **??** die amüsante Formel

$$C_n = \frac{1}{n+1}\binom{2n}{n}$$

Definition 3.1.9. Sei $(A, \top)$ eine Menge mit Verknüpfung. Ist $n \in \{1, 2, \ldots\}$ eine von Null verschiedene natürliche Zahl und $a \in A$, so schreiben wir

$$\underbrace{a \top a \top \ldots \top a}_{n\text{-mal}} = n^\top a$$

3.1.10. Ist m eine zweite von Null verschiedene natürliche Zahl, so erhalten wir für assoziative Verknüpfungen mithilfe unseres Lemmas 3.1.6 die Regeln $(n+m)^\top a = (n^\top a) \top (m^\top a)$ und $(nm)^\top a = n^\top (m^\top a)$. Ist unsere Verknüpfung auch noch kommutativ, so gilt zusätzlich $n^\top (a \top b) = (n^\top a) \top (n^\top b)$.

3.1.11. Sei $(A, \top)$ eine Menge mit Verknüpfung. Eine Teilmenge $B \subset A$ heißt **abgeschlossen unter der Verknüpfung** genau dann, wenn aus $a, b \in B$ folgt $a \top b \in B$. Natürlich ist in diesem Fall auch $(B, \top)$ eine Menge mit Verknüpfung, man spricht dann von der **auf B induzierten Verknüpfung**. Zum Beispiel ist $\mathbb{N} \subset \mathbb{Z}$ abgeschlossen unter der Addition, aber $\mathbb{Z}_{\neq 0} \subset \mathbb{Q}_{\neq 0}$ ist nicht abgeschlossen unter der durch die Division gegebenen Verknüpfung $(a, b) \mapsto a/b$.

Definition 3.1.12. Sei $(A, \top)$ eine Menge mit Verknüpfung. Ein Element $e \in A$ heißt **neutrales Element** von $(A, \top)$ genau dann, wenn gilt

$$e \top a = a \top e = a \quad \forall a \in A$$

3.1.13. In einer Menge mit Verknüpfung $(A, \top)$ kann es höchstens ein neutrales Element e geben, denn für jedes weitere Element e' mit $e' \top a = a \top e' = a \quad \forall a \in A$ haben wir $e' = e' \top e = e$. Wir dürfen also den bestimmten Artikel verwenden und in einer Menge mit Verknüpfung von **dem** neutralen Element reden.

Definition 3.1.14. Ein **Monoid** ist eine Menge mit einer assoziativen Verknüpfung, in der es ein neutrales Element gibt. Ist $(A, \top)$ ein Monoid, so erweitern wir unsere Notation $n^\top a$ auf alle natürlichen Zahlen $n \in \mathbb{N}$, indem wir $0^\top a$ als das neutrale Element von A verstehen, für alle $a \in A$.

Bemerkung 3.1.15. Das Wort "Monoid" ist wohl von griechisch "$\mu o \nu o \varsigma$" für "allein" abgeleitet: Ein Monoid besitzt nur eine einzige Verknüpfung.

3.1.16. In Monoiden gelten die Regeln 3.1.10 für alle $n, m \in \mathbb{N}$. Die natürlichen Zahlen bilden mit der Addition ein Monoid $(\mathbb{N}, +)$ mit neutralem Element 0. Sie bilden auch unter der Multiplikation ein Monoid $(\mathbb{N}, \cdot)$ mit neutralem Element 1. Für jede Menge X ist die Menge $\mathrm{Ens}(X)$ der Abbildungen von X in sich selbst ein Monoid. Die leere Menge ist *kein* Monoid, ihr fehlt das neutrale Element.

3.2 Gruppen

3.2.1. Ich empfehle, bei der Lektüre dieses Abschnitts die Tabelle auf Seite 54 gleich mitzulesen, die die Bedeutungen der nun folgenden Formalitäten in den zwei gebräuchlichsten Notationssystemen angibt. In diesen Notationssystemen sollten alle Formeln aus der Schulzeit vertraut sein. Wir erinnern uns an die Definition eines Monoids 3.1.14.

Definition 3.2.2. 1. Ist $(A, \top)$ ein Monoid und $a \in A$ ein Element, so nennen wir ein weiteres Element $\bar{a} \in A$ **invers zu** a genau dann, wenn

gilt $a\top\bar{a} = e = \bar{a}\top a$ für $e \in A$ das neutrale Element unseres Monoids. Ein Element, das ein Inverses besitzt, heißt **invertierbar**.

2. Eine **Gruppe** ist ein Monoid, in dem jedes Element ein Inverses besitzt.

3. Eine **kommutative Gruppe** oder **abelsche Gruppe** ist eine Gruppe, deren Verknüpfung kommutativ ist.

3.2.3. Der Begriff einer "Gruppe" wurde von Évariste Galois (1811-1832) in die Mathematik eingeführt. Er verwendetet den Begriff "Gruppe von Transformationen" sowohl in der Bedeutung einer "Menge von bijektiven Selbstabbildungen einer gegebenen Menge" als auch in der Bedeutung einer "Menge von bijektiven Selbstabbildungen einer gegebenen Menge, die abgeschlossen ist unter Verknüpfung und Inversenbildung", und die damit in der Tat ein Beispiel für eine Gruppe im Sinne der obigen Definition liefert. Die obige Definition konnte Galois beim besten Willen nicht geben, er starb ein gutes halbes Jahrhundert, bevor Cantor die Sprache der Mengenlehre entwickelte. Die Terminologie "abelsche Gruppe" wurde zu Ehren des norwegischen Mathematikers Niels Hendrik Abel eingeführt.

Lemma 3.2.4. *Jedes Element eines Monoids besitzt höchstens ein Inverses.*

Beweis. Aus $a\top\bar{a} = e = \bar{a}\top a$ und $a\top b = e = b\top a$ folgt durch Anwenden von $b\top$ auf die erste Gleichung mit dem Assoziativgesetz sofort $\bar{a} = b$. □

3.2.5. Wir dürfen also den bestimmten Artikel benutzen und von nun an von *dem* Inversen eines Elements einer Gruppe reden. Offensichtlich ist das Inverse des Inversen stets das ursprüngliche Element, in Formeln $\bar{\bar{a}} = a$.

Lemma 3.2.6. *Sind a und b Elemente einer Gruppe, so wird das Inverse von $a\top b$ wird gegeben durch die Formel $\overline{(a\top b)} = \bar{b}\top\bar{a}$.*

Beweis. In der Tat rechnen wir schnell $(a\top b)\top(\bar{b}\top\bar{a}) = e$. Diese Formel ist auch aus dem täglichen Leben vertraut: Wenn man sich morgends zuerst die Strümpfe anzieht und dann die Schuhe, so muß man abends zuerst die Schuhe ausziehen und dann die Strümpfe. □

Beispiele 3.2.7. Von unseren Beispielen 3.1.2 für Verknüpfungen oben ist nur $(\mathbb{Z}, +)$ eine Gruppe, und diese Gruppe ist kommutativ. Ein anderes Beispiel für eine kommutative Gruppe ist die Menge $\mathbb{Q}\backslash\{0\}$ der von Null verschiedenen rationalen Zahlen mit der Multiplikation als Verknüpfung.

Übung 3.2.8. Die invertierbaren Elemente eines Monoids bilden stets eine Gruppe. Ein Element a eines Monoids A ist invertierbar genau dann, wenn es $b, c \in A$ gibt mit $b\top a = e = a\top c$ für e das neutrale Element.

	123	213	312	321	132	231
123	123	213	312	321	132	231
213	213	123	321	312	231	132
312	312	132	231	213	321	123
321	321	231	132	123	312	213
132	132	312	213	231	123	321
231	231	321	123	132	213	312

Die Verknüpfungstafel der Gruppe aller Permutationen der Menge $\{1, 2, 3\}$. Eine solche Permutation σ habe ich dargestellt durch das geordnete Zahlentripel $\sigma(1)\sigma(2)\sigma(3)$, und im Kästchen aus der Zeile τ und der Spalte σ steht $\tau \circ \sigma$.

Definition 3.2.9. Ist $(A, \top)$ eine Gruppe, so erweitern wir unsere Notation $n^\top a$ auf alle $n \in \mathbb{Z}$, indem wir setzen $n^\top a = (-n)^\top \bar{a}$ für $n \in \{-1, -2, \ldots\}$.

3.2.10. In einer Gruppe gelten offensichtlich sogar für alle ganzen Zahlen $n \in \mathbb{Z}$ die Regeln $(n+m)^\top a = (n^\top a)\top(m^\top a)$ und $(nm)^\top a = n^\top(m^\top a)$. Ist die Gruppe kommutativ, so gilt zusätzlich $n^\top(a \top b) = (n^\top a)\top(n^\top b)$ für alle $n \in \mathbb{Z}$.

3.2.11. Verknüpfungen werden meist additiv oder multiplikativ geschrieben, also $a{+}b$ oder $a{\cdot}b$, wobei die additive Schreibweise kommutativen Verknüpfungen vorbehalten ist und die Bruchnotation $1/a$ und b/a aus nebenstehender Tabelle kommutativen multiplikativ geschriebenen Verknüpfungen. Bei additiv geschriebenen Gruppen bezeichnet man das Inverse von a meist als das **Negative** von a. Bei nichtkommutativen und multiplikativ notierten Gruppen benutzt man für das Inverse von a nur die von der allgemeinen Notation a^n abgeleitete Notation a^{-1}. Die Tabelle aus Seite 54 fasst die üblichen Notationen für unsere abstrakten Begriffsbildungen in diesem Kontext zusammen und gibt unsere allgemeinen Resultate und Konventionen in diesen Notationen wieder. Diejenigen Formeln und Konventionen, die keine Inversen brauchen, benutzt man auch allgemeiner für beliebige Monoide. Für die Gruppe der invertierbaren Elemente eines multiplikativ notierten Monoids A verwenden wir die Notation $A^\times$. Zum Beispiel haben wir $\mathbb{Z}^\times = \{1, -1\}$.

Beispiel 3.2.12. Für jede Menge X ist die Menge aller Bijektionen von X auf sich selbst eine Gruppe, mit der Komposition von Abbildungen als Verknüpfung. Wir notieren diese Gruppe $\mathrm{Ens}^\times(X)$ in Übereinstimmung mit unserer Konvention 3.2.11, schließlich handelt es sich um die Gruppe der invertierbaren Elemente des Monoids $\mathrm{Ens}(X)$. Ihre Elemente heißen die **Permutationen von** X. Die Gruppe der Permutationen einer Menge X ist für $|X| > 2$ nicht kommutativ. Das Inverse einer Bijektion ist ihre Umkehrabbildung.

Übung 3.2.13. Sind a, b, c Elemente einer Gruppe, so folgt aus $a \top b = a \top c$ bereits $b = c$. Ebenso folgt auch aus $b \top a = c \top a$ bereits $b = c$.

Übung 3.2.14. Sei A ein Monoid und e sein neutrales Element. Man zeige: Unser Monoid ist genau dann eine Gruppe, wenn es für jedes $a \in A$ ein $\bar{a} \in A$ gibt mit $\bar{a} \top a = e$, und dies Element $\bar{a}$ ist dann notwendig das Inverse von a in A. Noch Mutigere zeigen: Ist A eine Menge mit assoziativer Verknüpfung und existiert ein $e \in A$ mit $e \top a = a \; \forall a \in A$ sowie für jedes $a \in A$ ein $\bar{a} \in A$ mit $\bar{a} \top a = e$, so ist A eine Gruppe.

3.2.15. Gegeben eine Abbildung $I \to A$, $i \mapsto a_i$ von einer endlichen Menge in ein kommutatives additiv bzw. multiplikativ notiertes Monoid vereinbaren wir die Notationen

$$\sum_{i \in I} a_i \qquad \text{bzw.} \qquad \prod_{i \in I} a_i$$

für die "Verknüpfung aller a_i". Ist I die leere Menge, so vereinbaren wir, daß dieser Ausdruck das neutrale Element von A bedeuten möge, also 0 bzw. 1. Wir haben diese Notation bereits verwendet in 2.1.17, und für die konstante Abbildung $I \to \mathbb{N}$, $i \mapsto 1$ hätten wir zum Beispiel

$$\sum_{i \in I} 1 = |I|$$

Unsere Konvention 1.1.11 für mit einem Laufindex notierte Summen bzw. Produkte verwenden wir bei Monoiden analog.

abstrakt	additiv	multiplikativ
$a \top b$	$a + b$	$a \cdot b$, $a \circ b$, ab
e	0	1
$\bar{b}$	$-b$	$1/b$
$a \top \bar{b}$	$a - b$	a/b
$n^\top a$	na	a^n
$e \top a = a \top e = a$	$0 + a = a + 0 = a$	$1 \cdot a = a \cdot 1 = a$
$a \top \bar{a} = e$	$a + (-a) = 0$	$a/a = 1$
$\bar{\bar{a}} = a$	$-(-a) = a$	$1/(1/a) = a$
$(-1)^\top a = \bar{a}$	$(-1)a = -a$	$a^{-1} = 1/a$
$\overline{(a \top b)} = \bar{b} \top \bar{a}$	$-(a + b) = (-b) + (-a)$	$(ab)^{-1} = b^{-1}a^{-1}$, $1/ab = (1/b)(1/a)$
$\overline{(a \top \bar{b})} = b \top \bar{a}$	$-(a - b) = b - a$	$1/(a/b) = b/a$
$n^\top (m^\top a) = (nm)^\top a$	$n(ma) = (nm)a$	$(a^m)^n = a^{nm}$
$(m + n)^\top a = (m^\top a) \top (n^\top a)$	$(m + n)a = (ma) + (na)$	$a^{(m+n)} = (a^m)(a^n)$
$\overline{n^\top a} = (-n)^\top a$	$-(na) = (-n)a$	$(a^n)^{-1} = a^{-n}$
$0^\top a = e$	$0a = 0$	$a^0 = 1$
$n^\top (a \top b) = (n^\top a) \top (n^\top b)$	$n(a + b) = (na) + (nb)$	$(ab)^n = (a^n)(b^n)$

Tabelle I.1: Konventionen und Formeln in verschiedenen Notationssystemen. Bereits diese Tabelle muß mit einigen Hintergedanken gelesen werden, weil die Symbole $+, -, 0, 1$ darin in zweierlei Bedeutung vorkommen: Manchmal meinen sie konkrete Operationen und Elemente von $\mathbb{Z}$, manchmal stehen sie für Verknüpfung und Inversenbildung und neutrale Elemente in abstrakten Monoiden.

3.3 Körper

Definition 3.3.1. Ein **Körper** $(K, +, \cdot)$ (englisch **field**, französisch **corps**) ist eine Menge K mit zwei assoziativen und kommutativen Verknüpfungen $+$ und $\cdot$ derart, daß gilt

1. $(K, +)$ ist eine Gruppe, die **additive Gruppe** des Körpers.
2. Bezeichnet $0_K \in K$ das neutrale Element der Gruppe $(K, +)$, so folgt aus $a \neq 0_K \neq b$ schon $a \cdot b \neq 0_K$ und $(K \backslash \{0_K\}, \cdot)$ ist eine Gruppe, die **multiplikative Gruppe** des Körpers.
3. Es gilt das **Distributivgesetz**

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

3.3.2. Der Begriff "Körper" ist in diesem Zusammenhang wohl zu verstehen als "besonders gut unter den verschiedensten Rechenoperationen abgeschlossener Zahlbereich", in Analogie zu geometrischen Körpern wie Kugeln oder Zylindern, die man ja auch als "besonders gut in sich geschlossene Teilmengen des Raums" bezeichnen könnte.

3.3.3. Wenn wir mit Buchstaben rechnen, werden wir meist $a \cdot b = ab$ abkürzen. Zusätzlich vereinbaren wir zur Vermeidung von Klammern die Regel "Punkt vor Strich", so daß also zum Beispiel das Distributivgesetz kürzer in der Form $a(b + c) = ab + ac$ geschrieben werden kann. Die multiplikative Gruppe eines Körpers K notieren wir $K^\times = K \backslash \{0_K\}$ in Übereinstimmung mit unserer allgemeinen Notation 3.2.11, schließlich handelt es sich um die Menge der invertierbaren Elemente des multiplikativen Monoids K. Für das neutrale Element der Multiplikation vereinbaren wir die Bezeichnung $1_K \in K^\times$. Wir kürzen meist 0_K ab durch 0 und 1_K durch 1 in der Erwartung, daß man aus dem Kontext erschließt, ob mit 0 und 1 natürliche Zahlen oder Elemente eines speziellen Körpers gemeint sind. Meist kommt es darauf im Übrigen gar nicht an.

3.3.4. Für alle a, b in einem Körper und alle $n \geq 0$ gilt die binomische Formel

$$(a + b)^n = \sum_{\nu=0}^{n} \binom{n}{\nu} a^\nu b^{n-\nu}$$

Um das einzusehen prüft man, daß wir bei der Herleitung nach 1.1.20 nur Körperaxiome verwandt haben. Man beachte hierbei unsere Konvention $0_K^0 =$

1_K aus 3.1.14, angewandt auf das Monoid $(K, \cdot)$ in Verbindung mit der notationellen Konvention auf Seite 54. Die Multiplikation mit den Binomialkoeffizienten ist gemeint als wiederholte Addition im Sinne der Bezeichnungskonvention *na* auf Seite 54, angewandt auf den Spezialfall der additiven Gruppe unseres Körpers.

Beispiele 3.3.5. Ein Beispiel für einen Körper ist der Körper der rationalen Zahlen $(\mathbb{Q}, +, \cdot)$. Ein anderes Beispiel ist der zweielementige Körper mit den durch die Axiome erzwungenen Rechenregeln, der fundamental ist in der Informatik. Die ganzen Zahlen $(\mathbb{Z}, +, \cdot)$ bilden keinen Körper, da $(\mathbb{Z} \setminus \{0\}, \cdot)$ keine Gruppe ist, da es nämlich in $\mathbb{Z}\setminus\{0\}$ nur für 1 und -1 ein multiplikatives Inverses gibt. Es gibt keinen einelementigen Körper, da das Komplement seines Nullelements die leere Menge sein müßte: Dies Komplement kann dann aber unter der Multiplikation keine Gruppe sein, da es eben kein neutrales Element haben könnte.

Lemma 3.3.6 (Folgerungen aus den Körperaxiomen). *Sei K ein Körper. So gilt*

1. $a0 = 0 \quad \forall a \in K$.
2. $ab = 0 \Rightarrow a = 0$ *oder* $b = 0$.
3. $-a = (-1)a \quad \forall a \in K$.
4. $(-1)(-1) = 1$.
5. $(-a)(-b) = ab \quad \forall a, b \in K$.
6. $\frac{a}{b}\frac{c}{d} = \frac{ac}{bd}$ *für alle* $a, c \in K$ *und* $b, d \in K^\times$.
7. $\frac{ac}{bc} = \frac{a}{b}$ *für alle* $a \in K$ *und* $b, c \in K^\times$.
8. $\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$ *für alle* $a, c \in K$ *und* $b, d \in K^\times$.
9. $m(ab) = (ma)b$ *für alle* $m \in \mathbb{Z}$ *und* $a, b \in K$.

Beweis.

1. Man folgert das aus $a0 + a0 = a(0 + 0) = a0$ durch Hinzuaddieren von $-(a0)$ auf beiden Seiten.
2. In der Tat folgt aus ($a \neq 0$ und $b \neq 0$) schon ($ab \neq 0$) nach den Körperaxiomen.
3. In der Tat gilt $a + (-1)a = 1a + (-1)a = (1 + (-1))a = 0a = 0$, und $-a$ ist ja gerade definiert als das eindeutig bestimmte Element von K so daß $a + (-a) = 0$.
4. In der Tat gilt nach dem Vorhergehenden $(-1)(-1) = -(-1) = 1$.
5. Um das nachzuweisen ersetzen wir einfach $(-a) = (-1)a$ und $(-b) = (-1)b$ und verwenden $(-1)(-1) = 1$.
6. Das ist klar.
7. Das ist klar.
8. Das wird bewiesen, indem man die Brüche auf einen Hauptnenner bringt und das Distributivgesetz anwendet.
9. Das folgt durch wiederholtes Anwenden des Distributivgesetzes.

□

3.3.7. Die Frage, wie das Produkt zweier negativer Zahlen zu bilden sei, war lange umstritten. Mir scheint der vorhergehende Beweis das überzeugendste Argument für "Minus mal Minus gibt Plus": Es sagt salopp gesprochen, daß man diese Regel adoptieren muß, wenn man beim Rechnen das Ausklammern ohne alle Einschränkungen erlauben will.

Übung 3.3.8. Ist K ein Körper derart, daß es kein $x \in K$ gibt mit $x^2 = -1$, so kann man die Menge $K \times K = K^2$ zu einem Körper machen, indem man die Addition und Multiplikation definiert durch

$$\begin{aligned}(a,b) + (c,d) &= (a+c, b+d)\\ (a,b) \cdot (c,d) &= (ac - bd, ad + bc)\end{aligned}$$

Die Abbildung $K \to K^2$, $a \mapsto (a,0)$ ist dann verträglich mit Addition und Multiplikation. Kürzen wir $(a,0)$ mit a ab und setzen $(0,1) = \mathrm{i}$, so gilt $\mathrm{i}^2 = -1$ und $(a,b) = a + b\,\mathrm{i}$.

3.3.9. Auf die in der vorhergehenden Übung 3.3.8 erklärte Weise können wir etwa aus dem Körper $K = \mathbb{R}$ der "reellen Zahlen", sobald wir ihn kennengelernt haben, direkt den Körper $\mathbb{C}$ der "komplexen Zahlen" konstruieren. Man beachte, wie mühelos das in der Sprache der Mengenlehre zu machen ist. Als die komplexen Zahlen erfunden wurden, gab es noch keine Mengenlehre und beim Rechnen beschränkte man sich auf das Rechnen mit "reellen" Zahlen, ja selbst das Multiplizieren zweier negativer Zahlen wurde als eine fragwürdige Operation angesehen, und das Ziehen einer Quadratwurzel aus einer negativen Zahl als eine rein imaginäre Operation. In gewisser Weise ist es das ja auch geblieben, aber die Mengenlehre liefert eben unserer Imagination eine wunderbar präzise Sprache, in der wir uns auch über imaginierte Dinge unmißverständlich austauschen können. Man kann dieselbe Konstruktion auch allgemeiner durchführen, wenn man statt -1 irgendein anderes Element eines Körpers K betrachtet, das kein Quadrat ist. Noch allgemeinere Konstruktionen zur "Adjunktion höherer Wurzeln" oder sogar der "Adjunktion von Nullstellen polynomialer Gleichungen" können sie in der Algebra lernen, vergleiche etwa **??**.

Definition 3.3.10. Gegeben Mengen mit Verknüpfung $(A, \top)$ und $(B, \bot)$ verstehen wir unter einem **Homomorphismus** von A nach B eine Abbildung $\varphi : A \to B$ derart, daß gilt $\varphi(a \top a') = \varphi(a) \bot \varphi(a')$ für alle $a, a' \in A$. Sind unsere beiden Mengen mit Verknüpfung Gruppen, so sprechen wir von einem **Gruppenhomomorphismus**. Einen bijektiven Homomorphismus nennen wir einen **Isomorphismus**.

Bemerkung 3.3.11. Die Terminologie kommt von griechisch "$\mu o \rho \varphi \eta$" für "Gestalt" oder für uns besser "Struktur" und griechisch "$o \mu o \iota \varsigma$" für "gleich, ähn-

lich". Auf deutsch könnte man statt Homomorphismus auch "strukturerhaltende Abbildung" sagen. Das Wort "Isomorphismus" wird analog gebildet mit griechisch "$\iota\sigma o\varsigma$" für "gleich".

Übung 3.3.12. Gegeben Gruppen H und G bezeichne $\mathrm{Grp}(H,G)$ die Menge aller Gruppenhomomorphismen von H nach G. Man zeige, daß für jede Gruppe G die Vorschrift $\varphi \mapsto \varphi(1)$ eine Bijektion $\mathrm{Grp}(\mathbb{Z},G) \xrightarrow{\sim} G$ liefert. Hinweis: Man erinnere 3.2.10.

3.3.13. Dieselben Definitionen verwenden wir auch bei Mengen mit mehr als einer Verknüpfung. Zum Beispiel ist ein **Körperhomomorphismus** φ von einem Körper K in einen Körper L eine Abbildung $\varphi : K \to L$ derart, daß gilt $\varphi(a+b) = \varphi(a) + \varphi(b)$ und $\varphi(ab) = \varphi(a)\varphi(b)$ für alle $a, b \in K$, und ein **Körperisomorphismus** ist ein bijektiver Körperhomomorphismus. Ein Beispiel für einen Körperhomomorphismus ist unsere Abbildung $K \to K^2$ aus 3.3.8.

Übung 3.3.14. Ein Gruppenhomomorphismus $\varphi : G \to H$ bildet stets das neutrale Element von G auf das neutrale Element von H ab und vertauscht mit Inversenbildung, in Formeln $\varphi(a^{-1}) = (\varphi(a))^{-1}\ \forall a \in G$. Ein Körperhomomorphismus ist stets injektiv.

3.3.15. Von einem **Homomorphismus von Monoiden** fordern wir zusätzlich zur Verträglichkeit mit den Verknüpfungen auch noch, daß er das neutrale Element auf das neutrale Element abbilden soll. Anders als im Gruppenfall folgt das in dieser Allgemeinheit nicht automatisch, wie die Nullabbildung $(\mathbb{Z},+) \to (\mathbb{Z},\cdot)$ zeigt.

Kapitel II

Geschichtliches und Philosophisches

In diesem Kapitel steht allerhand über Mathematik, das nicht zum logisch kohärenten Aufbau beiträgt und teilweise auch stark persönlich gefärbt ist.

Inhalt

1	Zum Ursprung des Wortes Mathematik	62
2	Was ist Mathematik?	62
3	Zum Wesen der Mathematik	63
4	Herkunft einiger Symbole	64

1 Zum Ursprung des Wortes Mathematik

Das Wort “Mathematik” kommt vom griechischen “$\mu\alpha\vartheta\eta\mu\alpha\tau\iota\kappa o\varsigma$”, das sich hinwiederum ableitet von “$\mu\alpha\vartheta\eta\mu\alpha$” für “der Lerngegenstand, die Wissenschaft” nach dem Verb “$\mu\alpha\nu\vartheta\alpha\nu\omega$” für “lernen, verstehen”. Das Anhängen der Endung “-$\iota\kappa o\varsigma$” oder im Anschluß an einen Vokal “-$\tau\iota\kappa o\varsigma$” hat eine ähnliche Bedeutung wie im Deutschen das Anhängen von “-ig” oder “-lich” bzw. “-tlich”, etwa wie in Mut $\mapsto$ mutig, Haar $\mapsto$ haarig oder wohnen $\mapsto$ wohnlich, eigen $\mapsto$ eigentlich. In diesem Sinne wäre die wörtliche Übersetzung von “$\mu\alpha\vartheta\eta\mu\alpha\tau\iota\kappa o\varsigma$” also “das Lernige” oder “das Verständliche”. Platon verwendet den Begriff “$\tau o\ \mu\alpha\vartheta\eta\mu\alpha\tau\iota\kappa o\nu$” im Sinne von “Forschungsgegenstand” in Sophista 219c:2 und Timaeus 88c:1. In der hellenistischen Zeit verengte sich die Bedeutung ein erstes Mal und bezeichnete etwas, was wir heute eher als “Mathematik und Naturwissenschaften” bezeichnen würden. Erst in neuerer Zeit verengte sich die Bedeutung dann weiter auf das, was wir heute unter Mathematik verstehen.

2 Was ist Mathematik?

Ich denke, die heutige Mathematik mag man in einem ersten Ansatz beschreiben als die Wissenschaft von den einfachsten Begriffen: Wieviel einfacher sind doch Zahlen und ihre Rechenregeln, Geraden und Ebenen, oder auch Abbildungen zwischen Mengen im Vergleich zu Pflanzen und Menschen, Hass und Liebe, ja selbst Luft und Wasser! Diese einfachsten Begriffe müssen nun jedoch mit der größten Vorsicht und Präzision gehandhabt werden, damit uns unsere an den Umgang mit Pflanzen und Menschen, Hass und Liebe, Luft und Wasser gewöhnte Intelligenz nicht in die Irre führt. Die eigentliche Mathematik besteht dann darin, diese einfachsten Begriffe zu größeren Theorien zu kombinieren, und die eigentlichen Einsichten entstehen auch erst in dieser Gesamtschau. Ich sehe darin eine Affinität zur Musik, in der man ja auch von einfachsten Geräuschen, in der Klassik etwa von den Tönen der Tonleiter, ausgeht und diese einfachsten Grundbausteine zu Kompositionen kombiniert, deren Sinnhaftigkeit nur in der Gesamtschau erschlossen werden kann. Auf einen Gegensatz zur Musik will ich im nächsten Absatz noch ausführlicher zu sprechen kommen: Während auch die schönste Musik meines Erachtens vom Komponisten nicht entdeckt sondern vielmehr erschaffen wird, scheint es sich mir bei der Mathematik gerade umgekehrt zu verhalten. Sicher gibt es sozusagen “komponierte” mathematische Artikel, aber die mathematischen Inhalte selbst lassen sich von Menschenhand nicht formen und wollen entdeckt werden.

3 Zum Wesen der Mathematik

In diesem Zusammenhang will ich auf eine gerne diskutierte Frage eingehen: Wird Mathematik eigentlich entdeckt oder entwickelt? Aus meiner eigenen Erfahrung mit dieser widerspenstigen Materie und auch der Erfahrung beim Erklären von Beweisen scheint es mir offensichtlich, daß mathematische Inhalte "objektiv da sind", also unabhängig vom menschlichen Subjekt existieren und entdeckt werden. Was jedoch entwickelt werden muß ist eine Sprache, die es uns ermöglicht, uns über diese Inhalte zu verständigen und sie zu nutzen. Hier kam und kommt es durchaus zu parallelen Entwicklungen, man denke etwa an die beiden Notationen $\dot{x}$ und $\frac{\mathrm{d}x}{\mathrm{d}t}$ für die Ableitung oder an verschiedene Algorithmen zur Lösung linearer Gleichungssysteme oder an die verschiedenen Notationen für die natürlichen Zahlen.

Bildlich gesprochen scheint mir die Mathematik wie eine weitverzweigte Höhle, voller Wunder und wertvoller Mineralien, die wir Mathematiker einerseits erkunden und andererseits erschließen. Die Höhle selbst ist objektiv vorhanden und es gilt, immer weiter in sie vorzudringen und Neues zu entdecken. Wo und wie dann jedoch Treppen und Wege und Beleuchtung angebracht werden und eventuell sogar eine kleine Eisenbahn zum Transport der Mineralien, darin haben wir große Freiheit und in diesem Sinne wird Mathematik auch entwickelt. Natürlich sind diese beiden Aufgaben eng miteinander verwoben und wie weit wir selbst vordringen können hängt ganz wesentlich davon ab, wie weit unsere Vorläufer gekommen sind und wie weit sie die Höhle bereits zugänglich gemacht haben.

Diese Auffassung vom Sinn und Wesen der Mathematik bezeichnet man wegen ihrer engen Verwandtschaft mit Plato's Ideenlehre auch als "platonisch". Barry Mazur fordert in seinem Aufsatz [Maz08] die Platonisten auf, zu erklären, warum Beweise denn uns als Mathematikern so wichtig sein sollten, wenn es nur um das Erkennen einer unabhängig von uns existierenden Wirklichkeit geht. Dieser Aufforderung will ich gerne nachkommen. Ich fasse Beweise auf als Beiträge zum großen Projekt, die Welt der mathematischen Inhalte dem menschlichen Verstand zugänglich zu machen. Mir scheint es in diesem Sinne eine wichtige Aufgabe, auch für bereits bewiesene Erkenntnisse möglichst glatte und für menschliche Gehirne transparente Beweise zu finden, aufzuschreiben und öffentlich zugänglich zu machen. Was das Beweisen angeht, gibt es auch durchaus verschiedene Ansätze: Anschauliche Beweise aus der Schulgeometrie, etwa für den Satz des Pythagoras oder andere elementargeometrische Sätze, wären um im Bild zu bleiben eher ein Art Wegesystem für Fußgänger, wohingegen sich professionelle Mathematiker seit etwa 1900 meist auf einem aus Mengenlehre aufgebauten Schienennetz bewegen, das zwar große anfängliche Investitionen erfordert, danach aber dem Verstand

ein sehr schnelles, sicheres und tiefes Eindringen ermöglicht. Allerdings fällt es unseren durch die Bequemlichkeit dieses Zugangs verwöhnten Studenten meist bitter schwer, dann an interessanten und noch nicht erschlossenen Stellen wieder auszusteigen und sich zu Fuß weiter fortzubewegen oder gar selbst Schienen zu legen. Rechnergestützte Beweise sind für mich wie eine Erkundung mit Robotern nicht in derselben Weise befriedigend wie der persönliche Augenschein, aber wenn man an interessante Stellen partout nicht selbst hingelangen kann, sind doch schöne von Robotern geschossene Bilder allemal besser als gar nichts.

Die Mathematik wird insbesondere von Außenstehenden oft als eine tote Wissenschaft erlebt, in der "alles schon seit dreihundert Jahren bekannt sei". Dieser Eindruck mag damit zusammenhängen, daß Mathematik durchaus "verholzt" in dem Sinne, daß sie einen festen Stamm an Wissen und kodifizierter Sprache ausbildet. Das aber ermöglicht es unserer Wissenschaft auch gerade wieder, hoch hinaus zu wachsen.

Beim Erlernen dieser Wissenschaft denke ich, man soll versuchen, sich aller gedanklichen Kräfte zu bedienen, deren ein Mensch fähig ist. Geeignet für das Durchdringen mathematischer Sachverhalte scheinen mir insbesondere die räumliche oder noch besser die räumlich-zeitliche Anschauung, das abstrakte logische Denken und das formale Umformen von Zeichenketten auf Papier. Hilfreich kann auch unsere sprachliche Intelligenz sein: Bereits kleine Kinder lernen ja das Zählen, indem sie zunächst "Eins-Zwei-Drei-Vier-Fünf" memorieren wie "Abrakadabra Simsalabim", und ältere lernen ähnlich den Satz des Pythagoras oder die binomischen Formeln.

4 Herkunft einiger Symbole

Ich habe versucht, etwas über die Herkunft einiger mathematischer Symbole in Erfahrung zu bringen, die schon aus der Schule selbstverständlich sind. Das Gleichheitszeichen = scheint auf ein 1557 von Robert Recorde publiziertes Buch zurückzugehen und soll andeuten, daß das, was auf der linken und rechten Seite dieses Zeichens steht, so gleich ist wie die beiden Strichlein, die das uns heute so selbstverständliche Gleichheitszeichen bilden. Davor schrieb man statt einem Gleichheitszeichen meist *ae* für "äquivalent". Das Pluszeichen + ist wohl ein Auschnitt aus dem Symbol &, das hinwiederum enstanden ist durch Zusammenziehen der beiden Buchstaben im Wörtchen "et", lateinisch für "und".

Die Dezimaldarstellung der natürlichen Zahlen kam Mitte des vorigen Jahrtausends aus Indien über die Araber nach Italien. Bis dahin rechnete man in Europa in römischer Notation. Sie müssen nur versuchen, in dieser

Notation zwei größere Zahlen zu multiplizieren, um zu ermessen, welchen wissenschaftlichen und auch wirtschaftlichen Fortschritt der Übergang zur Dezimaldarstellung bedeutete. Das Beispiel der Dezimaldarstellung zeigt in meinen Augen auch, wie entscheidend das sorgfältige Einbeziehen trivialer Spezialfälle, manchmal als "Theorie der leeren Menge" verspottet, für die Eleganz der Darstellung mathematischer Sachverhalte sein kann: Sie wurde ja eben dadurch erst ermöglicht, daß man ein eigenes Symbol für "gar nichts" erfand! Ich denke, daß der Aufbau eines effizienten Notationssystems, obwohl er natürlich nicht denselben Stellenwert einnehmen kann wie die Entwicklung mathematischer Inhalte, dennoch in der Lehre ein wichtiges Ziel sein muß. In diesem Text habe ich mir die größte Mühe gegeben, unter den gebräuchlichen Notationen diejenigen auszuwählen, die mir am sinnvollsten schienen, und sie soweit wie möglich aufzuschlüsseln.

Die Herkunft der logischen Symbole $\exists$ und $\forall$ als umgedrehte E bzw. A haben wir bereits in I.2.3.3 erwähnt, sie wurden von Cantor in seiner Mengenlehre zuerst verwendet. Die Symbole $\mathbb{R}, \mathbb{Q}, \mathbb{N}, \mathbb{Z}$ wurden früher als fette Buchstaben gedruckt und zunächst nur beim Tafelanschrieb in dieser Gestalt wiedergegeben, da man fetten Druck an der Tafel nicht gut darstellen kann..

Teil B

Algebra

Kapitel III

Lineare Algebra

Inhalt

1 Gleichungssysteme und Vektorräume 71

1.1 Lösen linearer Gleichungssysteme 71

1.2 Ergänzungen zur Mengenlehre 77

1.3 Vektorräume und Untervektorräume 79

1.4 Lineare Unabhängigkeit und Basen 84

1.5 Lineare Abbildungen 92

1.6 Dimensionsformel 98

1.7 Lineare Abbildungen und Matrizen 101

1.8 Dualräume und transponierte Abbildungen 113

1.9 Affine Räume 119

2 Ringe, Polynome, Determinanten 129

2.1 Ringe 129

2.2 Untergruppen der ganzen Zahlen 133

2.3 Polynome 141

2.4 Äquivalenzrelationen 148

2.5 Quotientenkörper 149

2.6 Das Signum einer Permutation 151

2.7 Die Determinante 154

2.8 Eigenwerte und Eigenvektoren 165

3 Euklidische Vektorräume 170

3.1 Modellierung des Anschauungsraums 170

3.2 Geometrie in euklidischen Vektorräumen 173
3.3 Orthogonale und unitäre Abbildungen 180
3.4 Isometrien euklidischer affiner Räume 189
3.5 Winkel, Orientierung, Kreuzprodukt 191
3.6 Spektralsatz und Hauptachsentransformationen . . 201

4 Bilinearformen 209
4.1 Fundamentalmatrix 209
4.2 Definitheitseigenschaften 211
4.3 Klassifikation symmetrischer Bilinearformen 214
4.4 Alternierende Bilinearformen 219

5 Jordan'sche Normalform 221
5.1 Motivation durch Differentialgleichungen 221
5.2 Summen und Produkte von Vektorräumen 222
5.3 Hauptraumzerlegung 224
5.4 Jordan-Zerlegung 229
5.5 Jordan'sche Normalform 232

6 Algebra und Symmetrie 239
6.1 Gruppenwirkungen 239
6.2 Restklassen 245
6.3 Bahnformel 247
6.4 Normalteiler 249
6.5 Zyklische Gruppen 252
6.6 Endlich erzeugte abelsche Gruppen 256
6.7 Konjugationsklassen 263
6.8 Endliche Untergruppen der Drehgruppe 264
6.9 Skalarprodukte zu Drehgruppen 278

7 Universelle Konstruktionen 285
7.1 Quotientenvektorräume 285
7.2 Tensorprodukte von Vektorräumen 289
7.3 Kanonische Injektionen bei Tensorprodukten . . . 298
7.4 Alternierende Tensoren und Determinante 300
7.5 Das kanonische Skalarprodukt 306

1 Gleichungssysteme und Vektorräume

1.1 Lösen linearer Gleichungssysteme

1.1.1. Sei k ein Körper im Sinne von I.3.3.1. Gegeben seien n Gleichungen in m Unbekannten in der Form

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + & \ldots & + a_{1m}x_m = b_1 \\ a_{21}x_1 + a_{22}x_2 + & \ldots & + a_{2m}x_m = b_2 \\ \vdots & & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + & \ldots & + a_{nm}x_m = b_n \end{array}$$

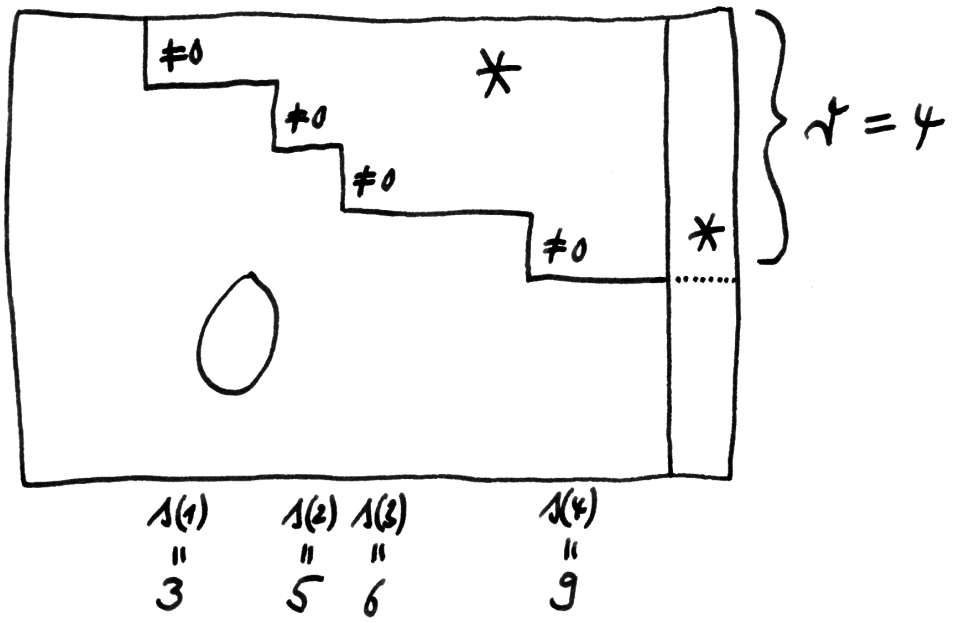
mit $a_{ij}, b_i \in k$ fest vorgegeben. Man spricht dann auch von einem **linearen Gleichungssystem**. Linear heißt es, weil darin keine komplizierteren Terme wie x_1^2 oder x_1x_2 vorkommen. Gleichungssysteme mit Termen dieser Art studieren wir erst in der algebraischen Geometrie. Sind alle b_i Null, so heißt unser System **homogen**, und das lineare Gleichungssystem, das entsteht, wenn wir alle b_i zu Null setzen, heißt das zugehörige **homogenisierte** Gleichungssystem. Gesucht ist eine Beschreibung aller m-Tupel $(x_1, \ldots, x_m)$ von Elementen von k derart, daß alle n Gleichungen gleichzeitig erfüllt sind. In der Begrifflichkeit und Notation, wie wir sie gleich in 1.2.2 einführen, bildet die Gesamtheit aller m-Tupel $(x_1, \ldots, x_m)$ von Elementen von k eine neue Menge k^m, und wir suchen eine möglichst explizite Beschreibung der Teilmenge $L \subset k^m$ aller derjenigen m-Tupel, die alle unsere n Gleichungen erfüllen, der sogenannten **Lösungsmenge** unseres Gleichungssystems.

1.1.2. Um die Lösungsmenge eines linearen Gleichungssystems zu bestimmen, kann man den **Gauß-Algorithmus** anwenden. Er basiert auf der elementaren Erkenntnis, daß sich die Lösungsmenge nicht ändert, wenn wir in einer der beiden folgenden Weisen zu einem neuen Gleichungssystem übergehen:

1. Wir ersetzen eine unserer Gleichungen durch ihre Summe mit einem Vielfachen einer anderen unserer Gleichungen;

2. Wir vertauschen zwei unserer Gleichungen.

Dann gehen wir mithilfe dieser beiden Operationen, also ohne die Lösungsmenge zu ändern, zu einem Gleichungssystem über, das **Zeilenstufenform** hat in dem Sinne, daß man ein $r \geq 0$ und Indizes $1 \leq s(1) < s(2) < \ldots < s(r) \leq m$ so angeben kann, daß in unserem transformierten Gleichungssystem gilt $a_{i,s(i)} \neq 0$ für $1 \leq i \leq r$, und daß zusätzlich $a_{\nu\mu} \neq 0$ nur gilt, wenn es ein i gibt mit $\nu \leq i$ und $\mu \geq s(i)$. Nebenstehendes Bild mag aufschlüsseln, welche Art von Gleichungssystemen diese Bedingungen beschreiben. Es ist



Ein System in Zeilenstufenform ist ein System der obigen Gestalt, bei dem im Teil mit den Koeffizienten a_{ij} wie angedeutet unterhalb solch einer "Treppe mit Stufenhöhe Eins aber mit variabler Breite der Stufen" nur Nullen stehen, vorn an den Stufenabsätzen aber von Null verschiedene Einträge. An die durch den senkrechten Strich abgetrennte letzte Spalte mit den gewünschten Lösungen b_i werden hierbei keinerlei Bedingungen gestellt.

üblich und erspart viel Schreibarbeit, die Symbole x_j sowie die Pluszeichen bei diesen Rechnungen wegzulassen und stattdessen ein Gleichungssystem der oben beschriebenen Art abzukürzen durch seine **erweiterte Koeffizientenmatrix**

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1m} & b_1 \\ a_{21} & a_{22} & & a_{2m} & b_2 \\ \vdots & & & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} & b_n \end{array}\right)$$

Die Spezifikation "erweitert" weist auf die letzte Spalte der b_i hin. Die Familie der a_{ij} für sich genommen heißt die **Koeffizientenmatrix** unseres Gleichungssystems.

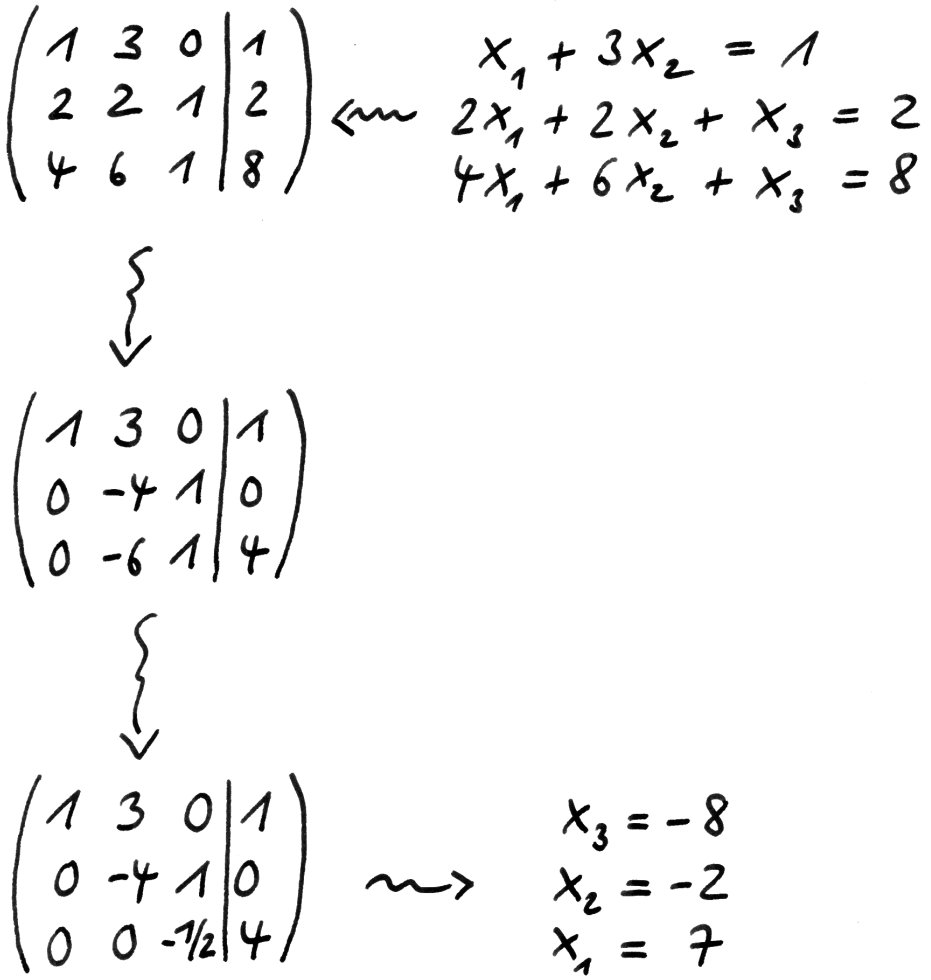
1.1.3. Der Gauß-Algorithmus funktioniert so: Sind alle Koeffizienten in der ersten Spalte Null, so ignorieren wir sie und machen mit der so entstehenden Matrix weiter. Ist ein Koeffizient in der ersten Spalte von Null verschieden, so bringen wir ihn durch eine Zeilenvertauschung an die oberste Stelle. Ziehen wir dann geeignete Vielfache der ersten Zeile von den anderen Zeilen ab, so gelangen wir zu einem System, bei dem wie angedeutet in der ersten Spalte unterhalb des obersten Eintrags nur Nullen stehen. Für das weitere ignorieren wir dann die erste Zeile und Spalte und machen mit der so entstehenden Matrix weiter.

1.1.4. Die Lösungsmenge eines linearen Gleichungssystems in Zeilenstufenform ist schnell bestimmt: Ist eine der Zahlen $b_{r+1}, \dots, b_n$ nicht Null, so besitzt es gar keine Lösung. Gilt dahingegen $b_{r+1} = \dots = b_n = 0$, können wir Zahlen x_μ für μ verschieden von den Spaltenindizes $s(1), \dots, s(r)$ der Stufen beliebig vorgehen und finden für jede solche Vorgabe der Reihe nach eindeutig bestimmte $x_{s(r)}, x_{s(r-1)}, \dots, x_{s(1)}$ derart, daß das entstehende m-Tupel $(x_1, \dots, x_m)$ eine Lösung unseres Gleichungssystems ist.

1.1.5. Eine Abbildung der Produktmenge $\{1, \dots, n\} \times \{1, \dots, m\}$ in eine Menge Z bezeichnet man als eine **$(n \times m)$-Matrix mit Koeffizienten in Z**. Gegeben solch eine Matrix A schreibt man meist A_{ij} oder a_{ij} statt $A(i,j)$ und veranschaulicht sich dieses Datum als ein quadratisches Arrangement von Elementen von Z wie eben im Fall $Z = k$. Das i heißt hierbei der **Zeilenindex**, da es angibt alias "indiziert", in welcher Zeile unser Eintrag a_{ij} steht, wohingegen man das j den **Spaltenindex** unseres Matrixeintrags nennt. Die Menge aller $(n \times m)$-Matrizen mit Koeffizienten in Z notieren wir

$$\mathrm{M}(n \times m; Z)$$

Im Fall $n = m$ sprechen wir von einer **quadratischen Matrix**. Manchmal werden wir sogar für beliebige Mengen X, Y, Z eine Abbildung $X \times Y \to Z$ als eine **$(X \times Y)$-Matrix mit Koeffizienten in Z** ansprechen.



Ein lineares Gleichungssystem mit drei Gleichungen und drei Unbekannten und seine Lösung mit dem Gauß-Algorithmus.

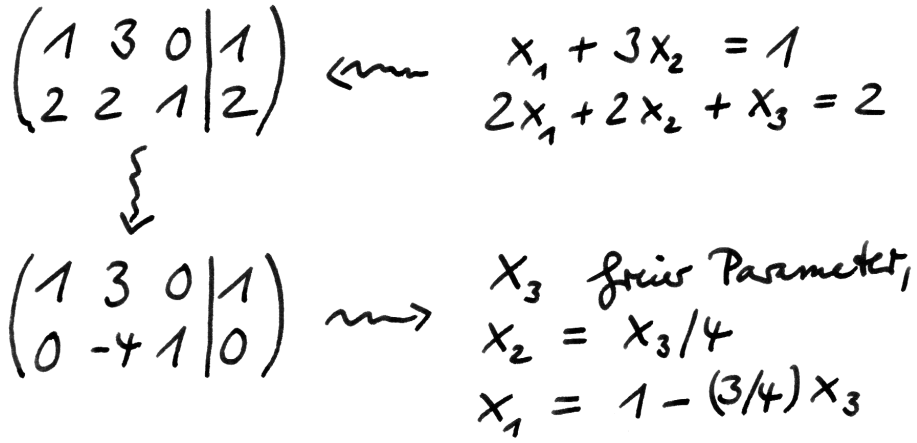
1.1.6.

Satz 1.1.7. *Ist die Lösungsmenge eines linearen Gleichungssystems nicht leer, so erhalten wir alle Lösungen, indem wir zu einer speziellen Lösung unseres Systems eine beliebige Lösung des zugehörigen homogenisierten Systems komponentenweise addieren.*

Beweis. Ist $c = (c_1, \ldots, c_m)$ eine Lösung unseres linearen Gleichungssystems und $d = (d_1, \ldots, d_m)$ eine Lösung des homogenisierten Systems, so ist offensichtlich die komponentenweise Summe $c \dot{+} d = (c_1 + d_1, \ldots, c_m + d_m)$ eine Lösung des ursprünglichen Systems. Ist andererseits $c' = (c'_1, \ldots, c'_m)$ eine weitere Lösung unseres linearen Gleichungssystems, so ist offensichtlich die komponentenweise Differenz $d = (c'_1 - c_1, \ldots, c'_m - c_m)$ eine Lösung des homogenisierten Systems, für die gilt $c' = c \dot{+} d$. □

1.1.8. Die vorstehenden Überlegungen zeigen, wie man die Lösungsmenge eines linearen Gleichungssystems bestimmen kann, und man erhält dabei im Fall einer nichtleeren Lösungsmenge sogar eine ausgezeichnete Bijektion von einem k^t mit besagter Lösungsmenge, für $t = m - r$, die eben jeder Vorgabe von x_j für $j \neq s(1), \ldots, s(r)$ die durch diese Vorgabe eindeutig bestimmte Lösung zuordnet. Der Gauß-Algorithmus gibt uns allerdings nicht vor, welche Zeilenvertauschungen wir unterwegs verwenden wollen, und damit stellt sich sofort die Frage, ob wir unabhängig von diesen Wahlen stets bei derselben Matrix in Zeilenstufenform ankommen. Das ist nun zwar nicht richtig, aber dennoch sind die "Breiten der einzelnen Stufen" alias die $s(i)$ unabhängig von allen Wahlen, denn sie lassen sich auch direkt beschreiben, indem wir im zugehörigen homogenisierten Gleichungssystem unsere Variablen von hinten durchgehen und jeweils fragen: Gibt es für jedes $(x_{j+1}, x_{j+2}, \ldots, x_m)$, das zu einer Lösung $(x_1, x_2, \ldots, x_m)$ ergänzbar ist, nur ein x_j derart, daß auch $(x_j, x_{j+1}, x_{j+2}, \ldots, x_m)$ zu einer Lösung $(x_1, x_2, \ldots, x_m)$ ergänzbar ist? Genau dann ist die Antwort "ja", wenn in der j-ten Spalte eine neue Stufe beginnt.

1.1.9. Nun könnten wir natürlich vor Anwendung des Gauss-Algorithmus auch zuerst unsere Variablen umnummerieren alias die Spalten unserer Koeffizientenmatrix vertauschen. Wir erhielten wieder eine Bijektion eines k^u mit der Lösungsmenge wie eben. Die Frage, der wir uns als nächstes zuwenden wollen, lautet nun: Gilt stets $u = t$, in anderen Worten, landen wir bei einer Zeilenstufenform mit derselben Zahl von Stufen, wenn wir zuerst die Spalten unseres Systems willkürlich vertauschen, bevor wir den Gauß-Algorithmus durchführen? Die Antwort lautet wieder "Ja", aber hierzu ist mir kein ganz elementares Argument mehr eingefallen, und darüber war ich sogar ganz froh: Diese Frage kann nun nämlich zur Motivation der Entwicklung der abstrakten Theorie der Vektorräume dienen, mit der wir an dieser Stelle beginnen.



Ein lineares Gleichungssystem mit zwei Gleichungen und drei Unbekannten, dessen Lösungsmenge nach unser allgemeinen Theorie für jedes x_3 genau einen Punkt (x_1, x_2, x_3) enthält, und zwar haben wir wegen der zweiten Gleichung $x_2 = x_3/4$ und dann wegen der ersten Gleichung $x_1 = 1 - (3/4)x_3$, so daß die allgemeine Lösung lautet $(1 - (3/4)\lambda, \lambda/4, \lambda)$ für variables λ.

Wir führen in diesem Rahmen den auch in vielen anderen Zusammenhängen äußerst nützlichen Begriff der "Dimension" eines "Vektorraums" ein, und zeigen in 1.5.8, daß die Stufenzahl in diesem Rahmen unabhängig von allen Wahlen beschrieben werden kann als die "Dimension des Lösungsraums" des zugehörigen homogenisierten Gleichungssystems. Zunächst jedoch führen wir weitere Begriffe der Mengenlehre ein, die wir dabei und auch darüber hinaus noch oft brauchen werden.

1.2 Ergänzungen zur Mengenlehre

1.2.1. Bis jetzt hatten wir nur das kartesische Produkt $X \times Y$ von zwei Mengen X und Y betrachtet. Ebenso kann man jedoch auch für Mengen $X_1, \ldots, X_n$ das kartesische Produkt

$$X_1 \times \ldots \times X_n = \{(x_1, \ldots, x_n) \mid x_i \in X_i \text{ für } 1 \leq i \leq n\}$$

einführen. Die Elemente von so einem Produkt bezeichnet man als n-**Tupel**.

1.2.2. Gegeben drei Mengen X, Y, Z kann man sich natürlich die Frage stellen, inwieweit die drei Mengen $(X \times Y) \times Z$, $X \times (Y \times Z)$ und $X \times Y \times Z$ übereinstimmen und auch allgemeiner, inwieweit "das kartesische Produkt $\times$ assoziativ ist". Wir werden derartige Fragen später im Rahmen der Kategorientheorie ausführlich diskutieren. Hier sei nur bemerkt, daß zum Beispiel alle unsere drei Tripelprodukte jedenfalls wohlbestimme Projektionen pr_X, pr_Y und pr_Z auf X, Y und Z haben und daß es eindeutig bestimmte Bijektionen zwischen ihnen gibt, die mit diesen drei Projektionen verträglich sind. Wir werden derartige Bijektionen meist nicht explizit machen. Es ist auch sinnvoll und allgemeine Konvention, das Produkt von Null Mengen als "die" einelementige Menge zu verstehen. Das kartesische Produkt von n Kopien einer Menge X kürzt man meist mit X^n ab, die Elemente von X^n sind also n-Tupel von Elementen aus X und X^0 besteht aus genau einem Element.

1.2.3. Für ein kartesisches Produkt hat man stets die **Projektionsabbildungen** oder **Projektionen**

$$\begin{array}{rccc} \mathrm{pr}_i : & X_1 \times \ldots \times X_n & \to & X_i \\ & (x_1, \ldots, x_n) & \mapsto & x_i \end{array}$$

Gegeben eine Menge Z und Abbildungen $f_i : Z \to X_i$ können wir eine Abbildung $f : Z \to X_1 \times \ldots \times X_n$ von Z in das kartesische Produkt der X_i bilden durch die Vorschrift mit $z \mapsto (f_1(z), \ldots, f_n(z))$, und jede Abbildung $f : Z \to X_1 \times \ldots \times X_n$ ist von dieser Form mit $f_i = \mathrm{pr}_i \circ f$. Wir schreiben dann kurz $f = (f_1, \ldots, f_n)$. In der exponentiellen Schreibweise geschrieben

haben wir also einen kanonische Bijektion $(X_1 \times \ldots \times X_n)^Z \xrightarrow{\sim} X_1^Z \times \ldots \times X_n^Z$. Besonders wichtig ist die **diagonale Einbettung** oder **Diagonale**

$$\begin{array}{rccc} \Delta = (\mathrm{id}, \mathrm{id}): & X & \to & X \times X \\ & x & \mapsto & (x, x) \end{array}$$

1.2.4. Ist ein weiteres Produkt von der Form $Y = Y_1 \times \ldots \times Y_n$ gegeben sowie Abbildungen $f_i : X_i \to Y_i$, so können wir auch die Abbildung

$$\begin{array}{ccc} X_1 \times \ldots \times X_n & \to & Y_1 \times \ldots \times Y_n \\ (x_1, \ldots, x_n) & \mapsto & (f_1(x_1), \ldots, f_n(x_n)) \end{array}$$

erklären. Wir notieren diese Abbildung $f_1 \times \cdots \times f_n$. Man beachte jedoch, daß keineswegs alle Abbildungen $X_1 \times \ldots \times X_n \to Y_1 \times \cdots \times Y_n$ von dieser Form sind. Man beachte allgemeiner, daß eine Abbildung $f : X_1 \times \ldots \times X_n \to Z$ von einem kartesischen Produkt in eine beliebige Menge Z sich keineswegs in ähnlicher Weise aus Abbildungen $X_i \to Z$ zusammensetzen läßt, wie wir das bei Abbildungen von einer beliebigen Menge in ein kartesisches Produkt gesehen hatten.

Definition 1.2.5. Gegeben eine Menge X erinnere ich an die Menge aller Teilmengen $\mathcal{P}(X) = \{U \mid U \subset X\}$ von X, die sogenannte **Potenzmenge** von X. Da es mich verwirrt, über Mengen von Mengen zu reden, werde ich Teilmengen von $\mathcal{P}(X)$ nach Möglichkeit als **Systeme von Teilmengen von** X ansprechen. Gegeben ein solches System $\mathcal{U} \subset \mathcal{P}(X)$ bildet man zwei neue Teilmengen von X, den **Schnitt** und die **Vereinigung** der Mengen aus unserem System $\mathcal{U}$, durch die Vorschrift

$$\begin{array}{rcl} \bigcup_{U \in \mathcal{U}} U & = & \{x \in X \mid \text{ Es gibt } U \in \mathcal{U} \text{ mit } x \in U\} \\ \bigcap_{U \in \mathcal{U}} U & = & \{x \in X \mid \text{ Für alle } U \in \mathcal{U} \text{ gilt } x \in U\} \end{array}$$

Insbesondere ist der Schnitt über das leere System von Teilmengen von X ganz X und die Vereinigung über das leere System von Teilmengen von X die leere Menge.

1.2.6. Wir würden nun gerne zum Beispiel die Erkenntnis, daß das Komplement eines derartigen Schnitts die Vereinigung der Komplemente ist, ausdrücken können in der Formel

$$X \setminus \left(\bigcap_{U \in \mathcal{U}} U \right) = \bigcup_{U \in \mathcal{U}} (X \setminus U)$$

Damit in dieser Formel auch die Bedeutung der rechten Seite unmißverständlich klar ist, führen wir weitere Notationen ein.

1.2.7. Gegeben Mengen A und I bezeichnet man eine Abbildung $I \to A$ ganz allgemein auch als eine **durch I indizierte Familie von Elementen von A** und benutzt die Notation

$$(a_i)_{i\in I}$$

Diese Sprechweise und Notation für Abbildungen verwendet man insbesondere dann, wenn man der Menge I eine untergeordnete Rolle zugedacht hat. Im Fall $I = \emptyset$ spricht man von der **leeren Familie** von Elementen von A.

Definition 1.2.8. Gegeben eine Familie $(X_i)_{i\in I}$ von Teilmengen einer Menge X erklärt man ihren **Schnitt** und ihre **Vereinigung** durch die Regeln

$$\begin{aligned} \textstyle\bigcap_{i\in I} X_i &= \{x \in X \mid \text{Für alle } i \in I \text{ gilt } x \in X_i\} \\ \textstyle\bigcup_{i\in I} X_i &= \{x \in X \mid \text{Es existiert ein } i \in I \text{ mit } x \in X_i\} \end{aligned}$$

Insbesondere ist der Schnitt über die leere Familie von Teilmengen von X ganz X und die Vereinigung über die leere Familie von Teilmengen von X ist die leere Menge.

Übung 1.2.9. Man verallgemeinere die Formeln aus I.2.1.13 auf diese Situation. Genauer schreibe man in Formeln und zeige, daß der Schnitt einer derartigen Vereinigung mit einer weiteren Menge die Vereinigung der Schnitte ist, die Vereinigung eines derartigen Schnitts mit einer weiteren Menge der Schnitt der Vereinigungen, das Komplement eines Schnitts die Vereinigung der Komplemente und das Komplement einer Vereinigung der Schnitt der Komplemente. Besonders Mutige versuchen, für eine durch ein Produkt indizierte Familie $(X_{ij})_{(i,j)\in I\times J}$ den Schnitt von Vereinigungen $\bigcap_{j\in J}\left(\bigcup_{i\in I} X_{ij}\right)$ als Vereinigung von Schnitten zu schreiben.

1.2.10. In 5.2 diskutieren wir allgemeiner Produkte und zusätzlich "disjunkte Vereinigungen" beliebiger nicht notwendig endlicher Mengensysteme.

1.3 Vektorräume und Untervektorräume

Definition 1.3.1. Ein **Vektorraum V über einem Körper k** ist ein Paar bestehend aus einer abelschen Gruppe $V = (V, \dot{+})$ und einer Abbildung

$$\begin{array}{ccc} k \times V & \to & V \\ (\lambda, \vec{v}) & \mapsto & \lambda\vec{v} \end{array}$$

derart, daß für alle $\lambda, \mu \in k$ und $\vec{v}, \vec{w} \in V$ die folgenden Formeln gelten:

$$\begin{aligned} \lambda(\vec{v} \dot{+} \vec{w}) &= (\lambda\vec{v}) \dot{+} (\lambda\vec{w}) \\ (\lambda + \mu)\vec{v} &= (\lambda\vec{v}) \dot{+} (\mu\vec{v}) \\ \lambda(\mu\vec{v}) &= (\lambda\mu)\vec{v} \\ 1\vec{v} &= \vec{v} \end{aligned}$$

1.3.2. Die Elemente eines Vektorraums nennt man meist die **Vektoren** des Vektorraums. Die Elemente des Körpers heißen in diesem Zusammenhang oft **Skalare** und die Abbildung $(\lambda, \vec{v}) \mapsto \lambda \cdot \vec{v}$ die **Multiplikation mit Skalaren** und ist nicht zu verwechseln mit dem "Skalarprodukt", das wir in 3.1.4 einführen und das aus zwei Vektoren einen Skalar macht. Ich habe oben aus didaktischen Gründen die Addition von Vektoren $\dot{+}$ notiert, um sie von der Addition und Multiplikation von Körperelementen zu unterscheiden, aber das werde ich nicht lange durchhalten. Mit der auch in diesem Zusammenhang allgemein üblichen Konvention "Punkt vor Strich" und der zu $+$ vereinfachten Notation für die Addition von Vektoren lauten unsere Vektorraumaxiome dann etwas übersichtlicher

$$\begin{array}{rcl} \lambda(\vec{v}+\vec{w}) & = & \lambda\vec{v}+\lambda\vec{w} \\ (\lambda+\mu)\vec{v} & = & \lambda\vec{v}+\mu\vec{v} \\ \lambda(\mu\vec{v}) & = & (\lambda\mu)\vec{v} \\ 1\vec{v} & = & \vec{v} \end{array}$$

Ich habe weiter aus didaktischen Gründen bis hierher Vektoren stets mit einem Pfeil notiert, das halte ich wohl etwas länger durch, aber auf Dauer werden Sie sich den Pfeil auch selbst dazudenken müssen. Die letzte Bedingung $1\vec{v} = \vec{v}$ schließt zum Beispiel den Fall aus, daß wir für V irgendeine von Null verschiedene abelsche Gruppe nehmen und dann einfach setzen $\lambda\vec{v} = \vec{0}$ für alle $\lambda \in k$ und $\vec{v} \in V$. Das neutrale Element der abelschen Gruppe V notieren wir $\vec{0}$ und nennen es den **Nullvektor**.

1.3.3. Die Bezeichnung "Vektor" kommt vom lateinischen "vehere" für fahren, transportieren, und rührt von unserem Beispiel I.1.2.5 der Gesamtheit aller Parallelverschiebungen der Ebene oder des Raums her, deren Elemente ja in gewisser Weise Punkte transportieren. Die Bezeichnung "Skalare" für Elemente des zugrundeliegenden Körpers kommt von dem lateinischen Wort "scala" für "Leiter" und hat sich von dort über das Metermaß zu einer Bezeichnung für das, was man auf einer Meßskala ablesen kann, als da heißt zu einer Bezeichnung für reelle Zahlen entwickelt. In der Mathematik werden nun aber nicht nur reelle Vektorräume betrachtet, und so überträgt man dann dieses Wort weiter und verwendet es auch im allgemeinen als Bezeichnung für die Elemente des zugrundeliegenden Körpers.

1.3.4. Gegeben ein Vektorraum V und ein Vektor $\vec{v} \in V$ gilt $0\vec{v} = \vec{0}$. In der Tat finden wir mit der zweiten Formel aus der Definition $0\vec{v} = (0+0)\vec{v} = 0\vec{v} \dot{+} 0\vec{v}$ und Subtraktion von $0\vec{v}$ auf beiden Seiten liefert $\vec{0} = 0\vec{v}$, in Worten "Null mal ein Vektor ist stets der Nullvektor".

1.3.5. Gegeben ein Vektorraum V und ein Vektor $\vec{v} \in V$ gilt $(-1)\vec{v} = -\vec{v}$, das Negative von $\vec{v}$ in der abelschen Gruppe V. In der Tat finden wir mit der

letzten und der zweiten Formel aus der Definition $\vec{v} \dot{+} (-1)\vec{v} = 1\vec{v} + (-1)\vec{v} = (1 + (-1))\vec{v} = 0\vec{v} = \vec{0}$ und damit ist $(-1)\vec{v}$ in der Tat das additive Inverse von $\vec{v}$, in Formeln $(-1)\vec{v} = -\vec{v}$.

Übung 1.3.6. Gegeben ein Vektorraum V über einem Körper k zeige man für alle $\lambda \in k$ die Identität $\lambda\vec{0} = \vec{0}$. Weiter zeige man, daß aus $\lambda\vec{v} = \vec{0}$ folgt $\lambda = 0$ oder $\vec{v} = \vec{0}$.

Übung 1.3.7. Eine vorgegebene abelsche Gruppe kann auf höchstens eine Weise mit der Struktur eines $\mathbb{Q}$-Vektorraums versehen werden.

Beispiele 1.3.8. Einige Beispiele für Vektorräume wurden bereits in I.1.2 diskutiert. Besonders wichtig ist das Beispiel des Vektorraums

$$V = k^n$$

über einem vorgegebenen Körper k mit den Operationen gegeben durch

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \dot{+} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{pmatrix}$$

$$\lambda \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{pmatrix}$$

für $\lambda, v_1, \ldots, v_n, w_1, \ldots, w_n \in k$. Wir haben unsere n-Tupel hier der Übersichtlichkeit untereinander geschrieben. Die erste dieser Gleichungen definiert die Summe zweier n-Tupel, also die Addition in unserem Vektorraum $V = k^n$, indem sie diese durch die Addition in k ausdrückt. Die zweite Gleichung leistet dasselbe für die Multiplikation mit Skalaren. Ich gebe nun einen Teil der didaktischen Notation auf und schreibe von hier an $+$ statt $\dot{+}$. Gegeben $\vec{v} \in k^n$ schreibe ich seine Komponenten $v_1, v_2, \ldots, v_n$ und versehe sie nicht mit Pfeilen, da sie ja Elemente des Grundkörpers sind. Wenn irgendwo einmal $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ stehen sollte, so sind nicht die n Komponenten eines n-Tupels $\vec{v}$ gemeint, sondern vielmehr n Vektoren eines Vektorraums. Sobald ich die Pfeil-Notation aufgegeben habe, muß der Leser aus dem Kontext erschließen, was jeweils gemeint ist.

Beispiel 1.3.9. Gegeben ein Körper k wird jede einelementige Menge V vermittels der offensichtlichen Operation zu einem k-Vektorraum. Wir sprechen dann von einem **Nullvektorraum**, weil er eben nur aus dem Nullvektor besteht, und verwenden oft auch den bestimmten Artikel und sprechen von *dem* Nullvektorraum, da er ja im Wesentlichen eindeutig bestimmt ist.

Übung 1.3.10. Gegeben ein Körper k und k-Vektorräume $V_1, \ldots, V_n$ können wir das kartesische Produkt $V_1 \times \ldots \times V_n$ zu einem k-Vektorraum machen, indem wir die Addition sowie die Multiplikation mit Skalaren komponentenweise definieren. Den so entstehenden Vektorraum notieren wir auch

$$V_1 \oplus \ldots \oplus V_n$$

und nennen ihn die **direkte Summe** oder noch genauer die **externe direkte Summe** von Vektorräumen, wenn wir sie von der in 1.5.17 diskutierten "internen Summe von Untervektorräumen" abgrenzen wollen.

Definition 1.3.11. Eine Teilmenge U eines Vektorraums V heißt ein **Untervektorraum** oder **Teilraum** genau dann, wenn U den Nullvektor enthält und wenn aus $\vec{u}, \vec{v} \in U$ und $\lambda \in k$ folgt $\vec{u} + \vec{v} \in U$ sowie $\lambda\vec{u} \in U$.

1.3.12. Die vom höheren Standpunkt aus "richtige" Definition eines Untervektorraums lautet eigentlich anders, und zwar so: Sei k ein Körper. Eine Teilmenge eines k-Vektorraums heißt ein Untervektorraum genau dann, wenn sie so mit der Struktur eines k-Vektorraums versehen werden kann, daß die Einbettung ein "Homomorphismus k-Vektorräumen" wird. Ich kann diese "bessere" Definition hier noch nicht geben, da wir Homomorphismen von k-Vektorräumen noch nicht kennengelernt haben. Sie scheint mir deshalb besser, da man in derselben Weise auch korrekte Definitionen von Untermonoiden, Untergruppen, Unterkörpern und Unter-was-nicht-noch-all-für-Strukturen erhält, die sie erst später kennenlernen werden.

Beispiel 1.3.13. Unter einem homogenen linearen Gleichungssystem über einem gegebenen Körper k versteht man, wie bereits erwähnt, ein System von Gleichungen der Gestalt

$$\begin{array}{ccccc} a_{11}x_1 + a_{12}x_2 + & \ldots & +a_{1m}x_m & = 0 \\ a_{21}x_1 + a_{22}x_2 + & \ldots & +a_{2m}x_m & = 0 \\ \vdots & & \vdots & \\ a_{n1}x_1 + a_{n2}x_2 + & \ldots & +a_{nm}x_m & = 0 \end{array}$$

bei dem also rechts nur Nullen stehen. Die Lösungsmenge eines solchen homogenenen Gleichungssystems ist offensichtlich ein Untervektorraum $L \subset k^m$.

Beispiel 1.3.14. Wir verlassen für dieses Beispiel unser kristallklar aus reiner Mengenlehre aufgebautes mathematisches Paradies und denken uns die Gesamtheit aller Parallelverschiebungen des Anschauungsraums I.1.2.6 als Vektorraum, obwohl mir nicht so klar scheint, ob es sich bei diesen Parallelverschiebungen auch wirklich um im Cantor'schen Sinne "wohlunterschiedene

Objekte unseres Denkens oder unserer Anschauung" handelt. Die Untervektorräume dieses Vektorraums können wir uns dann denken als (0) die einelementige Teilmenge, die nur aus dem Nullvektor alias der "Verschiebung um den Abstand Null" besteht, (1) jede Teilmenge, die aus allen Verschiebungen in einer festen Richtung oder ihre Gegenrichtung besteht, den Nullvektor eingeschlossen, (2) die Teilmenge aller "horizontalen" Verschiebungen, und allgemeiner die Teilmenge aller Verschiebungen in Richtungen die auf einem fest vorgegebenen von Null verschiedenen Vektor senkrecht stehen, sowie (3) die Menge überhaupt aller Parallelverschiebungen des Anschauungsraums.

1.3.15. Jeder Schnitt von Untervektorräumen eines Vektorraums ist wieder ein Untervektorraum. Betrachten wir für eine Teilmenge T eines Vektorraums V über einem Körper k den Schnitt aller Untervektorräume von V, die T umfassen, so erhalten wir offensichtlich den kleinsten Untervektorraum von V, der T umfasst. Insbesondere existiert stets solch ein kleinster Untervektorraum. Wir notieren ihn

$$\langle T\rangle = \langle T\rangle_k \subset V$$

und bezeichnen ihn als den **von** T **erzeugten** Untervektorraum oder den **von** T **aufgespannten** Untervektorraum von V oder auch das **Erzeugnis von** T oder den **Spann von** T. Er kann auch beschrieben werden als die Menge

$$\langle T\rangle = \left\{ \alpha_1\vec{v}_1 + \ldots + \alpha_r\vec{v}_r \;\middle|\; \begin{array}{l} \alpha_1,\ldots,\alpha_r \in k, \\ \vec{v}_1,\ldots,\vec{v}_r \in T, \\ r \geq 0 \end{array} \right\}$$

wobei die leere Summe mit $r = 0$ den Nullvektor meint. In der Tat ist die auf der rechten Seite dieser Gleichung beschriebene Menge offensichtlich ein Untervektorraum von V, und jeder Untervektorraum von V, der T umfaßt, muß auch die auf der rechten Seite dieser Gleichung beschriebene Menge umfassen. Einen Vektor der Gestalt $\alpha_1\vec{v}_1+\ldots+\alpha_r\vec{v}_r$ bezeichnen wir auch als eine **Linearkombination** der Vektoren $\vec{v}_1,\ldots,\vec{v}_r$. Die Elemente von $\langle T\rangle$ bezeichnen wir auch als **Linearkombinationen von Vektoren aus** T. Gemeint sind damit also immer endliche Linearkombinationen, auch wenn die Menge T selbst unendlich sein sollte.

Definition 1.3.16. Eine Teilmenge eines Vektorraums heißt ein **Erzeugendensystem** unseres Vektorraums genau dann, wenn ihr Erzeugnis der ganze Vektorraum ist. Ein Vektorraum, der ein endliches Erzeugendensystem besitzt, heißt **endlich erzeugt**. Manche Autoren verwenden auch gleichbedeutend die vielleicht präzisere Terminologie **endlich erzeugbar**.

Definition 1.3.17. Sei X eine Menge und k ein Körper. Die Menge $\mathrm{Ens}(X,k)$ aller Abbildungen $f : X \to k$ mit der punktweisen Addition und Multiplikation mit Skalaren ist offensichtlich ein k-Vektorraum. Darin bilden alle Abbildungen, die nur an endlich vielen Stellen von Null verschiedene Werte annehmen, einen Untervektorraum

$$kX \subset \mathrm{Ens}(X,k)$$

Dieser Untervektorraum heißt der **freie Vektorraum über der Menge** X. Ein Element $a \in kX$ fassen wir als "formale Linearkombination von Elementen von X" auf und notieren es statt $(a_x)_{x\in X}$ suggestiver $\sum_{x\in X} a_x x$. Im Fall der Menge $X = \{\sharp, \flat, \natural\}$ wäre ein typisches Element von $\mathbb{Q}X$ etwa der Ausdruck

$$\frac{1}{2}\,\sharp - \frac{7}{5}\,\flat + 3\,\natural$$

Übung 1.3.18. Man zeige, daß für eine unendliche Menge X weder der Vektorraum $\mathrm{Ens}(X,k)$ noch der freie Vektorraum kX über X endlich erzeugt sind. Hinweis: Für den Fall $\mathrm{Ens}(X,k)$ braucht man die Resultate des folgenden Abschnitts.

Übung 1.3.19. Gegeben eine Menge X und ein k-Vektorraum V ist auch die Menge $\mathrm{Ens}(X,V)$ aller Abbildungen von $X \to V$ ein k-Vektorraum, wenn man sie mit der Addition gegeben durch $(f+g)(x) = f(x) + g(x)$ und mit der Multiplikation mit Skalaren gegeben durch $(\lambda f)(x) = \lambda(f(x))$ versieht.

Übung 1.3.20. Eine Teilmenge eines Vektorraums heißt eine **Hyperebene** oder genauer **lineare Hyperebene** genau dann, wenn unsere Teilmenge ein echter Untervektorraum ist, der zusammen mit einem einzigen weiteren Vektor unseren ursprünglichen Vektorraum erzeugt. Man zeige, daß eine Hyperebene zusammen mit jedem Vektor außerhalb besagter Hyperebene unseren ursprünglichen Vektorraum erzeugt.

1.4 Lineare Unabhängigkeit und Basen

Definition 1.4.1. Eine Teilmenge L eines Vektorraums V heißt **linear unabhängig** genau dann, wenn für beliebige paarweise verschiedene Vektoren $\vec{v}_1, \ldots, \vec{v}_r \in L$ und beliebige Skalare $\alpha_1, \ldots, \alpha_r \in k$ aus $\alpha_1\vec{v}_1 + \ldots + \alpha_r\vec{v}_r = \vec{0}$ bereits folgt $\alpha_1 = \ldots = \alpha_r = 0$.

Definition 1.4.2. Eine Teilmenge L eines Vektorraums V heißt **linear abhängig** genau dann, wenn sie nicht linear unabhängig ist, wenn es also ausgeschrieben paarweise verschiedene Vektoren $\vec{v}_1, \ldots, \vec{v}_r \in L$ und Skalare $\alpha_1, \ldots, \alpha_r \in k$ gibt derart, daß nicht alle α_i Null sind und dennoch gilt $\alpha_1\vec{v}_1 + \ldots + \alpha_r\vec{v}_r = \vec{0}$.

Beispiel 1.4.3. Die leere Menge ist in jedem Vektorraum linear unabhängig. Eine einelementige Teilmenge ist linear unabhängig genau dann, wenn sie nicht aus dem Nullvektor besteht: Für den Nullvektor gilt nämlich $1\vec{0} = \vec{0}$ und nach unseren Annahmen gilt in einem Körper stets $1 \neq 0$.

Übung 1.4.4. Eine zweielementige Teilmenge eines Vektorraums ist linear unabhängig genau dann, wenn keiner ihrer beiden Vektoren ein Vielfaches des anderen ist.

Definition 1.4.5. Eine **Basis eines Vektorraums** ist ein linear unabhängiges Erzeugendensystem.

Übung 1.4.6. Sei $(X, \leq)$ eine partiell geordnete Menge und k ein Körper. Seien für alle $x \in X$ Abbildungen $f_x : X \to k$ gegeben mit $f_x(x) \neq 0$ und $f_x(y) \neq 0 \Rightarrow y \geq x$. Man zeige, daß dann die Familie $(f_x)_{x \in X}$ linear unabhängig ist im Vektorraum $\mathrm{Ens}(X, k)$ aller Abbildungen von X nach k.

1.4.7. Manchmal ist es praktisch und führt zu einer übersichtlicheren Darstellung, Varianten unserer Begriffe zu verwenden, die sich statt auf Teilmengen unseres Vektorraums auf Familien von Vektoren $(\vec{v}_i)_{i \in I}$ beziehen. Eine derartige Familie heißt ein Erzeugendensystem genau dann, wenn die Menge $\{\vec{v}_i \mid i \in I\}$ ein Erzeugendensystem ist. Sie heißt **linear unabhängig** oder ganz pedantisch **linear unabhängig als Familie** genau dann, wenn für beliebige paarweise verschiedene Indizes $i_1, \ldots, i_r \in I$ und beliebige Skalare $\alpha_1, \ldots, \alpha_r \in k$ aus $\alpha_1 \vec{v}_{i_1} + \ldots + \alpha_r \vec{v}_{i_r} = \vec{0}$ bereits folgt $\alpha_1 = \ldots = \alpha_r = 0$. Der wesentliche Unterschied zur Begrifflichkeit für Teilmengen liegt darin, daß bei einer Familie ja für verschiedene Indizes die zugehörigen Vektoren durchaus gleich sein könnten, was aber durch die Bedingung der linearen Unabhängigkeit dann doch wieder ausgeschlossen wird. Eine Familie von Vektoren, die nicht linear unabhängig ist, nennen wir eine **linear abhängige Familie**. Eine erzeugende und linear unabhängige Familie nennt man wieder eine **Basis** oder ausführlicher eine **durch $i \in I$ indizierte Basis**.

1.4.8. Besonders oft werden wir später Basen betrachten, die durch die Menge $\{1, \ldots, n\}$ indiziert sind. Hier ist dann der wesentliche Unterschied zu einer Basis im Sinne von 1.4.5, daß wir zusätzlich festlegen, welcher Basisvektor der erste, welcher der zweite und so weiter sein soll. In der Terminologie aus **??** bedeutet das gerade, daß wir eine Anordnung auf unserer Basis festlegen. Wollen wir das besonders hervorheben, so sprechen wir von einer **angeordneten Basis**.

Beispiel 1.4.9. Sei k ein Körper und $n \in \mathbb{N}$. Wir betrachten in k^n die Vektoren

$$\vec{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)$$

mit einer Eins an der i-ten Stelle und Nullen sonst. Dann bilden $\vec{e}_1, \ldots, \vec{e}_n$ eine angeordnete Basis von k^n, die sogenannte **Standardbasis** des k^n.

Satz 1.4.10 (über Linearkombinationen von Basiselementen). *Sei k ein Körper, V ein k-Vektorraum und $(\vec{v}_i)_{i\in I}$ eine Familie von Vektoren aus unserem Vektorraum V. So sind gleichbedeutend:*

1. *Die Familie $(\vec{v}_i)_{i\in I}$ ist eine Basis von V;*
2. *Für jeden Vektor $\vec{v} \in V$ gibt es genau eine Familie $(a_i)_{i\in I}$ von Elementen unseres Körpers k, in der für höchstens endlich viele i das a_i von Null verschieden ist und für die gilt*

$$\vec{v} = \sum_{i\in I} a_i \vec{v}_i$$

Beweis. $1 \Rightarrow 2$) Ist unsere Familie ein Erzeugendensystem, so gibt es schon einmal für jeden Vektor $\vec{v} \in V$ ein Darstellung als Linearkombination $\vec{v} = \sum_{i\in I} a_i\vec{v}_i$ in der für höchstens endlich viele i das a_i von Null verschieden ist. Ist unsere Familie linear unabhängig, so muß diese Darstellung eindeutig sein, denn ist $\vec{v} = \sum_{i\in I} b_i\vec{v}_i$ eine weitere derartige Darstellung von $\vec{v}$, so folgt $\vec{0} = \sum_{i\in I}(a_i - b_i)\vec{v}_i$ und dann wegen der linearen Unabhängigkeit unserer Familie $a_i - b_i = 0$ für alle $i \in I$.

$2 \Rightarrow 1$) Aus 2 folgt sofort, daß die $(\vec{v}_i)_{i\in I}$ ein Erzeugendensystem bilden. Wenn sich weiter jeder Vektor nur auf genau eine Weise als Linearkombination der $\vec{v}_i$ schreiben läßt, so gilt das insbesondere auch für den Nullvektor, für den also $\vec{0} = \sum_{i\in I} 0\vec{v}_i$ die einzig mögliche Darstellung als Linearkombination der $\vec{v}_i$ ist. Das zeigt dann die lineare Unabhängigkeit der $\vec{v}_i$. □

1.4.11. Mit dem Begriff des freien Vektorraums kX über einer Menge X aus 1.3.17 können wir den vorhergehenden Satz 1.4.10 auch wie folgt umformulieren: Gegeben ein k-Vektorraum V ist eine Familie $(\vec{v}_i)_{i\in I}$ von Vektoren eine Basis genau dann, wenn das "Auswerten formaler Ausdrücke"

$$\begin{array}{rccc} \Phi: & kI & \to & V \\ & (a_i)_{i\in I} & \mapsto & \sum_{i\in I} a_i\vec{v}_i \end{array}$$

eine Bijektion ist. Ausführlicher gilt für diese Abbildung Φ sogar

$$\begin{array}{lcll} ((\vec{v}_i)_{i\in I} \text{ ist Erzeugendensystem}) & \Leftrightarrow & (\Phi \text{ ist eine Surjektion} & kI \twoheadrightarrow V) \\ ((\vec{v}_i)_{i\in I} \text{ ist linear unabhängig}) & \Leftrightarrow & (\Phi \text{ ist eine Injektion} & kI \hookrightarrow V) \\ ((\vec{v}_i)_{i\in I} \text{ ist eine Basis}) & \Leftrightarrow & (\Phi \text{ ist eine Bijektion} & kI \xrightarrow{\sim} V) \end{array}$$

Hier folgt die erste Äquivalenz direkt aus den Definitionen. Um bei der zweiten Äquivalenz die Implikation $\Leftarrow$ einzusehen, muß man nur bemerken, daß Φ den Nullvektor auf Null wirft und folglich kein anderer Vektor aus kI von Φ auf Null geworfen werden kann. Um bei der zweiten Äquivalenz die Implikation $\Rightarrow$ einzusehen, argumentiert man wie im Beweis von 1.4.10. Die letzte Äquivalenz schließlich ist eine direkte Konsequenz der ersten beiden. Als Variante bemerken wir, daß wir für jede angeordnete Basis $\mathcal{B} = (\vec{v}_1, \ldots, \vec{v}_n)$ eines k-Vektorraums V eine Bijektion

$$\Phi_{\mathcal{B}} : k^n \xrightarrow{\sim} V$$

erhalten durch die Abbildungsvorschrift $(\lambda_1, \ldots, \lambda_n) \mapsto \lambda_1\vec{v}_1 + \ldots + \lambda_n\vec{v}_n$. Der Beweis ist mutatis mutandis derselbe.

Satz 1.4.12. *Jedes minimale alias unverkürzbare Erzeugendensystem eines Vektorraums ist eine Basis. Jede maximale alias unverlängerbare linear unabhängige Teilmenge eines Vektorraums ist eine Basis.*

1.4.13. Die Begriffe minimal und maximal sind hier zu verstehen im Sinne von **??** in Bezug auf Inklusionen zwischen Teilmengen unseres Vektorraums, nicht etwa in Bezug auf die Zahl ihrer Elemente.

Beweis. Ist $E \subset V$ ein Erzeugendensystem und ist E nicht linear unabhängig, so gilt eine Relation $\lambda_1\vec{v}_1 + \ldots + \lambda_n\vec{v}_n = \vec{0}$ mit $n \geq 1$, den $\vec{v}_i \in E$ paarweise verschieden und allen $\lambda_i \neq 0$. Nach Multiplikation mit λ_1^{-1} dürfen wir hier sogar $\lambda_1 = 1$ annehmen. Wir folgern

$$\vec{v}_1 = -\lambda_1^{-1}\lambda_2\vec{v}_2 - \ldots - \lambda_1^{-1}\lambda_n\vec{v}_n \in \langle E \backslash \vec{v}_1 \rangle$$

und damit ist auch $E\backslash\vec{v}_1$ bereits ein Erzeugendensystem und E war nicht minimal. Ist umgekehrt eine Teilmenge $L \subset V$ linear unabhängig und kein Erzeugendensystem, so ist für jedes $\vec{v} \in V \backslash \langle L \rangle$ auch $L \cup \{\vec{v}\}$ linear unabhängig und L war nicht maximal. □

Übung 1.4.14. Seien $L \subset E$ eine linear unabhängige Teilmenge in einem Erzeugendensystem eines Vektorraums. Ist A minimal unter allen Erzeugendensystemen unseres Vektorraums mit $L \subset A \subset E$, so ist A eine Basis. Ist A maximal unter allen linear unabhängigen Teilmengen unseres Vektorraums mit $L \subset A \subset E$, so ist A eine Basis.

1.4.15. Unser Satz 1.4.12 impliziert insbesondere, daß jeder endlich erzeugte Vektorraum eine endliche Basis besitzt: Wir lassen einfach aus einem endlichen Erzeugendensystem so lange Vektoren weg, bis wir bei einem unverkürzbaren Erzeugendensystem angekommen sind. Mit raffinierteren Methoden der Mengenlehre kann man sogar den sogenannten **Basisexistenzsatz**

zeigen, nach dem überhaupt jeder Vektorraum eine Basis besitzt: Wir diskutieren das in 1.4.31 und recht eigentlich erst in **??**.

Satz 1.4.16. *Ist V ein Vektorraum, $L \subset V$ eine linear unabhängige Teilmenge und $E \subset V$ ein Erzeugendensystem, so gilt*

$$|L| \leq |E|$$

1.4.17. Wir verwenden hier unsere Konvention, nach der wir für alle unendlichen Mengen X schlicht $|X| = \infty$ setzen. Der Satz gilt aber auch mit einer feineren Interpretation von $|X|$ als "Kardinalität", genauer folgt aus dem Zorn'schen Lemma die Existenz einer Injektion $L \hookrightarrow E$, wie in 1.4.19 in größerer Allgemeinheit diskutiert wird.

Erster Beweis. Durch Widerspruch. Nehmen wir an, wir hätten ein Erzeugendensystem $E = \{\vec{w}_1, \ldots, \vec{w}_m\}$ und eine linear unabhängige Teilmenge $L = \{\vec{v}_1, \ldots, \vec{v}_n\}$ mit $n > m$. Dann könnten wir natürlich $a_{ij} \in k$ finden mit

$$\begin{array}{ccccccccc} \vec{v}_1 & = & a_{11}\vec{w}_1 & + & a_{21}\vec{w}_2 & + & \cdots & + & a_{m1}\vec{w}_m \\ \vdots & & \vdots & & \vdots & & & & \vdots \\ \vec{v}_n & = & a_{1n}\vec{w}_1 & + & a_{2n}\vec{w}_2 & + & \cdots & + & a_{mn}\vec{w}_m \end{array}$$

Jetzt betrachten wir das "vertikal geschriebene" homogene lineare Gleichungssystem

$$\begin{array}{ccccc} x_1a_{11} & & x_1a_{21} & & x_1a_{m1} \\ + & & + & & + \\ \vdots & & \vdots & \cdots & \vdots \\ + & & + & & + \\ x_na_{1n} & & x_na_{2n} & \cdots & x_na_{mn} \\ = & & = & & = \\ 0 & & 0 & \cdots & 0 \end{array}$$

das in der üblichen Form geschrieben die Gestalt

$$\begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & 0 \\ a_{21}x_1 & + & \cdots & + & a_{2n}x_n & = & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & + & \cdots & + & a_{mn}x_n & = & 0 \end{array}$$

annimmt. Da unser Gleichungssystem weniger Gleichungen hat als Unbekannte, liefert der Gauß-Algorithmus dafür mindestens eine von Null verschiedene Lösung. Für jede solche Lösung gilt aber

$$x_1\vec{v}_1 + \cdots + x_n\vec{v}_n = 0$$

im Widerspruch zur linearen Unabhängigkeit der $\vec{v}_i$. □

Zweiter Beweis. Unser Satz folgt auch sofort aus dem Austauschsatz 1.4.18, den wir im Anschluß formulieren und beweisen. □

Satz 1.4.18 (Austauschsatz von Steinitz). *Ist V ein Vektorraum, $E \subset V$ ein Erzeugendensystem und $L \subset V$ eine endliche linear unabhängige Teilmenge, so gibt es eine Injektion $\varphi : L \hookrightarrow E$ derart, daß auch $(E \backslash \varphi(L)) \cup L$ ein Erzeugendensystem von V ist.*

1.4.19. Wir können in anderen Worten die Vektoren unserer linear unabhängigen Teilmenge so in unser Erzeugendensystem hineintauschen, daß es ein Erzeugendensystem bleibt. Mit raffinierteren Methoden der Mengenlehre kann unser Austauschsatz auch ohne die Voraussetzung L endlich gezeigt werden. Der Beweis in dieser Allgemeinheit wird in **??** skizziert.

Beweis. Sei $M \subset L$ eine maximale Teilmenge von L, für die es eine Injektion $\varphi : M \hookrightarrow E$ gibt mit der Eigenschaft, daß auch $(E \backslash \varphi(M)) \cup M$ ein Erzeugendensystem von V ist. Es gilt zu zeigen $M = L$. Sei sonst $\vec{w} \in L \setminus M$. Schreiben wir $\vec{w}$ als eine Linearkombination von Vektoren aus $(E \backslash \varphi(M)) \cup M$, genauer als

$$\vec{w} = \lambda_1 \vec{v}_1 + \ldots + \lambda_r \vec{v}_r + \mu_1 \vec{w}_1 + \ldots + \mu_s \vec{w}_s$$

mit $\vec{v}_1, \ldots, \vec{v}_r \in E \backslash \varphi(M)$ paarweise verschieden und $\vec{w}_1, \ldots, \vec{w}_s \in M$ paarweise verschieden, so muß in dieser Linearkombination mindestens ein Vektor $\vec{v}_i \in E \backslash \varphi(M)$ mit einem von Null verschiedenen Koeffizienten $\lambda_i \neq 0$ auftreten, da ja L linear unabhängig war und $\vec{w}$ auf die andere Seite gebracht mit dem Koeffizienten $(-1) \neq 0$ auftritt. Dann erhält man jedoch wieder ein Erzeugendensystem, wenn man zusätzlich diesen Vektor $\vec{v}_i$ durch unser $\vec{w}$ austauscht, denn es gilt ja

$$\vec{v}_i = -\lambda_i^{-1}\lambda_1 \vec{v}_1 - \ldots - \widehat{\lambda_i^{-1}\lambda_i \vec{v}_i} - \ldots - \lambda_i^{-1}\lambda_r \vec{v}_r - \lambda_i^{-1}\mu_1 \vec{w}_1 - \ldots - \lambda_i^{-1}\mu_s \vec{w}_s + \lambda_i^{-1}\vec{w}$$

mit einem Hut über dem wegzulassenden Summanden und damit liegt liegt $\vec{v}_i$ im Erzeugnis von $(E \backslash (\varphi(M) \cup \{\vec{v}_i\})) \cup M \cup \{\vec{w}\}$, das folglich immer noch der ganze Raum ist. Widerspruch zur Maximalität von M! □

Korollar 1.4.20 (Kardinalitäten von Basen). *Jeder endlich erzeugte Vektorraum besitzt eine endliche Basis, und je zwei seiner Basen haben gleich viele Elemente.*

1.4.21. In **??** wird mit raffinierteren Methoden der Mengenlehre gezeigt, daß es auch im Fall eines nicht notwendig endlich erzeugten Vektorraums für je zwei seiner Basen eine Bijektion zwischen der einen Basis und der anderen Basis gibt.

Beweis. Wie bereits in 1.4.15 erwähnt erhalten wir nach 1.4.12 eine endliche Basis, wenn wir ein beliebiges endliches Erzeugendensystem durch das Streichen von Vektoren zu einem minimalen Erzeugendensystem verkleinern. Gegeben zwei Basen B und B' eines Vektorraums haben wir nach 1.4.16 außerdem $|B| \leq |B'| \leq |B|$. □

Definition 1.4.22. Die Kardinalität einer und jeder Basis eines endlich erzeugten Vektorraums V heißt die **Dimension** von V und wird $\dim V$ notiert. Ist k ein Körper und wollen wir betonen, daß wir die Dimension als k-Vektorraum meinen, so schreiben wir

$$\dim V = \dim_k V$$

Ist der Vektorraum nicht endlich erzeugt, so schreiben wir $\dim V = \infty$ und nennen V **unendlichdimensional** ignorieren für gewöhnlich die durchaus möglichen feineren Unterscheidungen zwischen verschiedenen Unendlichkeiten.

1.4.23. Der Nullraum hat als Basis die leere Menge. Seine Dimension ist folglich Null. Allgemeiner haben wir nach 1.4.9 offensichtlich

$$\dim_k k^n = n$$

Korollar 1.4.24. *Sei V ein endlich erzeugter Vektorraum.*

1. *Jede linear unabhängige Teilmenge $L \subset V$ hat höchstens $\dim V$ Elemente und im Fall $|L| = \dim V$ ist L bereits eine Basis.*
2. *Jedes Erzeugendensystem $E \subset V$ hat mindestens $\dim V$ Elemente und im Fall $|E| = \dim V$ ist E bereits eine Basis.*

Beweis. Nach 1.4.16 haben wir für L eine linear unabhängige Teilmenge, B eine Basis und E ein Erzeugendensystem stets

$$|L| \leq |B| \leq |E|$$

Gibt es ein endliches Erzeugendensystem, so muß im Fall $|L| = |B|$ mithin L eine maximale linear unabhängige Teilmenge und damit nach 1.4.12 eine Basis sein. Im Fall $|B| = |E|$ muß E in derselben Weise ein minimales Erzeugendensystem und damit nach 1.4.12 eine Basis sein. □

Korollar 1.4.25. *Ein echter Untervektorraum eines endlichdimensionalen Vektorraums hat stets eine echt kleinere Dimension. Ist allgemeiner und in Formeln $U \subset V$ ein Untervektorraum, so gilt $\dim U \leq \dim V$ und aus $\dim U = \dim V < \infty$ folgt $U = V$.*

Beweis. Ist V nicht endlich erzeugt, so ist nichts zu zeigen. Ist V endlich erzeugt, so gibt es nach 1.4.24 in U eine maximale linear unabhängige Teilmenge, und jede derartige Teilmenge hat höchstens $\dim V$ Elemente. Jede derartige Teilmenge ist aber nach 1.4.12 notwendig eine Basis von U und das zeigt $\dim U \leq \dim V$. Gilt hier Gleichheit und ist V endlichdimensional, so ist wieder nach 1.4.24 jede Basis von U auch eine Basis von V und das zeigt $U = V$. □

Satz 1.4.26 (Dimensionssatz). *Gegeben ein Vektorraum V und Teilräume $U, W \subset V$ mit endlichdimensionalem Schnitt gilt*

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$$

Beweis. Ist U oder W unendlichdimensional, so ist das eh klar. Sonst wählen wir eine Basis $s_1, \ldots, s_d$ von $U \cap W$ und ergänzen sie erst durch $u_1, \ldots, u_r \in U$ zu einer Basis von U und dann weiter durch $w_1, \ldots, w_t \in W$ zu einer Basis von $U+W$. Wir haben gewonnen, wenn wir zeigen können, daß bei derartigen Wahlen bereits $s_1, \ldots, s_d, w_1, \ldots, w_t$ eine Basis von W ist. Dazu reicht es zu zeigen, daß diese Menge W erzeugt. Sicher können wir jedes $w \in W$ schreiben als Linearkombination

$$\begin{aligned} w \quad = \quad & \lambda_1 u_1 + \ldots + \lambda_r u_r \\ & + \mu_1 s_1 + \ldots + \mu_d s_d \\ & + \nu_1 w_1 + \ldots + \nu_t w_t \end{aligned}$$

Dabei gilt jedoch offensichtlich $\lambda_1 u_1 + \ldots + \lambda_r u_r \in W \cap U$. Dieser Ausdruck läßt sich damit auch als Linearkombination der s_i schreiben, so daß w selbst auch als Linearkombination der s_i und der w_j geschrieben werden kann, was zu zeigen war. Im übrigen muß dann auch bei der obigen Darstellung bereits gelten $\lambda_1 = \ldots = \lambda_r = 0$, aber das ist für unseren Beweis schon gar nicht mehr von Belang. □

Übung 1.4.27. Eine Gruppe, in der jedes Element sein eigenes Inverses ist, kann auf genau eine Weise mit der Struktur eines Vektorraums über dem Körper mit zwei Elementen versehen werden, und ihre Untergruppen sind dann genau die Untervektorräume.

Übung 1.4.28. Gegeben k-Vektorräume $V_1, \ldots, V_n$ zeige man für die Dimension ihres kartesischen Produkts im Sinne von **??** die Formel

$$\dim(V_1 \oplus \ldots \oplus V_n) = \dim(V_1) + \ldots + \dim(V_n)$$

Satz 1.4.29 (Basisergänzungssatz). *Ist in einem endlich erzeugten Vektorraum L eine linear unabhängige Teilmenge und E ein Erzeugendensystem, so läßt sich L durch Hinzunahme von Vektoren aus E zu einer Basis unseres Vektorraums ergänzen.*

Beweis. Nach 1.4.14 ist jede linear unabhängige Teilmenge B unseres Vektorraums, die maximal ist unter allen linear unabhängigen Teilmengen A mit $L \subset A \subset (L \cup E)$, bereits eine Basis. Nach 1.4.16 gibt es auch tatsächlich maximale Teilmengen B mit dieser Eigenschaft. □

1.4.30. Der Basisergänzungssatz gilt unverändert auch für nicht notwendig endlich erzeugte Vektorräume. Der Beweis in dieser Allgemeinheit wird in **??** gegeben. Er verwendet raffiniertere Methoden der Mengenlehre und paßt schlecht in eine Grundvorlesung. Um dennoch einige der im folgenden bewiesenen Sätze durch unnötig einschränkende Endlichkeitsannahmen zu verkomplizieren, werde ich für die Zwecke dieser Vorlesung den Basisergänzungssatz für nicht notwendig endlich erzeugte Vektorräume ohne Beweis hinnehmen. Damit Sie nicht den Überlick verlieren, was nun im allgemeinen und was nur für endlich erzeugte Vektorräume vollständig bewiesen ist, werde ich den Basisergänzungssatz in dieser Allgemeinheit sozusagen als Axiom behandeln. Ich verspreche, daß es das einzige Axiom dieser Art bleiben soll, so daß wir zumindest mit klaren Spielregeln weiterarbeiten. Ich werde auch mein möglichstes tun, an jeder Stelle klarzumachen, welche der im folgenden bewiesenen Aussagen im unendlichdimensionalen Fall auf der allgemeinen Form des Basisergänzungssatzes beruhen. Ein erstes Beispiel ist der bereits erwähnte Basisexistenzsatz.

Satz 1.4.31 (Basisexistenzsatz). *Jeder Vektorraum besitzt eine Basis.*

Beweis. Die leere Menge ist stets linear unabhängig, der ganze Vektorraum ist stets ein Erzeugendensystem. Nun wende man den allgemeinen Basisergänzungssatz 1.4.30 an. □

1.5 Lineare Abbildungen

Definition 1.5.1. Seien V, W Vektorräume über einem Körper k. Eine Abbildung $f : V \to W$ heißt **linear** oder genauer **k-linear** oder ein **Homomorphismus von k-Vektorräumen** genau dann, wenn für alle $\vec{v}, \vec{w} \in V$ und $\lambda \in k$ gilt

$$\begin{aligned} f(\vec{v} + \vec{w}) &= f(\vec{v}) + f(\vec{w}) \\ f(\lambda \vec{v}) &= \lambda f(\vec{v}) \end{aligned}$$

Eine bijektive lineare Abbildung heißt ein **Isomorphismus** von Vektorräumen. Gibt es zwischen zwei Vektorräumen einen Isomorphismus, so heißen sie **isomorph**. Ein Homomorphismus von einem Vektorraum in sich selber heißt ein **Endomorphismus** unseres Vektorraums. Ein Isomorphismus von einem Vektorraum in sich selber heißt ein **Automorphismus** unseres Vektorraums.

Beispiel 1.5.2. Die Projektionen auf die Faktoren $\mathrm{pr}_i : k^n \to k$ sind linear. Das Quadrieren $k \to k$ ist nicht linear, es sei denn, k ist ein Körper mit zwei Elementen.

Beispiel 1.5.3. Gegeben Vektorräume V, W sind die Projektionsabbildungen $\mathrm{pr}_V : (V \oplus W) \to V$ und $\mathrm{pr}_W : (V \oplus W) \to W$ linear. Dasselbe gilt allgemeiner für die Projektionen $\mathrm{pr}_i : V_1 \oplus \ldots \oplus V_n \to V_i$. Ebenso sind die **kanonischen Injektionen** $\mathrm{in}_V : V \to (V \oplus W)$, $v \mapsto (v, 0)$ und $\mathrm{in}_W : W \to (V \oplus W)$, $w \mapsto (0, w)$ linear und dasselbe gilt allgemeiner für die analog definierten Injektionen $\mathrm{in}_i : V_i \to V_1 \oplus \ldots \oplus V_n$.

Übung 1.5.4. Ist $f : V \to W$ ein Vektorraumisomorphismus, so ist auch die Umkehrabbildung $f^{-1} : W \to V$ ein Vektorraumisomorphismus. Insbesondere bilden die Automorphismen eines Vektorraums V eine Untergruppe seiner Permutationsgruppe. Sie heißt die **allgemeine lineare Gruppe** oder auch die **Automorphismengruppe** unseres Vektorraums V und wird notiert

$$\mathrm{GL}(V) = \mathrm{Aut}(V) \subset \mathrm{Ens}^\times(V)$$

nach der englischen Bezeichnung **general linear group**.

Übung 1.5.5. Jede Verknüpfung von Vektorraumhomomorphismen ist wieder ein Vektorraumhomomorphismus. Sind also in Formeln $f : V \to W$ und $g : U \to V$ Vektorraumhomomorphismen, so ist auch $f \circ g : U \to W$ ein Vektorraumhomomorphismus.

Übung 1.5.6. Das Bild eines Untervektorraums unter einer linearen Abbildung ist ein Untervektorraum. Das Urbild eines Untervektorraums unter einer linearen Abbildung ist ein Untervektorraum.

Satz 1.5.7. *Ein Vektorraum über einem Körper k ist genau dann isomorph zu k^n, wenn er die Dimension n hat.*

Beweis. Natürlich gehen unter einem Vektorraumisomorphismus Erzeugendensysteme in Erzeugendensysteme, linear unabhängige Teilmengen in linear unabhängige Teilmengen und Basen in Basen über. Sind also zwei Vektorräume isomorph, so haben sie auch dieselbe Dimension. Hat umgekehrt ein Vektorraum V eine angeordnete Basis $B = (\vec{v}_1, \ldots, \vec{v}_n)$ aus n Vektoren, so liefert die Vorschrift $(\lambda_1, \ldots, \lambda_n) \mapsto \lambda_1\vec{v}_1 + \ldots + \lambda_n\vec{v}_n$ etwa nach 1.4.11 einen Vektorraumisomorphismus $k^n \xrightarrow{\sim} V$. □

1.5.8. Nun können wir auch unsere Ausgangsfrage 1.1.9 lösen, ob die "Zahl der freien Parameter" bei unserer Darstellung der Lösungsmenge eines linearen Gleichungssystems eigentlich wohlbestimmt ist oder präziser, ob beim Anwenden des Gauss-Algorithmus dieselbe Zahl von Stufen entsteht, wenn

wir zuvor die Variablen umnummerieren alias die Spalten vertauschen. Wenn wir das für homogene Systeme zeigen können, folgt es offensichtlich für beliebige Systeme. Bei homogenen Systemen ist jedoch die Lösungsmenge $L \subset k^m$ ein Untervektorraum und wir erhalten einen Vektorraumisomorphismus $L \xrightarrow{\sim} k^{m-r}$ durch "Streichen aller Einträge, bei denen eine neue Stufe beginnt", also durch Weglassen von $x_{s(1)}, x_{s(2)}, \ldots, x_{s(r)}$ aus einem m-Tupel $(x_1, \ldots, x_m) \in L$. Damit erhalten wir für die Zahl r der Stufen die von allen Wahlen unabhängige Beschreibung als Zahl der Variablen abzüglich der Dimension des Lösungsraums, in Formeln $r = m - \dim_k L$.

1.5.9. Seien V, W Vektorräume über einem Körper k. Die Menge aller Homomorphismen von V nach W notieren wir

$$\mathrm{Hom}_k(V, W) = \mathrm{Hom}(V, W) \subset \mathrm{Ens}(V, W)$$

Lemma 1.5.10. *Seien V, W Vektorräume über einem Körper k und $B \subset V$ eine Basis. So liefert das Einschränken von Abbildungen eine Bijektion*

$$\mathrm{Hom}_k(V, W) \xrightarrow{\sim} \mathrm{Ens}(B, W)$$

Jede lineare Abbildung ist also in Worten festgelegt und festlegbar durch ihre Werte auf einer Basis.

Erster Beweis. Seien $f, g : V \to W$ linear. Gilt $f(\vec{v}) = g(\vec{v})$ für alle $\vec{v} \in B$, so folgt $f(\lambda_1\vec{v}_1 + \ldots + \lambda_r\vec{v}_r) = g(\lambda_1\vec{v}_1 + \ldots + \lambda_r\vec{v}_r)$ für alle $\lambda_1, \ldots, \lambda_r \in k$ und $\vec{v}_1, \ldots, \vec{v}_r \in B$ und damit $f(\vec{v}) = g(\vec{v})$ für alle $\vec{v}$ im Erzeugnis von B alias für alle $\vec{v} \in V$. Das zeigt die Injektivität der im Lemma betrachteten Einschränkungsabbildung. Ist umgekehrt eine Abbildung von Mengen $g : B \to W$ gegeben, so können wir sie zu einer linearen Abbildung $\tilde{g} : V \to W$ ausdehnen wie folgt: Jeder Vektor $\vec{v} \in V$ läßt sich ja nach 1.4.10 eindeutig als Linearkombination der Basisvektoren schreiben, etwa $\vec{v} = \lambda_1\vec{v}_1 + \ldots + \lambda_r\vec{v}_r$ mit paarweise verschiedenen $\vec{v}_i \in B$, und wenn wir nun schlicht

$$\tilde{g}(\vec{v}) = \lambda_1 g(\vec{v}_1) + \ldots + \lambda_r g(\vec{v}_r)$$

setzen, so erhalten wir die gesuchte lineare Ausdehnung von g. □

Übung 1.5.11. Man zeige, daß $\mathrm{Hom}_k(V, W)$ ein Untervektorraum der Menge aller Abbildungen $\mathrm{Ens}(V, W)$ von V nach W mit ihrer Vektorraumstruktur aus 1.3.17 ist. Man zeige für die Dimension von $\mathrm{Hom}_k(V, W)$ unter der Konvention $0 \cdot \infty = \infty \cdot 0 = 0$ die Formel

$$\dim \mathrm{Hom}_k(V, W) = (\dim V)(\dim W)$$

Diese Formel ist nur insofern mit Vorsicht zu genießen, da sie bei einer feineren Interpretation der Dimension als Kardinalität ihre Gültigkeit verliert. Hinweis: 1.5.10.

1.5.12. Der vorstehende Beweis befriedigt mich nicht vollständig, da wir darin doch einiges nicht ganz sauber ausgeführt haben: So muß eigentlich die Eindeutigkeit der Darstellung als Linearkombination von Basisvektoren sorgfältiger formuliert werden, und die Linearität unserer Ausdehnung $\tilde{g}$ könnte auch noch eine sorgfältigere Argumentation vertragen. Wir wiederholen deshalb unsere Argumentation nocheinmal in einer abstrakteren Sprache.

1.5.13. Wir erinnern nun an Begriff des freien k-Vektorraums kI über einer Menge I und erklären die **kanonische Einbettung** $\mathrm{can} : I \hookrightarrow kI$ dadurch, daß sie jedem Punkt $i \in I$ die charakteristische Funktion der einelementigen Menge $\{i\}$ zuordnen soll. In anderen Worten ist also $\mathrm{can}(i)$ die formale Linearkombination von Elementen von I, in der i selbst mit Koeffizient Eins auftritt und alle anderen $j \in I$ mit Koeffizient Null. Offensichtlich ist das Bild $\mathrm{can}(I)$ von ganz I unter dieser kanonischen Einbettung eine Basis des freien Vektorraums über I.

Satz 1.5.14 (**Universelle Eigenschaft freier Vektorräume**). *Sei I eine Menge und k ein Körper und kI der freie Vektorraum über I und* can $: I \to kI$ *die kanonische Einbettung. So liefert für jeden k-Vektorraum W das Vorschalten von* can *eine Bijektion*

$$\mathrm{Hom}_k(kI, W) \xrightarrow{\sim} \mathrm{Ens}(I, W)$$

Beweis. Stimmen zwei lineare Abbildungen auf einer Teilmenge eines Vektorraums überein, so auch auf dem von dieser Teilmenge erzeugten Untervektorraum. Da nun das Bild von can den ganzen freien Vektorraum kI erzeugt, folgt für lineare Abbildungen $f, g : kI \to W$ aus $f \circ \mathrm{can} = g \circ \mathrm{can}$ bereits $f = g$ und die Injektivität unserer Bijektion in spe ist gezeigt. Ist andererseits irgendeine Abbildung $\varphi : I \to W$ gegeben, etwa $\varphi : i \mapsto \vec{w}_i$, so können wir die lineare Abbildung

$$\begin{array}{rccc} \tilde{\phi} & : & kI & \to & W \\ & & (a_i)_{i\in I} & \mapsto & \sum a_i \vec{w}_i \end{array}$$

bilden, für die per definitionem gilt $\tilde{\phi} \circ \mathrm{can} = \varphi$. Das zeigt dann auch die Surjektivität unserer Bijektion in spe. □

Zweiter Beweis von Lemma 1.5.10. Da $B \subset V$ eine Basis ist, liefert das Auswerten formaler Ausdrücke nach 1.4.11 einen Vektorraumisomorphismus $\Phi : kB \xrightarrow{\sim} V$. Per definitionem ist seine Komposition mit $\mathrm{can} : B \to kB$ schlicht die Einbettung $i : B \hookrightarrow V$. Wir erhalten nun Bijektionen

$$\mathrm{Hom}_k(V, W) \overset{\circ\Phi}{\xrightarrow{\sim}} \mathrm{Hom}_k(kB, W) \overset{\circ\,\mathrm{can}}{\xrightarrow{\sim}} \mathrm{Ens}(B, W)$$

da Φ ein Isomorphismus ist und wegen 1.5.14. Das Lemma folgt. □

Übung 1.5.15. Man zeige: Gegeben Vektorräume $V_1, \ldots, V_n, V$ und lineare Abbildungen $f_i : V_i \to V$ erhalten wir auch eine lineare Abbildung $f : V_1 \oplus \ldots \oplus V_n \to V$ durch die Vorschrift $f(v_1, \ldots, v_n) = f_1(v_1) + \ldots + f_n(v_n)$. Auf diese Weise ergibt sich sogar einen Isomorphismus

$$\mathrm{Hom}(V_1, V) \oplus \ldots \oplus \mathrm{Hom}(V_n, V) \xrightarrow{\sim} \mathrm{Hom}(V_1 \oplus \ldots \oplus V_n, V)$$

Übung 1.5.16. Man zeige: Gegeben Vektorräume $V_1, \ldots, V_n, V$ und lineare Abbildungen $g_i : V \to V_i$ erhalten wir auch eine lineare Abbildung $g : V \to V_1 \oplus \ldots \oplus V_n$ durch die Vorschrift $g(v) = (g_1(v), \ldots, g_n(v))$. Auf diese Weise ergibt sich sogar einen Isomorphismus

$$\mathrm{Hom}(V, V_1) \oplus \ldots \oplus \mathrm{Hom}(V, V_n) \xrightarrow{\sim} \mathrm{Hom}(V, V_1 \oplus \ldots \oplus V_n)$$

Definition 1.5.17. Untervektorräume U, W eines Vektorraums V heißen **komplementär** genau dann, wenn die Addition eine Bijektion $U \times W \xrightarrow{\sim} V$ liefert. Nach 1.5.15 ist diese Abbildung dann unter Verwendung der in **??** eingeführten Notation sogar ein Vektorraumisomorphismus $U \oplus W \xrightarrow{\sim} V$. Man schreibt in diesem Fall auch abkürzend $V = U \oplus W$, und sagt dann, der Vektorraum V sei die **direkte Summe** oder genauer die **innere direkte Summe** der Teilräume U und W. Ebenso kürzt man auch für Teilräume $V_1, \ldots, V_n \subset V$ die Aussage, daß die Addition einen Isomorphismus $V_1 \oplus \ldots \oplus V_n \xrightarrow{\sim} V$ liefert, ab mit

$$V = V_1 \oplus \ldots \oplus V_n$$

und sagt dann, der Vektorraum V sei die **direkte Summe** oder genauer die **innere direkte Summe** der Teilräume V_i.

Übung 1.5.18. Man zeige, daß es in einem endlichdimensionalen Vektorraum zu jedem Untervektorraum einen, ja im allgemeinen sogar verschiedene komplementäre Untervektorräume gibt. Mutige zeigen es auch für nicht notwendig endlichdimensionale Vektorräume. Das benötigt jedoch den Basisergänzungssatz in seiner vollen Allgemeinheit 1.4.30, in der wir ihn nicht bewiesen, sondern als Axiom hingenommen haben.

Proposition 1.5.19.

1. *Für jede injektive lineare Abbildung $f : V \hookrightarrow W$ existiert ein* **Linksinverses**, *als da heißt eine lineare Abbildung $g : W \to V$ mit $g \circ f = \mathrm{id}_V$.*

2. *Für jede surjektive lineare Abbildung $f : V \twoheadrightarrow W$ existiert ein* **Rechtsinverses**, *als da heißt eine lineare Abbildung $g : W \to V$ mit $f \circ g = \mathrm{id}_W$.*

Bemerkung 1.5.20. Einen unabhängigen Beweis noch allgemeinerer Aussagen diskutieren wir in **??**.

Beweis. Der Beweis beider Aussagen benötigt den Basisergänzungssatz 1.4.29 bzw. 1.4.30, den wir im unendlichdimensionalen Fall nicht bewiesen, sondern als Axiom hingenommen haben. Um Teil 1 zu zeigen, wählen wir mit 1.5.18 ein Komplement $U \subset W$ von $f(V)$ und definieren $g : W \to V$ durch die Vorschrift $g(u + f(v)) = v \quad \forall u \in U, v \in V$: Das ist erlaubt, da nach unsern Annahmen die Abbildung $(u, v) \mapsto u + f(v)$ eine Bijektion $U \times V \xrightarrow{\sim} W$ induziert. Um Teil 1 zu zeigen, wählen wir mithilfe des Basisexistenzsatzes 1.4.31 eine Basis $B \subset W$, finden $\tilde{g} : B \to V$ mit $f(\tilde{g}(b)) = b$ für alle $b \in B$ und erklären $g : W \to V$ als die eindeutig bestimmte lineare Abbildung mit $g(b) = \tilde{g}(b) \quad \forall b \in B$. Dann folgt $f(g(b)) = b \quad \forall b \in B$ und damit sofort $f(g(w)) = w \quad \forall w \in W$. □

Übung 1.5.21. Jede lineare Abbildung von einem Untervektorraum U eines Vektorraums V in einen weiteren Vektorraum $f : U \to W$ läßt sich zu einer linearen Abbildung $\tilde{f} : V \to W$ auf dem ganzen Raum fortsetzen. Hinweis: 1.5.19.

1.5.22. Die folgenden Übungen sind dazu gedacht, die Diskussion der Determinante und allgemeinerer multilinearer Abbildungen vorzubereiten. Sie stehen nur deshalb an dieser Stelle, da sie eben nicht mehr als die bis hierher erworbenen Kenntnisse voraussetzen.

Definition 1.5.23. Seien V, X, U Vektorräume über einem Körper k. Eine Abbildung $F : V \times X \to U$ heißt **bilinear** genau dann, wenn sie für jedes feste $v \in V$ linear ist in $x \in X$ und für jedes feste $x \in X$ linear in $v \in V$, in Formeln

$$\begin{aligned} F(av + bw, x) &= aF(v, x) + bF(w, x) \\ F(v, cx + dy) &= cF(v, x) + dF(v, y) \end{aligned}$$

für alle $a, b, c, d \in k$ und $v, w \in V$ und $x, y \in X$. Die Menge aller solchen bilinearen Abbildungen notieren wir

$$\mathrm{Hom}_k^{(2)}(V \times X, U) \subset \mathrm{Ens}(V \times X, U)$$

Diese Notation befriedigt mich unter formalen Aspekten nicht vollständig, da das Symbol $\times$ darin nicht als kartesisches Produkt, sondern vielmehr als ein Trenner aufzufassen ist. Ich habe sie dennoch gewählt in der Hoffnung, daß sie sich leichter merken und lesen läßt als eine formal vielleicht bessere Notation wie etwa $\mathrm{Hom}_k^{(2)}(V, X; U)$.

Übung 1.5.24. Seien U, V, W Vektorräume und $A \subset U$ sowie $B \subset V$ jeweils eine Basis. So liefert die Einschränkung eine Bijektion

$$\mathrm{Hom}_k^{(2)}(U \times V, W) \xrightarrow{\sim} \mathrm{Ens}(A \times B, W)$$

In Worten ist also eine bilineare Abbildung festgelegt und festlegbar durch ihre Werte auf Paaren von Basisvektoren. Hinweis: Man orientiere sich am Beweis von 1.5.10.

Übung 1.5.25. Man zeige, daß für je drei Vektorräume U, V, W die Verknüpfung von linearen Abbildungen $\mathrm{Hom}(U, V) \times \mathrm{Hom}(V, W) \to \mathrm{Hom}(U, W)$ bilinear ist. Hier sind unsere Homomorphismenräume zu verstehen mit ihrer in 1.5.11 erklärten Vektorraumstruktur.

Übung 1.5.26. Gegeben Vektorräume U, V, W induziert die kanonische Identifikation $\mathrm{Ens}(U \times V, W) \xrightarrow{\sim} \mathrm{Ens}(U, \mathrm{Ens}(V, W))$ aus I.2.2.23 einen Isomorphismus $\mathrm{Hom}^{(2)}(U \times V, W) \xrightarrow{\sim} \mathrm{Hom}(U, \mathrm{Hom}(V, W))$ zwischen dem Raum der bilinearen Abbildungen $U \times V \to W$ und dem Raum der linearen Abbildungen $U \to \mathrm{Hom}(V, W)$.

1.6 Dimensionsformel

Definition 1.6.1. Das Bild einer linearen Abbildung $f : V \to W$ notiert man auch

$$\mathrm{im}(f) := f(V)$$

für französisch und englisch "image". Es ist nach 1.5.6 ein Untervektorraum von W. Das Urbild des Nullvektors unter einer linearen Abbildung $f : V \to W$ notiert man auch

$$\ker(f) := f^{-1}(0) = \{v \in V \mid f(v) = 0\}$$

und nennt es den **Kern** der linearen Abbildung f. Dieser Kern ist nach 1.5.6 ein Untervektorraum von V.

Übung 1.6.2. Der Kern einer von Null verschiedenen linearen Abbildung in den Grundkörper ist stets eine Hyperebene im Sinne von 1.3.20.

Lemma 1.6.3. *Eine lineare Abbildung $f : V \to W$ ist injektiv genau dann, wenn ihr Kern Null ist.*

Beweis. Liegen im Kern außer dem Nullvektor von V noch andere Vektoren, so werden verschieden Vektoren aus V unter f auf den Nullvektor von W abgebildet und unsere Abbildung ist nicht injektiv. Ist umgekehrt unsere Abbildung nicht injektiv, so gibt es $v \neq v_1$ in V mit $f(v) = f(v_1)$ und es folgt $f(v - v_1) = 0$ aber $v - v_1 \neq 0$. Mit $v - v_1$ liegt also ein von Null verschiedener Vektor im Kern, der folglich nicht der Nullraum sein kann. □

Übung 1.6.4. Gegeben ein Vektorraum V haben wir eine Bijektion

$$\begin{array}{ccc} \{f \in \operatorname{End} V \mid f^2 = f\} & \xrightarrow{\sim} & \left\{ (I,K) \in \mathcal{P}(V)^2 \;\middle|\; \begin{array}{c} I \text{ und } K \text{ sind Teilräume} \\ \text{von } V \text{ mit } I \oplus K = V \end{array} \right\} \\ f & \mapsto & (\operatorname{im} f, \ker f) \end{array}$$

Ein Endomorphismus f eines Vektorraums mit der Eigenschaft $f^2 = f$ heißt auch **idempotent**. In Worten ausgedrückt entsprechen also die idempotenten Endomorphismen eines Vektorraums eineindeutig seinen Zerlegungen in eine direkte Summe von zwei komplementären Teilräumen. Die Umkehrabbildung würde man in Worten so beschreiben, daß sie einer Zerlegung $V = I \oplus K$ die **Projektion von V auf I längs K** zuordnet.

Übung 1.6.5. Gegeben eine lineare Abbildung $f : V \to W$ gilt für alle $v \in V$ die Identität $f^{-1}(f(v)) = v + \ker f$ von Teilmengen von V.

Definition 1.6.6. Eine Teilmenge T eines Vektorraums V heißt ein **affiner Teilraum** genau dann, wenn es einen Vektor $v \in V$ und einen Untervektorraum $U \subset V$ gibt mit $T = v + U$.

1.6.7. Ist $f : V \to W$ eine lineare Abbildung, so ist also für alle $w \in W$ die Faser $f^{-1}(w)$ entweder leer oder aber ein affiner Teilraum von V. Wir diskutieren in 1.9.1 affine Teilräume beliebiger "affiner Räume". Der hier definierte Begriff wird sich dann als ein Spezialfall erweisen.

Übung 1.6.8. Ist $f : V \to W$ eine lineare Abbildung, so ist für jeden affinen Teilraum $A \subset W$ sein Urbild $f^{-1}(A)$ entweder leer oder aber ein affiner Teilraum von V.

Übung 1.6.9. Sei $p : V \twoheadrightarrow W$ eine surjektive lineare Abbildung. Genau dann ist ein Teilraum $U \subset V$ komplementär zu $\ker p$, wenn p einen Isomorphismus $p : U \xrightarrow{\sim} W$ induziert.

Satz 1.6.10. *Für jede lineare Abbildung $f : V \to W$ von Vektorräumen gilt die* **Dimensionsformel**

$$\dim V = \dim(\ker f) + \dim(\operatorname{im} f)$$

Beweis. Ist V endlich erzeugt, so ist auch $(\operatorname{im} f)$ endlich erzeugt, da ja für jedes Erzeugendensystems $E \subset V$ sein Bild $f(E)$ ein Erzeugendensystem von $f(V) = \operatorname{im} f$ ist. Ebenso ist mit V auch $(\ker f)$ endlich erzeugt, nach dem Korollar 1.4.25 ist ja sogar jeder Untervektorraum eines endlich erzeugten Vektorraums endlich erzeugt. Gilt also umgekehrt $\dim(\ker f) = \infty$ oder $\dim(\operatorname{im} f) = \infty$, so folgt $\dim V = \infty$ und unser Satz gilt. Wir brauchen ihn

also nur noch in dem Fall zu zeigen, daß $(\ker f)$ und $(\operatorname{im} f)$ beide endlichdimensional sind. In diesem Fall folgt er aus dem anschließenden präziseren Lemma 1.6.11, das uns sogar sagt, wie wir aus Basen von Kern und Bild eine Basis von V gewinnen können. □

Lemma 1.6.11. *Sei $f : V \to W$ eine lineare Abbildung. Ist A eine Basis ihres Kerns, B eine Basis ihres Bildes und $g : B \to V$ eine Wahl von Urbildern unserer Basis des Bildes, so ist $g(B) \cup A$ eine Basis von V.*

Beweis. Gegeben $\vec{v} \in V$ haben wir $f(\vec{v}) = \lambda_1\vec{w}_1 + \ldots + \lambda_r\vec{w}_r$ mit $\vec{w}_i \in B$. Offensichtlich liegt dann $\vec{v} - \lambda_1 g(\vec{w}_1) - \ldots - \lambda_r g(\vec{w}_r)$ im Kern von f und so folgt, daß $g(B) \cup A$ ganz V erzeugt. Um die lineare Unabhängigkeit zu zeigen nehmen wir an, es gelte

$$\lambda_1 g(\vec{w}_1) + \ldots + \lambda_r g(\vec{w}_r) + \mu_1\vec{v}_1 + \ldots + \mu_s\vec{v}_s = 0$$

mit den $\vec{v}_i \in A$ und $\vec{w}_j \in B$ paarweise verschieden. Wenden wir f an, so folgt $\lambda_1\vec{w}_1 + \ldots + \lambda_r\vec{w}_r = 0$ und damit $\lambda_1 = \ldots = \lambda_r = 0$ wegen der linearen Unabhängigkeit der $\vec{w}_i$. Setzen wir diese Erkenntnis in die ursprüngliche Gleichung ein, so folgt weiter $\mu_1 = \ldots = \mu_s = 0$ wegen der linearen Unabhängigkeit der Vektoren $\vec{v}_j$. □

Übung 1.6.12. Man zeige: Zwei Untervektorräume U, W eines Vektorraums V sind komplementär genau dann, wenn gilt $V = U + W$ und $U \cap W = 0$.

Übung 1.6.13. Man zeige: Zwei Untervektorräume U, W eines endlichdimensionalen Vektorraums V sind komplementär genau dann, wenn gilt $V = U + W$ und $\dim U + \dim V \geq \dim V$. Hinweis: 1.4.28.

Definition 1.6.14. Ein Punkt, der unter einer Abbildung auf sich selbst abgebildet wird, heißt ein **Fixpunkt** besagter Abbildung. Gegeben eine Abbildung $f : X \to X$ notiert man die Menge ihrer Fixpunkte auch

$$X^f = \{x \in X \mid f(x) = x\}$$

Übung 1.6.15. Gegeben ein Vektorraum V und ein Endomorphismus $f \in \operatorname{End} V$ bildet die Menge der von f festgehaltenen Vektoren einen Untervektorraum $V^f \subset V$.

Übung 1.6.16. Sei $\varphi : V \to V$ ein Endomorphismus eines endlichdimensionalen Vektorraums. Man zeige $(\ker(\varphi^2) = \ker\varphi) \Leftrightarrow (V = \ker\varphi \oplus \operatorname{im}\varphi)$.

1.7 Lineare Abbildungen und Matrizen

1.7.1. Wir bezeichnen in diesem Abschnitt unseren Körper mit K statt wie bisher mit k, weil das kleine k andere Aufgaben übernehmen soll.

Satz 1.7.2 (**Lineare Abbildungen und Matrizen**). *Gegeben ein Körper K und natürliche Zahlen $m, n \in \mathbb{N}$ erhalten wir eine Bijektion zwischen Homomorphismen und Matrizen*

$$\begin{array}{ccc} \mathrm{Hom}_K(K^m, K^n) & \xrightarrow{\sim} & \mathrm{M}(n \times m; K) \\ f & \mapsto & M(f) \end{array}$$

indem wir die **darstellende Matrix** $M(f)$ *unserer linearen Abbildung f erklären als die Matrix mit den Bildern der Vektoren der Standardbasis des K^m in den Spalten, in Formeln*

$$M(f) = (f(\mathrm{e}_1)|f(\mathrm{e}_2)|\ldots|f(\mathrm{e}_m))$$

Beweis. Das folgt unmittelbar aus unserer Erkenntnis 1.5.10, daß eine lineare Abbildung festgelegt und festlegbar ist durch ihre Werte auf den Vektoren einer Basis. □

1.7.3. In 1.7.9 werden wir sehen, wie auch die Umkehrung unserer Bijektion $f \mapsto M(f)$ explizit beschrieben werden kann, indem wir Vektoren des K^n bzw. K^m als **Spaltenvektoren** auffassen, als da heißt, als Elemente der Matrizenräume $\mathrm{M}(n \times 1; K)$ bzw. $\mathrm{M}(m \times 1; K)$, und jeder Matrix A die durch das "Davormultiplizieren von A" im Sinne der Matrixmultiplikation 1.7.5 gegebene lineare Abbildung zuordnen.

Beispiel 1.7.4. Die Matrix der Identität auf K^n ist die **Einheitsmatrix**

$$M(\mathrm{id}) = I = I_n = \begin{pmatrix} 1 & & 0 \\ & 1 & \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

mit Einträgen $I_{ij} = \delta_{ij}$ in der unter der Bezeichnung **Kroneckerdelta** bekannten und allgemein gebräuchlichen Konvention

$$\delta_{ij} = \begin{cases} 1 & i = j; \\ 0 & \text{sonst.} \end{cases}$$

Ist allgemeiner $m \geq n$, so ist die Matrix des "Weglassens der überzähligen Koordinaten" $f : (x_1, \ldots, x_m) \mapsto (x_1, \ldots, x_n)$ gerade

$$M(f) = \begin{pmatrix} 1 & & 0 & & 0 \ldots 0 \\ & \ddots & & & \\ & & \ddots & & \\ 0 & & & 1 & 0 \ldots 0 \end{pmatrix}$$

Die Matrix des "Vertauschens der Koordinaten" $g : K^2 \to K^2$ schließlich ist

$$M(g) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Definition 1.7.5. Gegeben ein Körper K und $m, n, l \in \mathbb{N}$ definieren wir die **Matrixmultiplikation**, eine Abbildung

$$\begin{array}{ccc} \mathrm{M}(n \times m; K) \times \mathrm{M}(m \times l; K) & \to & \mathrm{M}(n \times l; K) \\ (A \quad , \quad B) & \mapsto & AB \end{array}$$

durch die Formel

$$(AB)_{ik} = \sum_{j=1}^{m} A_{ij} B_{jk}$$

die den Eintrag der Produktmatrix AB in der i-ten Zeile und k-ten Spalte durch die Einträge der Matrizen A und B ausdrückt. In Worten gilt es, jeweils den j-ten Eintrag der i-ten Zeile von A mit dem j-ten Eintrag der k-ten Spalte von B zu multiplizieren, und die Summe dieser m Produkte ist dann der Eintrag der Produktmatrix AB in der i-ten Zeile und k-ten Spalte. Manchmal schreiben wir die Produktmatrix auch ausführlicher $AB = A \circ B$. Den Ursprung dieser auf den ersten Blick vielleicht absonderlich anmutenden Definition und unserer leicht mit dem Verknüpfen von Abbildungen zu verwechselnden Notation erklärt der folgende Satz.

Satz 1.7.6. *Gegeben lineare Abbildungen $g : K^l \to K^m$ und $f : K^m \to K^n$ ist die Matrix ihrer Verknüpfung das Produkt der zugehörigen Matrizen, in Formeln*

$$M(f \circ g) = M(f) \circ M(g)$$

Beweis. Sei (a_{ij}) die Matrix $M(f)$ und (b_{jk}) die Matrix $M(g)$. Wir notieren die Standardbasen von K^n, K^m und K^l als $\vec{u}_i, \vec{v}_j$ und $\vec{w}_k$ in der Hoffnung, daß die folgende Rechnung dadurch transparenter wird, daß wir nicht für die

Standardbasis in allen drei Räumen die sonst eigentlich übliche Notation $\vec{e}_r$ verwenden. In unserer Notation haben wir also

$$\begin{aligned} g(\vec{w}_k) &= (b_{*k}) &= b_{1k}\vec{v}_1 + \ldots + b_{mk}\vec{v}_m \\ f(\vec{v}_j) &= (a_{*j}) &= a_{1j}\vec{u}_1 + \ldots + a_{nj}\vec{u}_n \end{aligned}$$

und folgern

$$\begin{aligned} (f \circ g)(\vec{w}_k) &= f(b_{1k}\vec{v}_1 + \ldots + b_{mk}\vec{v}_m) \\ &= b_{1k}f(\vec{v}_1) + \ldots + b_{mk}f(\vec{v}_m) \\ &= \textstyle\sum_{j=1}^m b_{jk} f(\vec{v}_j) \\ &= \textstyle\sum_{j=1}^m b_{jk} \sum_{i=1}^n a_{ij}\vec{u}_i \\ &= \textstyle\sum_{i=1}^n \left(\sum_{j=1}^m a_{ij}b_{jk} \right) \vec{u}_i \end{aligned}$$

Andererseits sind ja die Einträge (c_{ik}) der Matrix $M(f \circ g)$ gerade definiert durch die Identität $(f \circ g)(\vec{w}_k) = c_{1k}\vec{u}_1 + \ldots + c_{nk}\vec{u}_n$, und durch einen Koeffizientenvergleich folgt für die Einträge c_{ik} von $M(f \circ g)$ wie gewünscht $c_{ik} = \sum_{j=1}^m a_{ij}b_{jk}$. □

Proposition 1.7.7. *Für die Matrixmultiplikation gelten die folgenden Rechenregeln:*

$$\begin{aligned} (A + A')B &= AB + A'B \\ A(B + B') &= AB + AB' \\ IB &= B \\ AI &= A \\ (AB)C &= A(BC) \end{aligned}$$

für beliebige $k, l, m, n \in \mathbb{N}$ *und* $A, A' \in \mathrm{M}(n \times m; K)$, $B, B' \in \mathrm{M}(m \times l; K)$, $C \in \mathrm{M}(l \times k; K)$ *und* $I = I_m$ *die* $(m \times m)$*-Einheitsmatrix.*

Erster Beweis. Stures Rechnen, ich führe nur zwei Teile beispielhaft aus. Wir haben $(AI)_{ij} = \sum A_{ik}I_{kj} = \sum A_{ik}\delta_{kj} = A_{ij}$ und das zeigt $AI = A$. Für die nächste Rechnung verwende ich einmal andere Notationen und nehme $\kappa, \lambda, \mu, \nu$ als Laufindizes. Dann haben wir

$$\begin{aligned} ((AB)C)_{\nu\kappa} &= \textstyle\sum_{\lambda=1}^l (AB)_{\nu\lambda} C_{\lambda\kappa} \\ &= \textstyle\sum_{\lambda=1}^l \left(\sum_{\mu=1}^m A_{\nu\mu}B_{\mu\lambda} \right) C_{\lambda\kappa} \\ &= \textstyle\sum_{\lambda,\mu=1}^{l,m} A_{\nu\mu}B_{\mu\lambda}C_{\lambda\kappa} \\ (A(BC))_{\nu\kappa} &= \textstyle\sum_{\mu=1}^m A_{\nu\mu}(BC)_{\mu\kappa} \\ &= \textstyle\sum_{\mu=1}^m A_{\nu\mu} \left(\sum_{\lambda=1}^l B_{\mu\lambda}C_{\lambda\kappa} \right) \\ &= \textstyle\sum_{\mu,\lambda=1}^{m,l} A_{\nu\mu}B_{\mu\lambda}C_{\lambda\kappa} \end{aligned}$$

und das zeigt $(AB)C = A(BC)$. □

Zweiter Beweis. Wir können unsere Rechenregeln für Matrizen auch mit 1.7.2 und 1.7.6 auf die entsprechenden Regeln für lineare Abbildungen zurückführen. Um zum Beispiel $(AB)C = A(BC)$ zu zeigen, betrachten wir die linearen Abbildungen a, b, c mit den entsprechenden Matrizen im Sinne von 1.7.2, finden mit 1.7.6 sofort

$$\begin{aligned}(AB)C &= (M(a)M(b))M(c)\\ &= M(a\circ b)M(c)\\ &= M((a\circ b)\circ c)\end{aligned}$$

$$\begin{aligned}A(BC) &= M(a)(M(b)M(c))\\ &= M(a)M(b\circ c)\\ &= M(a\circ(b\circ c))\end{aligned}$$

und die Behauptung ergibt sich aus der für die Verknüpfung von Abbildungen offensichtlichen Identität $(a\circ b)\circ c = a\circ(b\circ c)$. □

Übung 1.7.8. Gegeben eine Matrix $A \in \mathrm{M}(n\times m; K)$ definiert man die **transponierte Matrix** $A^\top \in \mathrm{M}(m\times n; K)$ durch die Vorschrift

$$(A^\top)_{ij} = A_{ji}$$

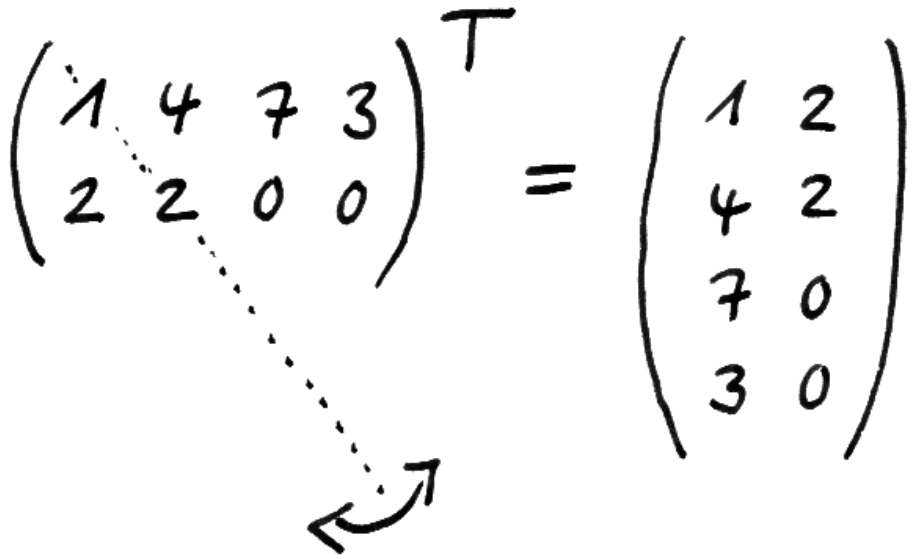
Anschaulich gesprochen entsteht also $A^\top$ aus A durch "Spiegeln an der Hauptdiagonalen". Zum Beispiel ist die Transponierte eines Spaltenvektors alias einer $(n\times 1)$-Matrix ein **Zeilenvektor** alias eine $(1\times n)$-Matrix. Natürlich gilt $(A^\top)^\top = A$. Man zeige $(AB)^\top = B^\top A^\top$.

1.7.9. Mit dem Formalismus der Matrixmultiplikation können wir auch die Umkehrung unserer Bijektion $\mathrm{Hom}_K(K^m, K^n) \xrightarrow{\sim} \mathrm{M}(n\times m; K)$, $f \mapsto M(f)$ aus 1.7.2 elegant beschreiben, indem wir die Elemente von K^m bzw. K^n als Spaltenvektoren auffassen und einer Matrix $A \in \mathrm{M}(n\times m; K)$ die durch Matrixmultiplikation gegebene Abbildung $(A\circ) : \mathrm{M}(m\times 1; K) \to \mathrm{M}(n\times 1; K)$ alias

$$(A\circ) : K^m \to K^n$$

zuordnen. Statt $A\circ x$ schreibt man dann einfacher auch schlicht Ax. Die Umkehrabbildung zu $f \mapsto M(f)$ kann mit diesen Konventionen also dargestellt werden in der Form $A \mapsto (x \mapsto Ax)$ für $x \in K^m$.

1.7.10. An dieser Stelle will ich kurz auf die Frage eingehen, ob denn Elemente eines K^m nun eigentlich Zeilenvektoren oder Spaltenvektoren sein sollen.



Die transponierte Matrix erhält man durch eine “Spiegelung an der Hauptdiagonalen”.

A priori sind Elemente eines K^m halt m-Tupel, und wie wir sie schreiben ist egal. Wenn wir eine Matrix davormultiplizieren wollen, ist es aber wichtig, unsere m-Tupel als Spaltenvektoren aufzufassen. Da das oft vorkommt, plädiere ich dafür, sich n-Tupel grundsätzlich als Spalten zu denken. Allerdings ist es in einen durchlaufenden Text sehr ungeschickt, Spaltenvektoren als solche zu schreiben. Da fügen sich Zeilenvektoren einfach viel besser ein, und wenn ich dennoch auf Spaltenvektoren bestehen will, schreibe ich sie im Text als "zu transponierende Zeilenvektoren", als da heißt, in der Form $(x_1, \ldots, x_m)^\top$. Oft schreibe ich aber auch einfach $(x_1, \ldots, x_m)$ und der Leser muß aus dem Kontext erschließen, was genau gemeint ist, wenn es denn darauf überhaupt ankommen sollte.

1.7.11. Eine Matrix A heißt **invertierbar** genau dann, wenn es weitere Matrizen B, C gibt mit $BA = I$ und $AC = I$. Das ist nach 1.7.6 gleichbedeutend dazu, daß die durch A gegebene lineare Abbildung $A : K^m \to K^n$ invertierbar alias ein Isomorphismus ist. Das ist nach 1.5.7 nur möglich für $n = m$, es können also nur **quadratische** Matrizen invertierbar sein. Für eine quadratische Matrix A sind des weiteren gleichbedeutend:

1. Es gibt eine quadratische Matrix B mit $BA = I$;
2. Es gibt eine quadratische Matrix C mit $AC = I$;
3. Die quadratische Matrix A ist invertierbar.

In der Tat folgt aus (1), daß die durch A gegebene lineare Abbildung injektiv ist, also ist sie bijektiv nach Dimensionsvergleich. Ebenso folgt aus (2), daß die durch A gegebene lineare Abbildung surjektiv ist, also ist sie bijektiv nach Dimensionsvergleich. Ist A invertierbar und $a : K^n \xrightarrow{\sim} K^n$ der zugehörige Vektorraumisomorphismus, so ist die Matrix $M(a^{-1})$ der Umkehrabbildung die einzige quadratische Matrix B mit $AB = I$ und auch die einzige quadratische Matrix B mit $BA = I$. Die invertierbaren $(n \times n)$-Matrizen sind insbesondere genau die invertierbaren Elemente des Monoids der $(n \times n)$-Matrizen mit der Matrixmultiplikation als Verknüpfung. Im Einklang mit unseren allgemeinen Konventionen für multiplikativ notierte Monoide notieren wir diese Matrix A^{-1} und nennen sie die **inverse Matrix zu** A. Die invertierbaren $(n \times n)$-Matrizen mit Einträgen in einem Körper K bilden mit der Matrixmultiplikation eine Gruppe, die **allgemeine lineare Gruppe der** $(n \times n)$**-Matrizen**, die man notiert als

$$\mathrm{M}(n \times n; K)^\times = \mathrm{GL}(n; K)$$

in Anlehnung an die englische Bezeichnung **general linear group**.

Übung 1.7.12. Die Automorphismengruppe eines zweidimensionalen Vektorraums über einem zweielementigen Körper ist isomorph zur Gruppe der Permutationen von drei Elementen, in Formeln $\mathrm{GL}(2;\mathbb{F}_2) \cong \mathcal{S}_3$.

1.7.13. Unser lineares Gleichungssystem

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + & \dots & + a_{1m}x_m = b_1 \\ a_{21}x_1 + a_{22}x_2 + & \dots & + a_{2m}x_m = b_2 \\ \vdots & & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + & \dots & + a_{nm}x_m = b_n \end{array}$$

können wir in unseren neuen Notationen zur Gleichung von Spaltenvektoren

$$Ax = b$$

abkürzen, wobei links das Produkt der Koeffizientenmatrix mit dem Spaltenvektor x gemeint ist. Gesucht ist das Urbild von $b \in K^n$ unter der linearen Abbildung $(A\circ) : K^m \to K^n$. Die Lösung des homogenisierten Systems ist genau der Kern dieser linearen Abbildung, und die Erkenntnis 1.1.7, nach der die allgemeine Lösung eines inhomogenen Systems die Summe einer speziellen Lösung des inhomogenen Systems mit einer allgemeinen Lösung des homogenisierten Systems ist, erweist sich als ein Spezialfall von 1.6.5. Die Operationen des Gauß-Algorithmus können wir in diesem Rahmen wie folgt interpretieren: Bezeichnet

$$E_{ij}$$

die **Basismatrix** mit dem Eintrag Eins in der i-ten Zeile und j-ten Spalte und Nullen sonst, so kann für $i \neq j$ das Gleichungssystem, das durch Addition des λ-fachen der j-ten Zeile zur i-ten Zeile entsteht, in Matrixschreibweise dargestellt werden als

$$(I + \lambda E_{ij})Ax = (I + \lambda E_{ij})b$$

Wegen $(I - \lambda E_{ij})(I + \lambda E_{ij}) = I$ hat es offensichtlich dieselbe Lösungsmenge wie das ursprüngliche System. Bezeichnet weiter P_{ij} für $i \neq j$ die Matrix zu der linearen Abbildung $K^m \xrightarrow{\sim} K^m$, die die i-te Koordinate mit der j-ten Koordinate vertauscht und sonst alles so läßt wie es ist, so kann das Gleichungssystem, das durch Vertauschen der i-ten Zeile mit der j-ten Zeile entsteht, in Matrixschreibweise dargestellt werden als

$$P_{ij}Ax = P_{ij}b$$

Wegen $P_{ij}P_{ij} = I$ hat es offensichtlich dieselbe Lösungsmenge wie das ursprüngliche System.

1.7.14. Unter einer **Elementarmatrix** verstehen wir eine quadratische Matrix, die sich in höchstens einem Eintrag von der Einheitsmatrix unterscheidet. Alle Elementarmatrizen mit Einträgen in einem Körper sind invertierbar mit Ausnahme der Matrizen, die entstehen, wenn man in der Einheitsmatrix eine Eins durch eine Null ersetzt.

Bemerkung 1.7.15. Es herrscht in der Literatur keine Einigkeit in der Frage, was genau unter einer Elementarmatrix zu verstehen sein soll. Manche Quellen bezeichnen zusätzlich zu unseren Elementarmatrizen auch noch die Permutationsmatrizen P_{ij} als Elementarmatrizen, andere Quellen hinwiederum lassen nur solche Matrizen zu, die sich von der Einheitsmatrix in höchstens einem Eintrag außerhalb der Diagonale unterscheiden.

Satz 1.7.16. *Jede quadratische Matrix mit Einträgen in einem Körper läßt sich als ein Produkt von Elementarmatrizen darstellen.*

Beweis. Zunächst einmal gilt das für die Permutationsmatrizen P_{ij}, die wir schreiben können als

$$P_{ij} = \operatorname{diag}(1, \ldots, 1, -1, 1, \ldots, 1)(I + E_{ij})(I - E_{ji})(I + E_{ij})$$

Hier soll die (-1) an der j-ten Stelle stehen und $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ meint die **Diagonalmatrix** mit Einträgen $a_{ij} = 0$ für $i \neq j$ und $a_{ii} = \lambda_i$. Nun beachten wir, daß das Inverse jeder invertierbaren Elementarmatrix wieder eine Elementarmatrix ist. Gegeben eine beliebige Matrix A finden wir nun nach dem Gauß-Algorithmus invertierbare Elementarmatrizen $S_1, \ldots, S_n$ derart, daß $S_n \ldots S_1 A$ Zeilenstufenform hat. Nun überzeugt man sich leicht, daß wir durch Daranmultiplizieren invertierbarer Elementarmatrizen von rechts alle Spaltenoperationen erhalten können, als da heißt, das Addieren des Vielfachen einer Spalte zu einer anderen, das Vertauschen zweier Spalten, sowie das Multiplizieren einer Spalte mit einem von Null verschiedenen Skalar. Wir können also weiter invertierbare Elementarmatrizen $T_1, \ldots, T_m$ finden derart, daß $S_n \ldots S_1 A T_1 \ldots T_m$ die Gestalt $\operatorname{diag}(1, \ldots, 1, 0, \ldots, 0)$ hat. Diese Matrix schreiben wir leicht als Produkt von nun nicht mehr invertierbaren diagonalen Elementarmatrizen $S_n \ldots S_1 A T_1 \ldots T_m = D_1 \ldots D_r$ und folgern

$$A = S_1^{-1} \ldots S_n^{-1} D_1 \ldots D_r T_m^{-1} \ldots T_1^{-1}$$ □

1.7.17. Eine Matrix, die nur auf der Diagonalen von Null verschiedene Einträge hat, und zwar erst einige Einsen und danach nur noch Nullen, nennen wir auch eine Matrix in **Smith-Normalform**.

Satz 1.7.18. *Für jede Matrix $A \in \mathrm{M}(n \times m; K)$ mit Einträgen in einem Körper K gibt es invertierbare Matrizen P, Q derart, daß PAQ eine Matrix in Smith-Normalform ist.*

Beweis. Wie beim Beweis von 1.7.16 finden wir nach dem Gauß-Algorithmus erst invertierbare Elementarmatrizen $S_1, \ldots, S_n$ derart, daß $S_n \ldots S_1 A$ Zeilenstufenform hat, und dann invertierbare Elementarmatrizen $T_1, \ldots, T_m$ derart, daß $S_n \ldots S_1 A T_1 \ldots T_m$ Smith-Normalform hat. □

Definition 1.7.19. Gegeben eine Matrix $A \in \mathrm{M}(n \times m; K)$ heißt die Dimension des von ihren Spaltenvektoren aufgespannten Untervektorraums von K^n der **Spaltenrang** unserer Matrix. Analog heißt die Dimension des von ihren Zeilenvektoren aufgespannten Untervektorraums von K^m der **Zeilenrang** unserer Matrix.

Satz 1.7.20. *Für jede Matrix stimmen Zeilenrang und Spaltenrang überein.*

1.7.21. Diese gemeinsame Zahl heißt dann der **Rang** unserer Matrix und wird $\operatorname{rk} A$ notiert nach der englischen Bezeichnung **rank**. Ist der Rang einer Matrix so groß wie für Matrizen derselben Gestalt möglich, sind also entweder die Spalten oder die Zeilen linear unabhängig, so sagt man, unsere Matrix habe **vollen Rang**.

Beweis. Der Spaltenrang einer Matrix $A \in \mathrm{M}(n \times m; K)$ kann interpretiert werden als die Dimension des Bildes von

$$(A\circ) : K^m \to K^n$$

Diese Interpretation zeigt sofort, daß PAQ denselben Spaltenrang hat wie A für beliebige invertierbare Matrizen P, Q. Durch Transponieren erkennen wir, daß PAQ auch denselben Zeilenrang hat wie A für beliebige invertierbare Matrizen P, Q. Nun finden wir jedoch nach 1.7.18 invertierbare Matrizen P, Q mit PAQ in Smith-Normalform. Dann stimmen natürlich Zeilenrang und Spaltenrang von PAQ überein, und dasselbe folgt für unsere ursprüngliche Matrix A. □

Definition 1.7.22. Ganz allgemein nennt man die Dimension des Bildes einer linearen Abbildung auch den **Rang** unserer linearen Abbildung. Dieser Rang kann unendlich sein, es gibt aber auch zwischen unendlichdimensionalen Vektorräumen durchaus von Null verschiedene Abbildungen endlichen Ranges.

Übung 1.7.23. Man gebe eine ganzzahlige (3×3)-Matrix vom Rang zwei ohne Eintrag Null an, bei der je zwei Spalten linear unabhängig sind.

Bemerkung 1.7.24. Um die Inverse einer $(n \times n)$-Matrix A zu berechnen, kann man wie folgt vorgehen: Man schreibt die Einheitsmatrix I daneben und wendet dann auf die $(n \times 2n)$-Matrix $(A|I)$ Zeilenoperationen an, einschließlich

des Multiplizierens einer Zeile mit einem von Null verschiedenene Skalar, bis man A erst in Zeilenstufenform und dann sogar zur Einheitsmatrix gemacht hat. Dann steht in der rechten Hälfte unserer $(n \times 2n)$-Matrix die Inverse zu A. In der Tat, sind unsere Zeilenumformungen etwa gegeben durch das Davormultiplizieren der Matrizen $S_1, S_2, \ldots, S_t$, so steht nach diesen Umformungen da

$$(S_t \ldots S_2 S_1 A | S_t \ldots S_2 S_1 I)$$

und wenn dann gilt $S_t \ldots S_2 S_1 A = I$, so folgt $S_t \ldots S_2 S_1 I = S_t \ldots S_2 S_1 = A^{-1}$. Dasselbe Verfahren funktioniert auch, wenn wir statt mit Zeilen- mit Spaltenumformungen arbeiten. Es ist nur nicht erlaubt, diese zu mischen, denn aus $S_t \ldots S_1 A T_1 \ldots T_r = I$ folgt noch lange nicht $S_t \ldots S_1 T_1 \ldots T_r = A^{-1}$.

1.7.25. Die im folgenden verwandten Notationen ${}_{\mathcal{B}}[v]$ und ${}_{\mathcal{A}}[f]_{\mathcal{B}}$ habe ich Urs Hartl abgeschaut. Ähnlich wie die geschickt gewählten Steckverbindungen, die man bei Computerzubehör gewohnt ist, sorgen sie auch hier dafür, daß man fast nichts mehr falsch zusammenstöpseln kann.

Satz 1.7.26 (**Lineare Abbildungen und Matrizen, Variante**). *Seien gegeben ein Körper k sowie k-Vektorräume V, W mit angeordneten Basen $\mathcal{A} = (\vec{v}_1, \ldots, \vec{v}_m)$ und $\mathcal{B} = (\vec{w}_1, \ldots, \vec{w}_n)$. Ordnen wir jeder linearen Abbildung $f : V \to W$ die* **darstellende Matrix** ${}_{\mathcal{B}}[f]_{\mathcal{A}}$ *zu mit Einträgen a_{ij} gegeben durch die Identitäten $f(\vec{v}_j) = a_{1j}\vec{w}_1 + \ldots + a_{nj}\vec{w}_n$, so erhalten wir eine Bijektion*

$$\begin{array}{ccc} \mathrm{Hom}_k(V,W) & \xrightarrow{\sim} & \mathrm{M}(n \times m; k) \\ f & \mapsto & {}_{\mathcal{B}}[f]_{\mathcal{A}} \end{array}$$

1.7.27. Wir nennen ${}_{\mathcal{B}}[f]_{\mathcal{A}}$ die Matrix der Abbildung f in Bezug auf die Basen $\mathcal{A}$ und $\mathcal{B}$. In Worten ausgedrückt stehen in ihren Spalten die Koordinaten der Bilder der Basis $\mathcal{A}$ des Ausgangsraums in Bezug auf die Basis $\mathcal{B}$ des Zielraums. Beliebt ist statt ${}_{\mathcal{B}}[f]_{\mathcal{A}}$ auch die alternative Notation $M_{\mathcal{B}}^{\mathcal{A}}(f)$.

Beweis. Wir könnten hier eine Variation unseres Beweises von 1.7.6 nochmal ausschreiben, aber stattdessen erinnern wir einfacher unsere Isomorphismen $\Phi_{\mathcal{A}} : k^m \xrightarrow{\sim} V$ und $\Phi_{\mathcal{B}} : k^n \xrightarrow{\sim} W$ und beachten die Identität ${}_{\mathcal{B}}[f]_{\mathcal{A}} = M(\Phi_{\mathcal{B}}^{-1} f \Phi_{\mathcal{A}})$, so daß wir unsere Abbildung schreiben können als die Komposition von Bijektionen

$$\begin{array}{ccccc} \mathrm{Hom}_k(V,W) & \xrightarrow{\sim} & \mathrm{Hom}_k(k^m,k^n) & \overset{M}{\xrightarrow{\sim}} & \mathrm{M}(n \times m; k) \\ f & \mapsto & \Phi_{\mathcal{B}}^{-1} f \Phi_{\mathcal{A}} & & \end{array}$$

□

Satz 1.7.28 (Darstellende Matrix einer Verknüpfung). *Gegeben ein Körper k und endlichdimensionale k-Vektorräume U, V, W mit angeordneten Basen $\mathcal{A}, \mathcal{B}, \mathcal{C}$ und lineare Abbildungen $f : U \to V$ und $g : V \to W$ gilt für die darstellenden Matrizen*

$$_{\mathcal{C}}[g \circ f]_{\mathcal{A}} = {}_{\mathcal{C}}[g]_{\mathcal{B}} \circ {}_{\mathcal{B}}[f]_{\mathcal{A}}$$

Beweis. Wir können die Behauptung nach Erinnern aller Notationen umschreiben zu $M(\Phi_{\mathcal{C}}^{-1} g f \Phi_{\mathcal{A}}) = M(\Phi_{\mathcal{C}}^{-1} g \Phi_{\mathcal{B}}) \circ M(\Phi_{\mathcal{B}}^{-1} f \Phi_{\mathcal{A}})$, und das folgt offensichtlich aus 1.7.6. □

Definition 1.7.29. Gegeben ein endlichdimensionaler Vektorraum V mit einer angeordneten Basis $\mathcal{A} = (\vec{v}_1, \ldots, \vec{v}_n)$ notieren wir die inverse Abbildung zu $\Phi_{\mathcal{A}} : k^n \xrightarrow{\sim} V$ in der Form $\vec{v} \mapsto {}_{\mathcal{A}}[\vec{v}]$.

Satz 1.7.30. *Gegeben endlichdimensionale Räume V, W mit angeordneten Basen $\mathcal{A}, \mathcal{B}$ und eine lineare Abbildung $f : V \to W$ gilt für jeden Vektor $v \in V$, wenn wir ${}_{\mathcal{A}}[v]$ und ${}_{\mathcal{B}}[f(v)]$ für die Zwecke der Matrixmultiplikation als Spaltenmatrizen auffassen, die Identität*

$${}_{\mathcal{B}}[f(v)] = {}_{\mathcal{B}}[f]_{\mathcal{A}} \circ {}_{\mathcal{A}}[v]$$

Beweis. Hier wird bei genauerer Betrachtung nur die Gleichheit von Spaltenvektoren $\Phi_{\mathcal{B}}^{-1}(f(v)) = M(\Phi_{\mathcal{B}}^{-1} f \Phi_{\mathcal{A}})(\Phi_{\mathcal{A}}^{-1} v)$ behauptet, die aus 1.7.9 folgt. □

1.7.31. Betrachtet man zu einem beliebigen Vektor $v \in V$ die lineare Abbildung $(\cdot v) : k \to V$, $\lambda \mapsto \lambda v$, und bezeichet mit (1) eben die angeordnete Basis (1) des k-Vektorraums k, so ergibt sich die Identität ${}_{\mathcal{A}}[v] = {}_{\mathcal{A}}[\cdot v]_{(1)}$. Wegen $(\cdot f(v)) = f \circ (\cdot v)$ können wir damit den vorhergehenden Satz 1.7.30 auffassen als den Spezialfall ${}_{\mathcal{B}}[\cdot f(v)]_{(1)} = {}_{\mathcal{B}}[f]_{\mathcal{A}} \circ {}_{\mathcal{A}}[\cdot v]_{(1)}$ von Satz 1.7.28 über die darstellende Matrix einer Verknüpfung.

Definition 1.7.32. Gegeben zwei angeordnete Basen $\mathcal{A} = (v_1, \ldots, v_n)$ und $\mathcal{B} = (w_1, \ldots, w_n)$ eines Vektorraums V nennt man die Matrix

$${}_{\mathcal{B}}[\mathrm{id}_V]_{\mathcal{A}}$$

auch die **Basiswechselmatrix**. Ihre Einträge a_{ij} werden per definitionem festgelegt durch die Gleichungen $w_j = \sum_{i=1}^{n} a_{ij} v_i$.

1.7.33. Offensichtlich ist ${}_{\mathcal{A}}[\mathrm{id}]_{\mathcal{A}} = I$ die Einheitsmatrix und nach 1.7.28 ist damit die Basiswechselmatrix ${}_{\mathcal{A}}[\mathrm{id}]_{\mathcal{B}}$ invers zur Basiswechselmatrix in der Gegenrichtung ${}_{\mathcal{B}}[\mathrm{id}]_{\mathcal{A}}$. Haben wir eine lineare Abbildung $f : V \to W$ und angeordnete Basen $\mathcal{A}, \mathcal{B}$ von V und angeordnete Basen $\mathcal{C}, \mathcal{D}$ von W, so folgt

aus 1.7.28 die Identität ${}_{\mathcal{D}}[f]_{\mathcal{B}} = {}_{\mathcal{D}}[\mathrm{id}_W]_{\mathcal{C}} \circ {}_{\mathcal{C}}[f]_{\mathcal{A}} \circ {}_{\mathcal{A}}[\mathrm{id}_V]_{\mathcal{B}}$. Sind noch spezieller $\mathcal{A}, \mathcal{B}$ zwei angeordnete Basen eines Vektorraums V und $f : V \to V$ ein Endomorphismus von V, so erhalten wir ${}_{\mathcal{B}}[f]_{\mathcal{B}} = {}_{\mathcal{B}}[\mathrm{id}]_{\mathcal{A}} \circ {}_{\mathcal{A}}[f]_{\mathcal{A}} \circ {}_{\mathcal{A}}[\mathrm{id}]_{\mathcal{B}}$ alias

$$N = T^{-1}MT$$

für $N = {}_{\mathcal{B}}[f]_{\mathcal{B}}$ und $M = {}_{\mathcal{A}}[f]_{\mathcal{A}}$ die darstellenden Matrizen sowie $T = {}_{\mathcal{A}}[\mathrm{id}]_{\mathcal{B}}$ die Basiswechselmatrix.

Übung 1.7.34. Gegeben ein K-Vektorraum V mit einer angeordneten Basis $\mathcal{A} = (v_1, \ldots, v_n)$ liefert die Zuordnung, die jeder weiteren angeordneten Basis $\mathcal{B}$ die Basiswechselmatrix von $\mathcal{A}$ nach $\mathcal{B}$ zuordnet, eine Bijektion

$$\begin{array}{ccc} \{\text{angeordnete Basen von } V\} & \xrightarrow{\sim} & \mathrm{GL}(n;K) \\ \mathcal{B} & \mapsto & {}_{\mathcal{B}}[\mathrm{id}]_{\mathcal{A}} \end{array}$$

Definition 1.7.35. Die **Spur** einer endlichen quadratischen Matrix ist definiert als die Summe ihrer Diagonaleinträge. Auf englisch und französisch sagt man **trace** und wir werden die Spur einer Matrix A notieren als

$$\mathrm{tr}(A)$$

Übung 1.7.36. Man zeige $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ wann immer A eine $(m \times n)$-Matrix ist und B eine $(n \times m)$-Matrix. Man folgere $\mathrm{tr}(BAB^{-1}) = \mathrm{tr}(A)$ wann immer A eine $(n \times n)$-Matrix ist und B eine invertierbare $(n \times n)$-Matrix. Insbesondere kann man jedem Endomorphismus f eines endlichdimensionalen Vektorraums V seine **Spur**

$$\mathrm{tr}(f) = \mathrm{tr}(f|V)$$

zuordnen als die Spur seiner Matrix in Bezug auf eine und jede Basis. Gegeben endlichdimensionale Vektorräume V, W und lineare Abbildungen $f : V \to W$ und $g : W \to V$ zeige man auch $\mathrm{tr}(fg) = \mathrm{tr}(gf)$. Eine vielleicht natürlichere Definition der Spur wird in 7.3.3 erklärt. Im Rahmen der Analysis werden wir die Spur in **??** als das Differential der Determinante an der Einheitsmatrix wiedersehen.

Übung 1.7.37. Ist L ein endlichdimensionaler k-Vektorraum und $A : L \to L$ eine k-lineare Abbildung, so gilt

$$\mathrm{tr}((A\circ)|\operatorname{End}_k L) = (\dim_k L)\,\mathrm{tr}(A|L)$$

Definition 1.7.38. Sei f ein Endomorphismus eines Vektorraums V. Ist f von endlichem Rang, so erklärt man die **Spur** $\mathrm{tr}\, f = \mathrm{tr}(f|V)$ **von** f als die Spur der Verknüpfung $\operatorname{im} f \hookrightarrow V \twoheadrightarrow \operatorname{im} f$ im Sinne unserer Definition 1.7.36 für die Spur eines Endomorphismus eines endlichdimensionalen Vektorraums.

1.7.39. Aus 1.7.36 folgt unmittelbar, daß diese Definition im Fall eines endlichdimensionalen Raums V dieselbe Spur liefert wie unsere ursprüngliche auf den endlichdimensionalen Fall beschränkte Definition 1.7.35.

Übung 1.7.40. Sind V, W Vektorräume und $f : V \to W$ und $g : W \to V$ lineare Abbildungen und ist eine unserer Abbildungen von endlichem Rang, so gilt $\operatorname{tr}(fg) = \operatorname{tr}(gf)$. Hinweis: Der endlichdimensionale Fall kann nach 1.7.36 vorausgesetzt werden.

Satz 1.7.41 (Smith-Normalform). *Gegeben eine lineare Abbildung endlichdimensionaler Vektorräume $f : V \to W$ existieren stets angeordnete Basen $\mathcal{A}$ von V und $\mathcal{B}$ von W derart, daß die darstellende Matrix ${}_{\mathcal{B}}[f]_{\mathcal{A}}$ nur auf der Diagonale von Null verschiedene Einträge hat, und zwar erst einige Einsen und danach nur noch Nullen.*

Beweis. Das folgt sofort aus 1.6.11: Wir wählen zunächst eine angeordnete Basis $(w_1, \ldots, w_r)$ des Bildes von f, dazu Urbilder $v_1, \ldots, v_r$ in V, ergänzen diese durch eine angeordnete Basis des Kerns von f zu einer angeordneten Basis $\mathcal{A} = (v_1, \ldots, v_n)$ von V, und ergänzen unsere angeordnete Basis des Bildes zu einer angeordneten Basis $\mathcal{B} = (w_1, \ldots, w_m)$ von W. In diesen Basen hat f offensichtlich die behauptete Gestalt. □

Übung 1.7.42. Sei $f : V \to V$ ein **nilpotenter** Endomorphismus eines endlichdimensionalen Vektorraums, als da heißt, es gebe $d \in \mathbb{N}$ mit $f^d = 0$. Man zeige, daß unser Vektorraum eine angeordnete Basis $\mathcal{B}$ besitzt derart, daß die Matrix ${}_{\mathcal{B}}[f]_{\mathcal{B}}$ von f in Bezug auf diese Basis eine obere Dreiecksmatrix ist mit Nullen auf der Diagonalen. Hinweis: Man betrachte die Teilräume $\ker(f) \subset \ldots \subset \ker(f^{d-1}) \subset \ker(f^d) = V$, beginne mit einer Basis von $\ker(f)$ und ergänze sie sukzessive zu einer Basis von V. Eine stärkere Aussage in dieser Richtung werden wir als 5.5.2 zeigen.

1.8 Dualräume und transponierte Abbildungen

Definition 1.8.1. Gegeben ein Körper K und ein K-Vektorraum V nennt man eine lineare Abbildung $V \to K$ auch eine **Linearform auf** V. Die Menge aller solchen Linearformen bildet nach 1.5.11 einen Untervektorraum $\operatorname{Hom}_K(V, K) \subset \operatorname{Ens}(V, K)$. Man nennt diesen Vektorraum aller Linearformen den **Dualraum von** V. Wir verwenden dafür die beiden Notationen

$$V^* = V^\top = \operatorname{Hom}_K(V, K)$$

Üblich ist die Notation V^*. Im Zusammenhang mir darstellenden Matrizen und dergleichen schien mir jedoch die Notation als $V^\top$ suggestivere Formeln

zu liefern, weshalb ich in diesem Zusammenhang die sonst eher unübliche Notation $V^\top$ vorziehe.

1.8.2. Die Bezeichnung als "Form" für Abbildungen mit Werten im Grundkörper ist allgemein üblich: Wir kennen bis jetzt nur Linearformen, später werden noch Bilinearformen und quadratische Formen hinzukommen. Über die Herkunft dieser Bezeichnungsweise weiß ich wenig, vermutlich steckt derselbe Wortstamm wie bei dem Wort "Formel" dahinter.

Beispiel 1.8.3. Denken wir uns die Gesamtheit aller Zeitspannen als reellen Vektorraum, so können wir uns den Dualraum dieses Vektorraums denken als die Gesamtheit aller "Frequenzen" oder vielleicht besser aller möglichen "Drehgeschwindigkeiten von Drehungen um eine feste Achse". Zeichnen wir genauer einen Drehsinn als positiv aus, so entpräche eine Drehgeschwindigkeit der Linearform, die jeder Zeitspanne die Zahl der in dieser Zeitspanne erfolgten Umdrehungen zuordnet. Die zur Basis "Minute" der Gesamtheit aller Zeitspannen "duale Basis", die wir gleich in allgemeinen Dualräumen einführen werden, bestünde dann aus dem Vektor "eine Umdrehung pro Minute in positivem Drehsinn", den man üblicherweise U/min notiert.

Beispiel 1.8.4. Denkt man sich den Raum der Richtungsvektoren des Anschauungsraums als reellen Vektorraum, so liefert jeder von Null verschiedene Vektor eine Linearform auf diesem Richtungsraum vermittels der Vorschrift "projeziere jeden weiteren Vektor orthogonal auf die Gerade durch den gegebenen Vektor und nimm die Zahl, mit der man den den gegebenen Vektor multiplizieren muß, um die Projektion zu erhalten". Diese Entsprechung hat nur den Nachteil, daß der doppelte Vektor die halbe Linearform liefert und daß überhaupt die Addition von Vektoren keineswegs der Addition von Linearformen entspricht. Wählt man eine feste Längeneinheit, so kann man den Raum der Linearformen auf dem Raum der Richtungsvektoren des Anschauungsraums identifizieren mit dem Raum der Richtungsvektoren selber, indem man jedem Vektor als Linearform dieselbe Linearform wie oben zuordnet, nur noch zusätzlich geteilt durch das Quadrat seiner Länge. In anderen Worten kann diese Linearform auch beschrieben werden als "beliebigem Vektor ordne zu Länge der Projektion mal Länge des gegebenen Vektors". Diese Identifikation ist dann ein Vektorraumisomorphismus, und es ist vielleicht die Möglichkeit dieser Identifikation, die es uns erschwert, eine allgemeine Vorstellung des Dualraums zu bilden. Sie benutzt jedoch die "euklidische Struktur" des Raums der Richtungsvektoren des Anschauungsraums, die das Reden über orthogonale Projektionen eigentlich erst ermöglicht und die wir in 3.1 mathematisch modellieren werden. Auf allgemeineren Vektorräumen stehen uns keine orthogonalen Projektionen zur Verfügung und der Dual-

raum kann dann nicht mehr in natürlicher Weise mit dem Ausgangsraum identifiziert werden.

1.8.5. Gegeben ein endlichdimensionaler Vektorraum stimmt seine Dimension etwa nach 1.5.11 mit der Dimension seines Dualraums überein, in Formeln

$$\dim V^\top = \dim V$$

Im Fall eines unendlichdimensionalen Vektorraums ist wieder nach 1.5.11 auch sein Dualraum unendlichdimensional, aber seine Dimension ist "noch unendlicher" als die Dimension des Ausgangsraums in einem Sinne, der in **??** präzisiert wird.

Übung 1.8.6. Sei k ein Körper und V ein k-Vektorraum. Eine endliche Familie von Linearformen $f_1, \ldots, f_n \in V^\top$ ist linear unabhängig genau dann, wenn sie eine Surjektion $(f_1, \ldots, f_n) : V \twoheadrightarrow k^n$ liefert.

Definition 1.8.7. Gegeben eine K-linare Abbildung $f : V \to W$ erklären wir die **duale** oder auch **transponierte** Abbildung

$$f^\top : W^\top \to V^\top$$

als das "Vorschalten von f", in Formeln $f^\top(\lambda) = \lambda \circ f : V \to K$ für jede Linearform $\lambda : W \to K$. Oft wird sie diese Abbildung auch $f^* : W^* \to V^*$ notiert.

1.8.8. Sicher gilt stets $\mathrm{id}_V^\top = \mathrm{id}_{V^\top} : V^\top \to V^\top$. Man prüft auch leicht für eine Veknüpfung $f \circ g$ von linearen Abbildungen die Identität

$$(f \circ g)^\top = g^\top \circ f^\top$$

In der Tat bedeutet das Vorschalten von $f \circ g$ nichts anderes, als erst f und dann g vorzuschalten.

Übung 1.8.9. Gegeben Vektorräume V, W liefern die transponierten Abbildungen zu den kanonischen Injektionen nach 1.5.3 auf den Dualräumen einen Isomorphismus $(\mathrm{in}_V^\top, \mathrm{in}_W^\top) : (V \oplus W)^\top \xrightarrow{\sim} V^\top \oplus W^\top$. Analoges gilt für allgemeinere endliche Summen.

1.8.10. Eine von Null verschiedene Linearform mag man sich veranschaulichen, indem man sich den affinen Teilraum vorstellt, auf dem sie den Wert Eins annimmt. In dieser Anschauung ist insbesondere für einen Automorphismus $f : \mathbb{R}^2 \xrightarrow{\sim} \mathbb{R}^2$ der Effekt des Inversen $(f^\top)^{-1}$ der transponierten Abbildung auf Linearformen gut verständlich.

1.8.11. Gegeben eine Basis $B \subset V$ erhalten wir im Dualraum $V^\top$ eine linear unabhängige Familie von Linearformen

$$(b^\top)_{b\in B}$$

indem wir $b^\top : V \to K$ erklären durch $b^\top(c) = \delta_{bc} \quad \forall c \in B$. Die $b^\top$ heißen die **Koordinatenfunktionen** oder kurz **Koordinaten** zur Basis B. Vielfach werden sie auch b^* notiert. Ist etwa $V = \mathbb{R}^n$ und $B = (\vec{e}_1, \ldots, \vec{e}_n)$ die Standardbasis, so wird $\vec{e}_i^{\,\top} : \mathbb{R}^n \to \mathbb{R}$ die "Projektion auf die i-te Koordinate" $\vec{e}_i^{\,\top} : (x_1, \ldots, x_n) \mapsto x_i$, die man oft auch einfach $x_i : \mathbb{R}^n \to \mathbb{R}$ notiert und die "i-te Koordinatenfunktion" nennt. Man beachte, daß die Koordinatenfunktion $b^\top$ keineswegs nur vom Basisvektor b abhängt, auch wenn die Notation das suggerieren mag, sondern vielmehr von der ganzen Basis B.

1.8.12. Für jeden endlichdimensionalen Vektorraum V hat der Dualraum, wie etwa aus 1.5.11 folgt, dieselbe Dimension wie V selber. Ist also $\mathcal{B}$ eine angeordnete Basis von V, so ist $\mathcal{B}^\top = (b^\top)_{b\in\mathcal{B}}$ als linear unabhängige Familie der richtigen Kardinalität auch eine angeordnete Basis des Dualraums $V^\top$. Man nennt dann $\mathcal{B}^\top$ die **duale Basis zur Basis** $\mathcal{B}$.

Proposition 1.8.13 (**Matrix der dualen Abbildung**). *Gegeben eine lineare Abbildung $f : V \to W$ von endlichdimensionalen Vektorräumen mit angeordneten Basen $\mathcal{A}$ bzw. $\mathcal{B}$ ist die darstellende Matrix der dualen Abbildung $f^\top : W^\top \to V^\top$ bezüglich der dualen Basen gerade die transponierte Matrix, in Formeln*

$$_{\mathcal{A}^\top}[f^\top]_{\mathcal{B}^\top} = (_{\mathcal{B}}[f]_{\mathcal{A}})^\top$$

1.8.14. Diese Identität ist auch der Grund dafür, daß ich hier das Dualisieren mit einem hochgestellten $\top$ notiert habe.

Beweis. Seien etwa $\mathcal{A} = (v_1, \ldots, v_n)$ und $\mathcal{B} = (w_1, \ldots, w_n)$. Die Matrixeinträge a_{ij} der darstellenden Matrix $_{\mathcal{B}}[f]_{\mathcal{A}}$ sind festgelegt durch die Identität von Vektoren $f(v_j) = \sum_i a_{ij} w_i$. Die Matrixeinträge b_{ji} der darstellenden Matrix $_{\mathcal{A}^\top}[f^\top]_{\mathcal{B}^\top}$ sind festgelegt durch die Identität von Linearformen $f^\top(w_i^\top) = \sum_j b_{ji} v_j^\top$. Es gilt zu zeigen $b_{ji} = a_{ij}$. Um das zu sehen, werten wir diese Identität von Linearformen auf den Vektoren v_k aus und erhalten

$$b_{ki} = \sum_j b_{ji} v_j^\top(v_k) = (f^\top(w_i^\top))(v_k) = w_i^\top(f(v_k)) = w_i^\top\left(\sum_l a_{lk} w_l\right) = a_{ik}$$

was zu zeigen war. □

1.8.15. Sei V ein endlichdimensionaler Vektorraum mit einer angeordneten Basis $\mathcal{A}$. Gegeben ein Vektor $v \in V$ und eine Linearform $\lambda \in V^\top$ kann man den Wert der Linearform auf dem Vektor auch darstellen als das Matrixprodukt $\lambda(v) = {}_{\mathcal{A}^\top}[\lambda]^\top \circ {}_{\mathcal{A}}[v]$ der Zeilenmatrix ${}_{\mathcal{A}^\top}[\lambda]^\top$ mit der Spaltenmatrix ${}_{\mathcal{A}}[v]$. Ist in der Tat $\mathcal{A} = (v_1, \ldots, v_n)$ und $v = a_1v_1 + \ldots + a_nv_n$ und $\lambda = b_1v_1^\top + \ldots + b_nv_n^\top$, so finden wir unmittelbar $\lambda(v) = b_1a_1 + \ldots + b_na_n$. Vereinbaren wir zusätzlich die Notation ${}_{\mathcal{A}^\top}[\lambda]^\top = [\lambda]_{\mathcal{A}}$, so nimmt diese Formel die besonders einfache Gestalt

$$\lambda(v) = [\lambda]_{\mathcal{A}} \circ {}_{\mathcal{A}}[v]$$

an. Diese Notation ist auch deshalb vernünftig, weil ja bezüglich der Standardbasis (1) des Grundkörpers K per definitionem gilt

$${}_{(1)}[\lambda]_{\mathcal{A}} = [\lambda]_{\mathcal{A}}$$

Erinnern wir dann noch für $v \in V$ an die lineare Abbildung $(\cdot v) : K \to V$ mit $\alpha \mapsto \alpha v$ und unsere Identität ${}_{\mathcal{A}}[\cdot v]_{(1)} = {}_{\mathcal{A}}[v]$, so kann obige Formel interpretiert werden als der Spezialfall

$${}_{(1)}[\lambda \circ (\cdot v)]_{(1)} = {}_{(1)}[\lambda]_{\mathcal{A}} \circ {}_{\mathcal{A}}[\cdot v]_{(1)}$$

der allgemeinen Formel 1.7.28 für die Matrix der Verknüpfung zweier linearer Abbildungen.

1.8.16. Gegeben ein Vektorraumisomorphismus $f : V \xrightarrow{\sim} W$ ist die duale Abbildung ein Vektorraumisomorphismus $f^\top : W^\top \xrightarrow{\sim} V^\top$ und ihre Inverse ist ein Vektorraumisomorphismus $(f^\top)^{-1} : V^\top \xrightarrow{\sim} W^\top$. Dieser Isomorphismus leistet, was man sich anschaulich vielleicht am ehesten unter dem "Transport einer Linearform" vorstellt: Gegeben $v \in V$ und $\lambda \in V^\top$ nimmt $(f^\top)^{-1}(\lambda)$ auf $f(v)$ denselben Wert an wie λ auf v. Betrachten wir etwa die Scherung $f : \mathbb{R}^2 \xrightarrow{\sim} \mathbb{R}^2$, $(x, y) \mapsto (x + y, y)$ mit der Matrix $M(f) = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$und $f(\vec{e}_1) = \vec{e}_1, f(\vec{e}_2) = \vec{e}_1 + \vec{e}_2$. Offensichtlich bleibt die y-Koordinate eines Punktes unter solch einer Scherung unverändert, $(f^\top)^{-1}(\vec{e}_2^\top) = \vec{e}_2^\top$ und die x-Koordinate des Urbildpunkts entspricht der Differenz zwischen x-Koordinate und y-Koordinate des Bildpunkts, $(f^\top)^{-1}(\vec{e}_1^\top) = \vec{e}_1^\top - \vec{e}_2^\top$. Das entspricht auch unseren Formeln, nach denen $f^\top$ bezüglich der Basis $(\vec{e}_1^\top, \vec{e}_2^\top)$ dargestellt wird durch die transponierte Matrix $\left(\begin{smallmatrix} 1 & 0 \\ -1 & 1 \end{smallmatrix}\right)$, was genau die Formel $(f^\top)^{-1} : \vec{e}_1^\top \mapsto \vec{e}_1^\top - \vec{e}_2^\top$ und $(f^\top)^{-1} : \vec{e}_2^\top \mapsto \vec{e}_2^\top$ beinhaltet.

Definition 1.8.17. Sei K ein Körper und V ein K-Vektorraum. Der Dualraum des Dualraums von V heißt sein **Bidualraum** und wird notiert $(V^\top)^\top =$

$V^{\top\top}$ oder in der Literatur meist V^{**}. Wir erklären die **kanonische Einbettung in den Bidualraum**

$$\mathrm{can} = \mathrm{can}_V : V \hookrightarrow V^{\top\top}$$

als die Vorschrift, die jedem Vektor $v \in V$ die "durch das Auswerten auf besagtem Vektor gegebene Linearform auf dem Raum der Linearformen" zuordnet. In Formeln ist $\mathrm{can}(v) \in V^{\top\top}$ also definiert als die lineare Abbildung $\mathrm{can}(v) : V^\top \to K$, $\lambda \mapsto \lambda(v)$.

1.8.18. Die Injektivität der kanonischen Abbildung $V \to V^{\top\top}$ ergibt sich aus der Erkenntnis, daß es für jeden von Null verschiedenen Vektor $v \neq 0$ eine Linearform $\lambda \in V^\top$ gibt mit $\lambda(v) \neq 0$. Man kann das etwa zeigen, indem man den Satz 1.5.21 über die Fortsetzbarkeit linearer Abbildungen bemüht oder auch, indem man v zu einer Basis B von V ergänzt und dann $\lambda = v^\top$ wählt. Im Fall unendlichdimensionaler Räume brauchen wir jedoch in jedem Fall den Basiserweiterungssatz in seiner vollen Allgemeinheit 1.4.30, in der wir ihn nicht bewiesen, sondern als Axiom hingenommen haben. Man kann ohne die ihm zugrundeliegenden raffinierteren Methoden der Mengenlehre noch nicht einmal zeigen, daß es auf einem beliebigen von Null verschiedenen Vektorraum überhaupt irgendeine von Null verschiedene Linearform gibt.

1.8.19. Im Fall eines endlichdimensionalen Vektorraums V zeigt ein Dimensionsvergleich unmittelbar, daß die kanonische Einbettung einen Isomorphismus $V \xrightarrow{\sim} V^{\top\top}$ liefern muß. Manchmal wird diese Erkenntnis als Gleichung $V = V^{\top\top}$ geschrieben, aber das ist dann mit einigen Hintergedanken zu lesen, denn gleich sind diese beiden Mengen ja keineswegs.

1.8.20. Gegeben eine lineare Abbildung $f : V \to W$ kommutiert das Diagramm

$$\begin{array}{ccc} V & \xrightarrow{\mathrm{can}_V} & V^{\top\top} \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f^{\top\top}} \\ W & \xrightarrow{\mathrm{can}_W} & W^{\top\top} \end{array}$$

als da heißt, es gilt die Identität $\mathrm{can}_W \circ f = f^{\top\top} \circ \mathrm{can}_V$ von Abbildungen $V \to W^{\top\top}$. Um das zu sehen, muß man nur für alle $v \in V$ die Identität $f^{\top\top}(\mathrm{can}_V(v)) = \mathrm{can}_W(f(v))$ in $W^{\top\top}$ prüfen. Dazu gilt es zu zeigen, daß beide Seiten auf allen $\lambda \in W^\top$ denselben Wert annehmen, daß also gilt

$$(f^{\top\top}(\mathrm{can}_V(v)))(\lambda) = (\mathrm{can}_W(f(v)))(\lambda)$$

alias $((\mathrm{can}_V\, v) \circ f^\top)(\lambda) = \lambda(f(v))$ alias $(\mathrm{can}_V\, v)(\lambda \circ f) = \lambda(f(v))$, und das ist klar.

Übung 1.8.21. Für endlichdimensionale Vektorräume V ist die kanonische Einbettung aus Dimensionsgründen stets ein Isomorphismus $V \xrightarrow{\sim} V^{\top\top}$. Gegeben ein endlichdimensionaler Vektorraum V zeige man, daß unter der kanonischen Identifikation $\mathrm{can}_V : V \xrightarrow{\sim} V^{\top\top}$ jede Basis B ihrer Bidualen entspricht, in Formeln

$$\mathrm{can}_V(b) = (b^\top)^\top \quad \forall b \in B$$

1.9 Affine Räume

Definition 1.9.1. Ein **affiner Raum** oder kurz **Raum** über einem Körper k ist ein Tripel

$$E = (E, \vec{E}, a)$$

bestehend aus einer nichtleeren Menge E, einer abelschen Gruppe $\vec{E} \subset \mathrm{Ens}^\times E$ von Permutationen von E, von der man fordert, daß für alle $e \in E$ das Anwenden auf e eine Bijektion $\vec{E} \xrightarrow{\sim} E$ liefert, sowie einer Abbildung $a : k \times \vec{E} \to \vec{E}$, die die abelsche Gruppe $\vec{E}$ zu einem k-Vektorraum macht. Die Elemente von $\vec{E}$ heißen die **Translationen** oder **Richtungsvektoren** unseres affinen Raums und den Vektorraum $\vec{E}$ selbst nennen wir den **Richtungsraum** unseres affinen Raums E. Die Operation von k auf $\vec{E}$ mag man die **Reskalierung von Translationen** nennen. Unter der **Dimension** unseres affinen Raums verstehen wir die Dimension seines Richtungsraums. Das Resultat der Operation von $\vec{u} \in \vec{E}$ auf $e \in E$ notieren wir $\vec{u} + e := \vec{u}(e)$.

1.9.2. Hier entsteht leider ein Konflikt mit der Notation aus **??**, nach der mit Pfeilen versehene Mannigfaltigkeiten orientierte Mannigfaltigkeiten andeuten sollen. Was im Einzelfall jeweils gemeint ist, muß der Leser aus dem Kontext erschließen. Die leere Menge kann in meinen Konventionen nie ein affiner Raum sein. Es gibt hier jedoch auch andere Konventionen.

1.9.3. Ein affiner Raum hat die Dimension Null genau dann, wenn er aus einem einzigen Punkt besteht. Affine Räume der Dimensionen Eins bzw. Zwei heißen **affine Geraden** bzw. **affine Ebenen**.

1.9.4. Ist E ein affiner Raum, so liefert nach Annahme für jedes $e \in E$ das Anwenden der Richtungsvektoren auf besagten Punkt eine Bijektion $\vec{E} \xrightarrow{\sim} E$, $\vec{u} \mapsto \vec{u} + e$ und es gilt $\vec{0} + e = e$ sowie $\vec{u} + (\vec{v} + e) = (\vec{u} + \vec{v}) + e$ für alle $\vec{u}, \vec{v} \in \vec{E}$ und $e \in E$. Flapsig gesprochen ist also ein affiner Raum ein "Vektorraum, bei dem man den Ursprung vergessen hat". Gegeben $e, e' \in E$ definieren wir $e - e'$ als den Richtungsvektor $\vec{u} \in \vec{E}$ mit $e = \vec{u} + e'$. In Schulbüchern verwendet man auch oft Großbuchstaben $A, B, C, \ldots$ für die Punkte eines affinen Raums und notiert $\overrightarrow{AB}$ den Richtungsvektor, der A nach B schiebt und den wir hier $B - A$ schreiben.

Beispiel 1.9.5. Jeder Vektorraum ist in offensichtlicher Weise auch ein affiner Raum. Es scheint mir besonders sinnfällig, den uns umgebenden Raum mathematisch als einen dreidimensionalen reellen affinen Raum

$$\mathbb{E}$$

zu modellieren. Der Buchstabe $\mathbb{E}$ soll an das französische Wort "éspace" für "Raum" erinnern. Manche Punkte von dieses Raums können wir uns direkt als Kirchturmspitzen, Zimmerecken und dergleichen denken, die Übrigen gilt es sich vorzustellen. Wir ignorieren dabei, daß die Erde sich um sich selber dreht und dabei gleichzeitig um die Sonne rast, die sich hinwiederum mit unvorstellbarer Geschwindigkeit um das Zentrum der Milchstraße bewegt, und damit ist es auch noch nicht zu Ende. In diesem Sinne meinen wir mit dem "Anschauungsraum" den Raum der klassischen Mechanik. Den zu unserem Anschauungsraum gehörigen Richtungsraum denkt man sich dann als die Gesamtheit aller "Parallelverschiebungen des Raums". In 1.9.28 werden wir lernen, in welchem Sinne die Bedingung, daß unsere Sichtlinien gerade die "affinen Geraden" sein sollen, die Struktur des Anschauungsraums als reeller affiner Raum bereits eindeutig festlegt. Unser Modell des Anschauungsraums ist allerdings hier noch unvollständig und wird erst in 3.1 fertig werden, wo wir auch das Messen mit Zollstöcken mathematisch modellieren, oder noch genauer in 3.5.12, wo wir den Begriff der Orientierung diskutieren. Daß wir hier als Grundkörper den Körper der reellen Zahlen nehmen, hat analytische Gründe: Im Kern liegen sie darin, daß für diesen Körper der Zwischenwertsatz **??** gilt. Deshalb modellieren reelle Vektorräume, insbesondere wenn es später auch um Drehungen, Winkel im Bogenmaß und dergleichen gehen wird, unsere geometrischen Anschauung besser als etwa Vektorräume über den rationalen Zahlen oder allgemeineren Teilkörpern von $\mathbb{R}$.

Beispiel 1.9.6. In derselben Weise mag man sich auch die Schreibfläche einer sich in jeder Richtung ins Unendliche erstreckenden Tafel als einen zweidimensionalen reellen affinen Raum denken.

Beispiel 1.9.7. Die Menge aller **Zeitpunkte** der klassischen Mechanik mag man sich als einen eindimensionalen reellen affinen Raum

$$\mathbb{T}$$

denken. Der Buchstabe $\mathbb{T}$ soll an das lateinische Wort "tempus" für "Zeit" erinnern. Eine mögliche Translation in diesem Raum wäre etwa die Vorschrift: Man warte von einem vorgegebenen Zeitpunkt sieben Ausschläge eines bestimmten Pendels, dann erreicht man den um diese Translation verschobenen Zeitpunkt. Die Elemente des Richtungsraums $\vec{\mathbb{T}}$ dieses affinen Raums hätte

man sich als "Zeitspannen" zu denken, wobei jedoch auch "negative Zeitspannen" zuzulassen wären. Die Flugbahn einer Fliege etwa würden wir durch eine Abbildung $\mathbb{T} \to \mathbb{E}$ oder genauer, da Fliegen ja sterblich sind, durch die Abbildung einer geeigneten Teilmenge $I \subset \mathbb{T}$ nach $\mathbb{E}$ beschreiben.

1.9.8. Vielfach findet man die begriffliche Variante eines **affinen Raums über einem vorgegebenen Vektorraum**: Darunter versteht man dann eine Menge E mit einer "freien transitiven Wirkung" des vorgegebenen Vektorraums. Ich ziehe die oben gegebene Definition vor, da sie jeden Bezug auf einen vorgegebenen Vektorraum vermeidet und den Anschauungsraum meines Erachtens besser modelliert.

Definition 1.9.9. Eine Abbildung $\varphi : E \to E'$ zwischen affinen Räumen heißt eine **affine Abbildung** genau dann, wenn es eine lineare Abbildung zwischen den zugehörigen Richtungräumen $\vec{\varphi} : \vec{E} \to \vec{E}'$ gibt mit

$$\varphi(\vec{u} + e) = \vec{\varphi}(\vec{u}) + \varphi(e) \quad \forall\, \vec{u} \in \vec{E},\ e \in E$$

Diese lineare Abbildung $\vec{\varphi}$ ist dann durch φ eindeutig bestimmt und heißt der **lineare Anteil** unserer affinen Abbildung. Eine bijektive affine Abbildung heißt auch ein **Isomorphismus von affinen Räumen**, ein Isomorphismus von einem affinen Raum auf sich selbst heißt ein **Automorphismus** von besagtem affinen Raum.

Beispiel 1.9.10. Eine Abbildung $\varphi : V \to W$ zwischen Vektorräumen ist affine genau dann, wenn es eine lineare Abbildung $\vec{\varphi} : V \to W$ und einen Punkt $w \in W$ gibt mit $\varphi(v) = w + \vec{\varphi}(v)$ für alle $v \in V$.

Übung 1.9.11. Die Verknüpfung affiner Abbildungen ist affin und der lineare Anteil einer Verknüpfung affiner Abbildungen ist die Verknüpfung ihrer linearen Anteile.

Übung 1.9.12. Beschreiben Sie in Worten eine affine Abbildung $\mathbb{T} \to \mathbb{E}$ des affinen Raums der Zeiten in den Anschauungsraum. Natürlich ist das mathematische Übung im eigentlichen Sinne!

1.9.13. Nach der reinen Lehre sollte eine Teilmenge eines affinen Raums ein "affiner Teilraum" heißen genau dann, wenn sie so mit der Struktur eines affinen Raums versehen werden kann, daß die Einbettung eine affine Abbildung wird. Da diese Definition jedoch für Anwendungen erst aufgeschlüsselt werden muß, nehmen wir als unsere Definition gleich die aufgeschlüsselte Fassung und überlassen dem Leser den Nachweis der Äquivalenz zur Definition aus der reinen Lehre als Übung 1.9.16.

Definition 1.9.14. Eine Teilmenge $F \subset E$ eines affinen Raums heißt ein **affiner Teilraum** genau dann, wenn es einen Punkt $p \in E$ und einen Untervektorraum $W \subset \vec{E}$ gibt mit

$$F = p + W$$

Die durch Restriktion gegebene Abbildung $W \to \mathrm{Ens}^\times F$ ist dann eine Injektion und wir erklären wir auf F die Struktur eines affinen Raums, indem wir als $\vec{F}$ das Bild von W in $\mathrm{Ens}^\times F$ nehmen und diese abelsche Gruppe mit derjenigen Struktur eines k-Vektorraums versehen, für die Restriktion $W \xrightarrow{\sim} \vec{F}$ ein Vektorraumisomorphismus ist.

Beispiel 1.9.15. Die affinen Teilräume des $\mathbb{R}^3$ sind genau: Alle einelementigen Teilmengen, alle Geraden $G = p+\mathbb{R}\vec{v}$ mit $\vec{v} \neq 0$, alle Ebenen $P = p+\mathbb{R}\vec{v}+\mathbb{R}\vec{w}$ mit $\vec{v}, \vec{w}$ linear unabhängig, und der ganze $\mathbb{R}^3$.

Übung 1.9.16. Sei E ein affiner Raum. Genau dann ist eine Teilmenge $F \subset E$ ein affiner Teilraum im Sinne von 1.9.14, wenn F eine Struktur als affiner Raum $(F, \vec{F}, b)$ besitzt derart, daß die Einbettung eine affine Abbildung ist. Diese affine Struktur ist dann eindeutig bestimmt.

1.9.17. Eine Teilmenge eines affinen Raums heißt eine **Gerade** oder genauer eine **affine Gerade** genau dann, wenn sie ein affiner Teilraum der Dimension Eins ist. Eine Teilmenge eines affinen Raums heißt eine **Ebene** oder genauer eine **affine Ebene** genau dann, wenn sie ein affiner Teilraum der Dimension Zwei ist.

1.9.18. Eine Teilmenge eines affinen Raums heißt eine **Hyperebene** oder genauer eine **affine Hyperebene** genau dann, wenn sie ein echter affiner Teilraum ist, dessen Richtungsraum im Sinne von 1.3.20 eine lineare Hyperebene im Richtungsraum unseres ursprünglichen affinen Raums ist.

1.9.19. Gegeben ein affiner Raum E mit einem affinen Teilraum $F \subset E$ verwenden wir von nun an das Symbol $\vec{F}$ auch für den Untervektorraum von $\vec{E}$, den wir als das Bild des Richtungsraums $\vec{F}$ von F unter dem linearen Anteil der Einbettung erhalten.

1.9.20. Ein nichtleerer Schnitt von affinen Teilräumen eines affinen Raums ist stets wieder ein affiner Teilraum, und der Richtungsraum des Schnitts ist der Schnitt der Richtungsräume, zumindest wenn wir alle diese Räume wie in 1.9.19 als Teilmengen des Richtungsraums unseres ursprünglichen Raums betrachten.

Definition 1.9.21. Zwei affine Teilräume $T, S \subset E$ eines affinen Raums E heißen **parallel** genau dann, wenn in $\vec{E}$ gilt $\vec{T} \subset \vec{S}$ oder $\vec{S} \subset \vec{T}$.

Definition 1.9.22. Gegeben eine nichtleere Teilmenge eines affinen Raums gibt es nach 1.9.20 einen kleinsten affinen Teilraum, der sie umfaßt. Wir bezeichnen ihn als den **von unserer Teilmenge erzeugten** affinen Teilraum. Ein **Erzeugendensystem** eines affinen Raums ist eine Teilmenge, die ihn erzeugt.

Übung 1.9.23. Durch je zwei verschiedene Punkte eines affinen Raums geht genau eine Gerade, als da heißt, es gibt genau einen affinen Teilraum der Dimension Eins, der unsere beiden Punkte enthält.

Übung 1.9.24. Durch je drei Punkte eines affinen Raums, die nicht auf einer Gerade liegen, geht genau eine Ebene.

Übung 1.9.25. Der von einer nichtleeren endlichen Teilmenge E eines affinen Raums erzeugte Teilraum hat höchstens die Dimension $|E| - 1$.

Übung 1.9.26. Gegeben zwei endlichdimensionale affine Teilräume A, B eines affinen Raums E gilt für die Dimension des affinen Erzeugnisses C ihrer Vereinigung die Formel

$$\dim C = \begin{cases} \dim A + \dim B - \dim(A \cap B) & \text{falls } A \cap B \neq \emptyset; \\ \dim A + \dim B - \dim(\vec{A} \cap \vec{B}) + 1 & \text{falls } A \cap B = \emptyset. \end{cases}$$

Hierbei ist der Schnitt in $\vec{E}$ zu verstehen.

Satz 1.9.27 (Geometrische Charakterisierung affiner Abbildungen). *Eine injektive Abbildung von einem mindestens zweidimensionalen reellen affinen Raum in einen weiteren reellen affinen Raum ist affin genau dann, wenn das Bild jeder Geraden unter unserer Abbildung wieder eine Gerade ist.*

1.9.28. Die affinen Geraden des Anschauungsraums denken wir uns als Sichtlinien. Der vorhergehende Satz 1.9.27 zeigt, daß im Fall reeller affiner Räume ab der Dimension Zwei die Kenntnis aller Geraden auch umgekehrt bereits die Struktur als reeller affiner Raum festlegt: Haben nämlich zwei Strukturen als affiner reeller Raum auf derselben Menge dieselben Geraden und ist nicht der ganze Raum eine Gerade, so ist nach 1.9.27 die Identität auf unserer Menge ein Morphismus zwischen ihr mit der einen Struktur als affiner Raum und ihr mit der anderen Struktur als affiner Raum. Dann aber müssen diese beiden Strukturen bereits übereinstimmen. Salopp gesprochen legt also insbesondere “die Kenntnis der Sichtlinien bereits fest, welche Abbildungen als Parallelverschiebungen anzusehen sind”. Explizit kann das so formuliert werden: Zunächst legt die Kenntnis der Sichtlinien alias Geraden fest, welche Teilmengen die Bezeichung als “Ebene” verdienen; Dann vereinbart man, zwei Geraden “parallel” zu nennen, wenn sie in einer Ebene liegen und sich nicht

schneiden; Und schließlich kann man dann Parallelverschiebungen charakterisieren als solche Abbildungen, die parallele Geraden in parallele Geraden überführen.

Beweis. Wir zeigen das zunächst unter der Annahme, daß sowohl unser Ausgangsraum als auch der Raum, in den abgebildet wird, beide die Dimension Zwei haben. Ohne Beschränkung der Allgemeinheit dürfen wir dann annehmen, daß es sich bei beiden Räumen um den $\mathbb{R}^2$ handelt, und indem wir unsere Abbildung noch mit einer geeigneten Verschiebung verknüpfen, dürfen wir auch annehmen, daß sie den Ursprung festhält. Diesen Fall behandeln wir als eigenständiges Lemma.

Lemma 1.9.29. *Eine injektive Abbildung $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ mit $\Phi(0) = 0$, unter der das Bild jeder affinen Geraden wieder eine affine Gerade ist, muß linear sein.*

Beweis. Halten wir eine geeignete lineare Abbildung dahinter, so erkennen wir, daß wir ohne Beschränkung der Allgemeinheit annehmen dürfen, daß unser Φ die Vektoren e_1 und e_2 der Standardbasis festhält. Unter dieser Zusatzannahme gilt es nun zu zeigen, daß Φ die Identität ist. Zunächst gibt es sicher Abbildungen $\psi_1, \psi_2 : \mathbb{R} \to \mathbb{R}$ mit $\Phi(a\mathrm{e}_i) = \psi_i(a)\,\mathrm{e}_i$. Da wir Φ injektiv angenommen haben, müssen unter Φ parallele alias sich nicht schneidende Geraden parallel bleiben. Die Gerade durch $a\mathrm{e}_1$ und $a\mathrm{e}_2$ für $a \neq 0, 1$ ist parallel zu der durch e_1 und e_2, also ist für $a \neq 0, 1$ auch die Gerade durch $\Phi(a\mathrm{e}_1) = \psi_1(a)\,\mathrm{e}_1$ und $\Phi(a\mathrm{e}_2) = \psi_2(a)\,\mathrm{e}_2$ parallel zu der durch $\Phi(\mathrm{e}_1) = \mathrm{e}_1$ und $\Phi(\mathrm{e}_2) = \mathrm{e}_2$. Es folgt $\psi_1(a) = \psi_2(a)$ für $a \neq 0, 1$. Für $a = 0, 1$ ist das eh klar und wir notieren diese Abbildung nun ψ. Natürlich gilt $\psi(0) = 0$ und $\psi(1) = 1$. Da man die Addition von linear unabhängigen Vektoren durch Parallelogramme darstellen kann, gilt $\Phi(v + w) = \Phi(v) + \Phi(w)$ falls v und w linear unabhängig sind. Wir erhalten für $a \in \mathbb{R}$ damit

$$\Phi(\mathrm{e}_1 + a\,\mathrm{e}_2) \;=\; \mathrm{e}_1 + \psi(a)\,\mathrm{e}_2$$

im Fall $a \neq 0$ wegen der linearen Unabhängigkeit und im Fall $a = 0$ wegen $\psi(0) = 0$. Daraus folgern wir

$$\begin{aligned} \Phi(\mathrm{e}_1 + (a+b)\,\mathrm{e}_2) &= \mathrm{e}_1 + \psi(a+b)\,\mathrm{e}_2 \\ \Phi(\mathrm{e}_1 + a\,\mathrm{e}_2 + b\,\mathrm{e}_2) &= \mathrm{e}_1 + \psi(a)\,\mathrm{e}_2 + \psi(b)\,\mathrm{e}_2 \end{aligned}$$

indem wir bei der zweiten Gleichung ohne Beschränkung der Allgemeinheit $b \neq 0$ annehmen und erst den letzten Summanden abspalten. Es folgt sofort $\psi(a+b) = \psi(a) + (b)$. Da für $a, b \in \mathbb{R}$ mit $a \neq 0$ und $b \neq 0, 1$ die Gerade durch e_1 und $a\mathrm{e}_2$ parallel ist zu der durch $b\mathrm{e}_1$ und $ab\mathrm{e}_2$ folgt auch $\psi(ab) = \psi(a)\psi(b)$

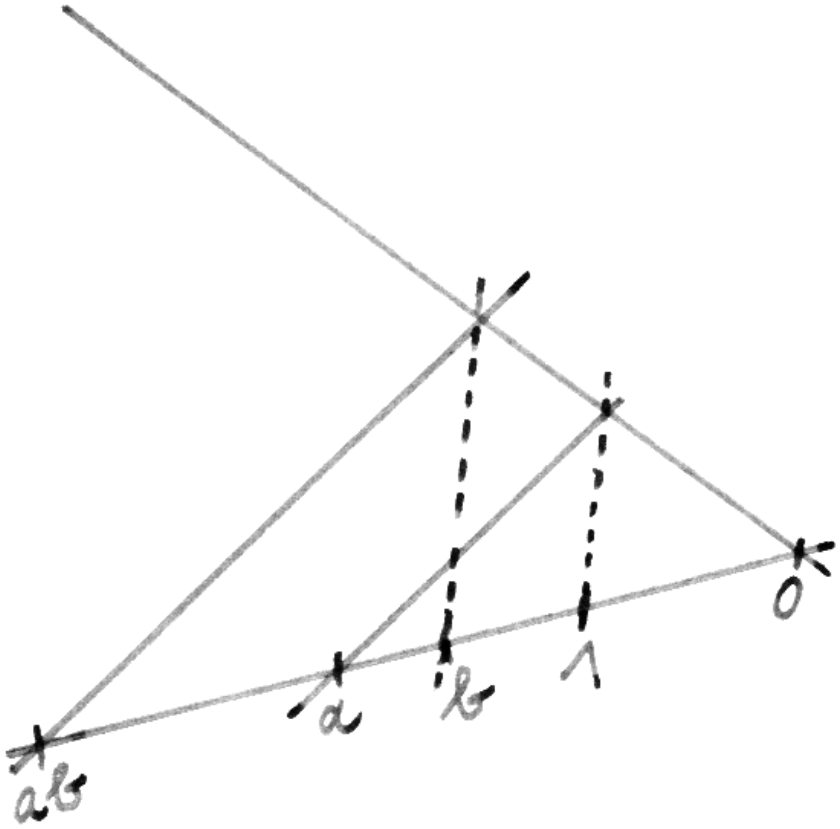
erst für alle $a, b \neq 0, 1$, dann aber wegen $\psi(0) = 0$ und $\psi(1) = 1$ sogar für alle $a, b \in \mathbb{R}$. Da nach **??.??** oder besser **??** die Identität der einzige Körperhomomorphismus $\psi : \mathbb{R} \to \mathbb{R}$ ist, folgt $\psi = \mathrm{id}$. Da wie bereits erwähnt gilt $\Phi(v+w) = \Phi(v)+\Phi(w)$ falls v und w linear unabhängig sind, folgt sofort $\Phi = \mathrm{id}$. □

Um nun Satz 1.9.27 zu zeigen, sei $\Phi : E \hookrightarrow F$ unsere injektive Abbildung von reellen affinen Räumen, unter der das Bild jeder Geraden eine Gerade ist. Sei ein Punkt $e \in E$ fest gewählt und seien $\vec{v}, \vec{w} \in \vec{E}$ linear unabhängig. Die Bilder von $e, e + \vec{v}$ und $e + \vec{w}$ können nicht auf einer Geraden liegen, da sonst zwei verschiedene Geraden auf dieselbe Gerade abgebildet würden im Widerspruch zur Injektivität von Φ. Folglich erzeugen diese Bilder eine affine Ebene. Die von $e, e + \vec{v}, e + \vec{w}$ aufgespannte affine Ebene kann beschrieben werden als die Vereinigung aller Geraden, die durch einen von e verschiedenen Punkt der Gerade $e + \mathbb{R}\vec{v}$ sowie durch einen Punkt der Geraden $e + \mathbb{R}\vec{w}$ laufen. Diese Ebene wird dann von Φ bijektiv abgebildet auf die von $\Phi(e), \Phi(e + \vec{w}), \Phi(e + \vec{w})$ aufgespannte Ebene, denn diese kann in derselben Weise beschrieben werden. Mit unserem Lemma 1.9.29 folgt, daß Φ eine affine Abbildung zwischen diesen Ebenen induzieren muß. Die Abbildung $\Psi : \vec{E} \to \vec{F}$ gegeben durch $\Phi(e + \vec{v}) = \Phi(e) + \Psi(\vec{v})$ ist nun linear, da nach dem Vorhergehenden ihre Restriktion auf jeden zweidimensionalen Teilraum von $\vec{E}$ linear ist. Das hinwiederum zeigt, daß Φ affin ist. □

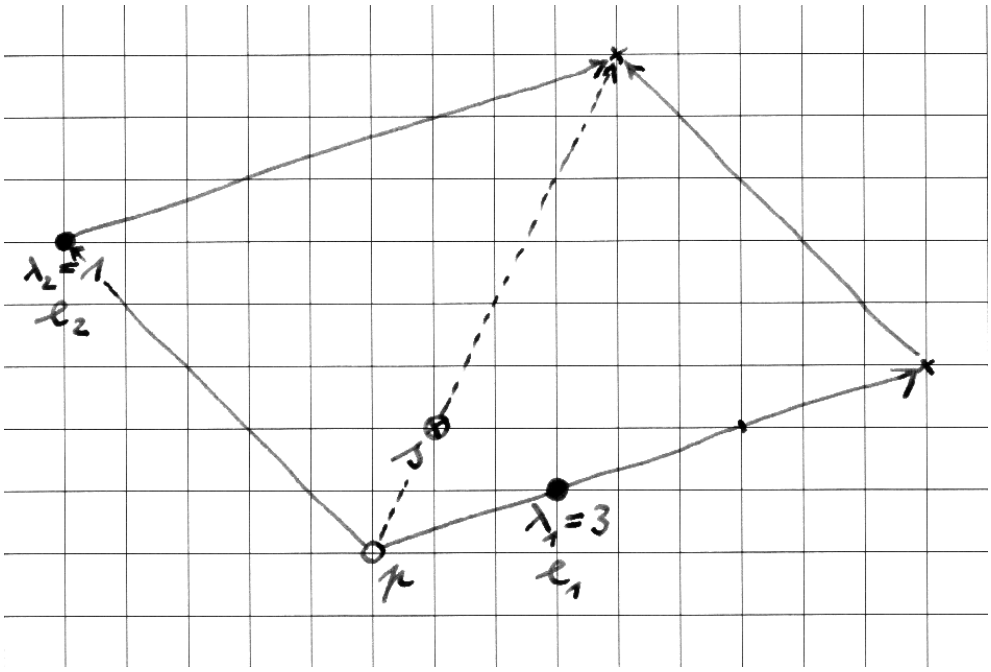
1.9.30. Geht man den Beweis von Lemma 1.9.29 nocheinmal durch, so erkennt man, daß er auch die folgende feinere Aussage zeigt: Sind K, L Körper und ist $\Phi : K^2 \hookrightarrow L^2$ eine Injektion mit $\Phi(0) = 0$, unter der das Bild jeder affinen Geraden wieder eine affine Gerade ist, so ist Φ ein Gruppenhomomorphismus und es gibt einen Körperisomorphismus $\psi : K \xrightarrow{\sim} L$ mit $\Phi(\lambda\vec{v}) = \psi(\lambda)\Phi(\vec{v})$ für alle $\lambda \in K$ und $\vec{v} \in K^2$. Salopp gesprochen ist also unsere Abbildung Φ "linear bis auf einen Körperisomorphismus".

1.9.31. Geht man den Beweis 1.9.27 im Lichte von 1.9.30 nocheinmal durch, so erkennt man, daß er auch die folgende feinere Aussage zeigt: Haben zwei Strukturen $(E, \vec{E}, a)$ und $(E, \vec{E}', a')$ auf ein- und derselben Menge E als zweidimensionaler affiner Raum über Körpern k bzw. k' dieselben Geraden, so gilt $\vec{E} = \vec{E}'$ und es gibt genau einen Körperisomorphismus $\varphi : k \xrightarrow{\sim} k'$ mit $a(\lambda, \vec{v}) = a'(\varphi(\lambda), \vec{v})$ für alle $\lambda \in k$ und $\vec{v} \in \vec{E}$. Salopp gesprochen kennt also ein weißes Blatt Papier zusammen mit einem Lineal bereits den Körper $\mathbb{R}$ der reellen Zahlen!

1.9.32. Gegeben ein affiner Raum E über einem Körper k und darin Punkte $e_1, \ldots, e_n \in E$ und Skalare $\lambda_1, \ldots, \lambda_n \in k$ mit $\lambda_1 + \ldots + \lambda_n \neq 0$ definiert man



Wie man auf einer Gerade der Papierebene mit zwei verschiedenen als Null und Eins ausgezeichneten Punkten zwei beliebige Punkte multipliziert, wenn man nur ein Lineal zur Verfügung hat, das aber "unendlich lang" ist in dem Sinne, daß man durch einen gegebenen Punkt die zu einer gegebenen Gerade parallele Gerade zeichnen kann.



Zwei fette Punkte der Gewichte 3 und 1 und ihr Schwerpunkt s nebst seiner Bestimmung mithilfe eines beliebigen weiteren Punktes p.

den **Schwerpunkt** s **der** e_i **mit den Gewichten** λ_i durch die Bedingung

$$\lambda_1(e_1 - s) + \ldots + \lambda_n(e_n - s) = \vec{0}$$

Daß höchstens ein Punkt s diese Bedingung erfüllen kann folgt daraus, daß für jedes weitere s', das unsere Bedingung erfüllt, gelten muß

$$(\lambda_1 + \ldots + \lambda_n)(s - s') = \vec{0}$$

Daß es überhaupt ein s gibt, das unsere Bedingung erfüllt, erkennt man, indem man einen beliebigen Punkt $p \in E$ wählt und $\lambda = \lambda_1 + \ldots + \lambda_n$ setzt und den Punkt

$$s = p + \frac{\lambda_1}{\lambda}(e_1 - p) + \ldots + \frac{\lambda_n}{\lambda}(e_n - p)$$

betrachtet. Für diesen Punkt $s \in E$ gilt ja

$$\lambda(s - p) = \lambda_1(e_1 - p) + \ldots + \lambda_n(e_n - p)$$

und daraus folgt dann leicht

$$\vec{0} = \lambda_1(e_1 - s) + \ldots + \lambda_n(e_n - s)$$

Übung 1.9.33. Ist E ein n-dimensionaler affiner Raum und $e_0, \ldots, e_n$ ein Erzeugendensystem von E, so gibt es für jeden Punkt $s \in E$ genau ein Tupel von Gewichten $(\lambda_0, \ldots, \lambda_n) \in k^{n+1}$ so daß gilt $\lambda_0 + \ldots + \lambda_n = 1$ und daß s der Schwerpunkt der e_i mit den Gewichten λ_i ist. Die λ_i heißen dann die **baryzentrischen Koordinaten von** s **in Bezug auf die** e_i, nach griechisch "$\beta\alpha\rho\upsilon\varsigma$" für "schwer".

2 Ringe, Polynome, Determinanten

2.1 Ringe

Definition 2.1.1. Ein **Ring** ist eine Menge mit zwei assoziativen Verknüpfungen $(R, +, \cdot)$ derart, daß gilt

1. $(R, +)$ ist eine kommutative Gruppe;
2. Es gelten die Distributivgesetze, d.h. für alle $a, b, c \in R$ gilt
$$\begin{aligned} a \cdot (b + c) &= (a \cdot b) + (a \cdot c) \\ (a + b) \cdot c &= (a \cdot c) + (b \cdot c) \end{aligned}$$
3. Es gibt ein Element $1 = 1_R \in R$ mit $1 \cdot a = a \cdot 1 = a \quad \forall a \in R$.

Die beiden Verknüpfungen heißen die **Addition** und die **Multiplikation** in unserem Ring. Wir könnten gleichbedeutend einen Ring auch definieren als eine Menge mit zwei Verknüpfungen $(R, +, \cdot)$ derart, daß $(R, +)$ eine kommutative Gruppe ist, daß $(R, \cdot)$ ein Monoid ist, und daß die Distributivgesetze gelten. Das Element $1 \in R$ aus unserer Definition ist wohlbestimmt als das neutrale Element des Monoids $(R, \cdot)$, es heißt das **Eins-Element** oder kurz die **Eins** unseres Rings. Ein Ring, dessen Multiplikation kommutativ ist, heißt ein **kommutativer Ring** und bei uns in unüblicher Verkürzung ein **Kring**.

2.1.2. Wir schreiben meist kürzer $a \cdot b = ab$ und vereinbaren die Regel "Punkt vor Strich", so daß zum Beispiel das erste Distributivgesetz auch in der Form $a(b + c) = ab + ac$ geschrieben werden kann.

2.1.3. Der Begriff "Ring" soll zum Ausdruck bringen, daß diese Struktur nicht in demselben Maße "geschlossen" ist wie ein Körper, da wir nämlich nicht die Existenz von multiplikativen Inversen fordern. Er wird auch im juristischen Sinne für gewisse Arten von weniger geschlossenen Körperschaften verwendet, so gibt es etwa den "Ring deutscher Makler" oder den "Ring deutscher Bergingenieure". Eine Struktur wie in der vorhergehenden Definition, bei der nur die Existenz eines Einselements nicht gefordert wird, bezeichnen wir als **Rng**. In der Literatur wird jedoch auch diese Struktur oft als "Ring" bezeichnet. Die Ringe, die eine Eins besitzen, heißen in dieser Terminologie "unitäre Ringe".

Beispiele 2.1.4. Die einelementige Menge mit der offensichtlichen Addition und Multiplikation ist ein Ring, der **Nullring**. Jeder Körper ist ein Ring. Die ganzen Zahlen $\mathbb{Z}$ bilden einen Ring. Ist R ein Ring und X eine Menge, so ist die Menge $\mathrm{Ens}(X, R)$ aller Abbildungen von X nach R ein Ring unter

punktweiser Multiplikation und Addition. Ist R ein Ring und $n \in \mathbb{N}$, so bilden die $(n \times n)$-Matrizen mit Einträgen in R einen Ring $\mathrm{M}(n \times n; R)$ unter der üblichen Addition und Multiplikation von Matrizen; im Fall $n = 0$ erhalten wir den Nullring, im Fall $n = 1$ ergibt sich R selbst. Ist A eine abelsche Gruppe, so bilden die Gruppenhomomorphismen von A in sich selbst, die sogenannten **Endomorphismen** von A, einen Ring mit der Verknüpfung von Abbildungen als Multiplikation und der punktweisen Summe als Addition. Wir notieren diesen Ring

$$\mathrm{Ab}\, A$$

und nennen ihn den **Endomorphismenring** der abelschen Gruppe A. Vielfach notiert man diesen Ring auch $\mathrm{End}\, A$ oder sogar $\mathrm{End}_{\mathbb{Z}} A$ aus Gründen, die erst in **??** erklärt werden.

Übung 2.1.5. Auf der abelschen Gruppe $\mathbb{Z}$ gibt es genau zwei Verknüpfungen, die als Multiplikation genommen die Addition zu einer Ringstruktur ergänzen.

2.1.6. Allgemeiner als in 1.6.4 erklärt heißt ein Element a eines beliebigen Rnges **idempotent** genau dann, wenn gilt $a^2 = a$. Allgemeiner als in 1.7.42 erklärt heißt ein Element a eines beliebigen Rnges **nilpotent** genau dann, wenn es $d \in \mathbb{N}$ gibt mit $a^d = 0$.

Beispiel 2.1.7. Gegeben eine ganze Zahl $m \in \mathbb{Z}$ konstruieren wir den **Restklassenring** $\mathbb{Z}/m\mathbb{Z}$ wie folgt: Seine Elemente sind diejenigen Teilmengen T von $\mathbb{Z}$, die in der Form $T = a + m\mathbb{Z}$ mit $a \in \mathbb{Z}$ dargestellt werden können. Die Teilmenge $a + m\mathbb{Z}$ heißt auch die **Restklasse von** a **modulo** m, da zumindest im Fall $a \geq 0$ ihre nichtnegativen Elemente genau alle natürlichen Zahlen sind, die beim Teilen durch m denselben Rest lassen wie a. Wir notieren diese Restklasse auch $\bar{a}$. Natürlich ist $\bar{a} = \bar{b}$ gleichbedeutend zu $a - b \in m\mathbb{Z}$. Gehören a und b zur selben Restklasse, in Formeln $a + m\mathbb{Z} = b + m\mathbb{Z}$, so nennen wir sie **kongruent modulo** m und schreiben

$$a \equiv b \pmod{m}$$

Offensichtlich gibt es für $m \in \mathbb{N}_{\geq 1}$ genau m Restklassen modulo m, in Formeln $|\mathbb{Z}/m\mathbb{Z}| = m$, und wir haben genauer

$$\mathbb{Z}/m\mathbb{Z} = \{\bar{0}, \bar{1}, \ldots, \overline{m-1}\}$$

Für alle $m \in \mathbb{Z}$ bilden die Restklassen ein Mengensystem $\mathbb{Z}/m\mathbb{Z} \subset \mathcal{P}(\mathbb{Z})$, das stabil ist unter der von der Addition auf $\mathbb{Z}$ im Sinne von I.3.1.2 6 induzierten Verknüpfung. Mit dieser Verknüpfung gilt $\bar{a} + \bar{b} = \overline{a+b} \quad \forall a, b \in \mathbb{Z}$ und

$\mathbb{Z}/m\mathbb{Z}$ wird eine abelsche Gruppe. Diese Gruppe wird sogar zu einem Ring vermittels der Multiplikation

$$T \odot S = T \cdot S + m\mathbb{Z} = \{ab + mr \mid a \in T, b \in S, r \in \mathbb{Z}\}$$

In der Tat prüft man für die so erklärte Multiplikation mühelos die Formeln

$$\bar{a} \odot \bar{b} = \overline{ab}$$

und damit folgen die Distributivgesetze für $\mathbb{Z}/m\mathbb{Z}$ unmittelbar aus den Distributivgesetzen im Ring $\mathbb{Z}$. Wir geben wir die komische Notation $\odot$ nun auch gleich wieder auf und schreiben stattdessen $\bar{a} \cdot \bar{b}$ oder noch kürzer $\bar{a}\bar{b}$.

Beispiel 2.1.8. Modulo $m = 2$ gibt es genau zwei Restklassen: Die Elemente der Restklasse von 0 bezeichnet man üblicherweise als **gerade Zahlen**, die Elemente der Restklasse von 1 als **ungerade Zahlen**. Der Ring $\mathbb{Z}/2\mathbb{Z}$ mit diesen beiden Elementen $\bar{0}$ und $\bar{1}$ ist offensichtlich sogar ein Körper.

Beispiel 2.1.9. Den Ring $\mathbb{Z}/12\mathbb{Z}$ könnte man als "Ring von Uhrzeiten" ansehen. Er hat die zwölf Elemente $\{\bar{0}, \bar{1}, \ldots, \overline{11}\}$ und wir haben $\overline{11} + \bar{5} = \overline{16} = \bar{4}$ alias "5 Stunden nach 11 Uhr ist es 4 Uhr." Weiter haben wir in $\mathbb{Z}/12\mathbb{Z}$ etwa auch $\bar{3} \cdot \bar{8} = \overline{24} = \bar{0}$. In einem Ring kann es also durchaus passieren, daß ein Produkt von zwei von Null verschiedenen Faktoren Null ist.

Proposition 2.1.10 (**Teilbarkeitskriterien über Quersummen**)**.** *Eine natürliche Zahl ist genau dann durch drei bzw. durch neun teilbar, wenn ihre Quersumme durch drei bzw. durch neun teilbar ist.*

Beweis. Wir erklären das Argument nur an einem Beispiel. Per definitionem gilt

$$1258 = 1 \cdot 10^3 + 2 \cdot 10^2 + 5 \cdot 10 + 8$$

Offensichtlich folgt

$$1258 \equiv 1 \cdot 10^3 + 2 \cdot 10^2 + 5 \cdot 10 + 8 \pmod 3$$

Da 10 kongruent ist zu 1 modulo 3 erhalten wir daraus

$$1258 \equiv 1 + 2 + 5 + 8 \pmod 3$$

Insbesondere ist die rechte Seite durch drei teilbar genau dann, wenn die linke Seite durch drei teilbar ist. Das Argument für neun statt drei geht genauso. □

Übung 2.1.11. Eine natürliche Zahl ist durch 11 teilbar genau dann, wenn ihre "alternierende Quersumme" durch 11 teilbar ist.

2.1.12. In $\mathbb{Z}/12\mathbb{Z}$ gilt zum Beispiel $\bar{3}\cdot\bar{5} = \bar{3}\cdot\bar{1}$. In allgemeinen Ringen dürfen wir also nicht kürzen. Dies Phänomen werden wir nun begrifflich fassen.

Definition 2.1.13.

1. Gegeben ein Kring R und Elemente $a, b \in R$ sagen wir, a **teilt** b oder a ist ein **Teiler** von b und schreiben $a|b$ genau dann, wenn es $d \in R$ gibt mit $ad = b$.

2. Natürlich teilt jedes Element eines Krings die Null. Ein Element a eines Rings R heißt ein **Nullteiler** von R genau dann, wenn es die Null "in nicht-trivialer Weise teilt", wenn es genauer und in Formeln $d \in R\backslash 0$ gibt mit $ad = 0$ oder $da = 0$.

3. Ein Ring heißt **nullteilerfrei** genau dann, wenn er außer der Null keine Nullteiler besitzt, wenn also das Produkt von je zwei von Null verschiedenen Elementen auch wieder von Null verschieden ist.

4. Ein Ring heißt ein **Integritätsbereich** genau dann, wenn er nullteilerfrei und ausserdem nicht der Nullring ist.

Bemerkung 2.1.14. Manche Autoren fordern von nullteilerfreien Ringen zusätzlich, daß sie nicht der Nullring sein dürfen, benutzen also dieses Wort als Synonym für "Integritätsbereich".

Übung 2.1.15. Man bestimme alle Nullteiler im Restklassenring $\mathbb{Z}/12\mathbb{Z}$.

2.1.16 (**Kürzen in Ringen**). Sei R ein Ring. Ist $a \in R$ kein Nullteiler, so folgt aus $ax = ay$ schon $x = y$. In der Tat haben wir nämlich $ax = ay \;\Rightarrow\; a(x-y) = 0 \;\Rightarrow\; x - y = 0 \;\Rightarrow\; x = y$.

Definition 2.1.17. Eine Abbildung $\varphi : R \to S$ von einem Ring in einen weiteren Ring heißt ein **Ringhomomorphismus** genau dann, wenn gilt $\varphi(1) = 1$ und $\varphi(a+b) = \varphi(a)+\varphi(b)$ sowie $\varphi(ab) = \varphi(a)\varphi(b)$ für alle $a, b \in R$. In anderen Worten ist ein Ringhomomorphismus also eine Abbildung, die sowohl für die Addition als auch für die Multiplikation ein Monoidhomomorphismus ist.

Übung 2.1.18. Gegeben eine abelsche Gruppe V und ein Körper k induziert die kanonische Identifikation $\mathrm{Ens}(k \times V, V) \xrightarrow{\sim} \mathrm{Ens}(k, \mathrm{Ens}(V, V))$ aus I.2.2.23 eine Bijektion

$$\left\{ \begin{array}{c} \text{Strukturen als } k\text{-Vektorraum} \\ \text{auf der abelschen Gruppe } V \end{array} \right\} \xrightarrow{\sim} \left\{ \begin{array}{c} \text{Ringhomomorphismen} \\ k \to \mathrm{Ab}\, V \end{array} \right\}$$

2.1.19. Von Homomorphismen zwischen Rngen können wir natürlich nicht fordern, daß sie das Einselement auf das Einselement abbilden. Wir sprechen

dann von **Rnghomomorphismen**. In der Terminologie, in der unsere Rnge als Ringe bezeichnet werden, werden unsere Ringhomomorphismen "unitäre Ringhomomorphismen" genannt.

Übung 2.1.20. Für jeden Ring R gibt es genau einen Ringhomomorphismus $\mathbb{Z} \to R$. Im Fall $R = \mathbb{Z}/m\mathbb{Z}$ wird dieser Ringhomomorphismus gegeben durch die Vorschrift $a \mapsto \bar{a}$. Hinweis: Man erinnere I.3.3.12.

Definition 2.1.21. Ein Element a eines Rings R heißt **invertierbar** oder auch eine **Einheit** genau dann, wenn es bezüglich der Multiplikation invertierbar ist im Sinne von I.3.2.2, wenn es also $b \in R$ gibt mit $ab = ba = 1$. Die Menge der invertierbaren Elemente eines Rings bildet unter der Multiplikation eine Gruppe, die man die **Gruppe der Einheiten von** R nennt und gemäß unserer allgemeinen Konventionen I.3.2.11 mit $R^\times$ bezeichnet.

2.1.22. A priori meint eine Einheit in der Physik das, was ein Mathematiker eine Basis eines eindimensionalen Vektorraums nennen würde. So wäre etwa die Sekunde s eine Basis des reellen Vektorraums $\vec{T}$ aller Zeitspannen aus **??**. In Formeln ausgedrückt bedeutet das gerade, daß das Daranmultiplizieren von s eine Bijektion $\mathbb{R} \xrightarrow{\sim} \vec{T}$ liefert. Mit den Einheiten eines kommutativen Ringes R verhält es sich nun genauso: Genau dann ist $u \in R$ eine Einheit, wenn das Daranmultiplizieren von u eine Bijektion $R \xrightarrow{\sim} R$ liefert. Daher rührt dann wohl auch die Terminologie.

2.1.23. Ein Körper ist in dieser Terminologie also ein kommutativer Ring, der nicht der Nullring ist und in dem jedes von Null verschiedene Element eine Einheit ist.

Übung 2.1.24. Jeder Ringhomomorphismus macht Einheiten zu Einheiten. Jeder Ringhomomorphismus von einem Körper in einen von Null verschiedenen Ring ist injektiv.

Übung 2.1.25. Ein Nullteiler kann nur im Nullring eine Einheit sein.

2.2 Untergruppen der ganzen Zahlen

2.2.1. Nach der reinen Lehre sollte eine Teilmenge einer Gruppe eine "Untergruppe" heißen genau dann, wenn sie so mit der Struktur einer Gruppe versehen werden kann, daß die Einbettung ein Gruppenhomomorphismus wird. Da diese Definition jedoch für Anwendungen erst aufgeschlüsselt werden muß, nehmen wir gleich die aufgeschlüsselte Fassung als unsere Definition und überlassen den Nachweis der Äquivalenz zur Definition nach der reinen Lehre dem Leser zur Übung.

Definition 2.2.2. Eine Teilmenge einer Gruppe heißt eine **Untergruppe** genau dann, wenn sie abgeschlossen ist unter der Verknüpfung und der Inversenbildung und wenn sie zusätzlich das neutrale Element enthält. Ist G eine multiplikativ geschriebene Gruppe, so ist demnach eine Teilmenge $U \subset G$ eine Untergruppe genau dann, wenn in Formeln gilt: $a, b \in U \Rightarrow ab \in U$, $a \in U \Rightarrow a^{-1} \in U$ sowie $1 \in U$.

Beispiele 2.2.3. In jeder Gruppe ist die einelementige Teilmenge, die nur aus dem neutralen Element besteht, eine Untergruppe. Wir nennen sie die **triviale Untergruppe**. Ebenso ist natürlich die ganze Gruppe stets eine Untergruppe von sich selber.

Übung 2.2.4. Eine endliche Teilmenge einer Gruppe, die mit je zwei Elementen auch ihr Produkt enthält, ist notwendig bereits eine Untergruppe.

Übung 2.2.5. Sind $H, K \subset G$ zwei Untergruppen einer multiplikativ gedachten Gruppe mit $H \cap K = 1$, so definiert die Multiplikation eine Injektion $H \times K \hookrightarrow G$.

2.2.6. Der Schnitt über eine beliebige Familie von Untergruppen einer gegebenen Gruppe ist selbst wieder eine Untergruppe. Für eine Teilmenge T einer Gruppe G definieren wir die **von T erzeugte Untergruppe**

$$\langle T \rangle \subset G$$

als die kleinste Untergruppe von G, die T enthält. Natürlich gibt es so eine kleinste Untergruppe, nämlich den Schnitt über alle Untergruppen von G, die T enthalten. Für $T \neq \emptyset$ können wir $\langle T \rangle$ konkret beschreiben als die Menge aller endlichen Produkte von Elementen aus T und deren Inversen. Für $T = \emptyset$ besteht $\langle T \rangle$ nur aus dem neutralen Element. Ist T durch einen Ausdruck in Mengenklammern gegeben, so lassen wir diese meist weg und schreiben also zum Beispiel kürzer $\langle a_1, \ldots, a_n \rangle$ statt $\langle \{a_1, \ldots, a_n\} \rangle$. Ob der Ausdruck $\langle T \rangle$ in einem speziellen Fall die von einer Menge T erzeugte Untergruppe oder vielmehr die von der einelementigen Menge mit einzigem Element T erzeugte Untergruppe meint, muß der Leser meist selbst aus dem Kontext erschließen. Schreiben wir $\langle\!| T \rangle$, so ist jedoch stets gemeint, daß T eine Menge von Erzeugern und nicht einen einzelnen Erzeuger bezeichnet.

2.2.7. Ist V ein k-Vektorraum und $T \subset V$ eine Teilmenge, so muß der Leser von nun an aus dem Kontext erschließen, ob mit $\langle T \rangle$ die von T erzeugte Untergruppe oder der von T erzeugte Untervektorraum gemeint ist. Zur Unterscheidung schreiben wir manchmal $\langle T \rangle_{\mathbb{Z}}$ für die von T erzeugte Untergruppe und $\langle T \rangle_k$ für den von T erzeugten Untervektorraum.

Lemma 2.2.8. *Das Bild einer Untergruppe unter einem Gruppenhomomorphismus ist stets eine Untergruppe. Das Urbild einer Untergruppe unter einem Gruppenhomomorphismus ist stets eine Untergruppe.*

Beweis. Dem Leser überlassen. □

Definition 2.2.9. Sei $\varphi : G \to H$ ein Gruppenhomomorphismus. Das Urbild der trivialen Untergruppe von H heißt der **Kern** von φ und wird bezeichnet mit

$$\ker \varphi = \{g \in G \mid \varphi(g) = e\}$$

Definition 2.2.10. Das **Bild** $\varphi(G)$ von ganz G unter φ wird nach der englischen und französischen Bezeichnung **image** bezeichnet mit

$$\operatorname{im} \varphi = \{\varphi(g) \mid g \in G\}$$

2.2.11. Nach 2.2.8 sind Kern und Bild eines Gruppenhomomorphismus stets Untergruppen im Definitionsbereich bzw. Wertebereich unseres Gruppenhomomorphismus.

Lemma 2.2.12. *Ein Gruppenhomomorphismus ist injektiv genau dann, wenn sein Kern trivial ist.*

Beweis. Sei $\varphi : G \to H$ unser Gruppenhomomorphismus. Wir argumentieren durch Widerspruch: Besteht $\ker \varphi$ aus mehr als einem Element, so kann φ natürlich nicht injektiv sein. Gibt es umgekehrt $x \neq y$ mit $\varphi(x) = \varphi(y)$, so liegen $x^{-1}y \neq e$ beide in $\ker \varphi$. □

Satz 2.2.13 (Untergruppen von $\mathbb{Z}$). *Jede Untergruppe $H \subset \mathbb{Z}$ ist von der Form $H = m\mathbb{Z}$ für genau ein $m \in \mathbb{N}$.*

Beweis. Im Fall $H = \{0\}$ ist $m = 0$ die einzige natürliche Zahl mit $H = m\mathbb{Z}$. Gilt $H \neq \{0\}$, so enthält H echt positive Elemente. Sei dann $m \in H$ das kleinste echt positive Element von H. Wir behaupten $H = m\mathbb{Z}$. Die Inklusion $H \supset m\mathbb{Z}$ ist hier offensichtlich. Aber gäbe es $n \in H \setminus m\mathbb{Z}$, so könnten wir n mit Rest teilen durch m und also schreiben $n = ms + r$ für geeignete $s, r \in \mathbb{Z}$ mit $0 < r < m$ und hätten $r = n - ms \in H$ im Widerspruch zur Minimalität von m. □

Definition 2.2.14. Seien $a, b \in \mathbb{Z}$. Wir sagen a **teilt** b und schreiben $a|b$ genau dann, wenn es $d \in \mathbb{Z}$ gibt mit $ad = b$. Sind zwei ganze Zahlen a, b nicht beide Null, so gibt es eine größte ganze Zahl c, die sie beide teilt. Diese Zahl heißt der **größte gemeinsame Teiler** von a und b. Zwei ganze Zahlen a und b heißen **teilerfremd** genau dann, wenn sie nicht beide Null sind und 1 ihr größter gemeinsamer Teiler ist.

Satz 2.2.15 (über den größten gemeinsamen Teiler). *Sind zwei ganze Zahlen $a, b \in \mathbb{Z}$ nicht beide Null und ist c ihr größter gemeinsamer Teiler, so gilt:*

1. *Es gibt $r, s \in \mathbb{Z}$ mit $c = ra + sb$.*
2. *Teilt $d \in \mathbb{Z}$ sowohl a als auch b, so teilt d auch den größten gemeinsamen Teiler von a und b.*

2.2.16. Der zweite Teil dieses Satzes ist einigermaßen offensichtlich, wenn man die Eindeutigkeit der Primfaktorzerlegung als bekannt voraussetzt. Da wir besagte Eindeutigkeit der Primfaktorzerlegung jedoch erst aus dem zweite Teil dieses Satzes ableiten werden, ist es wichtig, auch für den zweiten Teil des Satzes einen eigenständigen Beweis zu geben.

Beweis. Man betrachte die Teilmenge $a\mathbb{Z}+b\mathbb{Z} \subset \mathbb{Z}$. Sie ist offensichtlich eine von Null verschiedene Untergruppe von $\mathbb{Z}$. Also ist sie nach 2.2.13 von der Form $a\mathbb{Z} + b\mathbb{Z} = \hat{c}\mathbb{Z}$ für genau ein $\hat{c} > 0$ und es gilt

i. $\hat{c}$ teilt a und b; In der Tat haben wir ja $a, b \in \hat{c}\mathbb{Z}$;

ii. $\hat{c} = ra + sb$ für geeignete $r, s \in \mathbb{Z}$; In der Tat haben wir ja $\hat{c} \in a\mathbb{Z} + b\mathbb{Z}$;

iii. (d teilt a und b) $\Rightarrow$ (d teilt $\hat{c}$);

Daraus folgt aber sofort, daß $\hat{c}$ der größte gemeinsame Teiler von a und b ist, und damit folgt dann der Satz. □

2.2.17. Gegeben $a_1, \ldots, a_n \in \mathbb{Z}$ können wir mit der Notation 2.2.6 kürzer schreiben

$$a_1\mathbb{Z} + \ldots + a_n\mathbb{Z} = \langle a_1, \ldots, a_n \rangle$$

Üblich ist hier auch die Notation $(a_1, \ldots, a_n)$, die jedoch oft auch n-Tupel von ganzen Zahlen bezeichnet, also Elemente von $\mathbb{Z}^n$, und in der Analysis im Fall $n = 2$ meist ein offenes Intervall. Es gilt dann aus dem Kontext zu erschließen, was jeweils gemeint ist. Sind a und b nicht beide Null und ist c ihr größter gemeinsamer Teiler, so haben wir nach dem Vorhergehenden $\langle a, b \rangle = \langle c \rangle$. Wir benutzen von nun an diese Notation. Über die Tintenersparnis hinaus hat sie den Vorteil, auch im Fall $a = b = 0$ sinnvoll zu bleiben.

Definition 2.2.18. Eine **Primzahl** ist eine natürliche Zahl $p > 1$ derart, daß aus $p = ab$ mit $a, b \in \mathbb{N}$ schon folgt $a = 1$ oder $b = 1$.

2.2.19. Eine Möglichkeit, alle Primzahlen zu finden, ist das sogenannte **Sieb des Eratosthenes**: Man beginnt mit der kleinsten Primzahl, der Zwei. Streicht man alle Vielfachen der Zwei, d.h. alle geraden Zahlen, so ist die erste Zahl unter den Übrigen die nächste Primzahl, die Drei. Streicht man nun auch noch alle Vielfachen der Drei, so ist die erste Zahl unter den Übrigen die nächste Primzahl, die Fünf, und so weiter. "Der Erste" heißt auf lateinisch "Primus" und auf griechisch ähnlich und es könnte sein, daß die Bezeichnung "Primzahl" daher rührt.

Satz 2.2.20 (**Primfaktorzerlegung**)**.** *Jede natürliche Zahl $n \geq 2$ kann geschrieben werden als ein Produkt von Primzahlen $n = p_1 p_2 \dots p_r$, und diese Darstellung ist eindeutig bis auf die Reihenfolge der Faktoren.*

Beweis. Die Existenz der Primfaktorzerlegung ist klar mit vollständiger Induktion. Ihre Eindeutigkeit folgt mit vollständiger Induktion aus dem anschließenden Lemma. □

Lemma 2.2.21. *Teilt eine Primzahl ein Produkt von ganzen Zahlen, so teilt sie einen der Faktoren.*

2.2.22. Wenn wir die Eindeutigkeit der Primfaktorzerlegung als bekannt voraussetzen, so ist dies Lemma offensichtlich. Diese Argumentation hilft aber hier nicht weiter, da sie voraussetzt, was wir gerade erst beweisen wollen. Sicher ist Ihnen die Eindeutigkeit der Primfaktorzerlegung aus der Schule und ihrer Rechenerfahrung wohlvertraut. Um die Schwierigkeit zu sehen, sollten Sie vielleicht einmal selbst versuchen, einen Beweis dafür anzugeben. Im übrigen werden wir in **??** sehen, daß etwa im Ring $\mathbb{Z}[\sqrt{-5}]$ das Analogon zur Eindeutigkeit der Primfaktorzerlegung gar nicht mehr richtig ist.

Beweis. Sei p unsere Primzahl und seien $a, b \in \mathbb{Z}$ gegeben mit $p|ab$. Teilt p nicht a, so folgt $\langle p, a\rangle = \langle 1\rangle$, denn die Primzahl p hat nur die Teiler ± 1 und $\pm p$. Der größte gemeinsame von p und a kann aber nicht p sein und muß folglich 1 sein. Nach 2.2.15 gibt es also $r, s \in \mathbb{Z}$ mit $1 = rp + sa$. Es folgt $b = rpb + sab$ und damit $p|b$, denn p teilt natürlich rpb und teilt nach Annahme auch sab. □

2.2.23. Aus der Existenz der Primfaktorzerlegung folgt insbesondere, daß es unendlich viele Primzahlen geben muß: Für jede endliche Menge von Primzahlen können wir nämlich ihr Produkt bilden. Zählen wir zu diesem Produkt noch 1 hinzu, so kann keine Primzahl aus unserer endlichen Menge ein Primfaktor der neu entstandenen Zahl sein. Also ist jeder Primfaktor der neu entstandenen Zahl eine Primzahl außerhalb unserer vorgegebenen endlichen Menge von Primzahlen.

2.2.24. Noch offen (2009) ist die Frage, ob es auch unendlich viele **Primzahlzwillinge** gibt, d.h. Paare von Primzahlen mit der Differenz Zwei, wie zum Beispiel $5, 7$ oder $11, 13$ oder $17, 19$. Ebenso offen ist die Frage, ob jede gerade Zahl $n > 2$ die Summe von zwei Primzahlen ist. Diese Vermutung, daß das richtig sein sollte, ist bekannt als **Goldbach-Vermutung**.

2.2.25. Ich erkläre am Beispiel $a = 160$, $b = 625$ den sogenannten **euklidischen Algorithmus**, mit dem man den größten gemeinsamen Teiler c bestimmen kann nebst einer Darstellung $c = ra + rb$. In der linken Spalte von Gleichungen wird jeweils geteilt mit Rest, und will man nur den größten gemeinsamen Teiler kennen, so kann man die rechte Spalte ignorieren. Die oberste Zeile der rechten Spalte unserer Tabelle ist eine Trivialität, die Zweitoberste entsteht in offensichtlicher Weise aus der Zeile links daneben, und jede weitere Gleichung der rechten Spalte erhält man als eine Linearkombination der beiden darüberstehenden Gleichungen mit Koeffizienten, die sich aus der Gleichung links daneben ableiten lassen.

$$
\begin{array}{rcrcrcrcrcrcr}
 & & & & & & 0\cdot & 625 & + & 1\cdot & 160 & = & 160 \\
625 = 3\cdot & 160 & + & 145 & \Rightarrow & & 1\cdot & 625 & - & 3\cdot & 160 & = & 145 \\
160 = 1\cdot & 145 & + & 15 & \Rightarrow & & -1\cdot & 625 & + & 4\cdot & 160 & = & 15 \\
145 = 9\cdot & 15 & + & 10 & \Rightarrow & & 10\cdot & 625 & - & 39\cdot & 160 & = & 10 \\
15 = 1\cdot & 10 & + & 5 & \Rightarrow & & -11\cdot & 625 & + & 43\cdot & 160 & = & 5 \\
10 = 2\cdot & 5 & + & 0 & & & & & & & & &
\end{array}
$$

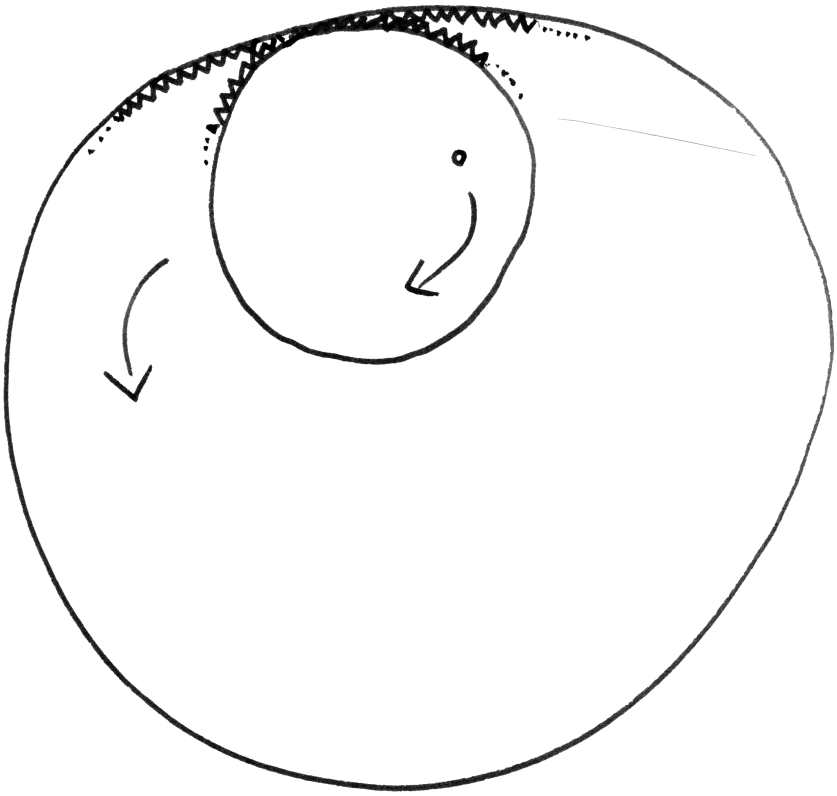
Aus der linken Spalte folgt für den größten gemeinsamen Teiler $\langle 625, 160\rangle = \langle 160, 145\rangle = \langle 145, 15\rangle = \langle 15, 10\rangle = \langle 10, 5\rangle = \langle 5, 0\rangle = \langle 5\rangle$ und wir finden mit der rechten Spalte für den größten gemeinsamen Teiler die Darstellung $-11 \cdot 625 + 43 \cdot 160 = 5$.

Übung 2.2.26. Beim sogenannten "Spirographen", einem Zeichenspiel für Kinder, kann man an einem innen mit 105 Zähnen versehenen Ring ein Zahnrad mit 24 Zähnen entlanglaufen lassen. Steckt man dabei einen Stift durch ein Loch außerhalb des Zentrums des Zahnrads, so entstehen dabei die köstlichsten Figuren. Wie oft muß man das Zahnrad auf dem inneren Zahnkranz umlaufen, bevor solch eine Figur fertig gemalt ist?

Proposition 2.2.27 (**Endliche Primkörper**). *Sei $m \in \mathbb{N}$.*

1. *Genau dann ist der Restklassenring $\mathbb{Z}/m\mathbb{Z}$ ein Integritätsbereich, wenn m eine Primzahl ist oder wenn gilt $m = 0$.*

2. *Genau dann ist $\mathbb{Z}/m\mathbb{Z}$ ein Körper, wenn m eine Primzahl ist.*



Der Spirograph aus Übung 2.2.26

2.2.28. Die Körper $\mathbb{Z}/p\mathbb{Z}$ für Primzahlen p sowie der Körper $\mathbb{Q}$ sind die "kleinstmöglichen Körper" in einem Sinne, der in **??** präzisiert wird. Man nennt diese Körper deshalb auch **Primkörper**. Die endlichen Primkörper werden meist $\mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p$ notiert, mit einem F für "field" oder "finite". Die Notation $\mathbb{F}_q$ verwendet man allerdings auch allgemeiner mit einer Primzahlpotenz q im Index als Bezeichnung für "den endlichen Körper mit q Elementen", den wir erst in **??** kennenlernen werden, und der weder als Ring noch als abelsche Gruppe isomorph ist zu $\mathbb{Z}/q\mathbb{Z}$.

Beweis. 1. Für $m = 0$ ist $\mathbb{Z}/m\mathbb{Z} \cong \mathbb{Z}$ offensichtlich ein Integritätsbereich. Für m eine Primzahl ist $\mathbb{Z}/m\mathbb{Z}$ ein Integritätsbereich, da eine Primzahl nach 2.2.21 nur dann ein Produkt teilen kann, wenn sie bereits einen der Faktoren teilt. Für $m = 1$ ist $\mathbb{Z}/m\mathbb{Z}$ der Nullring und damit kein Integritätsbereich. Für $m > 1$ keine Primzahl faktorisieren wir $m = ab$ mit $1 < a, b < m$ und erhalten $0 = \bar{a}\bar{b}$ aber $\bar{a} \neq 0$, $\bar{b} \neq 0$. Mithin hat dann $\mathbb{Z}/m\mathbb{Z}$ von Null verschiedene Nullteiler, und diese können offensichtlich keine Einheiten sein.

2. Es muß nur noch gezeigt werden, daß für jede Primzahl p der Ring $\mathbb{Z}/p\mathbb{Z}$ ein Körper ist, daß also jedes von Null verschiedene Element $a \neq 0$ ein multiplikatives Inverses besitzt. Da $\mathbb{Z}/p\mathbb{Z}$ nullteilerfrei ist, muß jedoch die Multiplikation mit jedem Element $a \neq 0$ injektiv und damit bijektiv sein, also gibt es b mit $ab = 1$. □

Definition 2.2.29. Gegeben ein Ring R gibt es nach 2.1.20 genau einen Ringhomomorphismus $\mathbb{Z} \to R$. Dessen Kern ist nach 2.2.8 eine Untergruppe von $\mathbb{Z}$ und hat nach 2.2.13 folglich die Gestalt $m\mathbb{Z}$ für genau ein $m \in \mathbb{N}$. Diese natürliche Zahl m nennt man die **Charakteristik des Rings** R und notiert sie $m = \operatorname{char} R$. Wir haben also etwa $\operatorname{char} \mathbb{Q} = \operatorname{char} \mathbb{R} = \operatorname{char} \mathbb{C} = 0$ und $\operatorname{char}(\mathbb{Z}/p\mathbb{Z}) = p$.

2.2.30. Es ist leicht zu sehen, daß die Charakteristik eines Körpers, wenn sie nicht Null ist, stets eine Primzahl sein muß: Hätten wir sonst einen Körper der Charakteristik $m = ab > 0$ mit natürlichen Zahlen $a < m$ und $b < m$, so wären die Bilder von a und b in unserem Körper k von Null verschiedene Elemente mit Produkt Null. Widerspruch!

Übung 2.2.31. Sei R ein kommutativer Ring, dessen Charakteristik eine Primzahl p ist, für den es also einen Ringhomomorphismus $\mathbb{Z}/p\mathbb{Z} \to R$ gibt. Man zeige, daß dann die sogenannte **Frobenius-Abbildung** $F : R \to R$, $a \mapsto a^p$ ein Ringhomomorphismus von R in sich selber ist. Hinweis: Man verwende, daß die binomische Formel I.3.3.4 offensichtlich auch in jedem kommutativen Ring gilt, ja sogar für je zwei Elemente a, b eines Rings mit $ab = ba$.

Übung 2.2.32. Sei p eine Primzahl. Eine abelsche Gruppe G kann genau dann mit der Struktur eines $\mathbb{F}_p$-Vektorraums versehen werden, wenn in additiver Notation gilt $pg = 0$ für alle $g \in G$, und die fragliche Vektorraumstruktur ist dann durch die Gruppenstruktur eindeutig bestimmt.

Übung 2.2.33. Wieviele Untervektorräume hat ein zweidimensionaler Vektorraum über einem Körper mit fünf Elementen? Wieviele angeordnete Basen?

Übung 2.2.34. Gegeben ein Vektorraum über einem endlichen Primkörper sind seine Untervektorräume genau die Untergruppen der zugrundeliegenden abelschen Gruppe.

Übung 2.2.35. Man zeige: In jedem endlichen Körper ist das Produkt aller von Null verschiedenen Elemente (-1). Hinweis: Man zeige zunächst, daß nur die Elemente ± 1 ihre eigenen Inversen sind. Als Spezialfall erhält man $(p-1)! \equiv -1 \pmod{p}$ für jede Primzahl p. Diese Aussage wird manchmal auch als **Satz von Wilson** zitiert.

2.2.36. Sei $m \geq 1$ eine natürliche Zahl. Eine Restklasse modulo m heißt eine **prime Restklasse** genau dann, wenn sie aus zu m teilerfremden Zahlen besteht. Wir zeigen in **??**, daß es in jeder primen Restklasse unendlich viele Primzahlen gibt. Im Fall $m = 10$ bedeutet das zum Beispiel, daß es jeweils unendlich viele Primzahlen gibt, deren Dezimaldarstellung mit einer der Ziffern $1, 3, 7$ und 9 endet.

Übung 2.2.37. Gegeben $m \geq 1$ sind die Einheiten des Restklassenrings $\mathbb{Z}/m\mathbb{Z}$ genau die Restklassen derjenigen Zahlen a mit $0 \leq a < m$, die zu m teilerfremd sind, in anderen Worten die primen Restklassen. In Formeln haben wir also $(\mathbb{Z}/m\mathbb{Z})^\times = \{\bar{a} \mid 0 \leq a < m,\ \langle m, a\rangle = \langle 1\rangle\}$. Hinweis: 2.2.15.

2.3 Polynome

2.3.1. Ist k ein Ring, so bildet die Menge $k[X]$ aller "formalen Ausdrücke" der Gestalt $a_n X^n + \ldots + a_1 X + a_0$ mit $a_i \in k$ unter der offensichtlichen Addition und Multiplikation einen Ring, den **Polynomring** über k in einer Veränderlichen X, und wir haben eine offensichtliche Einbettung $\mathrm{can} : k \hookrightarrow k[X]$. Die Herkunft der Bezeichnung diskutieren wir in **??**. Unsere Beschreibung ist hoffentlich verständlich, sie ist aber nicht so exakt, wie eine Definition es sein sollte. Deshalb geben wir auch noch eine exakte Variante.

Definition 2.3.2. Sei k ein Ring. Wir bezeichnen mit $k[X]$ die Menge aller Abbildungen $\varphi : \mathbb{N} \to k$, die nur an endlich vielen Stellen von Null verschiedene Werte annehmen, und definieren auf $k[X]$ eine Addition und eine

Multiplikation durch die Regeln

$$\begin{aligned}(\varphi+\psi)(n) &= \varphi(n)+\psi(n)\\ (\varphi\cdot\psi)(n) &= \textstyle\sum_{i+j=n}\varphi(i)\psi(j)\end{aligned}$$

Mit diesen Verknüpfungen wird $k[X]$ ein Ring, und ordnen wir jedem $a \in k$ die Abbildung $\mathbb{N} \to k$ zu, die bei 0 den Wert a annimmt und sonst den Wert Null, so erhalten wir eine Einbettung $\text{can} : k \hookrightarrow k[X]$, die wir schlicht $a \mapsto a$ notieren. Bezeichnen wir mit X die Abbildung $\mathbb{N} \to k$, die bei 1 den Wert 1 annimmt und sonst nur den Wert Null, so können wir jede Abbildung $\varphi \in k[X]$ eindeutig schreiben in der Form $\varphi = \sum_\nu \varphi(\nu)X^\nu$ und sind auf einem etwas formaleren Weg wieder am selben Punkt angelangt.

2.3.3. Die wichtigste Eigenschaft eines Polynomrings ist, daß man "für die Variable etwas einsetzen darf". Das wollen wir nun formal korrekt aufschreiben. Wir sagen, zwei Elemente a und b eines Rings **kommutieren** genau dann, wenn gilt $ab = ba$.

Proposition 2.3.4 (**Einsetzen in Polynome**)**.** *Seien k und B Ringe und $\varphi : k \to B$ ein Ringhomomorphismus und $b \in B$ ein Element, das mit jedem Element $a \in \varphi(k)$ kommutiert. So gibt es genau eine Erweiterung $\tilde{\varphi} = \tilde{\varphi}_b$ von φ zu einem Ringhomomorphismus $\tilde{\varphi} : k[X] \to B$ mit $\tilde{\varphi}(X) = b$.*

Beweis. Diese eindeutig bestimmte Abbildung $\tilde{\varphi}$ ist eben gegeben durch die Vorschrift $\tilde{\varphi}(a_nX^n + \ldots + a_1X + a_0) = \varphi(a_n)b^n + \ldots + \varphi(a_1)b + \varphi(a_0)$. □

2.3.5. Es ist üblich, das Bild unter $\tilde{\varphi}_b$ eines Polynoms $P \in k[X]$ abzukürzen als $\tilde{\varphi}_b(P) = P(b)$. So schreiben wir im Fall eines kommutativen Rings k zum Beispiel $P(A)$ für die Matrix, die ensteht beim Einsetzen einer quadratischen Matrix A in das Polynom P. In diesem Fall hätten wir $B = \mathrm{M}(n \times n; k)$ und φ wäre der Ringhomomorphismus, die jedem $a \in k$ das a-fache der Einheitsmatrix zuordnet.

Übung 2.3.6. Welche Matrix entsteht beim Einsetzen der quadratischen Matrix $\left(\begin{smallmatrix}0 & 1\\ -1 & 0\end{smallmatrix}\right)$ in das Polynom $X^2 + 1$?

Definition 2.3.7. Sei k ein Kring und $P \in k[X]$ ein Polynom. Ein Element $a \in k$ heißt eine **Nullstelle** oder auch eine **Wurzel** von P genau dann, wenn gilt $P(a) = 0$.

Definition 2.3.8. Sei k ein Ring. Jedem Polynom $P \in k[X]$ ordnen wir seinen **Grad** (engl. degree, franz. degré) $\operatorname{grad} P \in \mathbb{N} \cup \{-\infty\}$ zu durch die Vorschrift

$$\begin{aligned}\operatorname{grad} P &= n && \text{falls } P = a_nX^n + \ldots + a_0 \text{ mit } a_n \neq 0;\\ \operatorname{grad} P &= -\infty && \text{für } P \text{ das Nullpolynom.}\end{aligned}$$

Für ein von Null verschiedenes Polynom $P = a_nX^n + \ldots + a_1X + a_0$ mit $n = \operatorname{grad} P$ nennt man $a_n \in k\backslash 0$ seinen **Leitkoeffizienten**. Den Leitkoeffizienten des Nullpolynoms definieren wir als die Null von k. Ein Polynom heißt **normiert** genau dann, wenn sein Leitkoeffizient 1 ist. Das Nullpolynom ist demnach nur über dem Nullring normiert.

Lemma 2.3.9. *Ist k ein nullteilerfreier Ring, so ist auch der Polynomring $k[X]$ nullteilerfrei und es gilt* $\operatorname{grad}(PQ) = \operatorname{grad} P + \operatorname{grad} Q$.

Beweis. Ist k nullteilerfrei, so ist offensichtlich der Leitkoeffizient von PQ das Produkt der Leitkoeffizienten von P und Q. □

Übung 2.3.10. Ist k ein Integritätsbereich, so induziert die kanonische Einbettung $k \hookrightarrow k[X]$ auf den Einheitengruppen eine Bijektion $k^\times \xrightarrow{\sim} (k[X])^\times$. Im Ring $\mathbb{Z}/4\mathbb{Z}[X]$ ist aber auch $\bar{1} + \bar{2}X$ eine Einheit.

Lemma 2.3.11 (**Teilen mit Rest in Polynomringen**). *Sei k ein Ring. Gegeben Polynome $P, Q \in k[X]$ mit $Q = X^d + \ldots + a_1X + a_0$ für ein $d \geq 0$ gibt es Polynome A, R mit $P = AQ + R$ und* $\operatorname{grad} R < d$. *Ist k nullteilerfrei, so sind diese Polynome A und R sogar eindeutig bestimmt.*

2.3.12. Ein explizites Beispiel wird in **??** ausgearbeitet.

Beweis. Wir suchen A mit $\operatorname{grad}(P - AQ)$ kleinstmöglich. Gälte dann noch $\operatorname{grad}(P - AQ) \geq d$, sagen wir $P - AQ = aX^r + \ldots$ mit $a \neq 0$ und $r \geq d$, so hätte $P - (A + aX^{r-d})Q$ echt kleineren Grad als R, im Widerspruch zur Wahl von A. Das zeigt die Existenz. Für den Nachweis der Eindeutigkeit gehen wir aus von einer weiteren Gleichung $P = A'Q + R'$ mit $\operatorname{grad} R' < d$. Es folgt zunächst $(A - A')Q = R' - R$ und mit 2.3.9 weiter $A - A' = 0$ und dann auch $R' - R = 0$. □

Korollar 2.3.13 (**Abspalten von Linearfaktoren bei Nullstellen**). *Sei k ein Kring und $P \in k[X]$ ein Polynom. Genau dann ist $a \in k$ eine Nullstelle des Polynoms P, wenn $(X - a)$ das Polynom P teilt.*

Beweis. Nach 2.3.11 finden wir ein Polynom $A \in k[X]$ und eine Konstante $b \in k$ mit $P = A(X - a) + b$. □

Satz 2.3.14 (**Zahl der Nullstellen eines Polynoms**). *Ist k ein Körper oder allgemeiner ein kommutativer Integritätsbereich, so hat ein von Null verschiedenes Polynom $P \in k[X]$ höchstens* $\operatorname{grad} P$ *Nullstellen in k.*

Beweis. Ist $a \in k$ eine Nullstelle, so haben wir $P = A(X - a)$ mit $\operatorname{grad} A = \operatorname{grad} P - 1$. Eine von a verschiedene Nullstelle von P ist für k nullteilerfrei notwendig eine Nullstelle von A und der Satz folgt mit Induktion. □

2.3.15. Ist k ein Körper oder allgemeiner ein kommutativer Integritätsbereich, $P \in k[X]$ ein Polynom und $\lambda \in k$ eine Nullstelle von P, so nennen wir das Supremum über alle $n \in \mathbb{N}$ mit $(X-\lambda)^n | P(X)$ die **Vielfachheit der Nullstelle** λ oder auch ihre **Ordnung**.. Ganz genauso wie eben zeigt man auch, daß die Zahl der mit ihren Vielfachheiten gezählten Nullstellen eines von Null verschiedenen Polynoms beschränkt ist durch seinen Grad.

Definition 2.3.16. Ein Körper k heißt **algebraisch abgeschlossen** genau dann, wenn jedes nichtkonstante Polynom $P \in k[X] \setminus k$ mit Koeffizienten in unserem Körper auch eine Nullstelle in unserem Körper hat.

Beispiel 2.3.17. Der Körper $\mathbb{C}$ der komplexen Zahlen ist algebraisch abgeschlossen. Das ist die Aussage des sogenannten "Fundamentalsatzes der Algebra". den wir in der Analysis als **??** beweisen.

Satz 2.3.18. *Ist k ein algebraisch abgeschlossener Körper, so hat jedes Polynom $P \in k[X]$ eine* **Zerlegung in Linearfaktoren** *der Gestalt*

$$P(X) = c(X-\lambda_1)\dots(X-\lambda_n)$$

mit $n \geq 0$ und $c, \lambda_1, \dots, \lambda_n \in k$. Ist P nicht das Nullpolynom, so ist diese Zerlegung eindeutig bis auf die Reihenfolge der Faktoren.

2.3.19. Gegeben eine Nullstelle μ von P heißt die Zahl der Indizes i mit $\lambda_i = \mu$ die **Vielfachheit der Nullstelle** μ.

Beweis. Ist P ein konstantes Polynom, so ist nichts zu zeigen. Ist P nicht konstant, so gibt es nach Annahme eine Nullstelle $\lambda \in k$ von P und wir finden genau ein Polynom $\tilde{P}$ mit $P(X) = (X-\lambda)\tilde{P}(X)$. Der Satz folgt durch vollständige Induktion über den Grad von P. □

Korollar 2.3.20 (**Faktorisierung reeller Polynome**)**.** *Jedes Polynom P mit reellen Koeffizienten der Gestalt $P(x) = a_n x^n + \dots + a_1 x + a_0$ mit $n \geq 0$ und $a_n \neq 0$ besitzt eine Zerlegung in Faktoren der Gestalt*

$$P(x) = c(x-\lambda_1)\dots(x-\lambda_r)(x^2+\mu_1 x+\nu_1)\dots(x^2+\mu_s x+\nu_s)$$

mit $c, \lambda_1, \dots, \lambda_r, \mu_1, \dots, \mu_s, \nu_1, \dots, \nu_s \in \mathbb{R}$ derart, daß die quadratischen Faktoren keine reellen Nullstellen haben. Diese Zerlegung ist eindeutig bis auf die Reihenfolge der Faktoren.

Beweis. Da unser Polynom stabil ist unter der komplexen Konjugation, müssen sich seine (mit ihren Vielfachheiten genommenen) komplexen Nullstellen so durchnummerieren lassen, daß $\lambda_1, \dots, \lambda_r$ reell sind und daß eine gerade Zahl nicht reeller Nullstellen übrigbleibt mit $\lambda_{r+2i} = \bar{\lambda}_{r+2i+1}$. Die Produkte $(x-\lambda_{r+2i})(x-\lambda_{r+2i+1})$ haben dann reelle Koeffizienten, da sie ja stabil sind unter der komplexen Konjugation, haben jedoch keine reellen Nullstellen. □

Übung 2.3.21. Ein reelles Polynom hat bei $\lambda \in \mathbb{R}$ eine mehrfache Nullstelle genau dann, wenn auch seine Ableitung bei λ verschwindet.

Übung 2.3.22. Gegeben ein reelles Polynom, dessen komplexe Nullstellen bereits sämtlich reell sind, ist jede Nullstelle seiner Ableitung, die keine Nullstelle der Funktion selbst ist, eine einfache Nullstelle der Ableitung. Hinweis: Zwischen je zwei Nullstellen unserer Funktion muß mindestens eine Nullstelle ihrer Ableitung liegen.

Übung 2.3.23. Man zeige: Die rationalen Nullstellen eines normierten Polynoms mit ganzzahligen Koeffizienten $P \in \mathbb{Z}[X]$ sind bereits alle ganz, in Formeln folgt aus $P(\lambda) = 0$ für $\lambda \in \mathbb{Q}$ also bereits $\lambda \in \mathbb{Z}$.

2.3.24. Ähnlich wie den Polynomring in einer Variablen 2.3.2 konstruiert man auch Polynomringe in mehr Variablen. Ist die Zahl der Variablen endlich, so kann man induktiv definieren

$$R[X_1, \ldots, X_n] = (R[X_1, \ldots, X_{n-1}])[X_n]$$

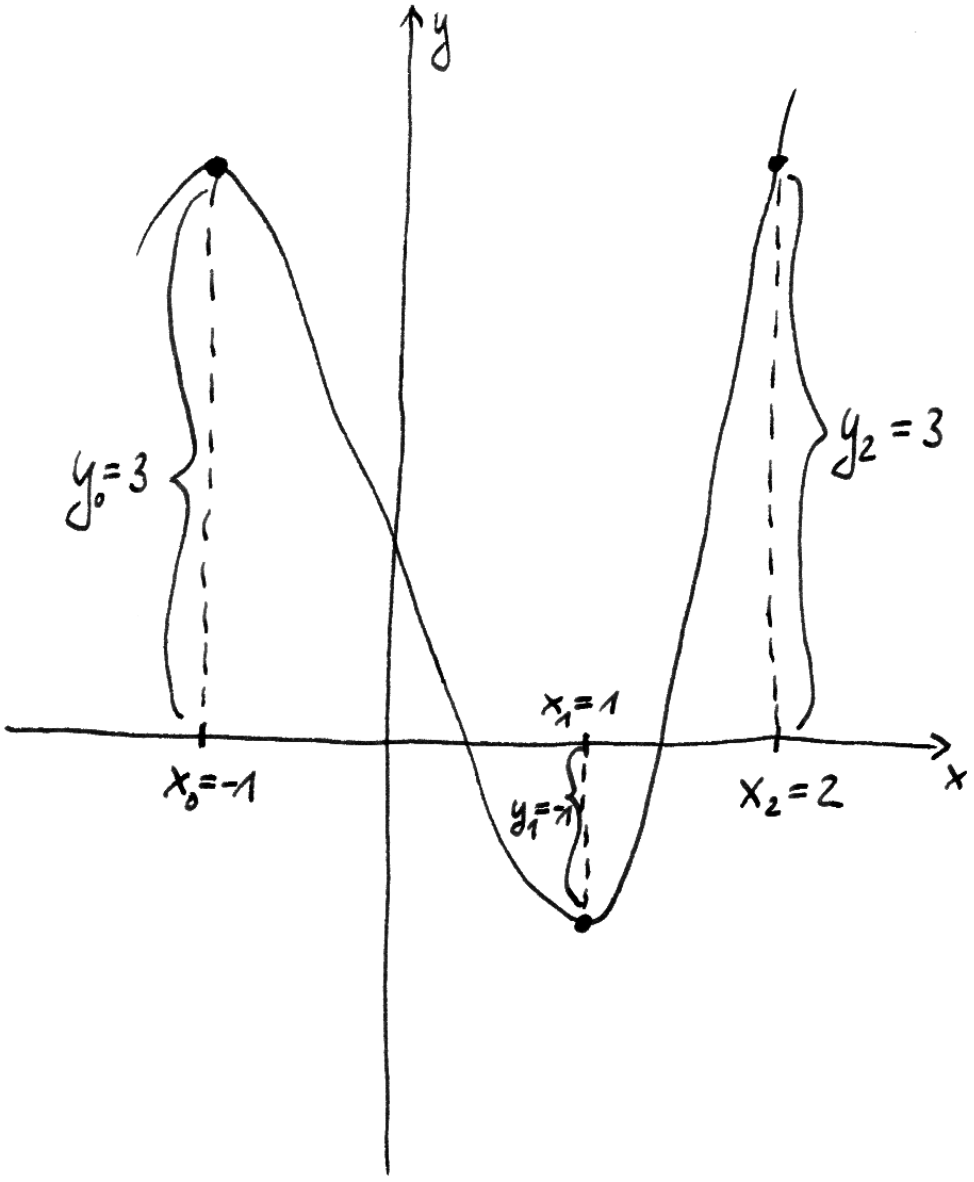
Man kann aber auch für eine beliebige Menge I den Polynomring $R[X_i]_{i\in I}$ bilden als die Menge aller "endlichen formalen Linearkombinationen mit Koeffizienten aus R von endlichen Monomen in den X_i". Ich verzichte an dieser Stelle auf eine formale Definition.

Übung 2.3.25. Gegeben ein Ring k bilden auch die **formalen Potenzreihen mit Koeffizienten in** k der Gestalt $\sum_{n\geq 0} a_n X^n$ mit $a_n \in k$ einen Ring, der meist $k[[X]]$ notiert wird. Man gebe eine exakte Definition dieses Rings und zeige, daß seine Einheiten genau diejenigen Potenzreihen sind, deren konstanter Term eine Einheit in k ist.

Übung 2.3.26. Gegeben ein Ring k bilden auch die **Laurentreihen mit Koeffizienten in** k der Gestalt $\sum_{n\geq -N} a_n X^n$ mit $a_n \in k$ und $N \in \mathbb{N}$ einen Ring, der meist $k((X))$ notiert wird. Man gebe eine exakte Definition dieses Rings und zeige, daß seine Einheiten genau diejenigen Potenzreihen sind, bei denen der Koeffizient der kleinsten mit von Null verschiedenem Koeffizienten auftauchenden Potenz von X eine Einheit in k ist. Insbesondere ist im Fall eines Körpers k auch $k((X))$ ein Körper.

Lemma 2.3.27 (Interpolation durch Polynome). *Seien k ein Körper, $x_0, \ldots, x_n \in k$ paarweise verschiedene "Stützstellen" und $y_0, \ldots, y_n \in k$ beliebig vorgegeben. So gibt es genau ein Polynom $P \in k[X]$ vom Grad $\leq n$ mit $P(x_0) = y_0, \ldots, P(x_n) = y_n$.*

Beweis. Zunächst ist sicher $(X - x_1) \ldots (X - x_n) = A_0(X)$ ein Polynom vom Grad n, das bei $x_1, \ldots, x_n$ verschwindet und an allen anderen Stellen von Null



Das Polynom $P(X) = X^3 - 3X + 1$ mit reellen Koeffizienten, das die an den Stützstellen $-1, 1, 2$ vorgegebenen Werte $3, -1, 3$ interpoliert.

verschieden ist, insbesondere auch bei x_0. Dann ist $L_0(X) = A_0(X)/A_0(x_0)$ ein Polynom vom Grad n, das bei x_0 den Wert Eins annimmt und bei $x_1, \ldots, x_n$ verschwindet. In derselben Weise konstruieren wir auch Polynome $L_1(X), \ldots, L_n(X)$ und erhalten ein mögliches Interpolationspolynom als

$$P(X) = y_0 L_0(X) + \ldots + y_n L_n(X) = \sum_{i=0}^{n} y_i \frac{\prod_{j\neq i}(X - x_j)}{\prod_{j\neq i}(x_i - x_j)}$$

Das zeigt die Existenz. Ist Q eine weitere Lösung derselben Interpolationsaufgabe vom Grad $\leq n$, so ist $P - Q$ ein Polynom vom Grad $\leq n$ mit $n+1$ Nullstellen, eben bei den Stützstellen $x_0, \ldots, x_n$. Wegen **??** muß dann $P - Q$ das Nullpolynom sein, und das zeigt die Eindeutigkeit. □

2.3.28. Um die bisher eingeführten algebraischen Konzepte anschaulicher zu machen, will ich sie in Bezug setzen zu geometrischen Konzepten. Ist k ein Kring, so können wir jedem Polynom $f \in k[X_1, \ldots, X_n]$ die Funktion $\tilde{f} : k^n \to k$, $(x_1, \ldots, x_n) \mapsto f(x_1, \ldots, x_n)$ zuordnen. Wir erhalten so einen Ringhomomorphismus

$$k[X_1, \ldots, X_n] \to \mathrm{Ens}(k^n, k)$$

Dieser Homomorphismus ist im Allgemeinen weder injektiv noch surjektiv. Schon für $n = 1$, $k = \mathbb{R}$ läßt sich ja keineswegs jede Abbildung $\mathbb{R} \to \mathbb{R}$ durch ein Polynom beschreiben, und im Fall eines endlichen Körpers k kann für $n \geq 1$ unsere k-lineare Auswertungsabbildung vom unendlichdimensionalen k-Vektorraum $k[X_1, \ldots, X_n]$ in den endlichdimensionalen k-Vektorraum $\mathrm{Ens}(k^n, k)$ unmöglich injektiv sein. Wir haben jedoch:

Satz 2.3.29. *1. Ist k ein unendlicher Körper oder allgemeiner ein unendlicher nullteilerfreier Kring, so ist $k[X_1, \ldots, X_n] \to \mathrm{Ens}(k^n, k)$ injektiv.*

2. Ist k ein endlicher Körper, so ist $k[X_1, \ldots, X_n] \to \mathrm{Ens}(k^n, k)$ surjektiv.

2.3.30. Den Kern der Surjektion in Teil 2 beschreibt Übung 6.5.8.

Beweis. 1. Durch Induktion über n. Der Fall $n = 0$ ist eh klar. Für $n = 1$ folgt die Behauptung aus der Erkenntnis, das jedes von Null verschiedene Polynom in $k[X]$ nur endlich viele Nullstellen in k haben kann. Der Kern der Abbildung

$$k[X] \to \mathrm{Ens}(k, k)$$

besteht also nur aus dem Nullpolynom. Für den Induktionsschritt setzen wir $X_n = Y$ und schreiben unser Polynom in der Gestalt

$$P = a_d Y^d + \ldots + a_1 Y + a_0$$

mit $a_i \in k[X_1, \ldots, X_{n-1}]$. Halten wir $(x_1, \ldots, x_{n-1}) = x \in k^{n-1}$ fest, so ist $a_d(x)Y^d + \ldots + a_1(x)Y + a_0(x) \in k[Y]$ das Nullpolynom nach dem Fall $n = 1$. Also verschwinden $a_d(x), \ldots, a_1(x), a_0(x)$ für alle $x \in k^{n-1}$, mit Induktion sind somit alle a_i schon das Nullpolynom und wir haben $P = 0$.

2. Das bleibt dem Leser überlassen. Man mag sich beim Beweis an 2.3.27 orientieren. Wir folgern in **??** eine allgemeinere Aussage aus dem abstrakten chinesischen Restsatz. □

Übung 2.3.31. Sei k ein unendlicher Körper. Verschwindet ein Polynom im Polynomring in d Variablen über k auf einer affinen Hyperebene in k^d, so wird es von der, bis auf einen Skalar eindeutig bestimmten, linearen Gleichung besagter Hyperebene geteilt. Hinweis: Ohne Beschränkung der Allgemeinheit mag man unsere Hyperebene als eine der Koordinatenhyperebenen annehmen. Man zeige auch allgemeiner: Verschwindet ein Polynom in d Veränderlichen über einem unendlichen Körper auf der Vereinigung der paarweise verschiedenen affinen Hyperebenen $H_1, \ldots, H_n \subset k^d$, so wird es vom Produkt der linearen Gleichungen unserer Hyperebenen geteilt.

2.4 Äquivalenzrelationen

Definition 2.4.1. Eine Relation $R \subset X \times X$ auf einer Menge X im Sinne von **??** heißt eine **Äquivalenzrelation** genau dann, wenn für alle $x, y, z \in X$ gilt

1. **Transitivität:** $(xRy$ und $yRz) \Rightarrow xRz$;
2. **Symmetrie:** $xRy \Leftrightarrow yRx$;
3. **Reflexivität:** xRx.

2.4.2. Gegeben eine Äquivalenzrelation $\sim$ auf einer Menge X betrachtet man für $x \in X$ die Menge $A(x) = \{z \in X \mid z \sim x\}$ und nennt sie die **Äquivalenzklasse** von x. Ein Element einer Äquivalenzklasse nennt man auch einen **Repräsentanten** der Klasse. Eine Teilmenge $Z \subset X$, die aus jeder Äquivalenzklasse genau ein Element enthält, heißt ein **Repräsentantensystem**. Aufgrund der Reflexivität gilt $x \in A(x)$, und man sieht leicht, daß für $x, y \in X$ die folgenden drei Aussagen gleichbedeutend sind:

1. $x \sim y$;
2. $A(x) = A(y)$;
3. $A(x) \cap A(y) \neq \emptyset$.

2.4.3. Gegeben eine Äquivalenzrelation $\sim$ auf einer Menge X bezeichnen wir die Menge aller Äquivalenzklassen, eine Teilmenge der Potenzmenge $\mathcal{P}(X)$, mit

$$(X/\sim) = \{A(x) \mid x \in X\}$$

und haben eine kanonische Abbildung $\text{can} : X \to (X/\sim)$, $x \mapsto A(x)$. Ist weiter $f : X \to Z$ eine Abbildung mit $x \sim y \Rightarrow f(x) = f(y)$, so gibt es genau eine Abbildung $\bar{f} : (X/\sim) \to Z$ mit $f = \bar{f} \circ \text{can}$. Wir zitieren diese Eigenschaft manchmal als die **universelle Eigenschaft des Raums der Äquivalenzklassen**.

2.4.4. Die kanonische Abbildung $\text{can} : X \to (X/\sim)$ ist stets eine Surjektion. Ist umgekehrt $f : X \twoheadrightarrow Z$ eine Surjektion und betrachten wir auf X die Relation $x \sim y \Leftrightarrow f(x) = f(y)$, so ist besagte Relation eine Äquivalenzrelation und die kanonische Abbildung $\bar{f}$ liefert eine Bijektion $\bar{f} : (X/\sim) \xrightarrow{\sim} Z$.

Beispiel 2.4.5. Gegeben eine ganze Zahl $m \in \mathbb{Z}$ ist unser "kongruent modulo m" aus 2.1.7 eine Äquivalenzrelation $\sim$ auf $\mathbb{Z}$ und die zugehörigen Äquivalenzklassen sind genau unsere Restklassen von dort, so daß wir also $\mathbb{Z}/\sim = \mathbb{Z}/m\mathbb{Z}$ erhalten.

2.4.6. Sind $R \subset X \times X$ und $S \subset Y \times Y$ Äquivalenzrelationen, so auch das Bild von $(R \times S) \subset (X \times X) \times (Y \times Y)$ unter der durch Vertauschen der mittleren Einträge gegebenen Identifikation $(X \times X) \times (Y \times Y) \xrightarrow{\sim} (X \times Y) \times (X \times Y)$. Wir notieren diese Äquivalenzrelation auf dem Produkt kurz $R \times S$.

2.5 Quotientenkörper

Das war in der Vorlesung 2008/09 nicht dran.

Definition 2.5.1. Gegeben ein kommutativer Integritätsbereich R konstruieren wir seinen **Quotientenkörper**

$$\text{Quot}(R)$$

wie folgt: Wir betrachten die Menge $R \times (R \backslash 0)$ und definieren darauf eine Relation $\sim$ durch die Vorschrift

$$(a, s) \sim (b, t) \text{ genau dann, wenn gilt } at = bs.$$

Diese Relation ist eine Äquivalenzrelation, wir bezeichnen die Menge der Äquivalenzklassen mit $\text{Quot}(R)$ und die Äquivalenzklasse von (a, s) mit $\frac{a}{s}$ oder a/s. Dann definieren wir auf $\text{Quot}(R)$ Verknüpfungen $+$ und $\cdot$ durch die Regeln

$$\frac{a}{s} + \frac{b}{t} = \frac{at + bs}{st} \quad \text{und} \quad \frac{a}{s} \cdot \frac{b}{t} = \frac{ab}{st}$$

und überlassen dem Leser den Nachweis, daß diese Verknüpfungen wohldefiniert sind und $\mathrm{Quot}(R)$ zu einem Körper machen und daß die Abbildung $\mathrm{can} : R \to \mathrm{Quot}(R)$, $r \mapsto r/1$ ein Ringhomomorphismus ist. Sie heißt die **kanonische Einbettung** unseres Integritätsbereichs in seinen Quotientenkörper.

2.5.2. Auf Englisch bezeichnet man den Quotientenkörper als **fraction field** und auf Französisch als **corps de fractions**. Dort verwendet man folgerichtig statt unserer Notation $\mathrm{Quot}(R)$ die Notation $\mathrm{Frac}(R)$.

Beispiel 2.5.3. Der Körper der rationalen Zahlen $\mathbb{Q}$ ist formal definiert als der Quotientenkörper des Rings der ganzen Zahlen, in Formeln $\mathbb{Q} = \mathrm{Quot}\,\mathbb{Z}$. Sicher wäre es unter formalen Aspekten betrachtet eigentlich richtig gewesen, diese Definition schon viel früher zu geben. Es schien mir jedoch didaktisch ungeschickt, gleich am Anfang derart viel Zeit und Formeln auf die exakte Konstruktion einer Struktur zu verwenden, die den meisten Studenten bereits zu Beginn ihres Studiums hinreichend vertraut sein sollte.

Satz 2.5.4 (Universelle Eigenschaft des Quotientenkörpers). *Sei R ein kommutativer Integritätsbereich. Ist $\varphi : R \to A$ ein Ringhomomorphismus, unter dem jedes von Null verschiedene Element von R auf eine Einheit abgebildet wird, so faktorisiert φ eindeutig über* $\mathrm{Quot}\,R$*, es gibt also in Formeln genau einen Ringhomomorphismus $\tilde{\varphi} : \mathrm{Quot}\,R \to A$ mit $\varphi = \tilde{\varphi} \circ \mathrm{can}$.*

Beweis. Für jedes mögliche $\tilde{\varphi}$ muß gelten $\tilde{\varphi}(r/s) = \varphi(r)\varphi(s)^{-1}$, und das zeigt bereits die Eindeutigkeit von $\tilde{\varphi}$. Um auch seine Existenz zu zeigen, betrachten wir die Abbildung $\hat{\varphi} : R \times (R\backslash 0) \to A$ gegeben durch $\hat{\varphi}(r,s) = \varphi(r)\varphi(s)^{-1}$ und prüfen, daß sie konstant ist auf Äquivalenzklassen. Dann muß sie nach 2.4.3 eine wohlbestimmte Abbildung $\mathrm{Quot}\,R \to A$ induzieren, von der der Leser leicht selbst prüfen wird, daß sie ein Ringhomomorphismus ist. □

2.5.5. Ist k ein Körper, so bezeichnet man den Quotientenkörper des Polynomrings mit $\mathrm{Quot}\,k[X] = k(X)$ und nennt seine Elemente **rationale Funktionen**. Ähnlich schreibt man bei mehreren Veränderlichen

$$\mathrm{Quot}\,k[X_1, \ldots, X_n] = k(X_1, \ldots, X_n)$$

Ist k unendlich, so kann man sich die Elemente von $k(X)$ als "fast überall definierte k-wertige Funktionen auf k" vorstellen. Etwas formaler betrachten wir für eine beliebige Menge M auf dem Ring $\mathrm{Ens}(M,k)$ aller Abbildungen von M nach k die Äquivalenzrelation $\sim$ gegeben durch $f \sim g$ genau dann, wenn gilt $f(x) = g(x)$ an allen außer endlich vielen Stellen $x \in M$. Es ist hoffentlich klar, wie zwei Äquivalenzklassen zu addieren bzw. zu multiplizieren sind, und daß die Menge der Äquivalenzklassen $\mathrm{Ens}(M,k)/\sim$ so zu

einem Ring wird. Wir nennen ihn den Ring der "fast überall definierten k-wertigen Funktionen auf M". Dann liefert die universelle Eigenschaft 2.5.4 eine Einbettung

$$k(X) \hookrightarrow \{\text{fast überall definierte } k\text{-wertige Funktionen auf } k\}$$

Man definiert für jede rationale Funktion $f \in k(X)$ ihren **Definitionsbereich** $D(f) \subset k$ als die Menge aller Punkte $a \in k$ derart, daß f sich schreiben läßt als Quotient von zwei Polynomen $f = g/h$ mit $h(a) \neq 0$. Haben wir zwei solche Darstellungen $f = g/h = \hat{g}/\hat{h}$ mit $h(a) \neq 0 \neq \hat{h}(a)$, so gilt offensichtlich $g(a)/h(a) = \hat{g}(a)/\hat{h}(a)$ und wir definieren $f(a)$ als diesen gemeinsamen Wert. In diesem Sinne liefert also jedes $f \in k(X)$ eine Abbildung

$$f : D(f) \to k$$

Die endlich vielen Punkte außerhalb des Definitionsbereichs von f heißen die **Polstellen** von f.

Übung 2.5.6. Gegeben ein unendlicher Körper k und eine von Null verschiedene rationale Funktion $f \in k(X)^\times$ sind die Polstellen von f genau die Nullstellen von $(1/f)$, als da heißt, die Stellen aus dem Definitionsbereich von $(1/f)$, an denen diese Funktion den Wert Null annimmt.

Übung 2.5.7. Gegeben ein Körper k liefert die Einbettung $k[X] \hookrightarrow k[[X]] \hookrightarrow k((X))$ offensichtlich einen Ringhomomorphismus und mithin eine Einbettung $k(X) \hookrightarrow k((X))$. Man bestimme das Bild von $(1+X)^{-1}$ unter dieser Einbettung.

2.6 Das Signum einer Permutation

Definition 2.6.1. Die Gruppe aller Permutationen alias bijektiven Selbstabbildungen der Menge $\{1, 2, \ldots, n\}$ notieren wir

$$\mathcal{S}_n = \mathrm{Ens}^\times\{1, 2, \ldots, n\}$$

Nach I.2.2.24 hat diese Gruppe $n!$ Elemente, in Formeln $|\mathcal{S}_n| = n!$. Eine Permutation, die zwei Elemente unserer Menge vertauscht und ansonsten die Identität ist, nennt man eine **Transposition**.

Definition 2.6.2. Ein **Fehlstand** einer Permutation $\sigma \in \mathcal{S}_n$ ist ein Paar (i, j) mit $1 \leq i < j \leq n$ aber $\sigma(i) > \sigma(j)$. Die Zahl der Fehlstände heißt die **Länge** $l(\sigma)$ unserer Permutation, in Formeln

$$l(\sigma) = |\{(i, j) \mid i < j \text{ aber } \sigma(i) > \sigma(j)\}|$$

Das **Signum** einer Permutation ist definiert als die Parität der Zahl ihrer Fehlstände, in Formeln

$$\operatorname{sgn}(\sigma) = (-1)^{l(\sigma)}$$

Eine Permutation mit Signum $+1$ alias gerader Länge heißt eine **gerade Permutation**, eine Permutation mit Signum -1 alias ungerader Länge eine **ungerade Permutation**.

Beispiel 2.6.3. Die Identität von $\mathcal{S}_n$ ist jeweils die einzige Permutation der Länge Null. Die Transposition, die die Zahlen i und j vertauscht, hat die Länge $2|i-j|-1$, wie auch nebenstehendes Bild sofort zeigt, und ist also insbesondere stets ungerade.

Lemma 2.6.4. *Für jede natürliche Zahl n ist unser Signum ein Gruppenhomomorphismus* $\operatorname{sgn} : \mathcal{S}_n \to \{1,-1\}$ *von der symmetrischen Gruppe $\mathcal{S}_n$ in die zweielementige Gruppe der Vorzeichen, in Formeln gilt also*

$$\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau) \quad \forall \sigma,\tau \in \mathcal{S}_n$$

Beweis. Wir vereinbaren speziell für diesen Beweis für das Vorzeichen einer von Null verschiedenen ganzen Zahl $a \in \mathbb{Z}\backslash 0$ die Notation $[a] \in \{1,-1\}$. Damit können wir das Signum einer Permutation σ dann auch schreiben als $\operatorname{sgn}(\sigma) = \prod_{i<j}[\sigma(i)-\sigma(j)]$, und da für eine beliebige weitere Permutation τ auch die $\{\tau(i),\tau(j)\}$ für $i<j$ genau alle zweielementigen Teilmengen von $\{1,\ldots,n\}$ durchlaufen, gilt für eine beliebige weitere Permutation τ die Formel

$$\operatorname{sgn}(\sigma) = \prod_{i<j} \frac{[\sigma(\tau(i)) - \sigma(\tau(j))]}{[\tau(i)-\tau(j)]}$$

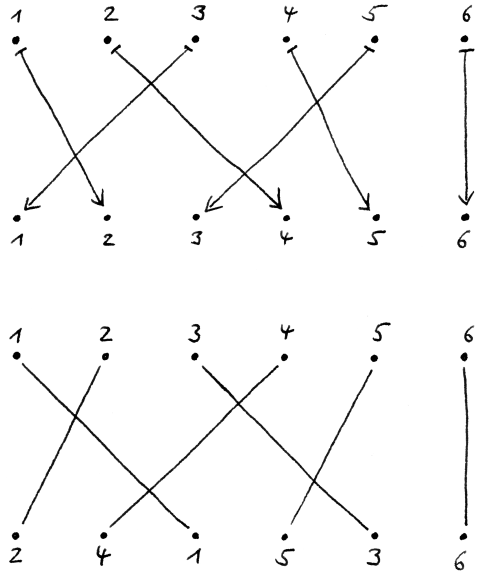
Mit diesen Erkenntnissen reduziert sich dann der Beweis auf die Rechnung

$$\prod_{i<j}[\sigma\tau(i)-\sigma\tau(j)] = \prod_{i<j} \frac{[\sigma(\tau(i)) - \sigma(\tau(j))]}{[\tau(i)-\tau(j)]} \prod_{i<j}[\tau(i)-\tau(j)] \qquad \square$$

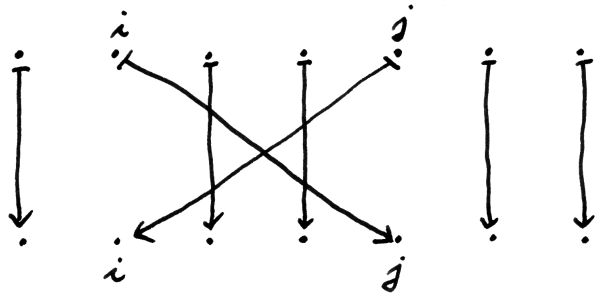
2.6.5. Für jedes n bilden die geraden Permutationen als Kern eines Gruppenhomomorphismus nach 2.2.11 eine Untergruppe von $\mathcal{S}_n$. Diese Gruppe heißt die **alternierende Gruppe** und wird A_n notiert.

Übung 2.6.6. Die Permutation $\sigma \in \mathcal{S}_n$, die i ganz nach vorne schiebt ohne die Reihenfolge der übrigen Elemente zu ändern, hat $(i-1)$ Fehlstände und folglich das Signum $\operatorname{sgn}(\sigma) = (-1)^{i-1}$.

Übung 2.6.7. Jede Permutation einer endlichen angeordneten Menge läßt sich darstellen als eine Verknüpfung von Transpositionen benachbarter Elemente.



Diese Bilder illustrieren zwei mögliche Anschauungen für die Länge einer Permutation, in diesem Fall der Permutation $\sigma \in \mathcal{S}_6$ mit $1 \mapsto 2$, $2 \mapsto 4$, $3 \mapsto 1$, $4 \mapsto 5$, $5 \mapsto 3$ und $6 \mapsto 6$: Im oberen Bild ist die Länge ganz offensichtlich die "Zahl der Kreuzungen von Abbildungspfeilen", in unserem Fall haben wir also $l(\sigma) = 4$. Im unteren Bild habe ich unter jede Zahl n jeweils $\sigma(n)$ geschrieben und dann gleiche Zahlen verbunden, und hier ist ähnlich $l(\sigma) = 4$ gerade die "Zahl der Kreuzungen solcher Verbindungslinien". Der Leser sei ermutigt, sich auch die Produktformel für das Signum 2.6.4 mithilfe dieser Bilder anschaulich zu machen.



Die Transposition, die i und j vertauscht, hat genau $2|i-j|-1$ Fehlstände. Insbesondere ist jede Transposition ungerade.

Übung 2.6.8. Ist T eine endliche Menge, so gibt es genau einen Gruppenhomomorphismus

$$\operatorname{sign} : \mathrm{Ens}^\times(T) \to \{1, -1\}$$

derart, daß für jede Bijektion $\beta : \{1, \ldots, n\} \xrightarrow{\sim} T$ und alle $\tau \in \mathrm{Ens}^\times(T)$ gilt $\operatorname{sign}(\tau) = \operatorname{sgn}(\beta^{-1} \circ \tau \circ \beta)$. Wir nennen ihn auch in dieser Allgemeinheit das **Signum** und kürzen ihn wieder mit sign = sgn ab. Auch in dieser Allgemeinheit nennen wir eine Permutation mit Signum $+1$ **gerade**, und eine Permutation mit Signum -1 **ungerade**. Es ist allerdings nicht mehr sinnvoll, in dieser Allgemeinheit von der "Länge" einer Permutation zu reden.

2.7 Die Determinante

Definition 2.7.1. Sei k ein Kring und $n \in \mathbb{N}$. Die **Determinante** ist die Abbildung $\det : \mathrm{M}(n \times n; k) \to k$ von den quadratischen Matrizen mit Einträgen in unserem Kring in besagten Kring selbst, die gegeben wird durch die Vorschrift

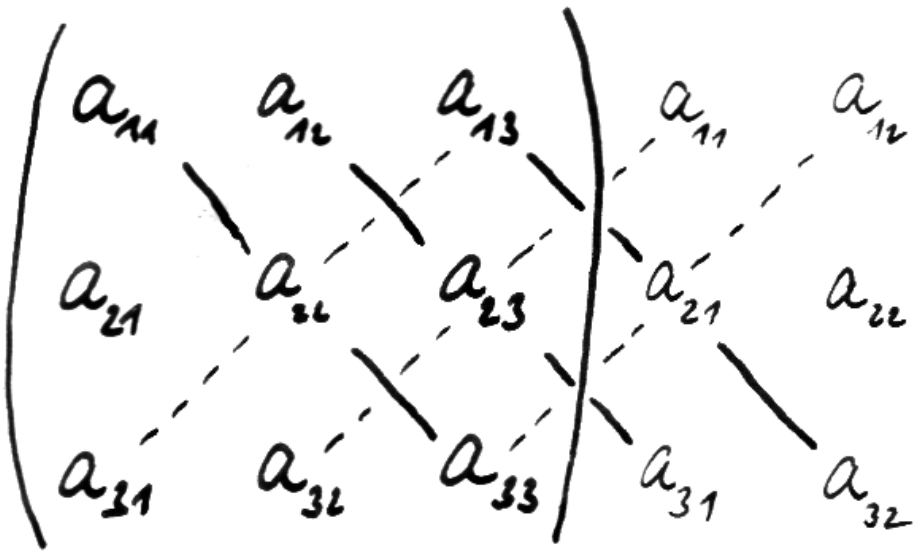
$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \quad \mapsto \quad \det A = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \ldots a_{n\sigma(n)}$$

Summiert wird über alle Permutationen von n, und der Vorfaktor $\operatorname{sgn}(\sigma)$ meint das Signum der Permutation σ. Unsere Formel heißt die **Leibniz-Formel**. Für den Extremfall $n = 0$ der "leeren Matrix" ist zu verstehen, daß ihr die Determinante 1 zugeordnet wird: Formal gibt es genau eine Permutation der leeren Menge, deren Signum ist Eins, und dies Signum wird multipliziert mit dem leeren Produkt, das nach unseren Konventionen den Wert Eins hat.

2.7.2. Wie wir in 2.7.16 sehen werden, bestimmt alias determiniert die Determinante, ob ein quadratisches lineares Gleichungssystem eindeutig lösbar ist. Daher rührt denn auch die Terminologie.

Beispiele 2.7.3. Wir erhalten etwa

$$\begin{aligned} \det(a) &= a \\ \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} &= ad - cb \\ \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= \begin{array}{l} a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ -a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12} \end{array} \end{aligned}$$



Um die Determinante einer (3×3)-Matrix zu berechnen mag man die erste und zweite Spalte danebenschreiben und dann die Produkte der drei Dreierdiagonalen nach rechts unten addieren und davon die Produkte der drei Dreierdiagonalen nach rechts oben abziehen. Diese Eselsbrücke heißt auch die "Jägerzaunformel". Für (4×4)-Matrizen ist es aber nicht mehr so einfach!

Im Fall der (3×3)-Matrizen heißt das manchmal die **Jägerzaunformel** aus einem Grund, den die nebenstehende Abbildung illustriert. Für $n \geq 4$ macht die Berechnung der Determinante anhand der Leibniz-Formel als Summe von $n! \geq 24$ Termen keinen Spaß mehr. Wir besprechen später, wie man in diesen Fällen geschickter vorgehen kann.

Beispiel 2.7.4. Die Determinante einer oberen Dreiecksmatrix ist das Produkt ihrer Diagonaleinträge. In der Tat ist die Identität die einzige Permutation σ mit $\sigma(i) \leq i$ für alle i, folglich trägt im Fall einer oberen Dreiecksmatrix nur der Summand mit $\sigma = \mathrm{id}$ zur Determinante bei. Dasselbe gilt für untere Dreiecksmatrizen.

Übung 2.7.5. Die Determinante einer Block-oberen-Dreiecksmatrix ist das Produkt der Determinanten ihrer Blöcke auf der Diagonalen. Hinweis: Man variiere das Argument für 2.7.4.

2.7.6 (**Betrag der Determinante und Volumen**). Vor der weiteren Entwicklung der Theorie will ich nun zuerst einmal die anschauliche Bedeutung der Determinante einer Matrix mit reellen Einträgen diskutieren. Ich beginne mit der anschauliche Bedeutung des Betrags der Determinante und beschränke mich dazu zunächst auf den Fall $n = 2$. Hoffentlich ist anschaulich klar, daß jede lineare Abbildung $L : \mathbb{R}^2 \to \mathbb{R}^2$ einen "Volumenverzerrungsfaktor" haben sollte, daß es also dazu eine reelle Konstante $c(L) \geq 0$ geben sollte derart, daß "das Bild unter L eines Flächenstücks U der Fläche $\mathrm{vol}(U)$ die Fläche $\mathrm{vol}(LU) = c(L)\,\mathrm{vol}(U)$ hat". Formal zeigt das die Transformationsformel **??**, die für besagte Konstante auch gleich die Formel

$$c(L) = |\det L|$$

liefert. Ich will diese Formel im folgenden heuristisch begründen. Anschaulich ist hoffentlich klar, daß unser $c : \mathrm{M}(2 \times 2; \mathbb{R}) \to \mathbb{R}_{\geq 0}$ die folgenden Eigenschaften haben sollte:

1. Es sollte "multiplikativ" sein, in Formeln $c(LM) = c(L)c(M)$;
2. Die Streckung einer Achse sollte die Fläche eines Flächenstücks genau durch Multiplikation mit dem Betrag des Streckfaktors ändern, in Formeln $c(\mathrm{diag}(a, 1)) = c(\mathrm{diag}(1, a)) = |a|$;
3. Scherungen sollten die Fläche eines Flächenstücks nicht ändern, in Formeln $c(D) = 1$ für D eine obere oder untere Dreiecksmatrix mit Einsen auf der Diagonale.

Da sich nun nach 1.7.16 jede Matrix als Produkt von Elementarmatrizen darstellen läßt, kann es höchstens eine Abbildung $c : \mathrm{M}(2 \times 2; \mathbb{R}) \to \mathbb{R}_{\geq 0}$ geben, die diese drei Eigenschaften hat. In 2.7.15 werden wir für unsere Determinante die "Multiplikationsformel" $\det(LM) = \det(L)\det(M)$ zeigen, und

zusammen mit unserer Formel 2.7.4 für die Determinante einer oberen oder unteren Dreiecksmatrix wird dann andererseits auch klar, daß $M \mapsto |\det M|$ eine Abbildung mit unseren drei Eigenschaften ist. Das beendet unsere heuristische Argumentation für die Stichhaltigkeit der Anschauung $|\det L| = c(L)$ für den Betrag der Determinante von (2×2)-Matrizen. In höheren Dimensionen liefert dieselbe Argumentation analoge Resultate, insbesondere kann der Betrag der Determinante einer (3×3)-Matrix aufgefaßt werden als der Faktor, um den die zugehörige lineare Abbildung Volumina ändert. Damit sollte auch anschaulich klar werden, warum $\det L \neq 0$ gleichbedeutend ist zur Invertierbarkeit von L, was wir im allgemeinen als 2.7.16 zeigen.

2.7.7 (**Vorzeichen der Determinante und Drehsinn**). Das Vorzeichen der Determinante einer invertierbaren reellen (2×2)-Matrix zeigt anschaulich gesprochen an, "ob die dadurch gegebene lineare Selbstabbildung der Ebene $\mathbb{R}^2$ den Drehsinn erhält oder umkehrt". Formal ist das vielleicht am ehesten in **??** enthalten, und im Fall allgemeiner angeordneter Körper wird diese anschauliche Erkenntnis ihrerseits unsere Definition 3.5.5 einer "Orientierung" auf einem Vektorraum über einem angeordneten Körper motivieren. Um die Beziehung zwischen Drehsinn und Determinante heuristisch zu begründen, können wir ähnlich argumentieren wie zuvor: Zunächst einmal führen wir ganz heuristisch eine angepaßte Notation ein und erklären für eine invertierbare lineare Abbildung $L : \mathbb{R}^2 \xrightarrow{\sim} \mathbb{R}^2$ das Vorzeichen $\varepsilon(L)$ durch die Vorschrift

$$\varepsilon(L) = \begin{cases} 1 & L \text{ erhält den Drehsinn;} \\ -1 & L \text{ kehrt den Drehsinn um.} \end{cases}$$

Vereinbaren wir speziell für diesen Beweis die Notation $[a]$ für das Vorzeichen einer von Null verschiedenen reellen Zahl, so können wir unsere Behauptung schreiben als die Formel

$$\varepsilon(L) = [\det L]$$

Ich will diese Formel im folgenden heuristisch begründen. Anschaulich ist hoffentlich klar, daß unser $\varepsilon : \mathrm{GL}(2;\mathbb{R}) \to \{1,-1\}$ die folgenden Eigenschaften haben sollte:

1. Es sollte "multiplikativ" sein, in Formeln $\varepsilon(LM) = \varepsilon(L)\varepsilon(M)$;

2. Die Streckung einer Achse sollte den Drehsinn genau durch die Multiplikation mit dem Vorzeichen des Streckfaktors ändern, in Formeln $\varepsilon(\mathrm{diag}(a,1)) = \varepsilon(\mathrm{diag}(1,a)) = [a]$;

3. Scherungen sollten den Drehsinn nicht ändern, in Formeln $\varepsilon(D) = 1$ für D eine obere oder untere Dreiecksmatrix mit Einsen auf der Diagonale.

Da sich nun nach 1.7.16 jede Matrix als Produkt von Elementarmatrizen darstellen läßt, kann es höchstens eine Abbildung $\varepsilon : \mathrm{GL}(2;\mathbb{R}) \to \{1,-1\}$ geben, die diese drei Eigenschaften hat. In 2.7.15 werden wir für unsere Determinante die "Multiplikationsformel" $\det(LM) = \det(L)\det(M)$ zeigen, und zusammen mit unserer Formel 2.7.4 für die Determinante einer oberen oder unteren Dreiecksmatrix wird dann andererseits auch klar, daß $M \mapsto [\det M]$ eine Abbildung mit unseren drei Eigenschaften ist. Das beendet unsere heuristische Argumentation für die Stichhaltigkeit der Anschauung $[\det L] = \varepsilon(L)$ für das Vorzeichen der Determinante von (2×2)-Matrizen. In höheren Dimensionen liefert dieselbe Argumentation analoge Resultate, etwa zeigt das Vorzeichen der Determinante einer invertierbaren Abbildung $L : \mathbb{R}^3 \to \mathbb{R}^3$ an, ob sie die "Händigkeit" erhält oder vielmehr "Rechtsgewinde und Linksgewinde vertauscht".

Bemerkung 2.7.8. Amüsant ist in diesem Zusammenhang die naive Frage, warum ein Spiegel "rechts und links vertauscht, aber nicht oben und unten". Die korrekte Antwort lautet, daß ein Spiegel ebensowenig rechts und links vertauscht wie oben und unten, sondern vielmehr vorne und hinten. Wir versuchen nur unbewußt, uns so gut wie möglich mit unserem Spiegelbild zu identifizieren, indem wir hinter den Spiegel treten, d.h. durch eine 180°-Drehung im Raum um eine geeignete vertikale Achse im Spiegel, und stellen dann fest, daß das zwar fast gelingt aber nicht ganz, und daß genauer die Verknüpfung der Spiegelung am Spiegel mit dieser Drehung gerade eine Spiegelung ist, die rechts und links vertauscht.

Definition 2.7.9. Seien V, U Vektorräume über einem Körper k. Eine bilineare Abbildung $F : V \times V \to U$ heißt **symmetrisch** genau dann, wenn gilt

$$F(v,w) = F(w,v) \quad \forall v,w \in V$$

Im Fall eines Grundkörpers k mit $1_k + 1_k \neq 0_k$ alias $\operatorname{char} k \neq 2$ heißt sie "alternierend" genau dann, wenn gilt $F(v,w) = -F(w,v) \quad \forall v,w \in V$. Um auch den Fall eines Grundkörpers der Charakteristik $\operatorname{char} k = 2$ korrekt einzubinden, müssen wir uns bei unserer Definition allerdings etwas von der ursprünglichen Bedeutung des Wortes "alternierend" entfernen und nennen im allgemeinen eine bilineare Abbildung $F : V \times V \to U$ **alternierend** genau dann, wenn gilt

$$F(v,v) = 0 \quad \forall v \in V$$

2.7.10. Gegeben eine bilineare Abbildung $F : V \times V \to U$ mit der Eigenschaft $F(v,v) = 0 \quad \forall v \in V$, die also im Sinne unserer Definition 2.7.9 alternierend

ist, folgern wir aus

$$\begin{aligned} 0 &= F(v+w, v+w) \\ &= F(v, v+w) + F(w, v+w) \\ &= F(v,v) + F(v,w) + F(w,v) + F(w,v) \\ &= F(v,w) + F(w,v) \end{aligned}$$

sofort $F(v,w) = -F(w,v) \quad \forall v, w \in V$. Gilt umgekehrt $F(v,w) = -F(w,v)$ $\forall v, w \in V$, so folgt $F(v,v) = -F(v,v)$ alias $(1_k + 1_k)F(v,v) = 0_k$ für alle $v \in V$, und ist $1_k + 1_k \neq 0_k$, so folgt daraus auch wieder $F(v,v) = 0$.

Definition 2.7.11. Sind $V_1, \ldots, V_n, W$ Vektorräume über einem Körper k, so heißt eine Abbildung $F : V_1 \times \ldots \times V_n \to W$ **multilinear** genau dann, wenn für alle j und alle beliebig aber fest gewählten $v_i \in V_i$ für $i \neq j$ die Abbildung $V_j \to W$, $v_j \mapsto F(v_1, \ldots, v_j, \ldots, v_n)$ linear ist.

Übung 2.7.12. Gegeben Vektorräume $V_1, V_2, \ldots, V_n, W$ über einem festen Körper bezeichne $\mathrm{Hom}^{(n)}(V_1 \times V_2 \times \ldots \times V_n, W)$ die Menge aller multilinearen Abbildungen $V_1 \times V_2 \times \ldots \times V_n \to W$. Man zeige: Ist $B_i \subset V_i$ jeweils eine Basis, so liefert die Restriktion eine Bijektion

$$\mathrm{Hom}^{(n)}(V_1 \times \ldots \times V_n, W) \xrightarrow{\sim} \mathrm{Ens}(B_1 \times \ldots \times B_n, W)$$

Jede multilineare Abbildung ist also festgelegt und festlegbar durch die Bilder aller Tupel von Basisvektoren. Den Spezialfall $n = 1$ kennen wir bereits aus 1.5.10, den Spezialfall $n = 2$ aus 1.5.24.

Definition 2.7.13. Seien V, W Vektorräume über einem Körper k. Eine multilineare Abbildung $F : V \times \ldots \times V \to W$ heißt **alternierend** genau dann, wenn sie auf jedem n-Tupel verschwindet, in dem zwei Einträge übereinstimmen, wenn also in Formeln gilt

$$(\exists i \neq j \text{ mit } v_i = v_j) \quad \Rightarrow \quad F(v_1, \ldots, v_i, \ldots, v_j, \ldots v_n) = 0$$

Das impliziert, daß sich das Vorzeichen von F ändert, wann immer man zwei Einträge vertauscht, in Formeln

$$F(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -F(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n)$$

und im Fall eines Grundkörpers einer von Zwei verschiedenen Charakteristik erhält man in derselben Weise, wie wir es in 2.7.10 für bilineare Abbildungen ausgeführt haben, auch die umgekehrte Implikation.

Satz 2.7.14 (**Charakterisierung der Determinante**). *Sei k ein Körper. Die Determinante*

$$\det : \mathrm{M}(n \times n; k) \to k$$

ist die einzige Abbildung, die (1) multilinear und alternierend ist als Funktion der n Spaltenvektoren und die (2) der Einheitsmatrix die Eins zuordnet.

Beweis. Daß unsere in 2.7.1 durch die Leibniz-Formel definierte Determinante multilinear ist und der Einheitsmatrix die Eins zuordnet, scheint mir offensichtlich. Stimmen weiter zwei Spalten einer Matrix überein, so verschwindet ihre Determinante, denn für $\tau \in \mathcal{S}_n$ die Transposition der entsprechenden Indizes gilt $a_{1\sigma(1)} \ldots a_{n\sigma(n)} = a_{1\tau\sigma(1)} \ldots a_{n\tau\sigma(n)}$ und $\operatorname{sgn}(\sigma) = -\operatorname{sgn}(\tau\sigma)$, so daß sich in der Leibniz-Formel die entsprechenden Terme gerade wegheben. Unsere durch die Leibniz-Formel gegebene Abbildung hat also die geforderten Eigenschaften, und es gilt nur noch zu zeigen, daß es keine weiteren Abbildungen $d : \mathrm{M}(n \times n; k) \to k$ mit den besagten Eigenschaften gibt. Nach 2.7.12 kennen wir aber unsere multilineare Abbildung d bereits, wenn wir ihre Werte

$$d(\mathrm{e}_{\sigma(1)}| \ldots |\mathrm{e}_{\sigma(n)})$$

kennen für alle Abbildungen $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$. Ist d zusätzlich alternierend, so gilt $d(\mathrm{e}_{\sigma(1)}| \ldots |\mathrm{e}_{\sigma(n)}) = 0$, falls σ nicht injektiv ist, und $d(\mathrm{e}_{\sigma\tau(1)}| \ldots |\mathrm{e}_{\sigma\tau(n)}) = -d(\mathrm{e}_{\sigma(1)}| \ldots |\mathrm{e}_{\sigma(n)})$ für jede Transposition τ. Mit 2.6.7 folgt daraus

$$d(\mathrm{e}_{\sigma(1)}| \ldots |\mathrm{e}_{\sigma(n)}) = \begin{cases} \operatorname{sgn}(\sigma) d(\mathrm{e}_1| \ldots |\mathrm{e}_n) & \sigma \in \mathcal{S}_n; \\ 0 & \text{sonst}, \end{cases}$$

und erfüllt d dann auch noch unsere Bedingung $d(\mathrm{e}_1| \ldots |\mathrm{e}_n) = 1$ für die Determinante der Einheitsmatrix, so folgt mit 2.7.12 sofort $d = \det$. □

Satz 2.7.15 (**Multiplikativität der Determinante**). *Sei k ein Kring. Gegeben quadratische Matrizen $A, B \in \mathrm{M}(n \times n; k)$ gilt*

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

Erster Beweis. Wir notieren $\mathcal{T}_n = \mathrm{Ens}(\{1, \ldots, n\})$ die Menge aller Abbildungen $\kappa : \{1, \ldots, n\} \to \{1, \ldots, n\}$ und rechnen

$$\begin{aligned}
\det(AB) &= \textstyle\sum_\sigma \operatorname{sgn}(\sigma) \prod_i (AB)_{i\sigma(i)} \\
&= \textstyle\sum_\sigma \operatorname{sgn}(\sigma) \prod_i \sum_j a_{ij} b_{j\sigma(i)} \\
&= \textstyle\sum_{\sigma \in \mathcal{S}_n,\, \kappa \in \mathcal{T}_n} \operatorname{sgn}(\sigma) a_{1\kappa(1)} b_{\kappa(1)\sigma(1)} \ldots a_{n\kappa(n)} b_{\kappa(n)\sigma(n)} \\
&= \textstyle\sum_{\kappa \in \mathcal{T}_n} a_{1\kappa(1)} \ldots a_{n\kappa(n)} \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) b_{\kappa(1)\sigma(1)} \ldots b_{\kappa(n)\sigma(n)} \\
&= \textstyle\sum_{\kappa \in \mathcal{T}_n} a_{1\kappa(1)} \ldots a_{n\kappa(n)} \det(B_\kappa)
\end{aligned}$$

wo B_κ diejenige Matrix bezeichnet, deren Zeilen der Reihe nach die Zeilen mit den Indizes $\kappa(1), \ldots, \kappa(n)$ der Matrix B sind. Aus 2.7.14 folgt aber $\det B_\kappa = 0$ falls $\kappa \notin \mathcal{S}_n$ und $(\det B_\kappa) = \operatorname{sgn}(\kappa)(\det B)$ falls $\kappa \in \mathcal{S}_n$, und damit erhalten wir dann $\det(AB) = (\det A)(\det B)$ wie gewünscht. □

Zweiter Beweis im Körperfall. Die Formel ist klar, wenn eine der beiden Matrizen eine Elementarmatrix ist, also eine Matrix, die sich von der Einheitsmatrix in höchstens einem Eintrag unterscheidet. Sie folgt im allgemeinen, da nach 1.7.16 jede Matrix ein Produkt von Elementarmatrizen ist. □

Dritter Beweis im Körperfall. Im Rahmen der Multilinearformen geben wir einen alternativen Beweis in **??** sowie in 7.4.15. □

Satz 2.7.16 (Determinantenkriterium für Invertierbarkeit). *Die Determinante einer quadratischen Matrix mit Einträgen in einem Körper ist von Null verschieden genau dann, wenn unsere Matrix invertierbar ist.*

Beweis. In Formeln behaupten wir für einen Körper K und eine beliebige Matrix $A \in \mathrm{M}(n \times n; K)$ also

$$\det A \neq 0 \quad \Leftrightarrow \quad A \text{ invertierbar}$$

Ist A invertierbar, so gibt es eine Matrix $B = A^{-1}$ mit $AB = I$. Mit der Multiplikationsformel folgt $(\det A)(\det B) = \det I = 1$ und folglich $\det A \neq 0$. Das zeigt die Implikation $\Leftarrow$. Ist A nicht invertierbar, so hat A nicht vollen Rang, die Familie der Spaltenvektoren von A ist demnach linear abhängig. Wir können also einen Spaltenvektor, ohne Beschränkung der Allgemeinheit den Ersten, durch die Anderen ausdrücken, etwa $a_{*1} = \lambda_2 a_{*2} + \ldots + \lambda_n a_{*n}$. Dann folgt jedoch unmittelbar

$$\begin{aligned} \det A &= \det(a_{*1}|a_{*2}|\ldots|a_{*n}) \\ &= \lambda_2 \det(a_{*2}|a_{*2}|\ldots|a_{*n}) + \ldots + \lambda_n \det(a_{*n}|a_{*2}|\ldots|a_{*n}) = 0 \end{aligned}$$

und damit ist auch die andere Implikation $\Rightarrow$ gezeigt. □

Übung 2.7.17. Man zeige die Formel für die **van-der-Monde-Determinante**

$$\det \begin{pmatrix} 1 & X_0 & X_0^2 & \ldots & X_0^n \\ \vdots & & & & \vdots \\ 1 & X_n & X_n^2 & \ldots & X_n^n \end{pmatrix} = \prod_{0 \leq j < i \leq n} (X_i - X_j)$$

Hinweis: Man mag von 2.3.31 und dem Fall des Grundkörpers $\mathbb{Q}$ ausgehen.

2.7.18. Aus der Multiplikationsformel folgt sofort $\det(A^{-1}) = (\det A)^{-1}$ für jede invertierbare Matrix A und damit ergibt sich für jede weitere quadratische Matrix B die Identität $\det(A^{-1}BA) = \det B$. Nach 1.7.33 hängt also für einen Endomorphismus $f : V \to V$ eines endlichdimensionalen Vektorraums die Determinante einer darstellenden Matrix nicht von der Wahl der zur Darstellung gewählten angeordneten Basis ab, in Formeln $\det({}_{\mathcal{B}}[f]_{\mathcal{B}}) = \det({}_{\mathcal{A}}[f]_{\mathcal{A}})$ für je zwei angeordnete Basen $\mathcal{A}$ und $\mathcal{B}$ von V. Diesen Skalar notieren wir

$$\det f$$

und nennen ihn die **Determinante des Endomorphismus** f. Dem einzigen Automorphismus des Nullraums ist insbesondere die Determinante 1 zuzuordnen.

Lemma 2.7.19. *Die Determinante einer Matrix ändert sich nicht beim Transponieren, in Formeln*

$$\det A^\top = \det A$$

Erster Beweis. Per definitionem gilt $\det A^\top = \sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \dots a_{\sigma(n)n}$. Ist nun $\tau = \sigma^{-1}$ die inverse Permutation, so haben wir $\operatorname{sgn}(\tau) = \operatorname{sgn}(\sigma)$ und darüber hinaus $a_{1\tau(1)} \dots a_{n\tau(n)} = a_{\sigma(1)1} \dots a_{\sigma(n)n}$, denn diese Produkte unterscheiden sich nur in der Reihenfolge ihrer Faktoren. Damit ergibt sich dann wie behauptet

$$\det A^\top = \sum_{\tau\in\mathcal{S}_n} \operatorname{sgn}(\tau) a_{1\tau(1)} \dots a_{n\tau(n)} \qquad \square$$

Zweiter Beweis. Arbeiten wir mit Koeffizienten in einem Körper, so können wir auch davon ausgehen, daß nach 1.7.16 jede quadratische Matrix A als ein Produkt von Elementarmatrizen $A = S_1 \dots S_r$ geschrieben werden kann. Für Elementarmatrizen S prüft man die Identität $\det S^\top = \det S$ leicht explizit, und dann liefert die Multiplikationsformel

$$\det A^\top = \det(S_r^\top \dots S_1^\top) = \det(S_r^\top) \dots \det(S_1^\top) = \det(S_r) \dots \det(S_1)$$

$$\det A = \det(S_1 \dots S_r) = \det(S_1) \dots \det(S_r)$$

und diese Produkte sind offensichtlich gleich. $\square$

Satz 2.7.20 (Laplace'scher Entwicklungssatz). *Gegeben eine $(n \times n)$-Matrix $A = (a_{ij})$ bezeichne $A\langle i,j\rangle$ die* **Streichmatrix***, die aus A durch Streichen der i-ten Zeile und j-ten Spalte entsteht. So gilt für jedes feste i die* **Entwicklung der Determinante nach der i-ten Zeile**

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A\langle i,j\rangle$$

und für jedes feste j die **Entwicklung nach der j-ten Spalte**

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A\langle i,j\rangle$$

2.7.21. Der folgende Beweis verwendet zwar die Sprache der Vektorräume, das Argument funktioniert jedoch ganz genauso statt für Matrizen mit Einträgen in einem Körper auch für Matrizen mit Einträgen in einem Kring.

Beweis. Wegen $\det A = \det A^\top$ reicht es, die erste unserer beiden Formeln zu zeigen. Wir wissen bereits, daß sich die Determinante einer quadratischen Matrix nur um den Faktor $(-1)^{j-1}$ ändert, wenn wir die j-te Spalte nach vorne schieben, ohne die Reihenfolge der übrigen Spalten zu ändern. Es reicht also, unsere Formel für die Entwicklung nach der ersten Spalte zu zeigen, was im folgenden Beweis insbesondere die Notation vereinfacht. Wir schreiben unsere Matrix als Folge von Spaltenvektoren $A = (a_{*1}|a_{*2}|\ldots|a_{*n})$ und schreiben den ersten Spaltenvektor als Linearkombination der Standardbasisvektoren

$$a_{*1} = a_{11}\mathrm{e}_1 + \ldots + a_{n1}\mathrm{e}_n$$

Die Multilinearität der Determinante liefert sofort

$$\det A = \sum_{i=1}^{n} a_{i1} \det(\mathrm{e}_i|a_{*2}|\ldots|a_{*n}) = \sum_{i=1}^{n} a_{i1}(-1)^{i-1} \det A\langle i,1\rangle$$

wo wir im zweiten Schritt die i-te Zeile der Matrix $(\mathrm{e}_i|a_{*2}|\ldots a_{*n})$ ganz nach oben geschoben haben, ohne die Reihenfolge der übrigen Zeilen zu ändern, um dann die Formel 2.7.5 für die Determinante von Block-oberen-Dreiecksmatrizen anzuwenden. □

Satz 2.7.22. *Bildet man zu einer quadratischen Matrix A die sogenannte* **adjungierte Matrix** $A^\sharp$ *mit den Einträgen $A^\sharp_{ij} = (-1)^{i+j} \det A\langle j,i\rangle$ für $A\langle j,i\rangle$ die entsprechende Streichmatrix nach 2.7.20, so gilt*

$$A \cdot A^\sharp = (\det A) \cdot I$$

Beweis. Es gilt zu zeigen

$$\sum_i (-1)^{i+j} a_{ki} \det A\langle j,i\rangle = \delta_{kj}(\det A)$$

Im Fall $k = j$ folgt das direkt aus unserer Entwicklung der Determinante nach der j-ten Zeile, im Fall $k \neq j$ steht die Formel für die Entwicklung nach der j-ten Zeile der Determinante der Matrix $\tilde{A}$ da, die aus A entsteht beim Ersetzen der j-ten Zeile durch die k-te Zeile. Da diese Matrix jedoch zwei gleiche Zeilen und damit Determinante Null hat, gilt unsere Formel auch in diesem Fall. □

Korollar 2.7.23. *Eine quadratische Matrix mit Einträgen in einem Kring besitzt genau dann eine Inverse mit Einträgen in besagtem Kring, wenn ihre Determinante in unserem Kring eine Einheit ist.*

2.7.24. Eine quadratische Matrix mit ganzzahligen Einträgen besitzt insbesondere genau dann eine Inverse mit ganzzahligen Einträgen, wenn ihre Determinante 1 oder -1 ist.

Beweis. Sei k unser Kring. Gegeben Matrizen $A, B \in \mathrm{M}(n\times n; k)$ mit $AB = I$ gilt natürlich $(\det A)(\det B) = \det I = 1$ und damit ist $\det A$ eine Einheit in k. Ist umgekehrt $\det A$ eine Einheit in k, so liefert nach der Cramer'schen Regel 2.7.22 die Formel $B = (\det A)^{-1}A^\sharp$ eine Matrix $B \in \mathrm{M}(n \times n; k)$ mit $AB = I$. Indem wir dasselbe Argument auf die transponierte Matrix anwenden und das Resultat wieder transponieren, finden wir auch $C \in \mathrm{M}(n\times n; k)$ mit $CA = I$. Daraus folgt sofort $B = C$ und folglich ist A invertierbar in $\mathrm{M}(n \times n; k)$. □

Übung 2.7.25. Es seien n^2 obere Dreiecksmatrizen $A_{11}, \ldots, A_{nn}$ mit m Zeilen und Spalten gegeben. Wir bilden die $(mn \times mn)$-Matrix

$$B = \begin{pmatrix} A_{11} & \ldots & A_{1n} \\ \vdots & & \vdots \\ A_{n1} & \ldots & A_{nn} \end{pmatrix}$$

Man zeige, daß gilt

$$\det B = \det \left(\sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\sigma) A_{1\sigma(1)} \ldots A_{n\sigma(n)} \right)$$

Hinweis: Wir betrachten diejenige Abbildung

$$f : \{1, \ldots, mn\} \to \{1, \ldots, m\}$$

die verträglich ist mit der Restklassenabbildung beider Mengen auf $\mathbb{Z}/m\mathbb{Z}$, und beachten, daß für eine Permutation $\sigma \in \mathcal{S}_{mn}$ mit $f(\sigma(i)) \le f(i) \quad \forall i$ notwendig Gleichheit gilt für alle i. Man zeige dieselbe Formel bei Koeffizienten in einem Körper auch für den Fall, daß die Matrizen A_{ij} paarweise kommutieren.

Übung 2.7.26. Man zeige dieselbe Formel auch für den Fall, daß die Matrizen A_{ij} paarweise kommutieren, ohne Koeffizienten in einem Körper vorauszusetzen. Vergleiche [Lorentz, LAII, S183, Aufgabe 63]. Hinweis: Ist A_{11} die Einheitsmatrix, so folgt die Behauptung durch Nullen der ersten Blockspalte und Induktion. Ist $\det A_{11}$ ein Nichtnullteiler unseres Rings R, so folgt

die Aussage durch Multiplizieren beider Seiten mit $\operatorname{diag}(A_{11}^\sharp, E, \ldots, E)$ für $A_{11}^\sharp$ die Komplementärmatrix zu A_{11}. Im allgemeinen kann man eine weitere Variable X einführen und A_{11} durch die Matrix $A_{11} + XE$ ersetzen, deren Determinante ein normiertes Polynom in $R[X]$ und deshalb kein Nullteiler ist. Nachher setze man dann $X = 0$.

2.8 Eigenwerte und Eigenvektoren

Definition 2.8.1. Sei $f : V \to V$ ein Endomorphismus eines Vektorraums über einem Körper k. Ein Skalar $\lambda \in k$ heißt ein **Eigenwert von** f genau dann, wenn es einen von Null verschiedenen Vektor $v \neq 0$ aus V gibt mit $f(v) = \lambda v$. Jeder derartige Vektor heißt dann ein **Eigenvektor von** f **zum Eigenwert** λ.

Beispiel 2.8.2. Die Drehung des Richtungsraums der Papierebene um den rechten Winkel im Uhrzeigersinn besitzt keinen Eigenwert. Für das Ableiten, aufgefaßt als Endomorphismus des Raums aller reellen polynomialen Funktionen, ist der einzige Eigenwert Null und die zugehörigen Eigenvektoren sind genau die von Null verschiedenen konstanten Polynome. Der lineare Anteil einer Drehung im Anschauungsraum besitzt stets Eigenvektoren zum Eigenwert Eins, nämlich alle Richtungsvektoren der Drehachse. Was aber überhaupt eine Drehung in der Sprache der Mengenlehre sein soll, werden wir noch diskutieren müssen.

Satz 2.8.3 (Existenz von Eigenwerten). *Jeder Endomorphismus eines von Null verschiedenen endlichdimensionalen Vektorraums über einem algebraisch abgeschlossenen Körper besitzt einen Eigenwert.*

2.8.4. Auf dem $\mathbb{C}$-Vektorraum $\mathbb{C}[X]$ der Polynome besitzt der Endomorphismus "Multipliziere mit X" keine Eigenwerte. Die Annahme endlicher Dimension ist also für die Gültigkeit des vorhergehenden Satzes wesentlich.

Beweis. Ohne Beschränkung der Allgemeinheit dürfen wir annehmen, unser Vektorraum sei der k^n und unser Endomorphismus sei gegeben durch die Multiplikation mit einer quadratischen Matrix $A \in \mathrm{M}(n \times n; k)$. Bezeichnet $I \in \mathrm{M}(n \times n; k)$ die Einheitsmatrix, so haben wir für $\lambda \in k$ die Äquivalenzen

$$\begin{aligned}
(\lambda \text{ ist Eigenwert von } A) &\Leftrightarrow \exists v \neq 0 \text{ mit } Av = \lambda v \\
&\Leftrightarrow \exists v \neq 0 \text{ mit } (A - \lambda I)v = 0 \\
&\Leftrightarrow \ker(A - \lambda I) \neq 0 \\
&\Leftrightarrow \det(A - \lambda I) = 0
\end{aligned}$$

Nun ist aber $\det(A - \lambda I)$ nach der Leibnizformel ein Polynom in λ mit dem Leitterm $(-\lambda)^n$ und besitzt folglich für $n \geq 1$ stets mindestens eine Nullstelle in unserem algebraisch abgeschlossenen Körper k. □

2.8.5. Sei k ein Körper und $A \in \mathrm{M}(n \times n; k)$ eine quadratische Matrix mit Koeffizienten in k. Das Polynom $\det(A - \lambda I)$ aus dem Polynomring $k[\lambda]$ heißt das **charakteristische Polynom** der Matrix A. Es wird auch notiert

$$\det(A - \lambda I) = \chi_A(\lambda)$$

Die Eigenwerte von A sind nach dem vorhergehenden Beweis genau die Nullstellen des charakteristischen Polynoms.

Übung 2.8.6. Sei k ein Körper und $A \in \mathrm{M}(n \times n; k)$ eine quadratische Matrix mit Koeffizienten in k. Man zeige, daß das charakteristische Polynom von A die Gestalt

$$\chi_A(\lambda) = (-\lambda)^n + \mathrm{tr}(A)(-\lambda)^{n-1} + \ldots + \det(A)$$

hat, also in Worten den Leitkoeffizienten $(-1)^n$, als nächsten Koeffizienten bis auf ein Vorzeichen die Spur von A, und als konstanten Term die Determinante von A.

Übung 2.8.7. Jeder Endomorphismus eines endlichdimensionalen reellen Vektorraums ungerader Dimension besitzt einen reellen Eigenwert. Ist die Determinante unseres Endomorphismus positiv, so besitzt er sogar einen positiven reellen Eigenwert.

2.8.8. Sei k ein Körper und $f : V \to V$ ein Endomorphismus eines endlichdimensionalen k-Vektorraums. Mit demselben Argument wie in 2.7.18 sehen wir, daß bezüglich jeder angeordneten Basis von V die darstellende Matrix von f dasselbe charakteristische Polynom hat, in Formeln $\det({}_{\mathcal{B}}[f]_{\mathcal{B}} - \lambda\,\mathrm{id}) = \det({}_{\mathcal{A}}[f]_{\mathcal{A}} - \lambda\,\mathrm{id})$ für je zwei angeordnete Basen $\mathcal{A}$ und $\mathcal{B}$ von V. Dies Polynom notieren wir dann

$$\chi_f = \chi_f(\lambda)$$

und nennen es das **charakteristische Polynom** des Endomorphismus f. Die Eigenwerte von f sind nach dem vorhergehenden Beweis genau die Nullstellen des charakteristischen Polynoms χ_f von f.

2.8.9. Das charakteristische Polynom einer Block-oberen-Dreiecksmatrix ist nach 2.7.5 das Produkt der charakteristischen Polynome ihrer Blöcke auf der Diagonalen.

Proposition 2.8.10. *Eine Matrix ist nilpotent genau dann, wenn ihr charakteristisches Polynom nur aus dem Leitterm besteht. In Formeln ist also* $A \in \mathrm{M}(n \times n; k)$ *nilpotent genau dann, wenn gilt* $\chi_A(\lambda) = (-\lambda)^n$.

Beweis. Ist unsere Matrix nilpotent, so ist sie nach 1.7.42 konjugiert zu einer oberen Dreiecksmatrix mit Nullen auf der Diagonalen und unsere Behauptung folgt aus 2.8.9. Besteht umgekehrt das charakteristische Polynom nur aus dem Leitterm, so ist unsere Matrix nilpotent nach dem gleich anschließenden Satz von Cayley-Hamilton. Ein alternatives Argument für die Rückrichtung liefert der Satz über die Hauptraumzerlegung 5.3.14. □

Satz 2.8.11 (Cayley-Hamilton). *Setzt man eine quadratische Matrix in ihr eigenes charakteristisches Polynom ein, so erhält man die Nullmatrix.*

Beweis. Gegeben eine quadratische Matrix A mit Koeffizienten in einem Kring gibt es nach 2.7.22 eine weitere Matrix $A^\sharp$ mit Koeffizienten in demselben Kring derart, daß im Ring der quadratischen Matrizen mit Einträgen in unserem Kring gilt

$$A^\sharp A = (\det A) \cdot E$$

für E die Einheitsmatrix. Nehmen wir speziell den Kring $k[t]$ und die Matrix $A = F - tE$ für eine vorgegebene Matrix $F \in \mathrm{M}(n \times n; k)$, so erhalten wir in $\mathrm{M}(n \times n; k[t])$ die Gleichung

$$A^\sharp(F - tE) = P_F(t) \cdot E$$

Bezeichne nun $f : k^n \to k^n$ die durch Multiplikation eines Spaltenvektors mit der Matrix F gegebene lineare Abbildung. Wenden wir auf beide Seiten unserer Gleichung von Matrizen den Ringhomomorphismus $k[t] \to \mathrm{End}_k\, k^n$ mit $t \mapsto f$ an, so erhalten wir in $\mathrm{M}(n \times n; \mathrm{End}_k\, k^n)$ alias $\mathrm{M}(n^2 \times n^2; k)$ die Gleichung

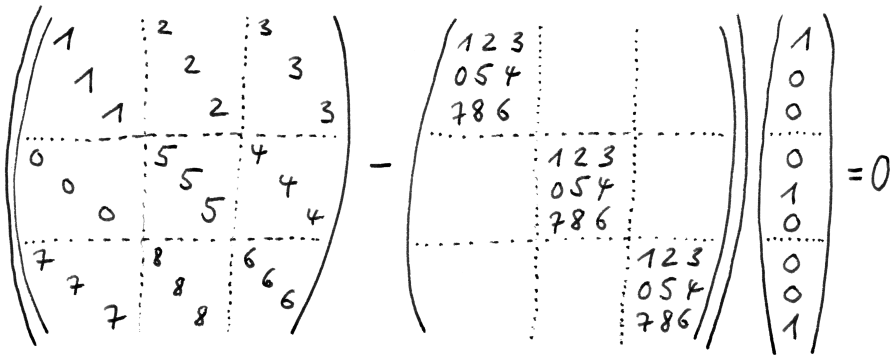
$$A^\sharp(F - fE) = P_F(f) \cdot E$$

Betrachten wir nun die Standardbasis $\mathrm{e}_1, \ldots, \mathrm{e}_n$ aus Spaltenvektoren des k^n und wenden beide Seiten dieser Gleichung an auf den Vektor $(\mathrm{e}_1^\top, \ldots, \mathrm{e}_n^\top)^\top$, aufgefaßt als Spaltenvektor in k^{n^2}, so ergibt auf der linken Seite schon die Multiplikation mit $(F - fE)$ den Nullvektor, denn bei

$$(F - fE)(\mathrm{e}_1^\top, \ldots, \mathrm{e}_n^\top)^\top$$

steht im i-ten Block von k^{n^2} genau $F_{i1}\,\mathrm{e}_1 + \ldots + F_{in}\,\mathrm{e}_n - f(\mathrm{e}_i) = 0$. Also wird die rechte Seite auch Null und es folgt $P_F(f)\,\mathrm{e}_1 = \ldots = P_F(f)\,\mathrm{e}_n = 0$. □

Definition 2.8.12. Ein Endomorphismus eines Vektorraums heißt **diagonalisierbar** genau dann, wenn unser Vektorraum von Eigenvektoren des besagten Endomorphismus erzeugt wird. Im Fall eines endlichdimensionalen Vektorraums ist das gleichbedeutend dazu, daß unser Vektorraum V eine angeordnete Basis $\mathcal{B} = (v_1, \ldots, v_n)$ besitzt derart, daß die Matrix unserer Abbildung $f : V \to V$ bezüglich dieser Basis Diagonalgestalt hat, in Formeln ${}_{\mathcal{B}}[f]_{\mathcal{B}} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. In der Tat bedeutet das ja gerade $f(v_i) = \lambda_i v_i$.



$(F - fE)(\mathrm{e}_1^\top, \ldots, \mathrm{e}_n^\top)^\top = 0$ am Beispiel einer Matrix F mit drei Zeilen und Spalten. Alle nicht ausgeschriebenen Einträge der obigen Matrizen sind als Null zu verstehen.

Lemma 2.8.13. *Die Restriktion eines diagonalisierbaren Endomorphismus auf einen unter besagtem Endomorphismus stabilen Teilraum ist wieder diagonalisierbar.*

Beweis. Sei $f : V \to V$ unser Endomorphismus. Gegeben $v \in W$ haben wir nach Annahme eine Darstellung $v = v_1 + \ldots + v_n$ mit $v_i \in V$ Eigenvektoren zu den paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_n \in k$. Dann gilt aber

$$(f - \lambda_2 \operatorname{id}) \ldots (f - \lambda_n \operatorname{id}) v = (\lambda_1 - \lambda_2) \ldots (\lambda_1 - \lambda_n) v_1 \in W$$

und folglich $v_1 \in W$. Ebenso zeigt man auch $v_2, \ldots, v_n \in W$, folglich wird auch W von Eigenvektoren erzeugt. □

Definition 2.8.14. Eine quadratische Matrix $A \in \mathrm{M}(n \times n; k)$ heißt **diagonalisierbar** genau dann, wenn der durch Multiplikation mit A gegebene Endomorphismus des k^n diagonalisierbar ist. Das hinwiederum ist gleichbedeutend zur Existenz einer invertierbaren Matrix $S \in \mathrm{GL}(n; k)$ mit $S^{-1}AS = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ diagonal alias $AS = S \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$: In den Spalten von S stehen dann die Vektoren einer Basis von k^n aus Eigenvektoren von A zu den Eigenwerten $\lambda_1, \ldots, \lambda_n$.

Beispiel 2.8.15. Eine nilpotente Matrix ist genau dann diagonalisierbar, wenn sie die Nullmatrix ist. Das folgende Lemma zeigt insbesondere, daß jede $(n \times n)$-Matrix, deren charakteristisches Polynom n paarweise verschiedene Nullstellen hat, diagonalisierbar sein muß.

Lemma 2.8.16. *Sei $f : V \to V$ ein Endomorphismus eines Vektorraums und seien $v_1, \ldots, v_n$ Eigenvektoren von f zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_n$. So sind unsere Eigenvektoren linear unabhängig.*

Beweis. Der Endomorphismus $(f - \lambda_2 \operatorname{id}) \ldots (f - \lambda_n \operatorname{id})$ macht $v_2, \ldots, v_n$ zu Null, aber nicht v_1. Gegeben $x_1, \ldots, x_n \in k$ mit $x_1 v_1 + \ldots + x_n v_n = 0$ folgt demnach durch Anwenden unseres Endomorphismus $x_1 = 0$. Ebenso zeigt man $x_2 = \ldots = x_n = 0$. □

Übung 2.8.17 (**Jordan'sche Normalform für** (2×2)**-Matrizen)**. Sei k ein algebraisch abgeschlossener Körper. Man zeige, daß es für jede Matrix $A \in \mathrm{M}(2 \times 2; k)$ eine invertierbare Matrix $P \in \mathrm{GL}(2; k)$ gibt derart, daß $P^{-1}AP$ eine der beiden Gestalten

$$\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \quad \text{oder} \quad \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \quad \text{hat.}$$

3 Euklidische Vektorräume

3.1 Modellierung des Anschauungsraums

3.1.1. Unter einem Automorphismus eines affinen Raums verstehen wir wie in 1.9.9 eine bijektive affine Abbildung unseres affinen Raums auf sich selbst. Es ist leicht zu sehen, daß die Umkehrabbildung jedes Automorphismus wieder ein Automorphismus ist. Folglich bilden die Automorphismen eines affinen Raums eine Untergruppe der Gruppe aller Bijektionen unseres affinen Raums auf sich selbst.

Definition 3.1.2. Eine **Bewegungsgruppe** eines dreidimensionalen reellen affinen Raums E ist eine alle Translationen umfassende Untergruppe

$$B \subset \operatorname{Aut} E$$

seiner Automorphismengruppe derart, daß es für je zwei Paare (H, L) von Teilmengen von E bestehend aus einer Halbebene und einer Halbgerade auf ihrem Rand genau einen Automorphismus aus B gibt, der sie ineinander überführt. In Formeln meinen wir hier Paare (H, L) von Teilmengen $L \subset H \subset E$, die in der Gestalt $L = p + \mathbb{R}_{\geq 0}\vec{v}$ und $H = p + \mathbb{R}\vec{v} + \mathbb{R}_{\geq 0}\vec{w}$ geschrieben werden können, mit $p \in E$ einem Punkt und $\vec{v}, \vec{w} \in \vec{E}$ linear unabhängigen Richtungsvektoren. Die Elemente einer ausgezeichneten Bewegungsgruppe B nennen wir **Bewegungen**.

3.1.3. Den uns umgebenden Raum modellieren wir mathematisch als einen dreidimensionalen reellen affinen Raum

$$\mathbb{E}$$

mit einer ausgezeichneten Bewegungsgruppe $B \subset \operatorname{Aut} \mathbb{E}$. Ich finde, daß es diese Struktur ist, die eigentlich die Bezeichnung als "euklidischer Raum" am ehesten verdient hätte, aber diese Bezeichnung ist leider schon anderweitig vergeben. Die Elemente von $\mathbb{E}$ denken wir uns als "alle möglichen Orte im Raum". Manche dieser Orte können direkt als Kirchturmspitzen, Zimmerecken und dergleichen angegeben werden, die Übrigen gilt es sich vorzustellen. Affine Geraden in $\mathbb{E}$ denken wir uns als Sichtlinien, wie in 1.9.5 und 1.9.28 besprochen. Bei Bewegungen denken wir an, nun, eben Bewegungen. Kippen wir etwa einen Stuhl um, so werden die Enden der Stuhlbeine, die Ecken der Sitzfläche, ja überhaupt alle seine Punkte jeweils in andere Punkte des Anschauungsraums überführt, und diese Abbildung läßt sich zu einer Abbildung $\mathbb{E} \to \mathbb{E}$ fortsetzen, die Sichtlinien in Sichtlinien überführt und die nach 1.9.27 folglich affin sein muß. Unsere ausgezeichnete Bewegungsgruppe B modelliert

die Menge aller derartigen Selbstabbildungen des Anschauungsraums. Unsere Bedingung an eine Bewegungsgruppe bedeutet anschaulich, daß man etwa jedes Messer aus einer festen Position heraus durch genau eine Bewegung in eine Position bringen kann, in der der Übergang vom Griff zur Klinge an einer vorgegebenen Stelle stattfindet, die Messerspitze in eine vorgegebene Richtung zeigt und der Schnitt den Raum entlang einer vorgegebenen Halbebene teilen würde.

Definition 3.1.4. Ein **Skalarprodukt** auf einem reellen Vektorraum V ist eine bilineare Abbildung $V \times V \to \mathbb{R}$, $(\vec{v}, \vec{w}) \mapsto \langle \vec{v}, \vec{w} \rangle$ derart, daß gilt $\langle \vec{v}, \vec{w} \rangle = \langle \vec{w}, \vec{v} \rangle$ für alle $\vec{v}, \vec{w} \in V$ und $\langle \vec{v}, \vec{v} \rangle \leq 0 \Rightarrow \vec{v} = \vec{0}$. Allgemeiner vereinbaren wir dieselbe Definition im Fall eines Vektorraums über einem beliebigen angeordneten Körper.

Satz 3.1.5 (Bewegungsgruppen und Skalarprodukte). *Sei E ein dreidimensionaler reeller affiner Raum mit einer ausgezeichneten Bewegungsgruppe B und einem ausgezeichneten von Null verschiedenen Richtungsvektor $\vec{m} \in \vec{E}$. So gibt es auf dem Richtungsraum $\vec{E}$ von E genau ein Skalarprodukt $\langle\ ,\ \rangle : \vec{E} \times \vec{E} \to \mathbb{R}$ mit den beiden folgenden Eigenschaften:*

1. *Die linearen Anteile aller unserer Bewegungen lassen besagtes Skalarprodukt invariant, in Formeln*

$$\langle \vec{\varphi}(\vec{v}), \vec{\varphi}(\vec{w}) \rangle = \langle \vec{v}, \vec{w} \rangle \quad \forall \vec{v}, \vec{w} \in \vec{E} \text{ und } \varphi \in B;$$

2. *Für unseren ausgezeichneten Richtungsvektor $\vec{m} \in \vec{E}$ gilt $\langle \vec{m}, \vec{m} \rangle = 1$.*

Unsere Bewegungsgruppe kann dann umgekehrt beschrieben werden als die Gruppe aller Automorphismen φ des affinen Raums E, deren lineare Anteile $\vec{\varphi}$ besagtes Skalarprodukt invariant lassen.

3.1.6. Als Richtungsvektor $\vec{m}$ mag man sich diejenige Parallelverschiebung denken, die das eine Ende des Urmeters in Paris auf sein anderes Ende schiebt. Der vorhergehende Satz 3.1.5 soll eine Brücke bilden zwischen der meines Erachtens intuitiv besonders gut zugänglichen Modellierung des Anschauungsraums als dreidimensionaler reeller affiner Raum mit einer ausgezeichneten Bewegungsgruppe und seiner algebraisch besonders eleganten wenn auch mit der Wahl eines ausgezeichneten Richtungsvektors belasteten Modellierung als dreidimensionaler reeller affiner Raum mit einem ausgezeichneten Skalarprodukt auf seinem Richtungsraum.

Beweis. Wir zeigen hier nur, daß es zu einer ausgezeichneten Bewegungsgruppe nicht mehr als ein Skalarprodukt mit den behaupteten Eigenschaften

geben kann. Die restlichen Aussagen des Satzes und insbesondere die Existenz eines Skalarprodukts mit den behaupteten Eigenschaften zeigen wir erst in 6.9.3. Die linearen Anteile von Bewegungen $\varphi \in B$ bilden sicher eine Untergruppe $D \subset \mathrm{GL}(\vec{E})$, deren Elemente wir **Drehungen im Richtungsraum** oder kurz **Drehungen** nennen. Gegeben $\vec{v} \in \vec{E}$ gibt es nach unseren Annahmen stets eine Drehung $d \in D$ und ein $\lambda \in \mathbb{R}$ mit $d\vec{v} = \lambda\vec{m}$, und dann haben wir notwendig

$$\langle \vec{v}, \vec{v} \rangle = \langle d\vec{v}, d\vec{v} \rangle = \langle \lambda\vec{m}, \lambda\vec{m} \rangle = \lambda^2$$

Für $\vec{v}, \vec{w}$ linear abhängig legen unsere Bedingung also $\langle \vec{v}, \vec{w} \rangle$ bereits fest. Sonst gibt es nach unseren Annahmen genau eine Drehung d, die die Halbgerade $\mathbb{R}_{\geq 0}\vec{v}$ auf sich selber abbildet und die Halbebene $\mathbb{R}\vec{v} + \mathbb{R}_{\geq 0}\vec{w}$ auf die "gegenüberliegende" Halbebene $\mathbb{R}\vec{v} + \mathbb{R}_{\geq 0}(-\vec{w})$. Da dann d^2 sowohl unsere Halbgerade als auch unsere Halbebene auf sich selber abbilden muß, folgt wieder aus unseren Annahmen $d^2 = \mathrm{id}$ und damit $d(\vec{v}) = \vec{v}$. Weiter gilt $d(\vec{w}) = \alpha\vec{v} + \beta\vec{w}$ mit $\beta < 0$ und folglich $\vec{r} := d(\vec{w}) - \vec{w} \neq 0$. Wir haben sicher $d(\vec{r}) = -\vec{r}$ und $\vec{v}$ und $\vec{r}$ sind linear unabhängig. Aus

$$\langle \vec{v}, \vec{r} \rangle = \langle d(\vec{v}), d(\vec{r}) \rangle = \langle \vec{v}, -\vec{r} \rangle = -\langle \vec{v}, \vec{r} \rangle$$

folgt sofort $\langle \vec{v}, \vec{r} \rangle = 0$. Damit gilt notwendig $\langle \vec{v}, \vec{w} \rangle = \gamma \langle \vec{v}, \vec{v} \rangle$ für die eindeutig bestimmte Zahl γ mit $\vec{w} = \gamma\vec{v} + \delta\vec{r}$. Das zeigt, daß es nicht mehr als ein Skalarprodukt mit den geforderten Eigenschaften geben kann. □

3.1.7. Unser Beweis enthält insbesondere die folgende Anleitung zur Konstruktion des Skalarprodukts, ausgehend von einem dreidimensionalen reellen Raum E mit einer ausgezeichneten Bewegungsgruppe $B \subset \mathrm{Aut}\, E$ und und einem ausgezeichneten von Null verschiedenen Richtungsvektor $\vec{m} \in \vec{E}$: Zunächst erkläre man die **Drehnorm** eines beliebigen Richtungsvektors $\vec{v}$ als diejenige nichtnegative reelle Zahl $\|\vec{v}\| = \lambda$, für die es eine Drehung d gibt mit $d(\vec{v}) = \lambda\vec{m}$. Ich vermeide, hier den Begriff "Länge" zu benutzen, da unsere Drehnorm schlicht reelle Zahlen als Werte annimmt. Natürlich muß noch gezeigt werden, daß es nicht mehr als ein solches λ geben kann, das geschieht beim Beweis unseres Satzes in 6.9.11. Dann vereinbare man für Richtungsvektoren $\vec{v}, \vec{r} \in \vec{E}$ die Sprechweise, $\vec{r}$ sei **drehsenkrecht zu** $\vec{v}$ genau dann, wenn es eine Drehung d gibt mit $d(\vec{v}) = \vec{v}$ und $d(\vec{r}) = -\vec{r}$. Gegeben Richtungsvektoren $\vec{v}, \vec{w}$ suche man schließlich eine Darstellung $\vec{w} = \gamma\vec{v} + \delta\vec{r}$ mit $\vec{r}$ drehsenkrecht zu $\vec{v}$ und setze

$$\langle \vec{v}, \vec{w} \rangle = \gamma \|\vec{v}\|^2$$

Anschaulich mag man sich $\gamma\vec{v}$ als die orthogonale Projektion von $\vec{w}$ auf die Gerade $\mathbb{R}\vec{v}$ denken und den Betrag des Skalarprodukts als das Produkt der

Drehnorm der orthogonalen Projektion von $\vec{w}$ auf $\mathbb{R}\vec{v}$ mit der Drehnorm von $\vec{v}$, in Formeln $|\langle \vec{v}, \vec{w}\rangle| = \|\gamma\vec{v}\| \cdot \|\vec{v}\|$. Damit scheint mir anschaulich klar, daß $\langle \vec{v}, \vec{w}\rangle$ bei festem $\vec{v}$ linear in $\vec{w}$ ist. Andererseits ist in dieser Anschauung auch die Identität $\langle \vec{v}, \vec{w}\rangle = \langle \vec{w}, \vec{v}\rangle$ zunächst für Vektoren gleicher Drehnorm aber dann auch für beliebige Vielfache derselben alias für beliebige Vektoren unmittelbar einleuchtend. Einen formal vollständigen Beweis geben wir jedoch erst in 6.9.11.

3.1.8. Für das solchermaßen aus einer Bewegungsgruppe nebst einem ausgezeichneten von Null verschiedenen Richtungsvektor konstruierte Skalarprodukt erkennt man unmittelbar, daß gegeben zwei Richtungsvektoren $\vec{r}$ und $\vec{v}$ der Vektor $\vec{r}$ drehsenkrecht ist zu $\vec{v}$ genau dann, wenn $\vec{r}$ **skalarproduktsenkrecht** zu $\vec{v}$ ist in dem Sinne, daß gilt $\langle \vec{v}, \vec{r}\rangle = 0$. Weiter erkennt man unmittelbar, daß für jeden Richtungsvektor $\vec{v}$ seine Drehnorm übereinstimmt mit seiner **Skalarproduktnorm** $\sqrt{\langle \vec{v}, \vec{v}\rangle}$. In Zukunft können wir uns also diese begrifflichen Feinheiten sparen und einfach nur von aufeinander **senkrecht** stehenden Vektoren und von der **Norm** eines Vektors reden.

3.1.9. Das zweidimensionale Analogon von 3.1.5 gilt nur unter der zusätzlichen Annahme, daß unsere Bewegungsgruppe im Sinne der Topologie "abgeschlossen" sein soll in der Gruppe aller affinen Automorphismen. Die Geometrie des Raums ist also unter dem Aspekt der Symmetrie leichter zu modellieren. Ich denke aber auch, daß unsere intuitive Vorstellung des Senkrechtstehens von Geraden, selbst wenn sie auf ein Papier gezeichnet sind, eigentlich räumlich ist und in etwa dem Konzept entspricht, das ich im vorhergehenden Beweis unter der Bezeichnung "drehsenkrecht" formalisiert habe.

Übung 3.1.10. Man zeige in einem beliebigen Vektorraum mit Skalarprodukt die **Parallelogrammregel**, nach der die Summe der Quadrate der vier Seiten eines Parallelogramms gleich der Summe der Quadrate der beiden Diagonalen ist, in Formeln

$$2\|v\|^2 + 2\|w\|^2 = \|v - w\|^2 + \|v + w\|^2$$

3.2 Geometrie in euklidischen Vektorräumen

3.2.1. Gegeben ein Körper k und ein k-Vektorraum V heißt eine bilineare Abbildung $b : V \times V \to k$ auch eine **Bilinearform auf** V. Wie in 2.7.9 heißt eine Bilinearform **symmetrisch** genau dann, wenn gilt $b(\vec{v}, \vec{w}) = b(\vec{w}, \vec{v})$. Ist k ein angeordneter Körper, so heißt eine Bilinearform **positiv definit** genau dann, wenn gilt $b(\vec{v}, \vec{v}) \leq 0 \Rightarrow \vec{v} = \vec{0}$. Mit diesen ganzen Begriffsbildungen kann man dann ein Skalarprodukt auch definieren als eine symmetrische positiv definite Bilinearform.

Wählen wir die durch die Kantenlängen unserer Kästchen gegebene Längeneinheit, so ist das zugehörige anschauliche Skalarprodukt der beiden als Pfeile eingezeichneten Vektoren $\langle \vec{v}, \vec{w} \rangle = \langle (4,0), (-2,3) \rangle = -8$ sowohl nach unserer Formel als auch nach der in 3.1.7 erklärten anschaulichen Interpretation, für die Sie sich allerdings noch eine dritte Koordinate hinzudenken müssen.

Beispiel 3.2.2. Auf $V = \mathbb{R}^n$ erhält man ein Skalarprodukt durch die Vorschrift $\langle \vec{v}, \vec{w} \rangle = v_1 w_1 + \ldots + v_n w_n$ für $\vec{v} = (v_1, \ldots, v_n)$ und $\vec{w} = (w_1, \ldots, w_n)$. Es heißt das **Standardskalarprodukt**. Man findet für das Standardskalarprodukt oft auch die alternative Notation $\vec{v} \cdot \vec{w}$. Mit dem Formalismus der Matrixmultiplikation können wir es auch schreiben als Produkt eines Zeilenvektors mit einem Spaltenvektor $\langle \vec{v}, \vec{w} \rangle = \vec{v}^\top \circ \vec{w}$, wo wir die offensichtliche Identifikation $\mathrm{M}(1 \times 1; \mathbb{R}) \xrightarrow{\sim} \mathbb{R}$ verwenden.

3.2.3. Sei E ein dreidimensionaler reeller affiner Raum mit einer ausgezeichneten Bewegungsgruppe B. Wir wählen einen ausgezeichneten von Null verschiedenen Richtungsvektor $\vec{m} \in \vec{E}$ und betrachten das zugehörige Skalarprodukt auf seinem Richtungsraum $\vec{E}$ nach 3.1.5. Wählen wir weiter in $\vec{E}$ eine Basis $\vec{v}_1, \vec{v}_2, \vec{v}_3$ von paarweise aufeinander drehsenkrecht stehenden Vektoren der Drehnorm Eins und identifizieren den Richtungsraum vermittels dieser Basis mit dem $\mathbb{R}^3$, so entspricht unser Skalarprodukt aus 3.1.5 genau dem Standardskalarprodukt des $\mathbb{R}^3$. In der Tat folgt mit der Bilinearität unseres Skalarprodukts aus 3.1.5 unmittelbar

$$\langle x\vec{v}_1 + y\vec{v}_2 + z\vec{v}_3, x'\vec{v}_1 + y'\vec{v}_2 + z'\vec{v}_3 \rangle = xx' + yy' + zz'$$

Definition 3.2.4. Ein **Skalarprodukt** auf einem komplexen Vektorraum V ist eine Abbildung $V \times V \to \mathbb{C}$, $(\vec{v}, \vec{w}) \mapsto \langle \vec{v}, \vec{w} \rangle$ derart, daß für alle $\vec{v}, \vec{w}, \vec{v}_1, \vec{w}_1 \in V$ und $\lambda, \mu \in \mathbb{C}$ gilt:

1. $\langle \vec{v}, \vec{w} + \vec{w}_1 \rangle = \langle \vec{v}, \vec{w} \rangle + \langle \vec{v}, \vec{w}_1 \rangle$, $\langle \vec{v}, \lambda \vec{w} \rangle = \lambda \langle \vec{v}, \vec{w} \rangle$;
2. $\langle \vec{v} + \vec{v}_1, \vec{w} \rangle = \langle \vec{v}, \vec{w} \rangle + \langle \vec{v}_1, \vec{w} \rangle$, $\langle \mu \vec{v}, \vec{w} \rangle = \bar{\mu} \langle \vec{v}, \vec{w} \rangle$;
3. $\langle \vec{v}, \vec{w} \rangle = \overline{\langle \vec{w}, \vec{v} \rangle}$, insbesondere also $\langle \vec{v}, \vec{v} \rangle \in \mathbb{R}$;
4. $\langle \vec{v}, \vec{v} \rangle \leq 0 \Rightarrow \vec{v} = 0$.

3.2.5. Nebenbei bemerkt folgt hier 2 schon aus 1 und 3, aber es kann auch nicht schaden, diese Formeln nochmal explizit hinzuschreiben. Eine Abbildung $V \times V \to \mathbb{C}$, die 1 und 2 erfüllt, nennt man eine **Sesquilinearform**. Gilt zusätzlich 3, so heißt die Sesquilinearform **hermitesch** nach dem französischen Mathematiker Hermite. Das Standardbeispiel ist $V = \mathbb{C}^n$ mit dem Skalarprodukt $\langle \vec{v}, \vec{w} \rangle = \bar{v}_1 w_1 + \ldots + \bar{v}_n w_n$ für $\vec{v} = (v_1, \ldots, v_n)$ und $\vec{w} = (w_1, \ldots, w_n)$. Mithilfe der Matrixmultiplikation kann dies Skalarprodukt auch geschrieben werden als

$$\langle \vec{v}, \vec{w} \rangle = \overline{\vec{v}}^\top \circ \vec{w}$$

wobei der Strich über einer Matrix mit komplexen Einträgen das komplexe Konjugieren aller Einträge meint. Viele Autoren verwenden auch die abweichende Konvention, nach der im komplexen Fall ein Skalarprodukt linear im ersten und schieflinear im zweiten Eintrag sein soll. Ich ziehe die hier gegebene Konvention vor, da dann bei der Interpretation von $\langle \vec{v}, \vec{w} \rangle$ als "$\vec{v}$ auf $\vec{w}$ angewendet" dieses Anwenden von $\vec{v}$ linear ist. Eine Anschauung für den komplexen Fall kann ich nicht anbieten, dafür wird er sich aber bei der weiteren Entwicklung der Theorie als außerordentlich nützlich erweisen.

Definition 3.2.6. Einen reellen bzw. komplexen Vektorraum mit Skalarprodukt nennt man auch einen reellen bzw. komplexen **euklidischen Vektorraum**. In einem euklidischen Vektorraum definiert man die **Länge** oder **euklidische Norm** oder kurz **Norm** $\|\vec{v}\| \in \mathbb{R}$ eines Vektors $\vec{v}$ durch $\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle}$. Daß das tatsächlich im Sinne von **??** eine Norm ist, werden wir gleich als 3.2.14 zeigen. Vektoren der Länge 1 heißen auch **normal**. Zwei Vektoren $\vec{v}, \vec{w}$ heißen **orthogonal** und man schreibt

$$\vec{v} \perp \vec{w}$$

genau dann, wenn gilt $\langle \vec{v}, \vec{w} \rangle = 0$. Man sagt dann auch, $\vec{v}$ und $\vec{w}$ **stehen senkrecht aufeinander**. Manchmal verwendet man das Symbol $\perp$ auch für allgemeinere Teilmengen S, T eines euklidischen Raums und schreibt $S \perp T$ als Abkürzung für $v \perp w \quad \forall v \in S, w \in T$.

3.2.7. Viele Autoren reservieren die Bezeichnung als euklidischer Vektorraum für reelle Vektorräume mit Skalarprodukt oder sogar für endlichdimensionale reelle Vektorräume mit Skalarprodukt. Ich verwende sie auch im Komplexen in der Hoffnung, durch die Verwendung dieses Begriffes die Übertragung unserer Anschauung ins Komplexe zu fördern. Üblich ist die Bezeichnung eines komplexen Vektorraums mit Skalarprodukt als **unitärer Raum** und im endlichdimensionalen Fall als **Hilbertraum** im Kontext der Definition von allgemeinen Hilberträumen die Bezeichnung als **Prä-Hilbertraum**.

Übung 3.2.8. In einem euklidischen Vektorraum gilt $\|\lambda \vec{v}\| = |\lambda| \|\vec{v}\|$ für alle Vektoren $\vec{v}$ und alle Skalare $\lambda \in \mathbb{R}$ bzw. $\lambda \in \mathbb{C}$.

3.2.9. Stehen zwei Vektoren $\vec{v}, \vec{w}$ eines euklidischen Vektorraums senkrecht aufeinander, so gilt der **Satz des Pythagoras**

$$\|\vec{v} + \vec{w}\|^2 = \|\vec{v}\|^2 + \|\vec{w}\|^2$$

In der Tat folgt ja aus $\vec{v} \perp \vec{w}$ schon

$$\langle \vec{v} + \vec{w}, \vec{v} + \vec{w} \rangle = \langle \vec{v}, \vec{v} \rangle + \langle \vec{v}, \vec{w} \rangle + \langle \vec{w}, \vec{v} \rangle + \langle \vec{w}, \vec{w} \rangle = \langle \vec{v}, \vec{v} \rangle + \langle \vec{w}, \vec{w} \rangle$$

Definition 3.2.10. Eine Familie $(\vec{v}_i)_{i\in I}$ von Vektoren eines euklidischen Vektorraums heißt ein **Orthonormalsystem** genau dann, wenn die Vektoren $\vec{v}_i$ alle die Länge 1 haben und paarweise aufeinander senkrecht stehen, wenn also mit dem Kroneckerdelta aus 1.7.4 in Formeln gilt

$$\langle \vec{v}_i, \vec{v}_j \rangle = \delta_{ij}$$

Ein Orthonormalsystem, das eine Basis ist, heißt eine **Orthonormalbasis**.

Lemma 3.2.11 (Orthogonale Projektion). *Ist V ein euklidischer Vektorraum und $\vec{v}_1, \ldots, \vec{v}_n$ ein endliches Orthonormalsystem, so kann man jeden Vektor $\vec{v} \in V$ in eindeutiger Weise schreiben als*

$$\vec{v} = \vec{p} + \vec{r}$$

mit $\vec{p}$ in dem von den $\vec{v}_i$ erzeugten Teilraum und $\vec{r}$ orthogonal zu allen $\vec{v}_i$.

3.2.12. Die Abbildung $\vec{v} \mapsto \vec{p}$ heißt die **orthogonale Projektion** auf den von den $\vec{v}_i$ aufgespannten Teilraum. Sie ist in der Terminologie von 1.6.4 die Projektion längs des Teilraums der auf allen $\vec{v}_i$ senkrechten Vektoren. Man beachte, daß die orthogonale Projektion von $\vec{v}$ genau derjenige Punkt $\vec{p}$ unseres Teilraums ist, der den kleinsten Abstand zu $\vec{v}$ hat: Für jeden Vektor $\vec{w} \neq \vec{0}$ aus unserem Teilraum gilt nämlich nach Pythagoras

$$\|(\vec{p} + \vec{w}) - \vec{v}\|^2 = \|\vec{p} - \vec{v}\|^2 + \|\vec{w}\|^2 > \|\vec{p} - \vec{v}\|^2$$

Beweis. Machen wir den Ansatz $\vec{p} = \sum \lambda_i \vec{v}_i$, so folgt $\langle \vec{v}_i, \vec{v} \rangle = \langle \vec{v}_i, \vec{p} \rangle = \lambda_i$ und damit die Eindeutigkeit von $\vec{p}$. Andererseits steht aber mit diesen λ_i der Vektor $\vec{r} = \vec{v} - \sum \lambda_i \vec{v}_i$ auch tatsächlich senkrecht auf allen $\vec{v}_i$, denn wir finden

$$\langle \vec{v}_j, \vec{r} \rangle = \langle \vec{v}_j, \vec{v} \rangle - \sum \lambda_i \langle \vec{v}_j, \vec{v}_i \rangle = \langle \vec{v}_j, \vec{v} \rangle - \lambda_j = 0 \qquad \square$$

3.2.13. Ist V ein euklidischer Vektorraum und $(\vec{v}_i)_{i\in I}$ eine Orthonormalbasis und $\vec{v} = \sum \lambda_i \vec{v}_i$ die Darstellung eines Vektors $\vec{v} \in V$, so erhalten wir durch Davormultiplizieren von $\vec{v}_j$ sofort $\lambda_j = \langle \vec{v}_j, \vec{v} \rangle$.

Satz 3.2.14. *1. Für beliebige Vektoren $\vec{v}, \vec{w}$ eines euklidischen Vektorraums gilt die* **Cauchy-Schwarz'sche Ungleichung**

$$|\langle \vec{v}, \vec{w} \rangle| \leq \|\vec{v}\| \|\vec{w}\|$$

mit Gleichheit genau dann, wenn $\vec{v}$ und $\vec{w}$ linear abhängig sind.

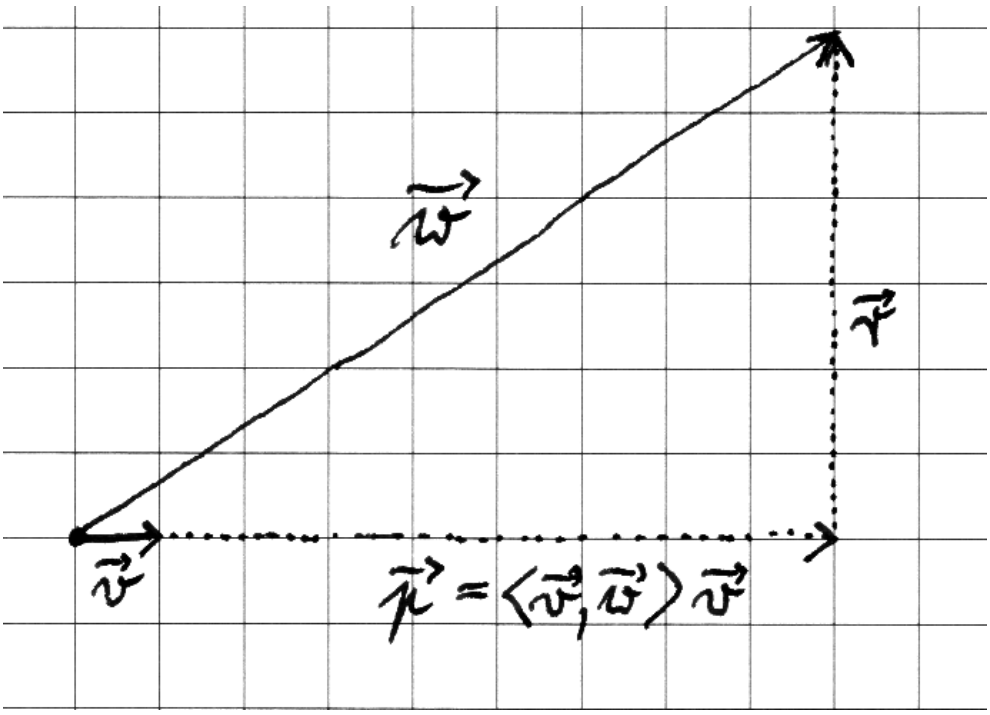


Illustration zum Beweis der Cauchy-Schwarz'schen Ungleichung. Wir haben darin $\vec{v} = (1, 0)$, $\vec{w} = (9, 6)$, $\langle \vec{v}, \vec{w} \rangle = 9$, $\vec{p} = (9, 0)$, $\vec{r} = (0, 6)$.

2. *Für beliebige Vektoren $\vec{v}, \vec{w}$ eines euklidischen Vektorraums gilt die* **Dreiecksungleichung**

$$\|\vec{v}+\vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$$

mit Gleichheit genau dann, wenn einer unserer Vektoren ein nichtnegatives Vielfaches des anderen ist, wenn also in Formeln gilt $\vec{v} \in \mathbb{R}_{\geq 0}\vec{w}$ oder $\vec{w} \in \mathbb{R}_{\geq 0}\vec{v}$.

Beweis. Um Teil 1 zu zeigen, nehmen wir zunächst $\|\vec{v}\| = 1$ an. Die orthogonale Projektion eines weiteren Vektors $\vec{w}$ auf die Gerade $\mathbb{R}\vec{v}$ wird dann nach 3.2.11 gegeben durch die Formel $\vec{p} = \langle \vec{v}, \vec{w}\rangle \vec{v}$ und mit Pythagoras erhalten wir $\|\vec{w}\|^2 = \|\vec{p}\|^2 + \|\vec{r}\|^2 \geq \|\vec{p}\|^2 = |\langle \vec{v}, \vec{w}\rangle|^2$ und folglich $|\langle \vec{v}, \vec{w}\rangle| \leq \|\vec{v}\|\|\vec{w}\|$ mit Gleichheit genau dann, wenn gilt $\vec{r} = \vec{0}$ alias wenn $\vec{w}$ ein Vielfaches von $\vec{v}$ ist. Diese Ungleichung muß aber offensichtlich erhalten bleiben, wenn wir darin $\vec{v}$ durch ein Vielfaches ersetzen, und so erhalten wir dann für beliebige Vektoren $\vec{v}, \vec{w}$ eines beliebigen euklidischen Vektorraums die Cauchy-Schwarz'sche Ungleichung $|\langle \vec{v}, \vec{w}\rangle| \leq \|\vec{v}\|\|\vec{w}\|$ mit Gleichheit genau dann, wenn $\vec{v}$ und $\vec{w}$ linear abhängig sind. Daraus hinwiederum ergibt sich sofort die Dreiecksungleichung $\|\vec{v}+\vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$, indem man beide Seiten quadriert und die Cauchy-Schwarz'sche Ungleichung anwendet. Insbesondere ist unsere euklidische Norm auch eine Norm im Sinne der in der Analysis in **??** gegebenen Definition. Der Beweis der letzten Aussage von Teil 2 sei dem Leser zur Übung überlassen. □

3.2.15. Ist V ein euklidischer Vektorraum und $\vec{v}_1, \ldots, \vec{v}_n$ ein endliches Orthonormalsystem, so ist für jeden Vektor $\vec{v} \in V$ seine orthogonale Projektion $\vec{p}$ auf den von unserem Orthonormalsystem erzeugten Teilraum höchstens so lang wie der Vektor selbst, in Formeln $\|\vec{v}\| \geq \|\vec{p}\|$ alias $\|\vec{v}\|^2 \geq \|\vec{p}\|^2$. Setzen wir hier unsere Darstellung von $\vec{p} = \sum \langle \vec{v}_i, \vec{v}\rangle \vec{v}_i$ aus dem Beweis von 3.2.11 ein, so ergibt sich die sogenannte **Bessel'sche Ungleichung**

$$\|\vec{v}\|^2 \geq \sum_{i=1}^{n} |\langle \vec{v}_i, \vec{v}\rangle|^2$$

Proposition 3.2.16. *Jeder endlichdimensionale reelle oder komplexe euklidische Vektorraum besitzt eine Orthonormalbasis.*

Beweis. Ist unser Raum der Nullraum, so tut es die leere Menge. Sonst finden wir einen von Null verschiedenen Vektor und erhalten, indem wir ihn mit dem Kehrwert seiner Länge multiplizieren, sogar einen Vektor $\vec{v}_1$ der Länge Eins. Die lineare Abbildung $\langle \vec{v}_1, \ \rangle$ hat als Kern einen Untervektorraum einer um eins kleineren Dimension. Eine offensichtliche Induktion beendet den Beweis. □

3.2.17. Gegeben ein euklidischer Vektorraum V und eine Teilmenge $T \subset V$ setzen wir

$$T^\perp = \{v \in V \mid v \perp t \ \ \forall t \in T\}$$

und nennen diese Menge den **Orthogonalraum** von T in V. Offensichtlich ist er stets ein Untervektorraum.

Proposition 3.2.18. *Gegeben ein euklidischer Vektorraum V und ein endlichdimensionaler Teilraum $U \subset V$ ist der Orthogonalraum von U in V auch ein Vektorraumkomplement, in Formeln*

$$V = U \oplus U^\perp$$

Beweis. Nach 3.2.16 besitzt U eine Orthonormalbasis, die dann natürlich auch ein endliches Orthonormalsystem in V ist. Die Proposition folgt so aus unserem Lemma 3.2.11 über orthogonale Projektionen. □

3.2.19. Gegeben ein euklidischer Vektorraum V und darin zwei Teilräume $U, W \subset V$ heißt W das **orthogonale Komplement von U in V** genau dann, wenn W sowohl ein Vektorraumkomplement als auch der Orthogonalraum zu U ist. Sagen wir von zwei Teilräumen eines euklidischen Raums, sie stünden aufeinander **orthogonal**, so ist gemeint, daß jeder Vektor des einen Teilraums auf jedem Vektor des anderen Teilraums senkrecht steht.

Übung 3.2.20. Gegeben ein euklidischer Vektorraum V und ein endlichdimensionaler Teilraum $U \subset V$ gilt $U = (U^\perp)^\perp$.

Übung 3.2.21. Man zeige, daß die Menge $\mathrm{L}^2_{\mathbb{R}}(\mathbb{N}) \subset \mathrm{Ens}(\mathbb{N}, \mathbb{R})$ aller reellen Folgen $a_0, a_1, \ldots$ mit $\sum a_i^2 < \infty$ einen Untervektorraum bildet und daß wir darauf durch die Vorschrift $\langle (a_i), (b_i) \rangle = \sum a_i b_i$ ein Skalarprodukt einführen können. Dann betrachte man in $\mathrm{L}^2_{\mathbb{R}}(\mathbb{N})$ den Untervektorraum U aller Folgen mit höchstens endlich vielen von Null verschiedenen Folgengliedern und zeige $U^\perp = 0$. Insbesondere ist in diesem Fall $U^\perp$ kein orthogonales Komplement zu U. Proposition **??** gilt also im allgemeinen nicht mehr, wenn wir unendlichdimensionale Teilräume U betrachten. Sie gilt jedoch wieder und sogar genau dann, wenn besagte Teilräume U zusätzlich "vollständig" sind, vergleiche **??**.

3.3 Orthogonale und unitäre Abbildungen

Definition 3.3.1. Eine lineare Abbildung $f : V \to W$ von euklidischen Vektorräumen heißt **orthogonal** im Reellen bzw. **unitär** im Komplexen genau dann, wenn sie das Skalarprodukt erhält, in Formeln

$$\langle f(v), f(w) \rangle = \langle v, w \rangle \quad \forall v, w \in V$$

Lemma 3.3.2. *Eine lineare Abbildung von euklidischen Vektorräumen ist orthogonal bzw. unitär genau dann, wenn sie die Längen aller Vektoren erhält, in Formeln*

$$\|f(v)\| = \|v\| \quad \forall v \in V$$

Beweis. Im Reellen folgt das aus der sogenannten **Polarisierungsidentität** $\langle v, w\rangle = \frac{1}{2}(\|v+w\|^2 - \|v\|^2 - \|w\|^2)$. Im Komplexen folgt aus der Variante $\mathrm{Re}\langle v, w\rangle = \frac{1}{2}(\|v+w\|^2 - \|v\|^2 - \|w\|^2)$ der Polarisierungsidentität zunächst einmal, daß f mit der Länge auch die Realteile aller Skalarprodukte erhalten muß. Wegen $\mathrm{Im}\langle v, w\rangle = -\mathrm{Re}\langle v, \mathrm{i}\, w\rangle$ erhält f dann natürlich auch die Imaginärteile aller Skalarprodukte. □

Proposition 3.3.3. *Gegeben reelle euklidische Vektorräume V, W ist eine Abbildung $f : V \to W$ orthogonal genau dann, wenn sie den Ursprung auf den Ursprung abbildet und alle Abstände erhält, in Formeln*

$$\|f(v) - f(w)\| = \|v - w\| \quad \forall v, w \in V$$

3.3.4. Man beachte, daß wir hier die Linearität von f nicht voraussetzen, sondern sie vielmehr aus schwächeren Voraussetzungen folgern.

Beweis. Ohne Beschränkung der Allgemeinheit dürfen wir V endlichdimensional annehmen. Dann hat V nach 3.2.16 eine endliche Orthonormalbasis $v_1, \ldots, v_n$. Gegeben Vektoren $v, w \in V$ mit $\|v\| = \|w\| = 1$ und $\|v - w\| = \sqrt{2}$ liefert die Polarisierungsidentität sofort $\langle v, w\rangle = 0$. Folglich bilden die Bilder $f(v_1), \ldots, f(v_n)$ unserer Orthonormalbasis von V ein Orthonormalsystem in W. Gegeben ein Vektor $v = x_1 v_1 + \ldots + x_x v_n$ können wir sein Bild nach 3.2.11 schreiben in der Form $f(v) = y_1 f(v_1) + \ldots + y_n f(v_n) + r$ mit $r \perp f(v_i)$ für alle i. Die Identität $\|v\| = \|f(v)\|$ liefert dann die Gleichung

$$x_1^2 + x_2^2 + \ldots + x_n^2 = y_1^2 + y_2^2 + \ldots + y_n^2 + \|r\|^2$$

und die Identität $\|v - v_1\| = \|f(v) - f(v_1)\|$ liefert zusätzlich die Gleichung

$$(x_1 - 1)^2 + x_2^2 + \ldots + x_n^2 = (y_1 - 1)^2 + y_2^2 + \ldots + y_n^2 + \|r\|^2$$

Die Subtraktion dieser beiden Gleichungen liefert $x_1 = y_1$. Analog erhalten wir $x_i = y_i$ für alle i und damit $r = 0$ und damit die Linearität sowie nach 3.3.5 auch die Orthogonalität von f. □

Lemma 3.3.5. *Seien V, W euklidische Räume und $\mathcal{B} \subset V$ eine Orthonormalbasis. Eine lineare Abbildung $f : V \to W$ ist orthogonal bzw. unitär genau dann, wenn sie die Orthonormalbasis $\mathcal{B}$ in ein Orthonormalsystem überführt, wenn also genauer die Familie $(f(v))_{v\in\mathcal{B}}$ ein Orthonormalsystem in W ist.*

Beweis. Es gilt zu zeigen $\langle f(v), f(w)\rangle = \langle v, w\rangle \quad \forall v, w \in V$. Wir wissen nach Annahme bereits, daß das gilt für alle $v, w \in \mathcal{B}$. Da beide Seiten bilinear bzw. sesquilinear sind als Abbildungen $V \times V \to \mathbb{R}$ bzw. $V \times V \to \mathbb{C}$, folgt es dann leicht für alle $v, w \in V$. □

Satz 3.3.6 (**Endlichdimensionale euklidische Vektorräume**). *Zwischen je zwei euklidischen reellen bzw. komplexen Vektorräumen derselben endlichen Dimension gibt es einen orthogonalen bzw. unitären Isomorphismus.*

Beweis. Wir wählen mit 3.2.16 in beiden Räumen jeweils eine Orthonormalbasis und erklären unseren Isomorphismus durch die Vorschrift, daß er von einer beliebig gewählten Bijektion zwischen den entsprechenden Basen herkommen soll. □

Definition 3.3.7. Gegeben ein reeller bzw. komplexer euklidischer Vektorraum V bilden die orthogonalen bzw. unitären Automorphismen von V jeweils eine Untergruppe der $\mathrm{GL}(V)$, die wir $\mathrm{O}(V)$ bzw. $\mathrm{U}(V)$ notieren.

Satz 3.3.8 (**Matrizen orthogonaler und unitärer Endomorphismen**).

1. *Eine Matrix $A \in \mathrm{M}(n \times n; \mathbb{R})$ beschreibt einen orthogonalen Endomorphismus des $\mathbb{R}^n$ mit dem Standardskalarpodukt genau dann, wenn gilt $A^\top A = I$ alias $A^\top = A^{-1}$, wenn also in Worten ihre Transponierte ihre Inverse ist.*

2. *Eine Matrix $A \in \mathrm{M}(n \times n; \mathbb{C})$ beschreibt einen unitären Endomorphismus des $\mathbb{C}^n$ mit dem Standardskalarpodukt genau dann, wenn gilt $\bar{A}^\top A = I$ alias $\bar{A}^\top = A^{-1}$, wenn also in Worten die Konjugierte ihrer Transponierten ihre Inverse ist.*

Beweis. Wir zeigen gleich den komplexen Fall. Die Identität $\langle Av, Aw\rangle = \langle v, w\rangle$ ist nach unserer Interpretation des Skalarprodukts in Termen der Matrixmultiplikation gleichbedeutend zu $(\overline{Av})^\top (Aw) = \bar{v}^\top w$ alias zu $\bar{v}^\top \bar{A}^\top A w = \bar{v}^\top w$. Gilt $\bar{A}^\top A = I$, so stimmt das natürlich für alle $v, w \in \mathbb{C}^n$. Stimmt es umgekehrt für alle $v, w \in \mathbb{C}^n$, so insbesondere auch für die Vektoren der Standardbasis $\mathrm{e}_i, \mathrm{e}_j$. Damit erhalten wir von der Mitte ausgehend die Gleichungskette $(\bar{A}^\top A)_{ij} = \mathrm{e}_i^\top \bar{A}^\top A \, \mathrm{e}_j = \mathrm{e}_i^\top \mathrm{e}_j = \delta_{ij}$ alias $\bar{A}^\top A = I$. □

Definition 3.3.9. Eine Matrix $A \in \mathrm{M}(n \times n; \mathbb{R})$ heißt **orthogonal** genau dann, wenn gilt $A^\top A = I$. Eine Matrix $A \in \mathrm{M}(n \times n; \mathbb{C})$ heißt **unitär** genau dann, wenn gilt $\bar{A}^\top A = I$. Der vorhergehende Satz 1 oder auch direkte Rechnung zeigt, daß diese Matrizen Untergruppen von $\mathrm{GL}(n; \mathbb{R})$ bzw. $\mathrm{GL}(n; \mathbb{C})$

bilden. Sie heißen die **orthogonale Gruppe** bzw. die **unitäre Gruppe** und werden notiert

$$\mathrm{O}(n) = \{A \in \mathrm{GL}(n;\mathbb{R}) \mid A^\top A = I\}$$
$$\mathrm{U}(n) = \{A \in \mathrm{GL}(n;\mathbb{C}) \mid \bar{A}^\top A = I\}$$

Beispiel 3.3.10. Die einzigen orthogonalen Endomorphismen von $\mathbb{R}$ sind die Identität und die Multiplikation mit (-1), in Formeln $\mathrm{O}(1) = \{1, -1\}$. Die einzigen orthogonalen Endomorphismen der Ebene $\mathbb{R}^2$ sind die Drehungen um den Ursprung und die Spiegelungen an Geraden durch den Ursprung,

$$\mathrm{O}(2) = \left\{ \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix}, \begin{pmatrix} \cos\vartheta & \sin\vartheta \\ \sin\vartheta & -\cos\vartheta \end{pmatrix} \middle| \, 0 \le \vartheta < 2\pi \right\}$$

In der Tat muß für A orthogonal die erste Spalte $A\mathrm{e}_1$ die Länge Eins haben, also auf dem Einheitskreis liegen, also hat sie die Gestalt $(\cos\vartheta, \sin\vartheta)^\top$ für wohlbestimmtes $\vartheta \in [0, 2\pi)$. Die zweite Spalte $A\,\mathrm{e}_2$ muß auch die Länge Eins haben und auf der ersten Spalte senkrecht stehen, und damit verbleiben nur noch die beiden beschriebenen Möglichkeiten. Die erste dieser Möglichkeiten beschreibt anschaulich gesprochen eine Drehung um den Winkel ϑ im Gegenuhrzeigersinn. Die Zweite beschreibt die Spiegelung an der Gerade, die mit der positiven x-Achse in der oberen Halbebene den Winkel $\vartheta/2$ einschließt. Diese Spiegelung hat im übrigen die Eigenwerte 1 und -1.

3.3.11. Die Determinante einer unitären oder orthogonalen Matrix hat stets den Betrag Eins. In der Tat folgt aus $\bar{A}^\top A = I$ sofort $1 = \det(\bar{A}^\top A) = \det(\bar{A}^\top)\det(A) = \det(\bar{A})\det(A) = \overline{\det(A)}\det(A)$.

3.3.12. Jeder Eigenwert eines unitären oder orthogonalen Endomorphismus eines euklidischen Vektorraums hat den Betrag Eins, da derartige Abbildungen die Länge von Vektoren erhalten.

Übung 3.3.13. Sei V ein reeller euklidischer Vektorraum. Man zeige:

1. Eine endliche Familie $v_1, \ldots, v_n$ von Vektoren von V ist orthonormal genau dann, wenn die zugehörige Abbildung $\Phi : \mathbb{R}^n \to V$ orthogonal ist für das Standard-Skalarprodukt auf $\mathbb{R}^n$.

2. Gegeben endliche angeordnete Basen $\mathcal{A}$, $\mathcal{B}$ von V mit $\mathcal{A}$ orthonormal ist $\mathcal{B}$ orthonormal genau dann, wenn die Basiswechselmatrix ${}_{\mathcal{B}}[\mathrm{id}]_{\mathcal{A}}$ orthogonal ist. Hinweis: Man betrachte das kommutative Diagramm

$$\begin{array}{ccc} V & \xrightarrow{\mathrm{id}} & V \\ {\scriptstyle \Phi_{\mathcal{A}}}\uparrow & & \uparrow{\scriptstyle \Phi_{\mathcal{B}}} \\ \mathbb{R}^n & \xrightarrow{{}_{\mathcal{B}}[\mathrm{id}]_{\mathcal{A}}} & \mathbb{R}^n \end{array}$$

Man formuliere und zeige auch die analogen Aussagen im Komplexen.

Definition 3.3.14. Die Elemente der orthogonalen bzw. unitären Gruppen mit Determinante Eins bilden jeweils Untergruppen. Sie heißen die die **spezielle orthogonale Gruppe** bzw. die **spezielle unitäre Gruppe** und werden notiert als

$$\mathrm{SO}(n) = \{A \in \mathrm{O}(n) \mid \det A = 1\}$$
$$\mathrm{SU}(n) = \{A \in \mathrm{U}(n) \mid \det A = 1\}$$

Ähnlich bezeichnen wir für einen endlichdimensionalen reellen bzw. komplexen Vektorraum V mit $\mathrm{SO}(V)$ bzw. $\mathrm{SU}(V)$ die Untergruppen von $\mathrm{GL}(V)$ aller orthogonalen bzw. unitären Automorphismen mit Determinante Eins.

Beispiele 3.3.15. Die Gruppe $\mathrm{SO}(2)$ besteht gerade aus allen Drehungen der Ebene um den Ursprung, in Formeln

$$\mathrm{SO}(2) = \left\{ \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix} \middle| \, 0 \le \vartheta < 2\pi \right\}$$

Die Gruppe $\mathrm{SO}(3)$ besteht aus allen Drehungen des Raums, wir diskutieren sie gleich noch ausführlicher.

3.3.16. Im Rahmen der Definition von Sinus und Cosinus zeigen wir in **??** folgende unter anderem auch, daß die Abbildung $\mathbb{R} \to \mathrm{SO}(2)$ gegeben durch

$$\vartheta \mapsto R_\vartheta = \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix}$$

ein Gruppenhomomorphismus von der additiven Gruppe der reellen Zahlen in die Gruppe $\mathrm{SO}(2)$ ist. Diese Aussage ist im übrigen gleichbedeutend zu den Additionstheoremen für Sinus und Cosinus. Aus **??** folgt sogar, daß jeder stetige Gruppenhomomorphismus $\mathbb{R} \to \mathrm{SO}(2)$ von der Form $\vartheta \mapsto R_{a\vartheta}$ ist für genau ein $a \in \mathbb{R}$.

Satz 3.3.17 (Satz vom Fußball). *Jede orthogonale Selbstabbildung mit Determinante Eins eines dreidimensionalen reellen euklidischen Vektorraums V hat einen Fixvektor. Anschaulich gesprochen ist unsere Abbildung also eine Drehung um eine Drehachse, eben um die von einem Fixvektor erzeugte Gerade, und formal hat jedes $D \in \mathrm{SO}(V)$ in einer geeigneten Orthonormalbasis eine Matrix der Gestalt*

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\vartheta & -\sin\vartheta \\ 0 & \sin\vartheta & \cos\vartheta \end{pmatrix}$$

3.3.18. Wird bei einem Fußballspiel der Ball also vor dem Anpfiff zur zweiten Halbzeit wieder in die Mitte gelegt, so befinden sich zwei gegenüberliegende Punkte auf dem Ball an derselben Stelle wie vor dem Anpfiff zur ersten Halbzeit.

Erster Beweis. Sei $D : V \to V$ unsere Abbildung. Das charakteristische Polynom von D hat den Grad 3 und damit nach **??** mindestens eine reelle Nullstelle. Alle komplexen Eigenwerte unserer Abbildung haben den Absolutbetrag 1 und ihr Produkt ist $\det D = 1$, wenn wir jeden Eigenwert mit seiner Vielfachheit als Nullstelle des charakteristischen Polynoms in unser Produkt eingehen lassen. Hat also unser Polynom mit Vielfachheiten gezählt drei reelle Nullstellen, so sind diese notwendig $1, 1, 1$ oder $1, -1, -1$. Gibt es dahingegen mit Vielfachheiten gezählt nur eine reelle Nullstelle, so müssen außerdem noch zwei echt komplexe Nullstellen der Gestalt $\lambda, \bar{\lambda}$ vorliegen, und als reelle Nullstelle kommt in diesem Fall nur die 1 in Betracht. In jedem Fall hat D einen Fixvektor. Auf der zu diesem Fixvektor orthogonalen Ebene induziert unsere orthogonale Abbildung wieder eine orthogonale Abbildung, die nach Übung 2.7.5 zur Determinante Block-diagonaler Matrizen wieder Determinante Eins hat. Wählen wir also als Orthonormalbasis einen Fixvektor der Länge Eins nebst den beiden Vektoren einer Orthonormalbasis seines orthogonalen Komplements, so hat die Matrix unserer Abbildung in Bezug auf diese Basis nach 3.3.15 die behauptete Gestalt. □

Zweiter Beweis. Nach 2.8.7 hat unsere Abbildung einen positiven reellen Eigenwert und nach 3.3.12 muß der Eins sein. Folglich hat unsere Abbildung einen Fixvektor. Der Rest des Beweises kann wie zuvor laufen. □

Übung 3.3.19. Jede orthogonale Selbstabbildung mit Determinante (-1) des dreidimensionalen reellen Raums ist die Verknüpfung einer Drehung um eine Achse mit einer Spiegelung an der zu dieser Achse senkrechten Hyperebene.

Übung 3.3.20. Jede bezüglich Inklusion maximale kommutative Untergruppe der Drehgruppe SO(3) ist entweder die Gruppe aller Drehungen um eine Achse oder konjugiert zur Gruppe aller Diagonalmatrizen aus SO(3).

Satz 3.3.21 (Spektralsatz für unitäre Endomorphismen). *Gegeben ein unitärer Endomorphismus eines endlichdimensionalen komplexen euklidischen Vektorraums gibt es stets eine Orthonormalbasis unseres Vektorraums, die aus Eigenvektoren unseres Endomorphismus besteht.*

Beweis. Ist unser Raum der Nullraum, so tut es die leere Menge. Sonst finden wir nach 2.8.3 einen Eigenvektor und durch Renormieren natürlich auch einen Eigenvektor der Länge Eins. Da unser Endomorphismus unitär ist,

erhält er auch den Orthogonalraum dieses Eigenvektors und induziert auf diesem Orthogonalraum eine unitäre Abbildung. Mit Induktion über die Dimension finden wir in unserem Orthogonalraum eine Orthonormalbasis aus Eigenvektoren, und durch Hinzunehmen unseres ursprünglichen Eigenvektors der Länge Eins erhalten wir daraus die gesuchte Orthonormalbasis aus Eigenvektoren des ganzen Raums. □

Übung 3.3.22. Ein Endomorphismus eines endlichdimensionalen komplexen euklidischen Vektorraums ist genau dann unitär, wenn unser Vektorraum eine Orthonormalbasis besitzt, die aus Eigenvektoren unseres Endomorphismus besteht, und wenn zusätzlich alle Eigenwerte Betrag Eins haben.

Korollar 3.3.23. *Für jede unitäre Matrix $A \in \mathrm{U}(n)$ gibt es eine weitere unitäre Matrix $B \in \mathrm{U}(n)$ mit $B^{-1}AB = \mathrm{diag}(z_1, \ldots, z_n)$ für $z_i \in S^1 \subset \mathbb{C}$ komplexe Zahlen der Länge Eins.*

Beweis. Man findet solch eine Matrix B, indem man eine Orthonormalbasis aus Eigenvektoren von $A : \mathbb{C}^n \to \mathbb{C}^n$, die es nach 3.3.21 ja geben muß, als Spaltenvektoren hintereinanderschreibt: Dann gilt ja ganz offensichtlich $AB = \mathrm{diag}(z_1, \ldots, z_n)B$ und die Matrix B ist unitär nach Lemma 3.3.5. □

Satz 3.3.24 (Normalform für orthogonale Matrizen). *Die Matrix einer orthogonalen Abbildung D von einem endlichdimensionalen reellen euklidischen Vektorraum V in sich selber hat stets bezüglich einer geeigneten angeordneten Orthonormalbasis eine blockdiagonale Gestalt der Form*

$$\mathrm{diag}\left(1, \ldots, 1, -1, \ldots, -1, \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix}, \ldots, \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}\right)$$

mit Winkeln $0 < \vartheta \leq \ldots \leq \varphi < \pi$, und unter den angegebenen Einschränkungen an die Winkel wird umgekehrt besagte blockdiagonale Matrix durch unsere orthogonale Abbildung D bereits eindeutig festgelegt.

3.3.25. Daraus folgt auch unmittelbar der Satz vom Fußball 3.3.17.

Beweis. Der Drehblock zu einem Winkel ϑ hat die komplexen Eigenwerte $\cos\vartheta + \mathrm{i}\sin\vartheta = \mathrm{e}^{\pm \mathrm{i}\vartheta}$ und das zeigt bereits die behauptete Eindeutigkeit. Die Existenz ist klar im Fall $\dim_{\mathbb{R}} V < 3$ wegen 3.3.10. Zu beachten ist hierbei, daß jede ebene Spiegelung in einer geeigneten Orthonormalbasis die darstellende Matrix $\mathrm{diag}(1, -1)$ hat und jede ebene Drehung in einer geeigneten Orthonormalbasis als darstellende Matrix entweder $\mathrm{diag}(1, 1)$ oder $\mathrm{diag}(-1, -1)$ oder einen Drehblock mit einem Winkel ϑ mit $0 < \vartheta < \pi$: Bei einer Drehung um einen Winkel ϑ mit $\pi < \vartheta < 2\pi$ nehmen wir dazu

als Orthonormalbasis des $\mathbb{R}^2$ die Standardbasis mit der umgekehrten Anordnung. Die Existenz folgt mit Induktion im allgemeinen, sobald wir zeigen, daß es unter der Voraussetzung $\dim_{\mathbb{R}} V \geq 3$ in V stets einen echten von Null verschiedenen unter D invarianten Teilraum gibt, indem wir nämlich die Induktionsannahme auf diesen Teilraum und sein orthogonales Komplement anwenden. Ohne Beschränkung der Allgemeinheit dürfen wir $V = \mathbb{R}^n$ annehmen. Nun hat D schon unter der Annahme $n \geq 1$ stets einen Eigenvektor $v = (v_1, \ldots, v_n)^\top \in \mathbb{C}^n$, sagen wir $Dv = \lambda v$ mit $\lambda \in \mathbb{C}$. Dann folgt für $\bar{v} = (\bar{v}_1, \ldots, \bar{v}_n)^\top$ sofort $D\bar{v} = \bar{\lambda}\bar{v}$ und das komplexe Erzeugnis $\langle v, \bar{v}\rangle_{\mathbb{C}}$ dieser beiden Vektoren ist ein echter D-stabiler Teilraum von $\mathbb{C}^n$. Der Schnitt $\langle v, \bar{v}\rangle_{\mathbb{C}} \cap \mathbb{R}^n$ ist also ein echter D-stabiler Teilraum von $\mathbb{R}^n$, und dieser Schnitt ist auch nicht Null, denn er enthält sowohl $v + \bar{v}$ als auch $\mathrm{i}(v - \bar{v})$, die wegen $v \neq 0$ nicht beide verschwinden können. □

Übung 3.3.26. Bezeichne $R^x_\varphi \in \mathrm{SO}(3)$ die Drehung um die x-Achse $\mathbb{R}\,\mathrm{e}_1$ mit dem Winkel φ, in Formeln

$$R^x_\varphi = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}$$

Bezeichne $R^z_\varphi \in \mathrm{SO}(3)$ die Drehung um die z-Achse $\mathbb{R}\,\mathrm{e}_3$ mit dem Winkel φ, in Formeln

$$R^z_\varphi = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Man zeige: Jede Drehung im Raum $D \in \mathrm{SO}(3)$ läßt sich darstellen als

$$D = R^z_\varphi R^x_\psi R^z_\vartheta$$

mit $\psi \in [0, \pi]$ und $\varphi, \vartheta \in [0, 2\pi)$. Gilt $\mathrm{e}_3 \neq \pm D(\mathrm{e}_3)$, so ist diese Darstellung sogar eindeutig. Die fraglichen Winkel heißen dann die **Euler'schen Winkel** unserer Drehung D. Hinweis: Aus der Anschauung, deren Formalisierung Ihnen überlassen bleiben möge, finden wir $\psi \in [0, \pi]$ und $\varphi \in [0, 2\pi)$ mit $R^z_\varphi R^x_\psi(\mathrm{e}_3) = D(\mathrm{e}_3)$. Es folgt $D^{-1} R^z_\varphi R^x_\psi = R^z_{-\vartheta}$ für geeignetes $\vartheta \in [0, 2\pi)$.

Satz 3.3.27 (Gram-Schmidt). *Seien $v_1, \ldots, v_k$ linear unabhängige Vektoren eines euklidischen Vektorraums. So existiert in unserem Vektorraum genau ein Orthonormalsystem $w_1, \ldots, w_k$ mit*

$$w_i \in \mathbb{R}_{>0} v_i + \langle v_{i-1}, \ldots, v_1\rangle \quad \forall i$$

Beweis. Nach **??** können wir v_i eindeutig zerlegen als $v_i = p_i + r_i$ mit p_i der orthogonalen Projektion von v_i auf $\langle v_{i-1}, \ldots v_1\rangle$ und r_i im orthogonalen Komplement dieses Teilraums. Die Vektoren $w_i = r_i/\|r_i\|$ bilden dann ein Orthonormalsystem mit den geforderten Eigenschaften. Daß es auch das einzige ist, mag sich der Leser zur Übung selber überlegen. □

3.3.28. Für unsere Basen gilt sicher $\langle w_{i-1}, \ldots, w_1\rangle \subset \langle v_{i-1}, \ldots, v_1\rangle$ und Dimensionsvergleich liefert sogar die Gleichheit dieser Erzeugnisse. Nach der Formel für othogonale Projektion aus dem Beweis von 3.2.11 können wir also die w_i induktiv bestimmen durch die Formeln

$$\begin{aligned} r_1 &= v_1 \\ w_1 &= r_1/\|r_1\| \\ &\vdots \\ r_i &= v_i - \textstyle\sum_{\nu=1}^{i-1} \langle w_\nu, v_i\rangle w_\nu \\ w_i &= r_i/\|r_i\| \\ &\vdots \end{aligned}$$

Dieses Vorgehen ist bekannt als das **Gram-Schmidt'sche Orthogonalisierungsverfahren**.

Korollar 3.3.29 (Iwasawa-Zerlegung für $\mathrm{GL}(n;\mathbb{R})$). *Bezeichne $A \subset \mathrm{GL}(n;\mathbb{R})$ die Diagonalmatrizen mit positiven Einträgen auf der Diagonale und $N \subset \mathrm{GL}(n;\mathbb{R})$ die oberen Dreiecksmatrizen mit Einsen auf der Diagonale. So definiert die Multiplikation eine Bijektion*

$$\mathrm{O}(n) \times A \times N \xrightarrow{\sim} \mathrm{GL}(n;\mathbb{R})$$

Beweis. Sicher gilt $A \cap N = \{I\}$, folglich definiert die Multiplikation eine Injektion

$$A \times N \hookrightarrow \mathrm{GL}(n;\mathbb{R})$$

Deren Bild AN ist nun offensichtlich eine Untergruppe, genauer die Gruppe der oberen Dreiecksmatrizen mit positiven Diagonaleinträgen, und wegen $\mathrm{O}(n) \cap AN = \{I\}$ definiert die Multiplikation schon mal eine Injektion $\mathrm{O}(n) \times AN \hookrightarrow \mathrm{GL}(n;\mathbb{R})$. Es bleibt, deren Surjektivität zu zeigen. Dazu betrachten wir in $\mathbb{R}^n$ die Standardbasis $\mathcal{S}$, eine beliebige angeordnete Basis $\mathcal{B}$ und die im Gram-Schmidt-Verfahren daraus entstehende angeordnete Basis $\mathcal{A}$. Unser Satz liefert für die zugehörige Basiswechselmatrix obere Dreiecksgestalt mit positiven Diagonaleinträgen, in Formeln

$${}_{\mathcal{B}}[\mathrm{id}]_{\mathcal{A}} \in AN$$

Aus ${}_{\mathcal{B}}[\mathrm{id}]_{\mathcal{A}} \circ {}_{\mathcal{A}}[\mathrm{id}]_{\mathcal{S}} = {}_{\mathcal{B}}[\mathrm{id}]_{\mathcal{S}}$ folgt dann die Surjektivität der Multiplikation $AN \times \mathrm{O}(n) \to \mathrm{GL}(n;\mathbb{R})$, und Invertieren liefert den Rest. □

Korollar 3.3.30 (**Iwasawa-Zerlegung für** $\mathrm{GL}(n;\mathbb{C})$**).** *Bezeichne* $A \subset \mathrm{GL}(n;\mathbb{C})$ *die Diagonalmatrizen mit reellen positiven Einträgen auf der Diagonale und* $N \subset \mathrm{GL}(n;\mathbb{C})$ *die oberen Dreiecksmatrizen mit Einsen auf der Diagonale. So definiert die Multiplikation eine Bijektion*

$$\mathrm{U}(n) \times A \times N \xrightarrow{\sim} \mathrm{GL}(n;\mathbb{C})$$

Beweis. Der Beweis geht analog wie im Reellen. □

3.4 Isometrien euklidischer affiner Räume

Definition 3.4.1. Ein **euklidischer reeller affiner Raum** ist ein Paar bestehend aus einem reellen affinen Raum und einem Skalarprodukt auf seinem Richtungsraum. Gegeben zwei Punkte p, q eines euklidischen reellen affinen Raums definieren wir ihren **Abstand** als die Länge des durch dieses Paar von Punkten erklärten Richtungsvektors, in Formeln

$$d(p,q) = \|p - q\|$$

Eine Abbildung $f : E \to E'$ zwischen reellen euklidischen affinen Räumen, die alle Abstände erhält, nennt man auch eine **Isometrie**. In Formeln fordern wir von einer Isometrie also

$$d(f(p), f(q)) = d(p,q) \quad \forall p, q \in E$$

Dieselbe Begriffsbildung verwendet man auch allgemeiner für Abbildungen zwischen sogenannten "metrischen Räumen", wie sie etwa in **??** erklärt werden. Ist eine Isometrie bijektiv, so spricht man von einem **isometrischen Isomorphismus**. Sprechen wir von einer **Isometrie eines Raums**, so meinen wir eine Abbildung dieses Raums in sich selber, die eine Isometrie ist.

Übung 3.4.2. Zwischen je zwei euklidischen reellen affinen Räumen derselben endlichen Dimension gibt es einen affinen isometrischen Isomorphismus.

Proposition 3.4.3. *Eine Abbildung zwischen euklidischen reellen affinen Räumen ist eine Isometrie genau dann, wenn sie affin ist mit orthogonalem linearen Anteil.*

Beweis. Sei $\varphi : E \to F$ unsere Abbildung und sei $p \in E$ beliebig gewählt. Erklären wir die Abbildung $\vec{\varphi} : \vec{E} \to \vec{F}$ durch die Vorschrift $\varphi(p + \vec{v}) =$

$\varphi(p) + \vec{\varphi}(\vec{v})$, so bildet $\vec{\varphi} : \vec{E} \to \vec{F}$ offensichtlich den Ursprung auf den Ursprung ab und erhält alle Abstände. Nach 3.3.3 ist folglich $\vec{\varphi}$ linear und orthogonal und damit φ affin mit orthogonalem linearem Anteil. Der Beweis der Gegenrichtung kann dem Leser überlassen bleiben. □

Satz 3.4.4 (**Isometrien affiner euklidischer Räume**). *Gegeben eine Isometrie φ eines endlichdimensionalen reellen affinen euklidischen Raums gibt es genau ein Paar $(d, \vec{w})$ bestehend aus einer Isometrie d mit mindestens einem Fixpunkt und einem Richtungsvektor $\vec{w}$ derart, daß gilt*

$$\varphi = (+\vec{w}) \circ d \quad \text{und} \quad \vec{d}(\vec{w}) = \vec{w}$$

3.4.5. In Worten läßt sich also jede Isometrie φ eines endlichdimensionalen reellen affinen euklidischen Raums eindeutig darstellen als Verknüpfung von einer Isometrie d mit Fixpunkt gefolgt von einer Verschiebung um einen unter dieser Isometrie invarianten Richtungsvektor. Natürlich haben dann d und φ denselben linearen Anteil, in Formeln $\vec{d} = \vec{\varphi}$, und unsere Isometrie kann dargestellt werden durch eine Abbildungsvorschrift der Gestalt $\varphi(x + \vec{u}) = x + \vec{\varphi}(\vec{u}) + \vec{w}$ mit $\vec{w}$ einem Fixvektor von $\vec{\varphi}$. Als x kann man dazu einen beliebigen Fixpunkt von d wählen.

Beweis. Gegeben ein orthogonaler Automorphismus f eines endlichdimensionalen euklidischen Vektorraums V ist $\ker(f - \mathrm{id}) = V^f$ das orthogonale Komplement von $\mathrm{im}(f - \mathrm{id})$ in V, in Formeln

$$V^f = \mathrm{im}(f - \mathrm{id})^\perp$$

In der Tat zeigt die Dimensionsformel 1.6.10 in Verbindung mit **??**, daß es ausreicht, die Inklusion $V^f \subset \mathrm{im}(f - \mathrm{id})^\perp$ zu zeigen. Aus $f(\vec{w}) = \vec{w}$ folgt aber offensichtlich $\langle \vec{w}, f(\vec{v}) - \vec{v} \rangle = \langle f(\vec{w}), f(\vec{v}) \rangle - \langle \vec{w}, \vec{v} \rangle = 0$ für alle $\vec{v} \in V$. Nun kann man ja für einen beliebigen Punkt $p \in E$ stets einen Vektor $\vec{v} \in \vec{E}$ finden mit $\varphi(p + \vec{u}) = p + \vec{\varphi}(\vec{u}) + \vec{v}$. Genau dann besitzt dann $(-\vec{w}) \circ \varphi$ einen Fixpunkt, wenn es $\vec{u} \in \vec{E}$ gibt mit

$$p + \vec{u} = p + \vec{\varphi}(\vec{u}) + \vec{v} - \vec{w}$$

alias $\vec{u} - \vec{\varphi}(\vec{u}) + \vec{w} = \vec{v}$. Wegen der Zerlegung $\vec{E} = \mathrm{im}(\vec{\varphi} - \mathrm{id}) \oplus \vec{E}^{\vec{\varphi}}$ vom Beginn des Beweises gibt es also genau ein $\vec{w} \in \vec{E}^{\vec{\varphi}}$ derart, daß $(-\vec{w}) \circ \varphi$ einen Fixpunkt hat. □

Beispiel 3.4.6 (**Isometrien der Gerade**). Jede abstandshaltende Selbstabbildung der reellen Zahlengeraden ist entweder eine Verschiebung $x \mapsto x + a$ oder eine Spiegelung $x \mapsto b - x$: In der Tat, ist der lineare Anteil unserer

Selbstabbildung die Identität, so handelt es sich nach 3.4.5 um eine Verschiebung; ist ihr linearer Anteil dahingegen das Negative der Identität, so muß in der Darstellung nach 3.4.5 der Vektor $\vec{w}$ der Nullvektor sein und wir haben eine Abbildung der Gestalt $x + \vec{u} \mapsto x - \vec{u}$ vor uns, die man wohl elementargeometrisch eine "Spiegelung am Punkt x" nennen würde.

Beispiel 3.4.7 (**Isometrien der Ebene**). Jede abstandserhaltende Selbstabbildung einer reellen euklidischen Ebene ist entweder (1) eine Verschiebung oder (2) eine Drehung um einen Punkt oder (3) eine Gleitspiegelung, d.h. eine Spiegelung an einer Gerade gefolgt von einer Verschiebung in Richtung eben dieser Gerade. In der Tat erhalten wir nach 3.4.5 Fall (1) für die Isometrien mit der Identität als linearem Anteil; Fall (2) für die Isometrien mit einer von der Identität verschiedenen Drehung als linearem Anteil; und Fall (3) für die Isometrien mit einer Spiegelung als linearem Anteil.

Beispiel 3.4.8 (**Isometrien des Raums**). Jede abstandserhaltende Selbstabbildung eines reellen dreidimensionalen euklidischen Raums ist entweder (1) eine "Verschraubung" alias eine Drehung um eine Achse gefolgt von einer Verschiebung in Richtung eben dieser Achse, oder (2) eine Drehung um eine Achse gefolgt von einer Spiegelung an einer Ebene senkrecht zu besagter Achse, oder (3) eine Verschiebung gefolgt von einer Spiegelung an einer unter besagter Verschiebung invarianten Ebene. In der Tat erhalten wir nach 3.4.5 Fall (1) für die Isometrien mit einer Drehung als linearem Anteil; Fall (2) für die Isometrien mit linearem Anteil bestehend im Sinne von 3.3.24 aus einem Drehblock und einem Eintrag (-1) auf der Diagonalen; und Fall (3) für die Isometrien mit einer Spiegelung an einer Ebene als linearem Anteil.

3.5 Winkel, Orientierung, Kreuzprodukt

Definition 3.5.1. Seien $\vec{v}, \vec{w}$ von Null verschiedene Vektoren eines reellen euklidischen Vektorraums. So ist der **von $\vec{v}$ und $\vec{w}$ eingeschlossene Winkel** $\vartheta \in [0, \pi]$ erklärt durch die Vorschrift

$$\cos\vartheta = \frac{\langle \vec{v}, \vec{w}\rangle}{\|\vec{v}\|\|\vec{w}\|}$$

Nach der Cauchy-Schwarz'schen Ungleichung 3.2.14 liegt der Quotient auf der rechten Seite dieser Gleichung stets im Intervall $[-1, 1]$ und nach **??** existiert stets genau ein Winkel $\vartheta \in [0, \pi]$ zu jedem in $[-1, 1]$ vorgegebenen Wert von $\cos\vartheta$. Man notiert diesen Winkel auch

$$\vartheta = \angle(\vec{v}, \vec{w}) = \arccos\left(\frac{\langle \vec{v}, \vec{w}\rangle}{\|\vec{v}\|\|\vec{w}\|}\right)$$

Ich verstehe stets cos im Sinne von **??** als die Abbildung $\cos : \mathbb{R} \to \mathbb{R}$, die von einem "Winkel im Bogenmaß" ausgeht, so daß etwa gilt $\cos(\pi/2) = 0$. Ich will im folgenden und insbesondere in 3.5.21 diskutieren, inwiefern diese Definition die "Richtige" ist.

Bemerkung 3.5.2. Gegeben $\lambda, \mu > 0$ gilt $\angle(\vec{v}, \vec{w}) = \angle(\lambda\vec{v}, \mu\vec{w})$. Zwei von Null verschiedene Vektoren stehen aufeinander senkrecht genau dann, wenn sie den Winkel $\pi/2$ einschließen.

Beispiel 3.5.3. Die drei Vektoren der Standardbasis des $\mathbb{R}^3$ bilden ja wohl ein gleichseitiges Dreieck und der Winkel an jeder Ecke sollte folglich $\pi/3$ sein. In der Tat finden wir für $\vec{v} = \vec{e}_3 - \vec{e}_1$ und $\vec{w} = \vec{e}_2 - \vec{e}_1$ als Skalarprodukt $\langle \vec{v}, \vec{w} \rangle = 1$ und wegen $\|\vec{v}\| = \|\vec{w}\| = \sqrt{2}$ ergibt sich für den Winkel $\cos\vartheta = 1/2$ alias $\vartheta = \pi/3$.

Übung 3.5.4. Gegeben zwei vom selben Punkt p ausgehende Halbgeraden L, R in einem euklidischen affinen Raum definiert man ihren Winkel $\angle(L, R)$ als $\angle(l - p, r - p)$ für beliebige $l \in L\backslash p$, $r \in R\backslash p$. Gegeben zwei Paare (L, R) und (L', R') von jeweils vom selben Punkt ausgehenden Halbgeraden in einem endlichdimensioanlen euklidischen affinen Raum zeige man, daß es genau dann eine Isometrie b von unserem Raum auf sich selber gibt mit $b(L) = L'$ und $b(R) = R'$, wenn gilt $\angle(L, R) = \angle(L', R')$.

Definition 3.5.5. Eine **Orientierung** eines endlichdimensionalen Vektorraums V über einem angeordneten Körper ist eine Vorschrift ε, die jeder angeordneten Basis B unseres Vektorraums ein Vorzeichen $\varepsilon(B) \in \{+1, -1\}$ zuordnet und zwar so, daß für je zwei angeordnete Basen B, B' die Determinante der Basiswechselmatrix das Vorzeichen $\varepsilon(B)\varepsilon(B')$ hat. Das Vorzeichen $\varepsilon(B)$ nennen wir dann die **Orientierung** der angeordneten Basis B unseres orientierten Vektorraums. Eine angeordnete Basis der Orientierung $+1$ nennen wir eine **orientierte Basis**. Sprechen wir von der **durch eine angeordnete Basis gegebene Orientierung**, so meinen wir diejenige Orientierung, die besagter Basis das Vorzeichen $+1$ zuordnet. Ein Isomorphismus von orientierten endlichdimensionalen Vektorräumen heißt **orientierungserhaltend** genau dann, wenn er die Orientierung von angeordneten Basen erhält. Andernfalls heißt er **orientierungsumkehrend**. Gegeben ein angeordneter Körper k bezeichnen wir diejenige Orientierung des k^n als die **Standardorientierung**, die der Standardbasis das Vorzeichen $+1$ zuordnet. Unter einer **Orientierung eines endlichdimensionalen affinen Raums** über einem angeordneten Körper verstehen wir eine Orientierung seines Richtungsraums.

3.5.6. In der Literatur findet man vielfach eine Variation der Definition der Orientierung, bei der eine Orientierung eines reellen Vektorraums als eine

Äquivalenzklasse von Basen unter einer geeigneten Äquivalenzrelation erklärt wird. Diese Definition liefert dasselbe in allen Fällen mit Ausnahme des Nullraums, und in diesem Fall scheint mir die hier gegebene Definition das sinnvollere Konzept zu liefern: Es erlaubt nämlich, den Hauptsatz der Differential- und Integralrechnung auch formal als Spezialfall des Satzes von Stokes anzusehen, vergleiche **??**.

3.5.7. Mit dieser Terminologie kann man etwa $\mathrm{SO}(n)$ auch als die Gruppe aller orientierungserhaltenden orthogonalen Automorphismen von $\mathbb{R}^n$ beschreiben.

3.5.8. Jeder endlichdimensionale Raum über einem angeordneten Körper besitzt genau zwei Orientierungen. Das gilt insbesondere auch für jeden einpunktigen Raum: Hier verwenden wir die Konvention, nach der der einzige Endomorphismus des Nullvektorraums die Determinante 1 hat. Der Nullvektorraum hat eine einzige angeordnete Basis, nämlich die leere Menge mit ihrer einzigen Anordnung, und eine Orientierung des Nullvektorraums zu wählen bedeutet schlicht, das Vorzeichen auszusuchen, das dieser Basis zugeordnet werden soll.

Beispiel 3.5.9. Eine Orientierung einer reellen Gerade anzugeben bedeutet anschaulich, auf dieser Gerade eine "Richtung" auszuwählen, eben die Richtung, in die diejenigen Vektoren zeigen, die positiv orientierte Basen ihres Richtungsraums bilden. Wir nennen diese Vektoren dann auch kurzerhand **positiv orientierte Vektoren** oder noch kürzer **positive Vektoren** und denken uns unsere Gerade mit der Anordnung versehen, für die die Addition positiver Vektoren Elemente vergrößert. Mit diesen Konventionen können wir für einen orientierten eindimensionalen Vektorraum L die Menge der positiven Vektoren mit $L_{>0}$ bezeichnen. Analog vereinbaren wir für die Elemente von $L_{<0}$ die Bezeichnung **negative Vektoren** und nennen folgerichtig die Elemente von $L_{\geq 0}$ die **nichtnegativen Vektoren**.

Beispiel 3.5.10. Denken wir uns die Tafelebene als einen zweidimensionalen reellen affinen Raum, so dürfen wir uns eine Orientierung der Tafelebene als die Auszeichnung eines Drehsinns denken, nämlich den Drehsinn derart, daß bei Drehung in diesem Drehsinn der erste Vektor einer positiv orientierten angeordneten Basis ihres Richtungsraums zuerst in ein positives Vielfaches des zweiten Vektors gedreht wird und erst dann in ein negatives Vielfaches. Wenn, wie etwa bei der Tafelebene oder bei einem vor uns liegenden Blatt Papier, zusätzlich festlegt ist, "von welcher Seite man auf eine Ebene gucken soll", so mag man diese beiden Orientierungen als "im Uhrzeigersinn" und "im Gegenuhrzeigersinn" ansprechen. Ist unsere Ebene dahingegen eine Glasscheibe und die Betrachter stehen auf beiden Seiten, so legt man eine

Orientierung besser fest, indem man einen Drehsinn mit einem Wachsstift einzeichnet.

Beispiel 3.5.11. Den eindimensionalen affinen Raum $\mathbb{T}$ aller Zeiten aus 1.9.7 denken wir uns stets mit der Orientierung versehen, für die jeder Richtungsvektor, der einen Zeitpunkt auf einen "späteren" Zeitpunkt schiebt, eine positiv orientierte Basis bildet. Den Richtungsraum $\vec{\mathbb{T}}$ bezeichnen wir als den Raum aller **Zeitspannen**, seine positiv orientierten Vektoren nennen wir **Zeiteinheiten**. Wir wählen für das folgende eine feste Zeiteinheit und nennen sie die **Sekunde** $s \in \vec{\mathbb{T}}$. Die Einteilung eines Tages in vierundzwanzig Stunden und die Einteilung dieser Stunden in je sechzig Minuten geht wohl auf die Babylonier zurück, die angeblich mit ihren Händen bis 60 zählten, indem sie mit jedem der 5 Finger der rechten Hand der Reihe nach die 12 Fingerglieder der linken Hand an den Fingern mit Ausnahme des Daumens berührten. Die Einteilung jeder Minute in wiederum 60 Sekunden bot sich dann als natürliche Verfeinerung an.

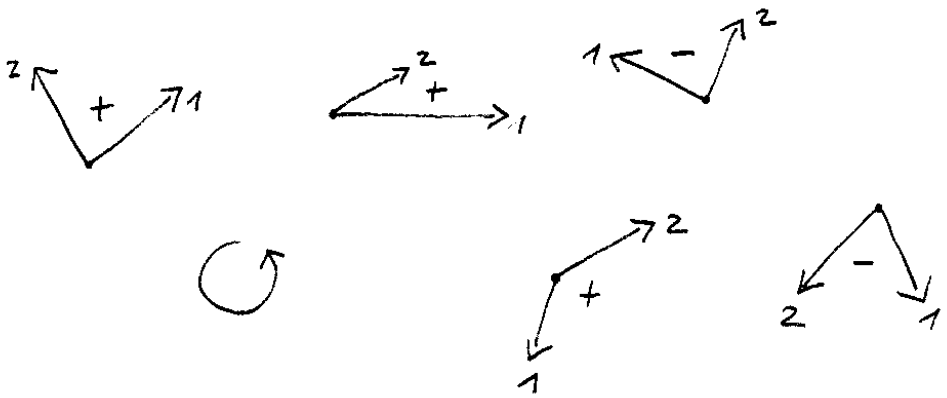
Beispiel 3.5.12. Auf unserem dreidimensionalen Anschauungsraum $\mathbb{E}$ können wir uns die "rechte-Hand-Orientierung" denken, in der die durch die Abfolge "Daumen-Zeigefinger-Mittelfinger" mit der rechten Hand angedeuteten angeordneten Basen positiv orientiert sind, und analog die davon verschiedene "linke-Hand-Orientierung".

3.5.13. Zwei angeordnete Basen eines endlichdimensionalen reellen Vektorraums liefern dieselbe Orientierung genau dann, wenn sie sich "stetig ineinander deformieren lassen" alias in derselben "Zusammenhangskomponente" des Raums aller angeordneten Basen liegen. Man kann sich davon etwa mithilfe der Iwasawa-Zerlegung 3.3.29 überzeugen. Auch die präzise Formulierung und der formale Beweis wird Ihnen davon ausgehend leicht gelingen, sobald Sie in der Analysis die Grundtatsachen über Stetigkeit in mehreren Veränderlichen kennengelernt haben. Eine äquivalente Aussage dürfen Sie in der Analysis als Übung **??** zeigen.

Definition 3.5.14. Gegeben ein orientierter zweidimensioaler reeller euklidischer Vektorraum V und ein Winkel ϑ bezeichne $R_\vartheta : V \to V$ diejenige lineare Abbildung, die in einer und jeder orientierten Orthonormalbasis von V die dargestellt wird durch die Matrix

$$\begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix}$$

3.5.15. Denken wir uns V als eine unendliche Tafel mit einem ausgezeichneten Ursprung und der Orientierung, für die "erst ein Pfeil nach rechts, dann ein Pfeil nach oben" eine orientierte Basis ist, so müssen wir uns R_ϑ denken



Angeordnete Basen des Raums der Richtungsvektoren der Papierebene mit den Vorzeichen, die der Orientierung “im Gegenuhrzeigersinn” entsprechen

als die "Drehung mit Zentrum im Ursprung um den Winkel ϑ im Gegenuhrzeigersinn".

Definition 3.5.16. Gegeben ein orientierter zweidimensionaler reeller euklidischer Vektorraum und zwei von Null ausgehende Halbgeraden G, H definieren wir ihren **orientierten Winkel**

$$\vartheta = \measuredangle(G, H) \in (-\pi, \pi]$$

als das eindeutig bestimmte $\vartheta \in (-\pi, \pi]$ mit $R_\vartheta(G) = H$. Gegeben zwei von Null verschiedene Vektoren v, w definieren wir ihren orientierten Winkel dann als $\measuredangle(v, w) = \measuredangle(\mathbb{R}_{\geq 0}v, \mathbb{R}_{\geq 0}w)$.

Übung 3.5.17. Gegeben ein von Null verschiedener Vektor $v \neq 0$ eines orientierten zweidimensionalen reellen euklidischen Vektorraums gilt für seinen orientierten Winkel mit seinem Negativen stets $\measuredangle(v, -v) = \pi$.

3.5.18. Gegeben von Null ausgehende Halbgeraden F, G, H in einem orientierten zweidimensionalen reellen euklidischen Vektorraum gilt stets die **Additivität der orientierten Winkel**

$$\measuredangle(F, H) \in \measuredangle(F, G) + \measuredangle(G, H) + 2\pi\mathbb{Z}$$

In der Tat gilt nach 3.3.16 ja $R_\vartheta R_\psi = R_{\vartheta+\psi}$ und $R_{2\pi} = \mathrm{id}$ ist eh klar.

Übung 3.5.19. Gegeben von Null verschiedene Vektoren v, w in einem orientierten zweidimensionalen reellen euklidischen Vektorraum haben wir stets $\measuredangle(v, w) + \measuredangle(w, -v) = \pm\pi$.

Übung 3.5.20. Der nichtorientierte Winkel kann aus dem orientierten Winkel berechnet werden vermittels der Beziehung

$$\angle(G, H) = |\measuredangle(G, H)|$$

3.5.21. Der eigentliche Grund für unsere Winkeldefinition 3.5.1 liegt in seinem engen Zusammenhang mit dem orientierten Winkel, dessen Definition hinwiederum durch die Additivität 3.5.18 motiviert ist. Natürlich könnten wir diese Additivität auch durch die Verwendung eines anderen Gruppenhomomorphismus $\mathbb{R} \to \mathrm{SO}(2)$ erreichen, und in der Tat sind hier auch viele andere Wahlen verbreitet. Die meisten sind von der Gestalt $\vartheta \mapsto R_{a\vartheta}$ für $a > 0$.

1. Auf der Schule wird der Gruppenhomomorphismus meist so gewählt, daß 360 die kleinste positive Zahl ist, die auf die Identität abgebildet wird: Das hat den Vorteil, daß die Winkel vieler einfacher geometrischer Figuren ganzen Zahlen entsprechen. Man deutet es bei der Winkeldarstellung durch ein hochgestelltes $^\circ$ an, wenn man mit dieser Wahl arbeitet, und spricht von "Grad".

2. Bei Vermessungsarbeiten wird der Gruppenhomomorphismus meist so gewählt, daß 400 die kleinste positive Zahl ist, die auf die Identität abgebildet wird: Das hat den Vorteil, daß rechte Winkel der Zahl 100 entsprechen, und ist dem Arbeiten mit computergesteuerten Geräten, die ja mit ihrem Bedienungspersonal üblicherweise im Zehnersystem kommunizieren, besonders gut angepaßt. Man deutet es bei der Winkeldarstellung durch ein nachgestelltes gon an, wenn man mit dieser Wahl arbeitet, und spricht von "Neugrad" oder "Gon".

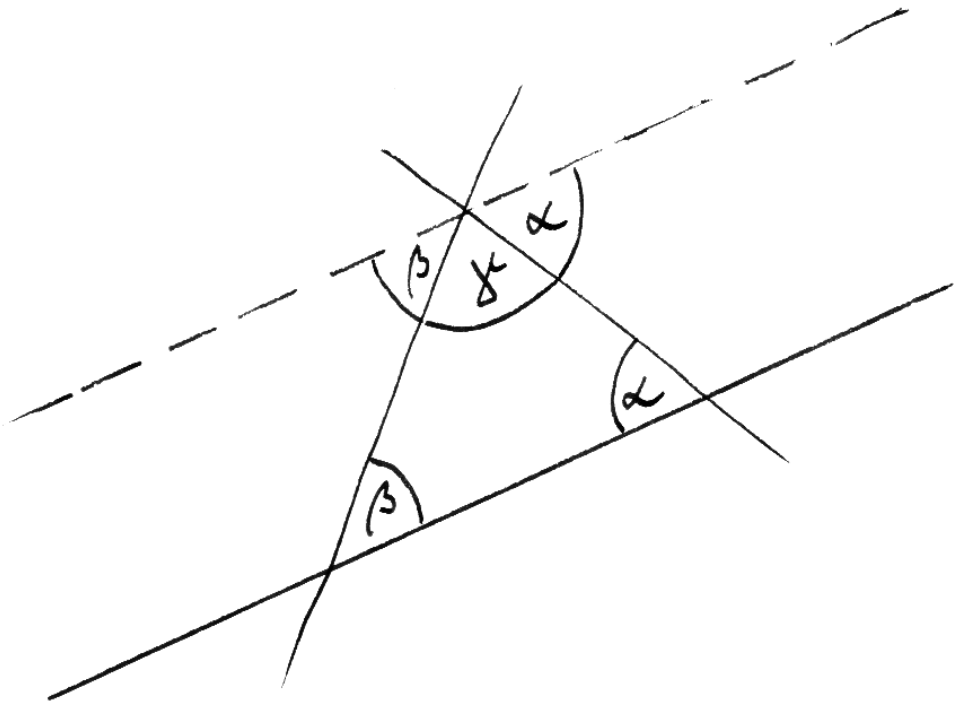
3. Mathematisch-abstrakt schiene es mir am natürlichsten, unsere orientierten Winkel schlicht als Elemente der Gruppe SO(2) aufzufassen, aber das wäre für die Anwender unpraktisch.

4. Die in diesem Text getroffene Wahl ist bei rechtem Licht betrachtet eigentlich die Wahl $a = \pi$: Wir drücken ja unsere Winkel in Wirklichkeit als Vielfache von π aus und kommen nicht ernsthaft auf die Idee, hier wirklich $\pi = 3,1415\ldots$ einzusetzen, auszumultiplizieren und die entstehende reelle Zahl mit einigen Nachkommastellen hinzuschreiben! Schreibt man diese reelle Zahl doch aus, so sollte man rad als Abkürzung für "Radian", zu deutsch "Bogenmaß", dahinterschreiben, um klarzumachen, welcher Winkel gemeint ist.

In gewisser Weise spielt das Symbol π bei unserer Winkelbezeichnung also eine ähnliche Rolle wie das hochgestellte $^\circ$ bei der auf der Schule üblichen Bezeichnungsweise. Ich halte es nicht für besonders glücklich, daß hier π nur für den halben und nicht für den ganzen Vollkreis steht, aber so ist die Notation nun einmal geschichtlich gewachsen, und die nachträgliche Einführung eines zusätzlichen eigenen Symbols für den Umfang eines Kreises mit Radius Eins will ich nun auch wieder nicht propagieren.

3.5.22. Wir zeigen nun auch in diesem Rahmen, daß die Winkelsumme im Dreieck 180° alias π ist. Ich will nicht behaupten, daß der anschließende Beweis klarer sei als der anschauliche Beweis, wie Sie Ihn vermutlich in der Schule kennengelernt haben. Ich will jedoch zeigen, wie dieser anschauliche Beweis in das "Paradies der Mengenlehre" hinübergerettet werden kann, in dem wir uns ja mittlerweile bewegen. Die ungeheure Eleganz und Effizienz der Sprache der Mengenlehre kommt in diesem Beispiel schlecht zur Geltung, in dem man eher den Eindruck gewinnen mag, mit Kanonen auf Spatzen zu schießen. Es handelt sich eben auch nicht um einen Ernstfall, sondern vielmehr um eine Kanonenprobe.

Proposition 3.5.23 (Winkelsumme im Dreieck). *Für drei Punkte p, q, r einer affinen euklidischen Ebene E, die nicht auf einer Geraden liegen, gilt*



Der auf der Schule übliche Beweis dafür, daß die Winkelsumme im Dreieck 180° ist

stets

$$\angle(q-p,r-p)+\angle(p-r,q-r)+\angle(r-q,p-q)=\pi$$

Beweis. Zunächst wählen wir eine Orientierung auf $\vec{E}$ und beachten, daß aufgrund unserer Definitionen für $\vec{v},\vec{w}\in\vec{E}$ linear unabhängig gilt

$$\measuredangle(\vec{v},\vec{w}) \;=\; \begin{cases} \angle(\vec{v},\vec{w}) & \text{falls } (\vec{v},\vec{w}) \text{ eine orientierte Basis von } \vec{E} \text{ ist;} \\ -\angle(\vec{v},\vec{w}) & \text{sonst.} \end{cases}$$

Jetzt kürzen wir die "Kantenvektoren" ab zu $\vec{v}_1=q-p,\vec{v}_2=p-r,\vec{v}_3=r-q$, so daß gilt $\vec{v}_1+\vec{v}_2+\vec{v}_3=\vec{0}$. Daraus folgt, daß $(\vec{v}_1,\vec{v}_2),(\vec{v}_2,\vec{v}_3)$ und $(\vec{v}_3,\vec{v}_1)$ alle drei gleich orientierte Basen sind, da nämlich die entsprechenden Basiswechselmatrizen alle positive Determinante haben. Für die orientierten Winkel wissen wir wegen der Additivität 3.5.18 bereits

$$\measuredangle(\vec{v}_1,\vec{v}_2)+\measuredangle(\vec{v}_2,\vec{v}_3)+\measuredangle(\vec{v}_3,\vec{v}_1)\in 2\pi\mathbb{Z}$$

Weiter gilt für $\vec{v},\vec{w}\neq\vec{0}$ nach 3.5.19 stets $\measuredangle(\vec{v},\vec{w})+\measuredangle(\vec{w},-\vec{v})=\pm\pi$ und damit folgt

$$\measuredangle(\vec{v}_1,-\vec{v}_2)+\measuredangle(\vec{v}_2,-\vec{v}_3)+\measuredangle(\vec{v}_3,-\vec{v}_1)\in\pi+2\pi\mathbb{Z}$$

Wir wissen aber bereits, daß diese drei orientierten Winkel alle positiv oder alle negativ sind und genauer, daß sie alle in $(0,\pi)$ oder $(-\pi,0)$ liegen. In beiden Fällen folgt unmittelbar

$$\angle(\vec{v}_1,-\vec{v}_2)+\angle(\vec{v}_2,-\vec{v}_3)+\angle(\vec{v}_2,-\vec{v}_1)=\pi$$ □

Definition 3.5.24. Als **Kreuzprodukt** bezeichnet man die bilineare Abbildung $\mathbb{R}^3\times\mathbb{R}^3\to\mathbb{R}^3$, die gegeben wird durch die Vorschrift

$$\left(\begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix},\begin{pmatrix}w_1\\w_2\\w_3\end{pmatrix}\right)\;\mapsto\;\begin{pmatrix}v_2w_3-v_3w_2\\v_2w_1-v_1w_3\\v_1w_2-v_2w_1\end{pmatrix}$$

Man notiert diese Abbildung mit einem Kreuz in der Form $(\vec{v},\vec{w})\mapsto\vec{v}\times\vec{w}$, daher die Bezeichnung als Kreuzprodukt.

3.5.25. Das Kreuzprodukt hat in Anbetracht der Jägerzaunformel 2.7.3 die Eigenschaft

$$\langle\vec{u},\vec{v}\times\vec{w}\rangle=\det(\vec{u}|\vec{v}|\vec{w})$$

und diese Eigenschaft für alle $\vec{u}\in\mathbb{R}^3$, ja sogar schon für $\vec{u}=\vec{e}_1,\vec{e}_2,\vec{e}_3$ charakterisiert auch bereits den Vektor $\vec{v}\times\vec{w}$. Für den Ausdruck $\langle\vec{u},\vec{v}\times\vec{w}\rangle=\det(\vec{u}|\vec{v}|\vec{w})$ findet man manchmal auch die Bezeichnung als **Spatprodukt**, die

darauf anspielt, daß diese Determinante ja nach 2.7.6 bis auf ein Vorzeichen gerade das Volumen des durch die fraglichen drei Vektoren gegebenen Parallelpipeds angibt. Die Kristalle des Feldspats haben aber nun oft die Gestalt eines Parallelpipeds, weswegen derartige Körper auch als **Spate** bezeichnet werden.

3.5.26. Die Charakterisierung 3.5.25 des Kreuzprodukts mithilfe der Determinante liefert unmittelbar eine Vielzahl von Eigenschaften.

1. Es gilt $\vec{v} \times \vec{w} = -\vec{w} \times \vec{v}$, denn Vertauschen zweier Spalten einer Matrix ändert das Vorzeichen der Determinante. Alternativ kann man das der definierenden Formel auch direkt ansehen.

2. $\vec{v} \times \vec{w}$ steht senkrecht auf $\vec{v}$ und $\vec{w}$, denn die Determinante jeder Matrix mit zwei gleichen Spalten ist Null.

3. Genau dann gilt $\vec{v} \times \vec{w} = \vec{0}$, wenn $(\vec{v}, \vec{w})$ ein linear abhängiges Paar von Vektoren ist. In der Tat, die Determinante jeder Matrix mit linear abhängigen Spalten ist Null, und je zwei linear unabhängige Vektoren $\vec{v}, \vec{w}$ des $\mathbb{R}^3$ lassen sich andererseits auch durch einen dritten Vektor $\vec{u}$ zu einer Basis ergänzen.

4. Ist $(\vec{v}, \vec{w})$ ein linear unabhängiges Paar von Vektoren, so muß das Tripel $(\vec{v} \times \vec{w}, \vec{v}, \vec{w})$ in dieser Anordnung eine orientierte Basis des $\mathbb{R}^3$ mit seiner Standardorientierung sein, denn die Determinante der Basiswechselmatrix von der Standardbasis ist ja dann das Längenquadrat des von Null verschiedenen Vektors $\vec{v} \times \vec{w}$.

3.5.27. Nehmen wir im Fall eines linear unabhängigen Paares $(\vec{v}, \vec{w})$ in obiger Formel $\vec{u} = (\vec{v} \times \vec{w})/\|\vec{v} \times \vec{w}\|$, so erkennen wir aus der anschaulichen Bedeutung 2.7.6 der Determinante als Volumen, daß wir uns die Länge von $\vec{v} \times \vec{w}$ gerade als das Volumen des von $\vec{u}, \vec{v}, \vec{w}$ aufgespannten Spats alias die Fläche des von $\vec{v}, \vec{w}$ aufgespannten Parallelograms denken dürfen.

3.5.28. Die Bedeutung des Kreuzprodukts liegt nun darin, daß es einerseits algebraisch leicht zu berechnen ist und andererseits die im Vorhergehenden erklärte einfache geometrische Interpretation hat, die ich hier nocheinmal zusammenfassen will: Für $\vec{v}, \vec{w}$ linear abhängig gilt $\vec{v} \times \vec{w} = \vec{0}$; Sonst ist $\vec{v} \times \vec{w}$ der Vektor, der senkrecht steht auf $\vec{v}$ und $\vec{w}$, dessen Länge der Fläche des von $\vec{v}$ und $\vec{w}$ aufgespannten Parallelogramms entspricht, und dessen Richtung dadurch festgelegt wird, daß $(\vec{v} \times \vec{w}, \vec{v}, \vec{w})$ dieselbe Orientierung hat wie die Standardbasis des $\mathbb{R}^3$.

3.5.29. Unsere geometrische Interpretation legt es nahe, daß für alle Drehungen $A \in \mathrm{SO}(3)$ gelten sollte

$$(A\vec{v}) \times (A\vec{w}) = A(\vec{v} \times \vec{w})$$

In der Tat, da für alle $A \in \mathrm{SO}(3)$ gilt $\langle A\vec{u}, A(\vec{v} \times \vec{w})\rangle = \langle \vec{u}, \vec{v} \times \vec{w}\rangle = \det(\vec{u}|\vec{v}|\vec{w}) = \det A(\vec{u}|\vec{v}|\vec{w}) = \det(A\vec{u}|A\vec{v}|A\vec{w}) = \langle A\vec{u}, (A\vec{v}) \times (A\vec{w})\rangle$, ergibt sich ohne Schwierigkeiten $A(\vec{v} \times \vec{w}) = (A\vec{v}) \times (A\vec{w})$ für alle $A \in \mathrm{SO}(3)$.

3.5.30. Allgemeiner kann man für einen beliebigen orientierten dreidimensionalen reellen euklidischen Vektorraum V das **Kreuzprodukt**

$$\begin{array}{ccc} V \times V & \to & V \\ (\vec{v}, \vec{w}) & \mapsto & \vec{v} \times \vec{w} \end{array}$$

dadurch erklären, daß es unter einem und jedem orientierungserhaltenden orthogonalen Isomorphismus $V \xrightarrow{\sim} \mathbb{R}^3$ dem eben definierten Kreuzprodukt entsprechen soll. Wir werden das nicht weiter vertiefen.

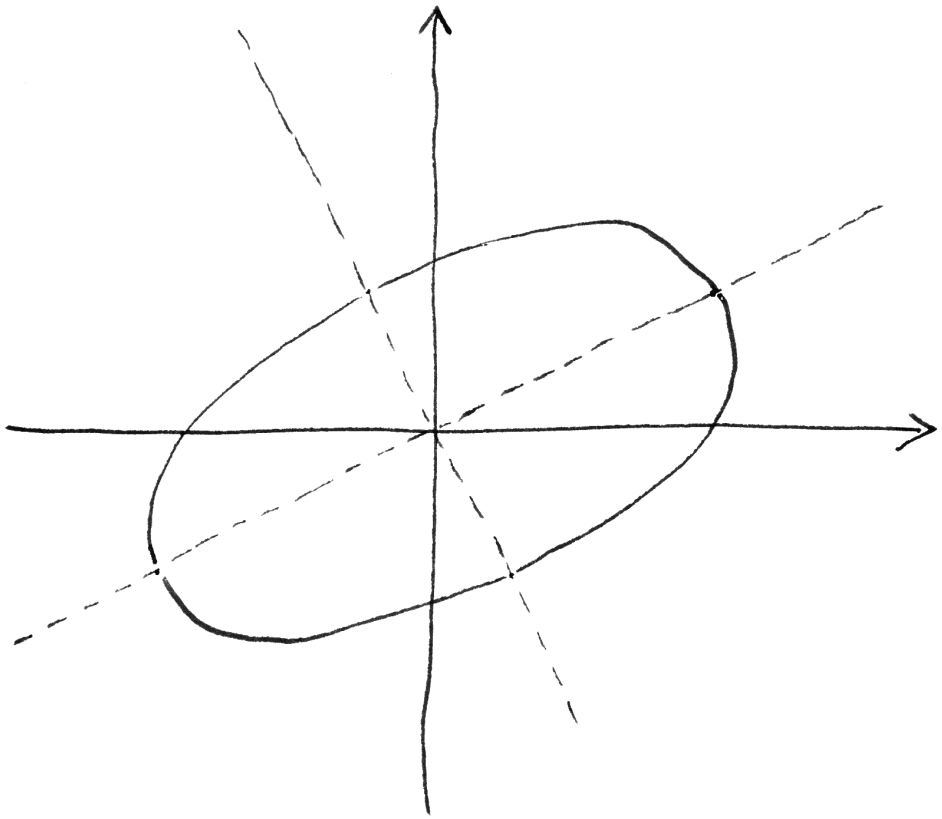
3.6 Spektralsatz und Hauptachsentransformationen

Satz 3.6.1 (**Hauptachsentransformation).** *Gegeben eine* **quadratische Form** *alias eine Funktion $q : \mathbb{R}^n \to \mathbb{R}$ der Gestalt $q(x_1, \ldots, x_n) = \sum_{i \le j} c_{ij} x_i x_j$ gibt es stets eine Drehung $D \in \mathrm{SO}(n)$ des $\mathbb{R}^n$ und Skalare $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ mit*

$$(q \circ D)(x_1, \ldots, x_n) = \lambda_1 x_1^2 + \ldots + \lambda_n x_n^2$$

3.6.2. Man kann sich die Bedeutung dieses Satzes auf zwei Weisen veranschaulichen: Entweder "aktiv" in dem Sinne, daß der Graph unserer Funktion q unter der Drehung D^{-1} oder präziser der Abbildung $D^{-1} \times \mathrm{id}$ in den Graphen unserer Linearkombination von Quadraten übergeht; Oder "passiv" in dem Sinne, daß unsere Funktion beim Einführen neuer Koordinaten mit Koordinatenachsen in Richtung der Spaltenvektoren von D in den neuen Koordinaten ausgedrückt die fragliche Form annimmt, in Formeln $q(y_1\vec{v}_1 + \ldots + y_n\vec{v}_n) = \lambda_1 y_1^2 + \ldots + \lambda_n y_n^2$ für $\vec{v}_i$ die Spalten von D, also für $D = (\vec{v}_1|\ldots|\vec{v}_n)$. Die von den Spalten der Matrix D erzeugten Geraden bilden dann ein System von "Hauptachsen" für unsere quadratische Form q.

3.6.3. In der Analysis, etwa in **??**, können Sie lernen, wie man eine hinreichend differenzierbare Funktion $\mathbb{R}^n \to \mathbb{R}$ etwa um den Ursprung bis zu zweiter Ordnung approximieren kann durch eine polynomiale Funktion vom Totalgrad höchstens Zwei alias eine Summe von einer Konstanten, einer Linearform und einer quadratischen Form. Ist die fragliche Linearform Null alias



Dieses Bild zeigt die Ellipse, auf der die positiv definite quadratische Form $17x^2 - 12xy + 8y^2$ bei einer geeigneten Wahl des Maßstabs den Wert Eins annimmt. Gestrichelt sind die Hauptachsen eingetragen, die in diesem Fall die Richtungsvektoren $(2, 1)$ und $(-1, 2)$ haben.

hat der Graph unserer Funktion am Ursprung eine horizontale Tangentialebene alias hat unsere Funktion am Ursprung eine "kritische Stelle", so wird sie dort bis zur Ordnung Zwei approximiert durch die fragliche quadratische Form plus die Konstante. So führt dann das Studium der Minima und Maxima von Funktionen mehrerer Veränderlichen auf das Studium quadratischer Formen.

Beweis. Eine Matrix A heißt **symmetrisch** genau dann, wenn sie mit ihrer eigenen Transponierten übereinstimmt, in Formeln $A^\top = A$. Wir finden eine symmetrische Matrix $A \in \mathrm{M}(n \times n; \mathbb{R})$ mit

$$q(x) = x^\top A x$$

für den Spaltenvektor $x = (x_1, \ldots, x_n)^\top$, indem wir als diagonale Matrixeinträge $a_{ii} = c_{ii}$ nehmen und außerhalb der Diagonalen $a_{ij} = a_{ji} = c_{ij}/2$ setzen. Nach 3.6.5 gibt es dann eine Drehung $D \in \mathrm{SO}(n)$ mit $D^{-1}AD = D^\top AD = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ für geeignete $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, nämlich die Eigenwerte von A mit ihren Vielfachheiten. Es folgt

$$q(Dx) = x^\top D^\top ADx = \lambda_1 x_1^2 + \ldots + \lambda_n x_n^2 \qquad \square$$

Übung 3.6.4. Man zeige: Gegeben eine Polynomfunktion vom Grad höchstens zwei mit reellen Koeffizienten, also eine Abbildung $q : \mathbb{R}^n \to \mathbb{R}$ der Gestalt

$$q(x_1, \ldots, x_n) = \sum_{i \le j} a_{ij} x_i x_j + \sum_i b_i x_i + c$$

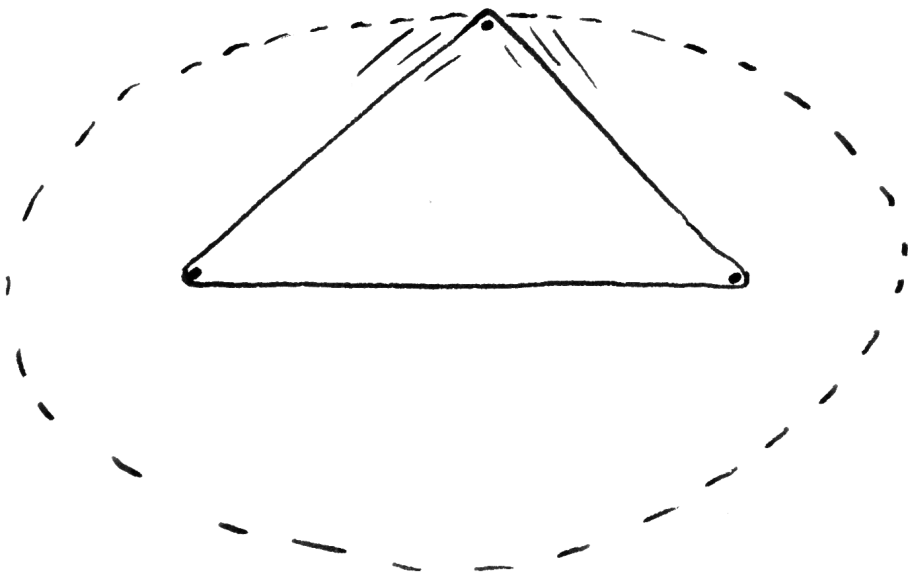
gibt es eine abstandserhaltende Selbstabbildung $D : \mathbb{R}^n \to \mathbb{R}^n$ mit

$$(q \circ D)(x_1, \ldots, x_n) = \lambda_1 x_1^2 + \ldots + \lambda_k x_k^2 + \lambda_{k+1} x_{k+1} + \ldots + \lambda_n x_n + \lambda_0$$

für geeignetes k und geeignete reelle λ_i. Man sagt dann, "unter unserer Bewegung D gehe unsere Quadrik in ihre Standardform über".

Proposition 3.6.5. *Gegeben eine symmetrische Matrix $A \in \mathrm{M}(n \times n; \mathbb{R})$ gibt es eine orthogonale Matrix mit Determinante Eins $D \in \mathrm{SO}(n)$ derart, daß $D^\top AD = D^{-1}AD$ diagonal ist.*

Beweis. Das folgt sofort aus dem Spektralsatz für "selbstadjungierte" Abbildungen 3.6.10, indem wir als Spalten von D die Vektoren einer Orthonormalbasis von $\mathbb{R}^n$ aus Eigenvektoren von $A : \mathbb{R}^n \to \mathbb{R}^n$ nehmen, und notfalls noch eine Spalte mit (-1) multiplizieren, um $\det D = 1$ zu erreichen. $\square$



Wie Gärtner Ellipsen zeichnen.

Definition 3.6.6. Seien V, W euklidische Vektorräume und $A : V \to W$ und $B : W \to V$ lineare Abbildungen. Unsere Abbildungen heißen zueinander **adjungiert** genau dann, wenn gilt

$$\langle Av, w\rangle = \langle v, Bw\rangle \quad \forall v \in V, w \in W$$

Jede lineare Abbildung A wie oben hat höchstens eine Adjungierte, denn sind B, C beide adjungiert zu A, so folgt $\langle v, Bw - Cw\rangle = 0 \quad \forall v, w$ und damit $Bw = Cw \quad \forall w$. Versehen wir $\mathbb{R}^n, \mathbb{R}^m$ jeweils mit dem Standardskalarprodukt, so wird für $A \in \mathrm{M}(m \times n; \mathbb{R})$ die adjungierte Abbildung zu $A : \mathbb{R}^n \to \mathbb{R}^m$ gegeben durch die transponierte Matrix als $A^\top : \mathbb{R}^m \to \mathbb{R}^n$. Ebenso ist im Komplexen $\bar{A}^\top : \mathbb{C}^m \to \mathbb{C}^n$ adjungiert zu $A : \mathbb{C}^n \to \mathbb{C}^m$. In der Tat finden wir mühelos

$$\langle Ax, y\rangle = (\overline{Ax})^\top y = \bar{x}^\top \bar{A}^\top y = \langle x, \bar{A}^\top y\rangle$$

für alle $x \in \mathbb{C}^m$, $y \in \mathbb{C}^n$. Wir folgern mit 3.3.6, daß jede lineare Abbildung von endlichdimensionalen reellen oder komplexen euklidischen Vektorräumen genau eine Adjungierte hat.

3.6.7. Im folgenden geben wir noch eine abstrakte Konstruktion der adjungierten Abbildung im endlichdimensionalen Fall. Zu jedem komplexen Vektorraum V bilden wir zunächst den **komplex konjugierten Vektorraum** $\overline{V}$, indem wir dieselbe unterliegende additive Gruppe nehmen, die Operation von $a \in \mathbb{C}$ auf $v \in V$ jedoch abändern zu einer Operation $a \cdot v$, die mit der ursprünglichen Operation av verknüpft ist durch die Formel $a \cdot v = \bar{a}v$. Für jede $\mathbb{C}$-lineare Abbildung $f : V \to W$ von komplexen Vektorräumen ist dieselbe Abbildung auch eine $\mathbb{C}$-lineare Abbildung $\overline{V} \to \overline{W}$. Wir bezeichnen diese Abbildung dennoch mit dem neuen Symbol $\bar{f} : \overline{V} \to \overline{W}$ und nennen sie die **konjugierte Abbildung**, weil ihre Matrix anders aussieht: Sind genauer $\mathcal{A}$ und $\mathcal{B}$ angeordnete Basen von V und W, so hat die konjugierte Abbildung die konjugierte Matrix, in Formeln

$${}_{\mathcal{B}}[\bar{f}]_{\mathcal{A}} = \overline{{}_{\mathcal{B}}[f]_{\mathcal{A}}}$$

Ich kann leider für das Konzept der adjungierten Abbildung keinerlei Anschauung anbieten. Eine koordinatenfreie Konstruktion der adjungierten Abbildung erhält man wie folgt: Jedes Skalarprodukt $g = \langle\ ,\ \rangle$ auf V liefert eine injektive lineare Abbildung

$$\begin{array}{rccc} \mathrm{can} = \mathrm{can}_g : & \overline{V} & \hookrightarrow & V^\top \\ & v & \mapsto & \langle v,\ \rangle \end{array}$$

in den Dualraum von V. Lineare Abbildungen A, B zwischen komplexen euklidischen Vektorräumen sind nun adjungiert genau dann, wenn das Diagramm

$$\begin{array}{ccc} \overline{V} & \xrightarrow{\text{can}} & V^\top \\ {\scriptstyle \bar{A}}\downarrow & & \downarrow{\scriptstyle B^\top} \\ \overline{W} & \xrightarrow{\text{can}} & W^\top \end{array}$$

kommutiert, mit $B^\top : V^\top \to W^\top$ der "transponierten" alias "dualen" Abbildung zu $B : W \to V$. Im endlichdimensionalen Fall sind unsere kanonischen Abbildungen can in den Horizontalen nach Dimensionsvergleich Isomorphismen, in diesem Fall liefert also das obige kommutative Diagramm auch einen alternativen Beweis für die Existenz und Eindeutigkeit adjungierter Abbildungen.

Definition 3.6.8. Ein Endomorphismus eines reellen oder komplexen euklidischen Vektorraums heißt **selbstadjungiert** genau dann, wenn er zu sich selbst adjungiert ist.

Beispiel 3.6.9. Eine reelle $(n \times n)$-Matrix A beschreibt eine selbstadjungierte Abbildung $A : \mathbb{R}^n \to \mathbb{R}^n$ genau dann, wenn sie symmetrisch ist, in Formeln $A = A^\top$. Eine komplexe $(n \times n)$-Matrix A beschreibt eine selbstadjungierte Abbildung $A : \mathbb{C}^n \to \mathbb{C}^n$ genau dann, wenn sie die Identität $A = \bar{A}^\top$ erfüllt. Solche Matrizen heißen auch **hermitesch**.

Satz 3.6.10 (Spektralsatz für selbstadjungierte Abbildungen). *Für jeden selbstadjungierten Endomorphismus eines endlichdimensionalen euklidischen Vektorraums besitzt unser Vektorraum eine Orthonormalbasis aus Eigenvektoren, und auch im komplexen Fall sind alle Eigenwerte unseres Endomorphismus reell.*

Erster Beweis. Sei V unser euklidischer Vektorraum und $A : V \to V$ selbstadjungiert. Gegeben $0 \neq v \in V$ und $\lambda \in \mathbb{C}$ mit $Av = \lambda v$ folgern wir von der Mitte ausgehend die Gleichungskette

$$\bar{\lambda}\langle v, v\rangle = \langle v, v\rangle = \langle Av, v\rangle = \langle v, Av\rangle = \langle v, \lambda v\rangle = \lambda\langle v, v\rangle$$

und daraus folgt bereits $\lambda \in \mathbb{R}$. Weiter ist das orthogonale Komplement $v^\perp$ eines Eigenvektors v stabil unter A, denn aus $\langle v, w\rangle = 0$ folgt $\langle v, Aw\rangle = \langle Av, w\rangle = \bar{\lambda}\langle v, w\rangle = 0$. Bis hierher brauchen wir nicht einmal V als endlichdimensional vorraussetzen. Nun können wir den Beweis im Komplexen mit Induktion beenden. Im Fall $V = 0$ ist der Satz klar. Sonst finden wir einen Eigenvektor v_1, den wir ohne Beschränkung der Allgemeinheit normiert annehmen dürfen. Dann wenden wir auf die auf seinem orthogonalen

Komplement induzierte Abbildung $A : v_1^\perp \to v_1^\perp$ die Induktionsvoraussetzung an und finden darin eine Orthonormalbasis $v_2, \ldots, v_n$ aus Eigenvektoren von A. Damit ist $v_1, \ldots, v_n$ die gesuchte Orthonormalbasis von V aus Eigenvektoren von A. Das erledigt den komplexen Fall. Im reellen Fall überlegen wir uns zunächst, daß die darstellende Matrix von A in Bezug auf eine Orthonormalbasis von V symmetrisch sein muß. Diese Matrix hat im Fall $\dim_{\mathbb{R}} V > 0$ in $\mathbb{C}$ mindestens einen Eigenwert, der dann aber nach unseren Überlegungen zu Beginn des Beweises sogar reell sein muß. Dazu finden wir dann wieder einen Eigenvektor aus V, und der Beweis läuft von hier an wie im komplexen Fall. □

Zweiter Beweis im Reellen. Man betrachte auf $V\backslash 0$ die Funktion

$$v \mapsto R(v) = \frac{\langle Av, v\rangle}{\langle v, v\rangle}$$

Sie heißt der **Raleigh-Quotient**, deshalb der Buchstabe R. Schränken wir diese Funktion ein auf die Einheitssphäre $\{v \mid \|v\| = 1\}$, so nimmt sie dort nach Heine-Borel **??** und **??** ihr Maximum an, etwa an einer Stelle v_+. Da unsere Funktion konstant ist auf jeder Geraden durch den Nullpunkt, muß sie an derselben Stelle auch als Funktion $V\backslash 0 \to \mathbb{R}$ ihr Maximum annehmen. Wir betrachten wir nun für $w \in V$ die für hinreichend kleines $t \in \mathbb{R}$ wohldefinierte Funktion $t \mapsto R_w(t) = R(v_+ + tw)$, ausgeschrieben

$$R_w(t) = \frac{\langle A(v_+ + tw), v_+ + tw\rangle}{\langle v_+ + tw, v_+ + tw\rangle}$$

Sie ist offensichtlich differenzierbar, folglich muß ihre Ableitung bei $t = 0$ verschwinden. Dann verschwindet also auch der Zähler, wenn wir diese Ableitung $R'_w(0)$ mithilfe der Quotientenregel berechnen, und wir folgern

$$(\langle Aw, v_+\rangle + \langle Av_+, w\rangle)\langle v_+, v_+\rangle - 2\langle Av_+, v_+\rangle\langle v_+, w\rangle = 0$$

für alle $w \in V$. Mithilfe der Selbstadjungiertheit von A folgern wir insbesondere

$$w \perp v_+ \quad \Rightarrow \quad w \perp Av_+$$

Das liefert offensichtlich $Av_+ \in \mathbb{R}v_+$ und wir haben einen Eigenvektor gefunden. Der Rest des Arguments läuft von da an wie beim ersten Beweis. Dies Argument vermeidet den Fundamentalsatz der Algebra, benutzt jedoch einen Teil der Resultate der reellen Analysis, aus denen wir in **??** den Fundamentalsatz der Algebra herleiten werden. □

3.6.11. Anschaulich ist die in diesem Beweis versteckte Erkenntnis auch recht klar: Durch den Punkt der Ellipse, der am nächsten am Ursprung liegt, geht in der Tat eine Hauptachse. Dasselbe gilt natürlich für den Punkt, der dem Ursprung am fernsten liegt, als da heißt, der kleinstmögliche Wert des Raleigh-Quotienten ist auch ein Eigenwert und jede Stelle, an der er angenommen wird, ist ein Eigenvektor unseres selbstadjungierten Operators zu diesem Eigenwert.

Übung 3.6.12. Man zeige: Ein Endomorphismus eines endlichdimensionalen euklidischen Vektorraums ist genau dann selbstadjungiert, wenn es dazu eine Orthonormalbasis aus Eigenvektoren gibt und alle Eigenwerte reell sind.

4 Bilinearformen

4.1 Fundamentalmatrix

Definition 4.1.1. Gegeben ein Körper k und ein k-Vektorraum V erinnern wir daran, daß wir in 3.2 bilineare Abbildungen $b : V \times V \to k$ auch Bilinearformen auf V genannt hatten. Die Menge aller Bilinearformen auf einem k-Vektorraum V notiere ich

$$\mathrm{Bil}_k(V) = \mathrm{Bil}(V)$$

Sie bilden einen Untervektorraum im Vektorraum $\mathrm{Ens}(V \times V, k)$ aller Abbildungen von $V \times V$ nach k.

Satz 4.1.2 (Fundamentalmatrix einer Bilinearform auf k^n). *Gegeben ein Körper k und eine natürliche Zahl $n \in \mathbb{N}$ erhalten wir eine Bijektion*

$$\begin{array}{ccc} \mathrm{Bil}(k^n) & \xrightarrow{\sim} & \mathrm{M}(n \times n; k) \\ b & \mapsto & F(b) \end{array}$$

indem wir die **Fundamentalmatrix** $F(b)$ *unserer Bilinearform erklären duch die Vorschrift $F(b)_{ij} = b(\mathrm{e}_i, \mathrm{e}_j)$. Die Umkehrabbildung kann beschrieben werden durch die Abbildungsvorschrift $F \mapsto b_F$ mit $b_F(x, y) = x^\top F y$.*

Beweis. Die erste Aussage folgt unmittelbar aus Übung 1.5.24, nach der eine bilineare Abbildung festgelegt und festlegbar ist durch ihre Werte auf Paaren von Basisvektoren. Um zu prüfen, daß unsere Beschreibung der Umkehrabbildung korrekt ist, reicht es aus, für jede Matrix F zu zeigen $F(b_F) = F$ alias $b_F(\mathrm{e}_i, \mathrm{e}_j) = F_{ij} \quad \forall i, j$. Das hinwiederum folgt jedoch unmittelbar aus $b_F(\mathrm{e}_i, \mathrm{e}_j) = \mathrm{e}_i^\top F \mathrm{e}_j$. □

Satz 4.1.3 (Fundamentalmatrix einer Bilinearform, Variante). *Gegeben ein Körper k und ein k-Vektorraum V liefert jede angeordnete Basis $\mathcal{B} = (v_1, \ldots, v_n)$ von V eine Bijektion*

$$\begin{array}{ccc} \mathrm{Bil}(V) & \xrightarrow{\sim} & \mathrm{M}(n \times n; k) \\ b & \mapsto & F_{\mathcal{B}}(b) \end{array}$$

indem wir die **Fundamentalmatrix** $F = F_{\mathcal{B}}(b)$ *unserer Bilinearform b bezüglich unserer Basis $\mathcal{B}$ erklären durch die Vorschrift $F_{ij} = b(v_i, v_j)$. Die Umkehrabbildung kann in diesem Fall beschrieben werden durch die Abbildungsvorschrift $F \mapsto b_F$ mit*

$$b_F(v, w) = {}_{\mathcal{B}}[v]^\top \circ F \circ {}_{\mathcal{B}}[w]$$

Beweis. Die erste Aussage folgt wieder unmittelbar aus der Erkenntnis 1.5.24, daß eine bilineare Abbildung festgelegt und festlegbar ist durch ihre Werte auf Paaren von Basisvektoren. Für die zweite Aussage zeigen wir nun zur Abwechslung einmal $b_{F(b)} = b$ alias $b_{F(b)}(v,w) = b(v,w)$ für alle v, w. Dazu müssen wir ja nur zeigen $b_{F(b)}(v_i, v_j) = b(v_i, v_j)$ für alle i, j alias

$$ {}_{\mathcal{B}}[v_i]^\top \circ F_{\mathcal{B}}(b) \circ {}_{\mathcal{B}}[v_j] = (F_{\mathcal{B}}(b))_{ij} $$

Das ist jedoch klar wegen ${}_{\mathcal{B}}[v_i] = \mathrm{e}_i$. □

4.1.4. Eine Bilinearform ist symmetrisch genau dann, wenn ihre Fundamentalmatrix bezüglich einer gegebenen Basis symmetrisch ist. Ist also in Formeln V ein k-Vektorraum und $\mathcal{B}$ eine angeordnete Basis von V und $b : V \times V \to k$ eine Bilinearform, so gilt

$$ b \text{ symmetrisch} \quad \Leftrightarrow \quad F_{\mathcal{B}}(b) \text{ symmetrisch} $$

In der Tat, ist $(v_1, \ldots, v_n)$ unsere angeordnete Basis, so gilt für symmetrisches b ja $b(v_i, v_j) = b(v_j, v_i)$ und damit die Identität $F_{ij} = F_{ji}$ für die Einträge $F_{ij} = b(v_i, v_j)$ der Fundamentalmatrix $F = F_{\mathcal{B}}(b)$. Bezeichnet ganz allgemein $\tau : V \times V \xrightarrow{\sim} V \times V$ das Vertauschen $\tau : (v,w) \mapsto (w,v)$, so haben wir für jede Bilinearform b offensichtlich die Identität $F_{\mathcal{B}}(b \circ \tau) = F_{\mathcal{B}}(b)^\top$. Ist also die Fundamentalmatrix symmetrisch, in Formeln $F_{\mathcal{B}}(b)^\top = F_{\mathcal{B}}(b)$, so folgt mit 4.1.3 sofort $b \circ \tau = b$ alias b symmetrisch.

Proposition 4.1.5 (Fundamentalmatrix und Basiswechsel). *Gegeben ein Körper k und ein endlichdimensionaler k-Vektorraum V mit zwei angeordneten Basen $\mathcal{A}, \mathcal{B}$ gilt zwischen den Fundamentalmatrizen einer Bilinearform b in Bezug auf unsere beiden Basen die Beziehung*

$$ {}_{\mathcal{A}}[\mathrm{id}]_{\mathcal{B}}^\top \circ F_{\mathcal{A}}(b) \circ {}_{\mathcal{A}}[\mathrm{id}]_{\mathcal{B}} = F_{\mathcal{B}}(b) $$

4.1.6. Man berechnet also die Fundamentalmatrix einer Bilinearform bezüglich einer Basis aus ihrer Fundamentalmatrix bezüglich einer anderen Basis, indem man von rechts die Basiswechselmatrix dranmultipliziert und von links ihre Transponierte.

Erster Beweis. Gegeben $v, w \in V$ gilt eben

$$ \begin{aligned} b(v,w) &= {}_{\mathcal{B}}[v]^\top \circ F_{\mathcal{B}}(b) \circ {}_{\mathcal{B}}[w] \\ b(v,w) &= {}_{\mathcal{A}}[v]^\top \circ F_{\mathcal{A}}(b) \circ {}_{\mathcal{A}}[w] \\ &= ({}_{\mathcal{A}}[\mathrm{id}]_{\mathcal{B}} \circ {}_{\mathcal{B}}[v])^\top \circ F_{\mathcal{A}}(b) \circ {}_{\mathcal{A}}[\mathrm{id}]_{\mathcal{B}} \circ {}_{\mathcal{B}}[w] \\ &= {}_{\mathcal{B}}[v]^\top \circ {}_{\mathcal{A}}[\mathrm{id}]_{\mathcal{B}}^\top \circ F_{\mathcal{A}}(b) \circ {}_{\mathcal{A}}[\mathrm{id}]_{\mathcal{B}} \circ {}_{\mathcal{B}}[w] \end{aligned} $$

Gilt für Matrizen $F, G \in \mathrm{M}(n \times m; k)$ jedoch $x^\top F y = x^\top G y$ für alle Spaltenvektoren $x \in k^n$, $y \in k^m$, so folgt $F = G$. Damit liefern unsere Gleichungen die gewünschte Identität ${}_{\mathcal{A}}[\mathrm{id}]_{\mathcal{B}}^\top \circ F_{\mathcal{A}}(b) \circ {}_{\mathcal{A}}[\mathrm{id}]_{\mathcal{B}} = F_{\mathcal{B}}(b)$. □

4.2 Definitheitseigenschaften

Definition 4.2.1. Sei k ein angeordneter Körper. Eine symmetrische quadratische Matrix $A \in M(n \times n; k)$ heißt

1. **positiv definit** genau dann, wenn gilt $x^\top A x \leq 0 \Rightarrow x = 0$;
2. **positiv semidefinit** genau dann, wenn gilt $x^\top A x \geq 0 \quad \forall x \in k^n$;
3. **negativ definit** genau dann, wenn gilt $x^\top A x \geq 0 \Rightarrow x = 0$;
4. **negativ semidefinit** genau dann, wenn gilt $x^\top A x \leq 0 \quad \forall x \in k^n$;
5. **indefinit** genau dann, wenn es $x, y \in k^n$ gibt mit $x^\top A x > 0$ und $y^\top A y < 0$;

Proposition 4.2.2 (Definitheits-Kriterien mit Eigenwerten).

1. *Eine reelle symmetrische Matrix ist positiv definit genau dann, wenn alle ihre Eigenwerte positiv sind.*
2. *Eine reelle symmetrische Matrix ist positiv semidefinit genau dann, wenn alle ihre Eigenwerte nichtnegativ sind.*
3. *Eine reelle symmetrische Matrix ist negativ definit genau dann, wenn alle ihre Eigenwerte negativ sind.*
4. *Eine reelle symmetrische Matrix ist negativ semidefinit genau dann, wenn alle ihre Eigenwerte nichtpositiv sind.*
5. *Eine reelle symmetrische Matrix ist indefinit genau dann, wenn sie positive und negative Eigenwerte hat.*

Beweis. Sei $A \in \mathrm{M}(n \times n; \mathbb{R})$ unsere reelle symmetrische Matrix. Gegeben $B \in \mathrm{GL}(n; \mathbb{R})$ hat natürlich $B^\top A B$ dieselbe Definitheit wie A. Nach dem Satz über Hauptachsentransformationen gibt es $D \in \mathrm{SO}(n)$ mit $D^\top A D = D^{-1} A D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Diese Diagonalmatrix hat folglich sowohl dieselbe Definitheit als auch dieselben Eigenwerte wie A. Die Proposition folgt unmittelbar. □

Satz 4.2.3 (**Hurwitz-Kriterium**). *Eine reelle symmetrische $(n\times n)$-Matrix ist positiv definit genau dann, wenn für alle $k < n$ die quadratische Untermatrix, die man durch Wegstreichen der letzten k Spalten und der untersten k Zeilen erhält, eine positive Determinante hat.*

Beweis. Ist eine reelle symmetrische quadratische Matrix nicht positiv definit und hat dennoch eine positive Determinante, so muß sie zwei negative Eigenwerte oder einen negativen Eigenwert einer Vielfachheit größer Eins haben. Damit finden wir leicht einen zweidimensionalen Teilraum, auf dem die durch unsere Matrix definierte symmetrische Bilinearform negativ definit ist. Damit kann sie aber auf keiner Hyperebene positiv definit sein, da eben eine Hyperebene mit einem zweidimensionalen Teilraum mehr als nur den Nullvektor gemeinsam haben muß. Eine offensichtliche Induktion beendet den Beweis. □

Übung 4.2.4. Gegeben zwei verschiedene Punkte der Ebene $p, q \in \mathbb{R}^2$ und eine positive Zahl $b < \|p - q\|$ zeige man, daß die Punkte $r \in \mathbb{R}^2$ mit $\|r - p\| - \|r - q\| = b$ einen Hyperbelast bilden, als da heißt, daß die Menge dieser Punkte unter einer geeigneten Bewegung in die Menge der Lösungen mit positiver x-Koordinate eines Gleichungssystems der Gestalt $x^2 - \mu y^2 = c$ mit μ, c positiv übergeht. Gegeben zwei verschiedene Punkte der Ebene $p, q \in \mathbb{R}^2$ und eine positive Zahl $a > \|p - q\|$ zeige man weiter, daß die Punkte $r \in \mathbb{R}^2$ mit $\|r - p\| + \|r - q\| = a$ eine Ellipse bilden, als da heißt, daß die Menge dieser Punkte unter einer geeigneten Bewegung in die Menge der Lösungen eines Gleichungssystems der Gestalt $x^2 + \mu y^2 = c$ mit μ, c positiv übergeht. So erstellen übrigends Gärtner elliptische Beete: Sie rammen zwei Pflöcke ein, legen um diese eine Seilschlaufe und fahren mit einem dritten Pflock in der Schlaufe um diese beiden Pflöcke herum, soweit außen wie möglich.

Satz 4.2.5 (**Satz über Hauptachsentransformationen, Variante**). *Sei V ein reeller Vektorraum mit zwei symmetrischen Bilinearformen s und b, von denen die erste ein Skalarprodukt ist. So besitzt V eine angeordnete Basis $(v_1, \ldots, v_n)$ mit $s(v_i, v_j) = \delta_{ij}$ und $b(v_i, v_j) = 0$ für $i \neq j$.*

4.2.6. Gegeben eine Bilinearform b auf einem Vektorraum V verstehen wir unter einer **Orthogonalbasis für** b eine Basis $(v_i)_{i\in I}$ mit $b(v_i, v_j) = 0$ falls $i \neq j$. In dieser Terminologie sagt der vorhergehende Satz also, daß es eine Orthonormalbasis für s gibt, die gleichzeitig eine Orthogonalbasis für b ist.

Beweis. Wir wählen eine Orthonormalbasis $\mathcal{B}$ in Bezug auf s. Die Fundamentalmatrix $A = F_{\mathcal{B}}(b)$ ist symmetrisch, nach dem Spektralsatz finden wir folglich $D \in \mathrm{SO}(n)$ mit $D^\top AD = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Jetzt können wir aber eine Basis $\mathcal{A}$ in V finden derart, daß $D = {}_{\mathcal{B}}[\mathrm{id}]_{\mathcal{A}}$ die Basiswechselmatrix von

$\mathcal{A}$ nach $\mathcal{B}$ ist, und mit $\mathcal{B}$ ist nach 3.3.13 dann auch $\mathcal{A}$ eine Orthonormalbasis von V in Bezug auf s. Es folgt

$$F_{\mathcal{A}}(b) = D^\top F_{\mathcal{B}}(b) D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$$

und wir haben in $\mathcal{A}$ unsere Basis gefunden, die für s orthonormal und für b orthogonal ist. □

Satz 4.2.7 (**Satz über Hauptachsentransformationen, Variante**). *Sei V ein komplexer Vektorraum mit zwei hermiteschen Bilinearformen s und b, von denen die erste ein Skalarprodukt ist. So besitzt V eine angeordnete Basis $(v_1, \ldots, v_n)$ mit $s(v_i, v_j) = \delta_{ij}$ und $b(v_i, v_j) = 0$ für $i \neq j$.*

Beweis. Analog wie im Reellen. □

Satz 4.2.8 (**Polar-Zerlegung in** $\mathrm{GL}(n;\mathbb{R})$). *Jede Matrix $A \in \mathrm{GL}(n;\mathbb{R})$ hat eine eindeutige Darstellung als Produkt $A = DP$ mit $D \in \mathrm{O}(n)$ orthogonal und P symmetrisch positiv definit.*

Beispiel 4.2.9. Im Fall $\mathrm{GL}(1;\mathbb{R})$ ist das die vielleicht noch nicht sehr aufregende Zerlegung $a = (a/|a|) \cdot |a|$ einer von Null verschiedenen reellen Zahl als das Produkt von einem Vorzeichen mit einer positiven reellen Zahl.

Beweis. Wir beginnen mit dem Nachweis der Eindeutigkeit. Gegeben eine Zerlegung $A = DP$ wie oben haben wir sicher $A^\top A = P^\top D^\top DP = P^\top P = P^2$ und folglich muß P die Matrix sein, die auf den Eigenräumen von $A^\top A$ zum Eigenwert λ jeweils operiert durch den Eigenwert $\sqrt{\lambda}$. Das zeigt die Eindeutigkeit unserer Zerlegung. Andererseits folgt aus $A^\top Av = \lambda v$ sofort $v^\top A^\top Av = \lambda \|v\|^2 = \|Av\|^2$ und somit $\lambda > 0$, so daß wir P symmetrisch und positiv definit finden können mit $P^2 = A^\top A$. Für $D = AP^{-1}$ folgt dann $D^\top D = P^{-1}A^\top AP = I$ und folglich ist D orthogonal. □

Bemerkung 4.2.10. Jede Matrix $A \in \mathrm{M}(n \times n;\mathbb{R})$ hat eine Darstellung als Produkt $A = DP$ mit $D \in \mathrm{O}(n)$ orthogonal und P symmetrisch positiv semidefinit. Allerdings ist D dann nicht mehr eindeutig bestimmt. *Wie zeigt man das?*

Definition 4.2.11. Eine hermitesche Matrix $A \in \mathrm{M}(n \times n;\mathbb{C})$ heißt **positiv definit** genau dann, wenn gilt $\bar{x}^\top Ax \leq 0 \Rightarrow x = 0$.

Satz 4.2.12 (**Polar-Zerlegung in** $\mathrm{GL}(n;\mathbb{C})$). *Jede Matrix $A \in \mathrm{GL}(n;\mathbb{C})$ hat eine eindeutige Darstellung als Produkt $A = DP$ mit $D \in \mathrm{U}(n)$ unitär und P hermitesch positiv definit.*

Beispiel 4.2.13. Im Fall $\mathrm{GL}(1;\mathbb{R})$ ist das die Zerlegung $a = (a/|a|) \cdot |a|$ einer von Null verschiedenen komplexen Zahl in eine Zahl der Länge Eins und eine positive reelle Zahl.

Beweis. Analog wie im Reellen. □

Beispiel 4.2.14. Im Fall $\mathrm{GL}(1;\mathbb{C})$ ist das die Zerlegung $a = (a/|a|) \cdot |a|$.

4.3 Klassifikation symmetrischer Bilinearformen

4.3.1 (**Physikalische Motivation**). In der speziellen Relativitätstheorie modelliert man die Welt, in der wir leben, als als einen vierdimensionalen reellen affinen Raum X aller "Raum-Zeit-Punkte" alias "Ereignisse". Wählen wir ein räumliches Koordinatensystem und einen Beginn der Zeitrechnung und eine Zeiteinheit, so können wir X mit dem $\mathbb{R}^4$ identifizieren und jedes Ereignis wird spezifiziert durch eine Zeitkoordinate und drei Raumkoordinaten, also ein Viertupel von reellen Zahlen (t, x, y, z). Das Licht breitet sich mit Lichtgeschwindigkeit aus, genau dann wird also eine Explosion am Raumzeitpunkt $p = (t, x, y, z)$ gesehen bei $p' = (t', x', y', z')$, wenn gilt

$$t' \geq t \text{ und } c^2(t'-t)^2 - (x'-x)^2 - (y'-y)^2 - (z'-z)^2 = 0$$

für c die Lichtgeschwindigkeit. Betrachten wir auf $\mathbb{R}^4$ die sogenannte **Lorentz-Metrik** alias die symmetrische Bilinearform l mit der Fundamentalmatrix

$$\mathrm{diag}(c^2, -1, -1, -1)$$

so kann die zweite unserer Bedingungen auch umgeschrieben werden zur Bedingung $l(\vec{v}, \vec{v}) = 0$ für $\vec{v} = p' - p$. Wenn Sie bereits die Definition einer Metrik kennen, seien Sie gewarnt, daß diese Lorentz-Metrik im Sinne der in der Mathematik üblichen Terminologie keine Metrik ist. Nun vergessen wir wieder unsere Koordinaten und modellieren die Welt, in der wir leben, als einen vierdimensionalen reellen affinen Raum X mitsamt einer symmetrischen Bilinearform

$$l : \vec{X} \times \vec{X} \to \mathbb{R}$$

auf seinem Richtungsraum. Wir fordern, daß deren Fundamentalmatrix bezüglich mindestens einer Basis die oben angegebene Gestalt hat und daß sie die Ausbreitung des Lichts in der Weise beschreibt, daß $l(\vec{v}, \vec{v}) = 0$ gleichbedeutend ist dazu, daß eine Explosion am Raumzeitpunkt $p \in X$ entweder bei $p + \vec{v}$ oder bei $p - \vec{v}$ gesehen werden kann. Wir werden später zeigen, daß jede weitere symmetrische Bilinearform l' mit der Eigenschaft

$l'(\vec{v},\vec{v}) = 0 \Leftrightarrow l(\vec{v},\vec{v}) = 0$ bereits ein Vielfaches von l sein muß. Die Wahl eines möglichen l bedeutet die Wahl einer Längeneinheit oder gleichbedeutend einer Zeiteinheit in der speziellen Relativitätstheorie. Das ist jedoch nicht, was an dieser Stelle diskutiert werden soll. Wir stellen uns die viel einfachere Frage, ob unsere Bilinearform nicht etwa bezüglich einer anderen Basis auch $\operatorname{diag}(1,1,-1,-1)$ als Fundamentalmatrix haben könnte. Das geht nun zwar nicht, aber wir wollen eben unter anderem verstehen, warum es nicht geht, und entwickeln dazu die Anfänge der allgemeinen Theorie der symmetrischen Bilinearformen.

4.3.2 (**Mathematische Motivation**). Gegeben ein Körper k interessiert man sich für die **Klassifikation der symmetrischen Bilinearformen über** k. Damit ist gemeint, daß wir eine Familie $(V_i, b_i)_{i\in I}$ von endlichdimensionalen k-Vektorräumen mit symmetrischer Bilinearform suchen mit der Eigenschaft, daß für jedes Paar (V, b) bestehend aus einem endlichdimensionalen k-Vektorraum V mit einer symmetrischen Bilinearform b genau ein $i \in I$ existiert derart, daß es einen Isomorphismus $V \xrightarrow{\sim} V_i$ gibt, unter dem unser b dem vorgegebenen b_i entspricht. Eine derartige Klassifikation ist eng mit der Struktur des Körpers verknüpft und im allgemeinen sehr schwierig zu erreichen. Wir geben zumindest im Fall eines algebraisch abgeschlossenen Körpers einer Charakteristik ungleich Zwei in 4.3.4 sowie im Fall $k = \mathbb{R}$ in 4.3.13 Klassifikationen an.

Satz 4.3.3 (**Existenz einer Orthogonalbasis**). *Sei V ein endlichdimensionaler Vektorraum über einem Körper k mit* $\operatorname{char} k \neq 2$. *So gibt es für jede symmetrische Bilinearform b auf V eine* **Orthogonalbasis** *alias eine Basis $\mathcal{B} = (v_1, \ldots, v_n)$ von V mit*

$$i \neq j \quad \Rightarrow \quad b(v_i, v_j) = 0$$

Beweis. Gilt für jeden Vektor $v \in V$ bereits $b(v,v) = 0$, so folgt $2b(v,w) = b(b+w, v+w) - b(v,v) - b(w,w) = 0$ für alle v, w und wegen $2 \neq 0$ in k folgt $b = 0$ und die Aussage des Satzes ist evident. Sonst gibt es einen Vektor $v_1 \in V$ mit $b(v_1, v_1) \neq 0$. Dann ist

$$\begin{array}{rccc} \varphi: & V & \to & k \\ & w & \mapsto & b(v_1, w) \end{array}$$

eine lineare Abbildung mit $v_1 \notin \ker\varphi$ und aus Dimensionsgründen gilt $kv_1 \oplus \ker\varphi = V$. Mit Induktion über die Dimension dürfen wir annehmen, daß $\ker\varphi$ eine Orthogonalbasis $(v_2, \ldots, v_n)$ besitzt. Dann ist aber $(v_1, v_2, \ldots, v_n)$ eine Orthogonalbasis von V. □

4.3.4. Ist $\operatorname{char} k \neq 2$ und k algebraisch abgeschlossen oder allgemeiner das Quadrieren eine Surjektion $k \twoheadrightarrow k$, $x \mapsto x^2$, so können wir die im Satz gefundene Basis sogar so abändern, daß die Fundamentalmatrix die Gestalt $\operatorname{diag}(1, \ldots, 1, 0, \ldots, 0)$ hat. Die Zahl der Einsen ist hierbei wohldefiniert, denn der Rang der der Fundamentalmatrix einer Bilinearform hängt von der gewählten Basis nach 4.1.5 überhaupt nicht ab.

Definition 4.3.5. Ist V ein endlichdimensionaler Vektorraum und b eine Bilinearform auf V, so erklären wir den **Rang** von b als den Rang einer Fundamentalmatrix

$$\operatorname{rk}(b) = \operatorname{rk} F_{\mathcal{B}}(b)$$

in Bezug auf eine und jede angeordnete Basis $\mathcal{B}$ von V. Nach 4.1.5 hängt diese Zahl in der Tat nicht von der Basis $\mathcal{B}$ ab.

Definition 4.3.6. Eine Bilinearform auf einem Vektorraum V oder allgemeiner eine **bilineare Paarung** alias eine bilineare Abbildung

$$b : V \times W \to k$$

vom Produkt zweier Vektorräume in den Grundkörper heißt **nichtausgeartet** genau dann, wenn es für jedes $v \in V \backslash 0$ ein $w \in W$ gibt mit $b(v, w) \neq 0$ und umgekehrt auch für jedes $w \in W \backslash 0$ ein $v \in V$ mit $b(v, w) \neq 0$. Andernfalls heißt unsere Bilinearform oder allgemeiner unsere Paarung **ausgeartet**.

Definition 4.3.7. Gegeben ein Körper k und ein k-Vektorraum V und eine Bilinearform $b : V \times V \to k$ erklären wir den **Ausartungsraum** alias das **Radikal** von V als den Untervektorraum

$$\operatorname{rad} b = \{v \in V \mid b(w, v) = 0 \quad \forall w \in V\}$$

Wir werden dieses Konzept im Wesentlichen nur für symmetrische oder alternierende Bilinearformen verwenden und verzichten deshalb darauf, unseren Ausartungsraum feiner "Rechtsausartungsraum" zu nennen und zusätzlich noch einen "Linksausartungsraum" einzuführen.

Satz 4.3.8 (Rang und Radikal). *Der Rang und das Radikal einer Bilinearform b auf einem endlichdimensionalen Vektorraum V sind verknüpft durch die Beziehung*

$$\operatorname{rk}(b) + \dim(\operatorname{rad}(b)) = \dim V$$

4.3.9. Eine Bilinearform auf einem endlichdimensionalen Vektorraum ist also insbesondere genau dann nichtausgeartet, wenn sie maximalen Rang hat.

Beweis. Ist $\mathcal{B}$ eine angeordnete Basis von V, so können wir $F_{\mathcal{B}}(b)$ auch verstehen als die Transponierte der Matrix der Abbildung $\hat{b} : V \to V^\top$, die jedem $w \in V$ die Linearform "paare mit w unter b" alias $(\hat{b}(w))(v) = b(w,v)$ zuordnet, genauer und in Formeln haben wir

$$ {}_{\mathcal{B}^\top}[\hat{b}]_{\mathcal{B}} = F_{\mathcal{B}}(b)^\top$$

In der Tat, setzen wir $\mathcal{B} = (v_1, \ldots, v_n)$ und machen den Ansatz $\hat{b}(v_i) = a_{1i}v_1^\top + \ldots + a_{ni}v_n^\top$, so liefert das Auswerten der Linearformen auf beiden Seiten dieser Gleichung auf dem Basisvektor v_j die Identität $b(v_i, v_j) = (\hat{b}(v_i))(v_j) = a_{ji}$ und damit die Gleichheit aller Einträge unserer beiden Matrizen. Insbesondere gilt $\operatorname{rk}(b) = \operatorname{rk}(\hat{b}) = \dim(\operatorname{im}\hat{b})$. Wegen $\operatorname{rad}(b) = \ker(\hat{b})$ folgt unsere Identität damit aus der Dimensionsformel 1.6.10. □

4.3.10. Die Identität ${}_{\mathcal{B}^\top}[\hat{b}]_{\mathcal{B}} = F_{\mathcal{B}}(b)^\top$ aus dem vorhergehenden Beweis liefert auch einen zweiten Zugang zu unserer Formel 4.1.5 über das Verhalten der Fundamentalmatrix unter Basiswechsel: Wir rechnen einfach

$$F_{\mathcal{B}}(b)^\top = {}_{\mathcal{B}^\top}[\hat{b}]_{\mathcal{B}} = {}_{\mathcal{B}^\top}[\mathrm{id}]_{\mathcal{A}^\top} \circ {}_{\mathcal{A}^\top}[\hat{b}]_{\mathcal{A}} \circ {}_{\mathcal{A}}[\mathrm{id}]_{\mathcal{B}} = ({}_{\mathcal{A}}[\mathrm{id}]_{\mathcal{B}})^\top \circ F_{\mathcal{A}}(b)^\top \circ {}_{\mathcal{A}}[\mathrm{id}]_{\mathcal{B}}$$

unter Verwendung unserer Formel 1.8.13 für die Matrix der transponierten Abbildung. Transponieren liefert dann ein weiteres Mal unsere Formel 4.1.5.

4.3.11. Übung 1.5.26 liefert uns für jeden Vektorraum V einen kanonischen Isomorphismus

$$\begin{array}{ccc} \mathrm{Bil}(V) & \xrightarrow{\sim} & \mathrm{Hom}(V, V^\top) \\ b & \mapsto & \hat{b} \end{array}$$

zwischen dem Raum der Bilinearformen auf V und dem Raum der linearen Abbildungen von V in seinen Dualraum $V^\top$, gegeben durch die Abbildungsvorschrift $b \mapsto \hat{b}$ mit $\hat{b} : v \mapsto b(v, \ )$ alias $(\hat{b}(v))(w) = b(v,w)$. In **??** verwende ich die alternative Notation $\hat{b} = \mathrm{can}_b^1$ und betrachte zusätzlich auch noch $\mathrm{can}_b^2 : V \to V^\top$ gegeben durch $v \mapsto b(\ , v)$.

Übung 4.3.12. Gegeben zwei endlichdimensionale Vektorräume V, W mit einer nichtausgearteten Paarung $b : V \times W \to k$ und ein Untervektorraum $U \subset W$ sei $U^\perp = \{v \in V \mid b(v,u) = 0 \quad \forall u \in U\}$. Man zeige die Formel $\dim U + \dim U^\perp = \dim V$.

Satz 4.3.13 (Sylvester'scher Trägheitssatz). *Gegeben ein endlichdimensionaler reeller Vektorraum V mit einer symmetrischen Bilinearform gibt es stets eine Basis $\mathcal{B} = (v_1, \ldots, v_n)$, in der die Fundamentalmatrix die Gestalt*

$$\mathrm{diag}(1, \ldots, 1, -1, \ldots, -1, 0, \ldots, 0)$$

hat. Die Zahl der Einsen, Minus-Einsen und Nullen wird hierbei durch besagte Bilinearform bereits eindeutig festgelegt.

4.3.14. Die Differenz zwischen der Zahl der Minus-Einsen und der Zahl der Einsen heißt in diesem Kontext die **Signatur** unserer symmetrischen Bilinearform.

Beweis. Die Existenz einer Basis mit den geforderten Eigenschaften folgt unmittelbar aus der Existenz einer Orthogonalbasis 4.3.3, indem wir deren Vektoren mit geeigneten Skalaren multiplizieren. Die Zahl der Nullen ist wohlbestimmt als die Dimension des Radikals V_0. Ich behaupte, daß die Zahl der Einsen bzw. Minus-Einsen beschrieben werden kann als die jeweils maximal mögliche Dimension für einen Teilraum, auf dem unsere Bilinearform positiv definit bzw. negativ definit ist. Sind in der Tat $V_+ \subset V$ und $V_- \subset V$ Teilräume der maximal möglichen Dimension mit dieser Eigenschaft, ja sogar irgendwelche Teilräume mit dieser Eigenschaft, so folgt $V_- \cap V_0 = 0$ und $V_+ \cap (V_- \oplus V_0) = 0$ und damit

$$\dim V_+ + \dim V_- + \dim V_0 \leq \dim V$$

Für jede Orthogonalbasis B bezeichne nun B_+, B_- und B_0 die Basisvektoren, deren Paarung mit sich selber eine positive Zahl bzw. eine negative Zahl bzw. Null ergibt. So haben wir natürlich

$$|B_+| + |B_-| + |B_0| = \dim V$$

Da aber wegen der Maximalität von $V_\pm$ offensichtlich gilt $|B_\pm| \leq \dim V_\pm$, und da $|B_0| \leq \dim V_0$ eh klar ist, muß überall Gleichheit gelten. □

Definition 4.3.15. Gegeben ein Körper k und ein k-Vektorraum V versteht man unter einer **quadratischen Form auf** V eine Abbildung

$$q : V \to k$$

die sich darstellen läßt in der Gestalt $q(v) = f_1(v)g_1(v) + \ldots + f_r(v)g_r(v)$ mit $f_i, g_i \in V^\top$ Linearformen auf V.

4.3.16. Man erhält für jeden Körper k eine Bijektion

$$\left\{ \begin{array}{c} \text{obere } (n \times n)\text{-Dreiecksmatrizen} \\ \text{mit Einträgen aus } k \end{array} \right\} \xrightarrow{\sim} \{\text{Quadratische Formen auf } k^n\}$$

$$(b_{ij}) \mapsto \sum_{i \geq j} b_{ij} x_i x_j$$

Für jeden Körper einer Charakteristik $\operatorname{char} k \neq 2$ erhält man darüberhinaus auch eine Bijektion

$$\left\{ \begin{array}{c} \text{symmetrische } (n \times n)\text{-Matrizen} \\ \text{mit Einträgen in } k \end{array} \right\} \xrightarrow{\sim} \{\text{Quadratische Formen auf } k^n\}$$

$$(a_{ij}) \mapsto \sum a_{ij} x_i x_j$$

und koordinatenfrei formuliert ergibt sich für jeden endlichdimensionalen Vektorraum V über einem Körper einer Charakteristik $\operatorname{char} k \neq 2$ eine Bijektion

$$\begin{array}{ccc} \left\{ \begin{array}{c} \text{symmetrische Bilinearformen} \\ \text{auf } V \end{array} \right\} & \xrightarrow{\sim} & \left\{ \begin{array}{c} \text{Quadratische Formen} \\ \text{auf } V \end{array} \right\} \\ b & \mapsto & (v \mapsto b(v,v)) \end{array}$$

Ich erwähne das hier im Wesentlichen deshalb, weil in der Literatur das bei uns als Frage nach der "Klassifikation der symmetrischen Bilinearformen" formulierte Ziel meist als die Frage nach der "Klassifikation der quadratischen Formen" formuliert wird. Wenn man vom Fall der Charakteristik Zwei einmal absieht, sind diese beiden Fragen also äquivalent.

Übung 4.3.17. Sei k ein Körper und V ein endlichdimensionaler k-Vektorraum. Eine Abbildung $q : V \to k$ ist eine quadratische Form auf V genau dann, wenn gilt $q(\alpha v) = \alpha^2 q(v) \quad \forall \alpha \in k, v \in V$ und wenn außerdem die Abbildung $V \times V \to k$, $(v, w) \mapsto q(v + w) - q(v) - q(w)$ bilinear ist.

4.4 Alternierende Bilinearformen

4.4.1. Wir erinnern daran, daß nach 2.7.9 eine Bilinearform **alternierend** heißt genau dann, wenn Null herauskommt, sobald wir an beiden Stellen denselben Vektor einsetzen. Ich kann für dieses Konzept leider keinerlei Anschauung anbieten.

Satz 4.4.2 (Klassifikation alternierender Bilinearformen). *Gegeben ein endlichdimensionaler Vektorraum V über einem Körper k und eine alternierende Bilinearform $\omega : V \times V \to k$ besitzt V stets eine Basis, bezüglich derer ω eine Fundamentalmatrix hat der Gestalt*

$$\begin{pmatrix} \boxed{\begin{matrix} 0 & -1 \\ 1 & 0 \end{matrix}} & & & & & \\ & \ddots & & & & \\ & & \boxed{\begin{matrix} 0 & -1 \\ 1 & 0 \end{matrix}} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

und die Zahl der Zweierblöcke hängt nicht von der Wahl der Basis ab.

Beweis. Ist unsere Form nicht Null, so finden wir $v, w \in V$ mit $\omega(v, w) \neq 0$. Durch Multiplikation von v mit einem Skalar erreichen wir sogar $\omega(v, w) = 1$ und damit $\omega(w, v) = -1$. Die Vektoren v, w können wir schon einmal als die ersten beiden Vektoren unserer Basis in spe festhalten. Wir betrachten nun die Linearformen $\omega(v, \ ) : V \to k$ und $\omega(w, \ ) : V \to k$. Sie sind beide nicht Null und ihre Kerne sind verschieden, genauer liegt v im Kern der ersten, nicht aber der zweiten Abbildung und w im Kern der zweiten, nicht aber der ersten. Für den Schnitt

$$S = \{u \in V \mid \omega(v, u) = 0 = \omega(w, u)\}$$

haben wir also $\dim S = \dim V - 2$ und $(kv \oplus kw) \cap S = 0$. Aus Dimensionsgründen folgt

$$V = (kv \oplus kw) \oplus S$$

und eine offensichtliche Induktion über die Dimension beendet den Beweis der Existenz. Die Zahl der Nullen nach den Zweierkästchen kann beschrieben werden als die Dimension des Radikals unserer Bilinearform und ist deshalb ebenso wie die Zahl der Zweierkästchen von der Wahl der Basis unabhängig. □

4.4.3. Eine alternierende Bilinearform auf einem Vektorraum heißt **nicht ausgeartet** genau dann, wenn ihr Radikal Null ist. Eine nichtausgeartete alternierende Bilinearform heißt auch eine **symplektische Form**, und ein mit einer symplektischen Form versehener Vektorraum heißt ein **symplektischer Vektorraum**. Symplektische Vektorräume spielen in der Hamilton'schen Mechanik eine wichtige Rolle. Nach 1.4.20 ist die Dimension eines endlichdimensionalen symplektischen Vektorraums stets gerade.

5 Jordan'sche Normalform

5.1 Motivation durch Differentialgleichungen

5.1.1. Wie etwa in **??** erklärt wird, kann man die Exponentialabbildung auf komplexen quadratischen Matrizen erklären durch die Exponentialreihe

$$\begin{array}{rccc} \exp: & \mathrm{M}(n\times n;\mathbb{C}) & \to & \mathrm{M}(n\times n;\mathbb{C}) \\ & A & \mapsto & \sum_{k=0}^{\infty} \frac{1}{k!} A^k \end{array}$$

Wie etwa in **??** erklärt wird, spielt diese Abbildung eine zentrale Rolle bei der Lösung von Systemen linearer Differentialgleichungen mit konstanten Koeffizienten. Ist genauer $A \in \mathrm{M}(n\times n;\mathbb{C})$ eine quadratische Matrix und $c \in \mathbb{C}^n$ ein Spaltenvektor, so gibt es genau eine differenzierbare Abbildung $\gamma : \mathbb{R} \to \mathbb{C}^n$ mit Anfangswert $\gamma(0) = c$ derart, daß gilt $\dot{\gamma}(t) = A\gamma(t)$ für alle $t \in \mathbb{R}$, und zwar die Abbildung

$$\gamma(t) = \exp(tA)c$$

Diese Erkenntnis soll dazu motivieren, nach einem möglichst guten Verständnis von $\exp A$ zu suchen. Die Formal $\exp(PAP^{-1}) = P(\exp A)P^{-1}$ für P invertierbar folgt ziemlich direkt aus der Definition, wie Sie in der Analysis als Übung **??** ausführen dürfen. Des weiteren erklären wir in **??**, warum für kommutierende quadratische Matrizen A, B stets gilt $\exp(A + B) = (\exp A)(\exp B)$. In 5.4.1 werden wir im folgenden unter der Überschrift "Jordan-Zerlegung" zeigen, daß sich jede komplexe quadratische Matrix A auf genau eine Weise zerlegen läßt als eine Summe $A = D + N$ mit D diagonalisierbar und N nilpotent und $DN = ND$. Ist dann noch P invertierbar mit $PDP^{-1} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, so folgt

$$\begin{array}{rcl} \exp A & = & \exp D \ \exp N \\ & = & P^{-1} \exp(\mathrm{diag}(\lambda_1, \ldots, \lambda_n))P \ \exp N \\ & = & P^{-1} \,\mathrm{diag}(\mathrm{e}^{\lambda_1}, \ldots, \mathrm{e}^{\lambda_n})P \ \exp N \\ \\ \exp tA & = & P^{-1} \,\mathrm{diag}(\mathrm{e}^{t\lambda_1}, \ldots, \mathrm{e}^{t\lambda_n})P \ \exp tN \end{array}$$

Hierbei bricht die Reihe für $\exp tN$ ab und wir erhalten so ein recht befriedigendes qualitatives Bild und mit etwas mehr Rechnen auch eine sehr explizite Beschreibung der Lösungen.

Übung 5.1.2. Die Exponentialabbildung von Matrizen liefert eine Bijektion

$$\exp : \{\text{symmetrische Matrizen}\} \xrightarrow{\sim} \{\text{positiv definite symmetrische Matrizen}\}$$

5.2 Summen und Produkte von Vektorräumen

5.2.1. Allgemeiner als in I.2.2.17 diskutiert kann man auch für eine beliebige Familie von Mengen $(X_i)_{i\in I}$ eine neue Menge, ihr **Produkt**, bilden als die Menge aller Tupel $(x_i)_{i\in I}$ mit $x_i \in X_i$ für alle $i \in I$. Dieses Produkt notiert man

$$\prod_{i\in I} X_i$$

und die Projektionsabbildungen werden mit $\mathrm{pr}_j : \left(\prod_{i\in I} X_i\right) \to X_j$ oder ähnlich bezeichnet. Wieder können wir für beliebige Abbildungen $f_i : Z \to X_i$ eine Abbildung $f = (f_i)_{i\in I} : Z \to \prod_{i\in I} X_i$ definieren durch die Vorschrift $f(z) = (f_i(z))_{i\in I}$ und jede Abbildung von einer Menge Z in ein Produkt ist von dieser Form mit $f_i = \mathrm{pr}_i \circ f$. In Formeln ausgedrückt definiert das Bilden der Kompositionen mit den Projektionen also für jede Menge Z eine Bijektion

$$\begin{array}{ccc} \mathrm{Ens}\left(Z, \prod_{i\in I} X_i\right) & \xrightarrow{\sim} & \prod_{i\in I} \mathrm{Ens}(Z, X_i) \\ f & \mapsto & (\mathrm{pr}_i \circ f) \end{array}$$

5.2.2. Dual kann man für eine beliebige Familie $(X_i)_{i\in I}$ von Mengen auch ihre **disjunkte Vereinigung**

$$\coprod_{i\in I} X_i = \bigcup_{i\in I} (X_i \times \{i\})$$

bilden. Das Anhängen der Indizes auf der rechten Seite ist hier eine Vorsichtsmaßnahme für den Fall, daß unsere Mengen nicht disjunkt gewesen sein sollten. Jede derartige disjunkte Vereinigung ist versehen mit Inklusionsabbildungen $\mathrm{in}_j : X_j \to \left(\coprod_{i\in I} X_i\right)$. Weiter können wir für beliebige Abbildungen $f_i : X_i \to Z$ in eine Menge Z die Abbildung $f : \coprod_{i\in I} X_i \to Z$ bilden durch die Vorschrift $f(x) = f_i(x)$ für $x \in X_i$, und jede Abbildung der disjunkten Vereinigung in eine Menge Z ist von dieser Form mit $f_i = f \circ \mathrm{in}_i$. In Formeln ausgedrückt definiert das Bilden der Kompositionen mit den Projektionen also für jede Menge Z eine Bijektion

$$\begin{array}{ccc} \mathrm{Ens}\left(\coprod_{i\in I} X_i, Z\right) & \xrightarrow{\sim} & \prod_{i\in I} \mathrm{Ens}(X_i, Z) \\ f & \mapsto & (f \circ \mathrm{in}_i) \end{array}$$

Definition 5.2.3. Gegeben eine Familie $(V_i)_{i\in I}$ von Vektorräumen über einem Körper k bilden wir zwei neue k-Vektorräume, ihr **Produkt** $\prod V_i$ und ihre **direkte Summe** oder kurz **Summe** $\bigoplus V_i$ durch die Regeln

$$\begin{array}{rcl} \prod_{i\in I} V_i & = & \{(v_i)_{i\in I} \mid v_i \in V_i\} \\ \bigoplus_{i\in I} V_i & = & \{(v_i)_{i\in I} \mid v_i \in V_i \text{ und nur endlich viele } v_i \text{ sind nicht null}\} \end{array}$$

mit der offensichtlichen komponentenweisen Addition und Multiplikation mit Skalaren aus k. Dieselben Konstruktionen sind auch im Fall vonm Gruppen sinnvoll, wenn wir "null" als das jeweilige neutrale Element verstehen, und wir werden beide Konstruktionen auch in diesem Kontext verwenden.

5.2.4. Für eine endliche Familie von Gruppen oder Vektorräumen $V_1, \ldots, V_s$ stimmen die direkte Summe und das Produkt überein. Wir benutzen dann alternativ die Notationen

$$V_1 \oplus \ldots \oplus V_s = V_1 \times \ldots \times V_s$$

Beispiel 5.2.5. Im Fall der konstanten Familie $(k)_{x\in X}$ erhalten wir einen Isomorphismus des freien Vektorraums über X im Sinne von 1.3.17 mit unserer direkten Summe

$$kX \xrightarrow{\sim} \bigoplus_{x\in X} k$$

vermittels der Abbildungsvorschrift $\sum_{x\in X} a_x x \mapsto (a_x)_{x\in X}$. Auch im Fall einer allgemeineren konstanten Familie $(V)_{x\in X}$ erhalten wir einen Isomorphismus

$$\mathrm{Ens}(X, V) \xrightarrow{\sim} \prod_{x\in X} V$$

vermittels der Abbildungsvorschrift $f \mapsto (f(x))_{x\in X}$.

5.2.6. Das Produkt bzw. die Summe haben im Fall von Vektorräumen oder allgemeiner von abelschen Gruppen die folgenden Eigenschaften: Die offensichtlichen Einbettungen und Projektionen sind Homomorphismen

$$\mathrm{in}_i : V_i \hookrightarrow \bigoplus_{i\in I} V_i \qquad \text{bzw.} \qquad \mathrm{pr}_i : \prod_{i\in I} V_i \twoheadrightarrow V_i$$

und ist V ein weiterer k-Modul, so induzieren die durch Vorschalten der in_i bzw. Nachschalten der pr_i gegebenen Abbildungen Bijektionen, ja sogar Isomorphismen

$$\begin{array}{rcl} \mathrm{Hom}_k\left(\bigoplus_{i\in I} V_i, V\right) & \xrightarrow{\sim} & \prod_{i\in I} \mathrm{Hom}_k(V_i, V) \\ f & \mapsto & (f \circ \mathrm{in}_i)_{i\in I} \end{array}$$

$$\begin{array}{rcl} \mathrm{Hom}_k\left(V, \prod_{i\in I} V_i\right) & \xrightarrow{\sim} & \prod_{i\in I} \mathrm{Hom}_k(V, V_i) \\ f & \mapsto & (\mathrm{pr}_i \circ f)_{i\in I} \end{array}$$

Im Fall nichtabelscher Gruppen ist nur die zweite dieser Abbildungen eine Bijektion. Ich gebe zu, daß das Symbol in_i nun in zweierlei Bedeutung verwendet wird: Einmal bei Mengen für die Einbettung in eine disjunkte

Vereinigung und ein andermal bei Vektorräumen für die Einbettung in eine direkte Summe. Was jeweils gemeint ist, muß aus dem Kontext erschlossen werden. Betrachten wir im Fall des ersten Isomorphismus speziell den Fall $V = k$, so erhalten wir einen Isomorphismus zwischen dem Dualraum einer direkten Summe und dem Produkt der Dualräume der Summanden.

5.2.7. Gegeben eine Familie $(V_i)_{i\in I}$ von Untervektorräumen eines Vektorraums V bezeichnet man den von ihrer Vereinigung erzeugten Untervektorraum auch als ihre **Summe** und notiert ihn $\sum_{i\in I} V_i$. Diese Summe kann auch interpretiert werden als das Bild des natürlichen Homomorphismus $\bigoplus_{i\in I} V_i \to V$ von der direkten Summe nach V. Ist dieser Homomorphismus injektiv, so sagen wir, die **Summe der Untervektorräume** V_i **sei direkt** und schreiben statt $\sum_{i\in I} V_i$ auch $\bigoplus_{i\in I} V_i$.

Lemma 5.2.8. *Gegeben eine Familie $(V_i)_{i\in I}$ von Untervektorräumen eines Vektorraums V ist der natürliche Homomorphismus $\bigoplus_{i\in I} V_i \hookrightarrow V$ eine Injektion genau dann, wenn für jede endliche Teilmenge $J \subset I$ und jedes $i \in I\backslash J$ gilt*

$$V_i \cap \sum_{j\in J} V_j = 0$$

Beweis. Ist der natürliche Homomorphismus eine Injektion, so folgt aus $i \in I\backslash J$ offensichtlich $V_i \cap \sum_{j\in J} V_j = 0$, und das sogar für beliebiges $J \subset I$. Ist der natürliche Homomorphismus keine Injektion, so liegt ein von Null verschiedener Vektor $v = (v_i)_{i\in I} \neq 0$ der direkten Summe in seinem Kern. Dieser hat nur in endlich vielen Summanden eine von Null verschiedene Komponente, die Menge $K = \{i \mid v_i \neq 0\}$ ist also endlich und nicht leer. Per definitionem gilt nun $\sum_{k\in K} v_k = 0$. Wählen wir $i \in K$ und nehmen $J = K\backslash i$, so folgt $0 \neq -v_i = \sum_{j\in J} v_j$ und damit $V_i \cap \sum_{j\in J} V_j \neq 0$. □

Übung 5.2.9. Ist $(V_i)_{i\in I}$ eine Familie von Vektorräumen und $B_i \subset V_i$ jeweils eine Basis, so ist die Vereinigung $\bigcup_{i\in I} \mathrm{in}_i(B_i)$ der Bilder ihrer Basen eine Basis der direkten Summe $\bigoplus_{i\in I} V_i$. Diese Basis ist auch in offensichtlicher Bijektion zur disjunkten Vereinigung von Basen $\coprod_{i\in I} B_i$.

5.3 Hauptraumzerlegung

Definition 5.3.1. Gegeben eine Abbildung $f : X \to X$ von einer Menge in sich selber nennen wir eine Teilmenge $Y \subset X$ **stabil unter** f genau dann, wenn gilt $x \in Y \Rightarrow f(x) \in Y$.

Definition 5.3.2. Gegeben ein Vektorraum V und dazu ein Endomorphismus $f : V \to V$ und ein Skalar λ aus dem Grundkörper erklären wir den **Eigenraum von f zum Eigenwert λ** durch

$$\mathrm{Eig}(f;\lambda) = \mathrm{Eig}(f|V;\lambda) = \ker(f - \lambda\,\mathrm{id})$$

und den **Hauptraum von f zum Eigenwert λ** durch

$$\mathrm{Hau}(f;\lambda) = \mathrm{Hau}(f|V;\lambda) = \bigcup_{n\geq 0} \ker(f - \lambda\,\mathrm{id})^n$$

Der Eigenraum zum Eigenwert λ besteht also genau aus allen Eigenvektoren zum Eigenwert λ und dem Nullvektor. Die von Null verschiedenen Elemente des Hauptraums zum Eigenwert λ heißen die **Hauptvektoren zum Eigenwert λ**.

5.3.3. Diese Räume sind beide Untervektorräume unseres ursprünglichen Vektorraums V. Sie sind auch offensichtlich stabil unter unserem Endomorphismus f, ja sogar unter jedem Endomorphismus $g : V \to V$, der mit f kommutiert, in Formeln impliziert $gf = fg$ also

$$g(\mathrm{Eig}(f;\lambda)) \subset \mathrm{Eig}(f;\lambda) \qquad \text{und} \qquad g(\mathrm{Hau}(f;\lambda)) \subset \mathrm{Hau}(f;\lambda).$$

Eine noch allgemeinere Aussage formuliert Übung 5.3.5.

Beispiel 5.3.4. Der Eigenraum zum Eigenwert Null einer linearen Abbildung $f : V \to V$ ist gerade ihr Kern $\mathrm{Eig}(f|V;0) = \ker f$. Der Eigenraum zum Eigenwert Eins einer linearen Abbildung $f : V \to V$ besteht genau aus allen Fixpunkten unserer Abbildung, in Formeln $\mathrm{Eig}(f|V;1) = V^f$. Der Hauptraum zum Eigenwert Null des durch Ableiten gegebenen Endomorphismus des Raums der Polynomfunktionen $\partial : \mathbb{R}[x] \to \mathbb{R}[x]$ ist der ganzen Raum, in Formeln $\mathrm{Hau}(\partial;0) = \mathbb{R}[x]$. Allgemeiner hat ein Endomorphismus $f : V \to V$ eines Vektorraums die Eigenschaft $\mathrm{Hau}(f;0) = V$ genau dann, wenn es für jeden Vektor $v \in V$ ein $N \in \mathbb{N}$ gibt mit $f^N(v) = 0$. Man sagt dann auch, f sei **lokal nilpotent**.

Übung 5.3.5. Gegeben ein kommutatives Diagramm von Vektorräumen der Gestalt

$$\begin{array}{ccc} V & \xrightarrow{g} & W \\ x\downarrow & & \downarrow y \\ V & \xrightarrow{g} & W \end{array}$$

alias Vektorräume V, W und lineare Abbildungen $g : V \to W$ und $x : V \to V$ und $y : W \to W$ mit $gx = yg$ bildet g Eigenräume in Eigenräume und Haupträume in Haupträume ab, in Formeln gilt also für alle λ aus dem jeweiligen Grundkörper $g(\mathrm{Eig}(x;\lambda)) \subset \mathrm{Eig}(y;\lambda)$ und $g(\mathrm{Hau}(x;\lambda)) \subset \mathrm{Hau}(y;\lambda)$.

5.3.6. Ist der Hauptraum zu einem Eigenwert λ nicht Null, so ist auch der zugehörige Eigenraum nicht Null: Ist in der Tat ein Vektor $v \neq 0$ gegeben mit $(f - \lambda\,\mathrm{id})^n v = 0$ für ein $n \in \mathbb{N}$, so gibt es auch ein kleinstmögliches derartiges $n \geq 1$, und dann ist $(f - \lambda\,\mathrm{id})^{n-1} v$ ein Eigenvektor zum Eigenwert λ.

5.3.7. Ist $U \subset V$ ein unter einer linearen Abbildung $f : V \to V$ stabiler Untervektorraum, so gilt für die Einschränkung von f auf U offensichtlich

$$\begin{array}{rcl} \mathrm{Eig}(f|U;\lambda) & = & \mathrm{Eig}(f;\lambda) \cap U \\ \mathrm{Hau}(f|U;\lambda) & = & \mathrm{Hau}(f;\lambda) \cap U \end{array}$$

5.3.8. Ist $V = U \oplus W$ die direkte Summe zweier unter f stabiler Untervektorräume, so gilt offensichtlich

$$\begin{array}{rcl} \mathrm{Eig}(f;\lambda) & = & \mathrm{Eig}(f|U;\lambda) \oplus \mathrm{Eig}(f|W;\lambda) \\ \mathrm{Hau}(f;\lambda) & = & \mathrm{Hau}(f|U;\lambda) \oplus \mathrm{Hau}(f|W;\lambda) \end{array}$$

Übung 5.3.9. Gegeben eine Menge von paarweise kommutierenden Endomorphismen eines von Null verschiedenen endlichdimensionalen Vektorraums über einem algebraisch abgeschlossenen Körper gibt es stets einen simultanen Eigenvektor. Hinweis: 5.3.3

Übung 5.3.10. Sei V ein Vektorraum und $T \subset \mathrm{End}\, V$ ein endlichdimensionaler Untervektorraum seines Endomorphismenraums, der aus diagonalisierbaren und paarweise kommutierenden Abbildungen besteht. So besitzt V unter T eine "simultane Eigenraumzerlegung"

$$V = \bigoplus_{\lambda \in T^*} V_\lambda$$

in die "simultanen Eigenräume" $V_\lambda = \{v \in V \mid xv = \lambda(x)v \ \forall x \in T\}$. Hinweis: Sei $x_0, \ldots, x_n$ eine Basis von T. Da x_0 diagonalisierbar ist, zerfällt V in Eigenräume unter x_0. Da die x_i für $i \geq 1$ mit x_0 kommutieren, stabilisieren sie dessen Eigenräume. Eine Induktion unter Verwendung von 2.8.13 beendet den Beweis.

Übung 5.3.11. Ein Endomorphismus eines euklidischen Vektorraums heißt **normal** genau dann, wenn er mit seinem Adjungierten kommutiert. Man zeige: Ein Endomorphismus eines endlichdimensionalen euklidischen Vektorraums ist genau dann normal, wenn es dazu eine Orthonormalbasis aus Eigenvektoren gibt. Hinweis: Kommutierende Endomorphismen stabilisieren die Eigenräume aller beteiligten Endomorphismen. Ist A^* adjungiert zu A, so sind $A + A^*$ und $\mathrm{i}(A - A^*)$ selbstadjungiert.

Proposition 5.3.12. *Die Summe der Haupträume ist stets direkt, d.h. für jeden Endomorphismus eines k-Vektorraums $f : V \to V$ liefern die Einbettungen der Haupträume eine Injektion*

$$\bigoplus_{\lambda\in k} \mathrm{Hau}(f;\lambda) \hookrightarrow V$$

Beweis. Wir zeigen zunächst, daß der Schnitt von je zwei Haupträumen zu verschiedenen Eigenwerten der Nullraum ist. Sonst müßte es nämlich nach 5.3.3, 5.3.6 und 5.3.7 in einem Hauptraum $\mathrm{Hau}(f;\lambda)$ auch einen Eigenvektor $v \neq 0$ zu einem Eigenwert $\mu \neq \lambda$ geben, und für diesen Vektor gälte $(f-\lambda\,\mathrm{id})^n v = (\mu-\lambda)^n v \neq 0$ für alle $n \geq 0$ im Widerspruch zu unserer Annahme $v \in \mathrm{Hau}(f;\lambda)$. Also ist der Schnitt von je zwei Haupträumen zu verschiedenen Eigenwerten der Nullraum. Wäre nun die Summe der Haupträume nicht direkt, so gäbe es nach 5.2.8 eine endliche direkte Summe

$$H = H_1 \oplus \ldots \oplus H_n$$

von Haupträumen, deren Bild in V von einem weiteren Hauptraum $\mathrm{Hau}(f;\mu)$ nichttrivial geschnitten wird. Da aber alle H_i stabil sind unter f, gilt nach 5.3.8

$$\mathrm{Hau}(f;\mu) \cap H = \mathrm{Hau}(f|H;\mu) = \mathrm{Hau}(f|H_1;\mu) \oplus \ldots \oplus \mathrm{Hau}(f|H_n;\mu)$$

und diese Summanden sind alle Null als Schnitte von Haupträumen zu verschiedenen Eigenwerten. Also ist der fragliche Schnitt doch Null und die Summe der Haupträume muß direkt sein. □

Beispiel 5.3.13. Wir zeigen, daß im $\mathbb{R}$-Vektorraum $\mathcal{C}^\infty(\mathbb{R},\mathbb{R})$ aller beliebig oft differenzierbaren Funktionen $\mathbb{R} \to \mathbb{R}$ die Funktionen $t \mapsto t^n \mathrm{e}^{\lambda t}$ eine linear unabhängige Familie $(t^n \mathrm{e}^{\lambda t})_{(n,\lambda)\in\mathbb{N}\times\mathbb{R}}$ bilden. In der Tat, betrachten wir den durch das Ableiten gegebenen Endomorphismus $D : \mathcal{C}^\infty(\mathbb{R},\mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R},\mathbb{R})$, so liegen alle $t^n \mathrm{e}^{\lambda t}$ für festes λ im λ-Hauptraum $\mathrm{Hau}(D;\lambda)$. Wegen 5.3.12 reicht es also, für jedes feste λ die lineare Unabhängigkeit der $t^n \mathrm{e}^{\lambda t}$ zu zeigen, und die folgt unmittelbar aus unserer Erkenntnis 2.3.14, daß ein reelles Polynom nur dann überall den Wert Null annimmt, wenn es das Nullpolynom ist. In derselben Weise zeigt man auch, daß im $\mathbb{C}$-Vektorraum $\mathcal{C}^\infty(\mathbb{R})$ aller beliebig oft differenzierbaren Funktionen $\mathbb{R} \to \mathbb{C}$ die Funktionen $t \mapsto t^n \mathrm{e}^{\lambda t}$ eine linear unabhängige Familie $(t^n \mathrm{e}^{\lambda t})_{(n,\lambda)\in\mathbb{N}\times\mathbb{C}}$ bilden, vergleiche **??**.

Satz 5.3.14 (Hauptraumzerlegung). *Ein endlichdimensionaler Vektorraum über einem algebraisch abgeschlossenen Körper zerfällt unter jedem*

Endomorphismus in die direkte Summe seiner Haupträume. In Formeln folgt über einem algebraisch abgeschlossenen Körper k aus $\dim_k V < \infty$ also

$$\bigoplus_{\lambda\in k} \mathrm{Hau}(f;\lambda) \xrightarrow{\sim} V$$

5.3.15. Der Satz gilt mit demselben Beweis auch, wenn wir statt der algebraischen Abgeschlossenheit des Grundkörpers nur voraussetzen, daß das charakteristische Polynom unseres Endomorphismus über unserem Körper vollständig in Linearfaktoren zerfällt.

Erster Beweis. Wir zeigen zunächst, daß der Hauptraum zum Eigenwert Null stets ein unter unserem Endomorphismus stabiles Komplement besitzt. Bezeichnet $f : V \to V$ unseren Endomorphismus, so behaupten wir genauer sogar, daß für hinreichend großes $n \gg 0$ unser Vektorraum V in die direkte Summe

$$V = (\ker f^n) \oplus (\operatorname{im} f^n)$$

zerfällt. Die Bilder der f^ν bilden in der Tat für wachsendes ν eine monoton fallende Folge von Untervektorräumen. Da V nach Annahme endliche Dimension hat, gibt es eine Stelle n, ab der diese Folge konstant wird. Für dieses n muß die Surjektion $f^n : \operatorname{im} f^n \twoheadrightarrow \operatorname{im} f^{2n}$ aus Dimensionsgründen ein Isomorphismus sein, also haben wir $(\ker f^n) \cap (\operatorname{im} f^n) = 0$ und nochmaliger Dimensionsvergleich mit der Dimensionsformel 1.6.10 zeigt über 1.6.13 die behauptete Zerlegung, die man auch als **Fitting-Zerlegung** bezeichnet. Die Hauptraumzerlegung ergibt sich nun leicht mit vollständiger Induktion über die Dimension: Ist unser Vektorraum Null, so ist eh nichts zu zeigen. Sonst gibt es einen Eigenwert λ. Die Fitting-Zerlegung von V zum Endomorphismus $(f - \lambda\,\mathrm{id})$ liefert dann zum λ-Hauptraum von f ein f-stabiles Komplement, und dieses Komplement können wir nach Induktionsannahme bereits in die Summe seiner Haupträume zerlegen. □

Zweiter Beweis. Das folgt mit 5.3.12 und Dimensionsvergleich auch unmittelbar aus der anschließenden Proposition 5.3.17, nach der die Dimension der Haupträume mit den Vielfachheiten der entsprechenden Eigenwerte als Nullstellen des charakteristischen Polynoms zusammenfallen. Man beachte jedoch, daß der hier gegebene Beweis dieser Proposition 5.3.17 auch auf der Fitting-Zerlegung beruht. □

Übung 5.3.16. Ein Vektorraum über einem algebraisch abgeschlossenen Körper zerfällt unter einem Endomorphismus in die direkte Summe seiner Haupträume genau dann, wenn unser Endomorphismus **lokal endlich** ist, als da heißt, jeder Vektor liegt in einem endlichdimensionalen unter unserem Endomorphismus stabilen Teilraum.

Proposition 5.3.17. *Gegeben ein Endomorphismus eines endlichdimensionalen Vektorraums stimmt die Dimension jedes Hauptraums überein mit der Vielfachheit des entsprechenden Eigenwerts als Nullstelle des charakteristischen Polynoms.*

Beweis. Sei $f : V \to V$ unser Endomorphismus und λ ein Skalar. Die Fitting-Zerlegung zu $(f - \lambda\,\mathrm{id})$ zerlegt V in die direkte Summe des λ-Hauptraums und eines f-stabilen Komplements

$$V = \mathrm{Hau}(f;\lambda) \oplus W$$

derart, daß λ kein Eigenwert von $f : W \to W$ ist. Auf dem Hauptraum ist $(f - \lambda\,\mathrm{id})$ nilpotent, nach 1.7.42 finden wir also darin eine Basis, bezüglich derer die Matrix von $(f - \lambda\,\mathrm{id})$ obere Dreiecksgestalt hat mit Nullen auf der Diagonalen. Bezüglich derselben Basis hat die Matrix von f obere Dreiecksgestalt mit lauter Einträgen λ auf der Diagonalen. Ergänzen wir diese Basis durch eine Basis von W zu einer Basis von V, so ist die zugehörige Matrix von $f : V \to V$ blockdiagonal und unsere Formel 2.8.9 liefert $\chi_f(T) = (\lambda - T)^d \chi_{f|W}(T)$ für $d = \dim \mathrm{Hau}(f;\lambda)$ die Dimension des λ-Hauptraums und $\chi_{f|W}(T)$ ohne Nullstelle bei λ. □

Bemerkung 5.3.18. Betrachten wir Vektorräume unendlicher Dimension, so besitzt der Hauptraum zum Eigenwert Null eines Endomorphismus im allgemeinen kein unter besagtem Endomorphismus stabiles Komplement mehr. Betrachten wir zum Beispiel den Vektorraum V aller Abbildungen von der Menge $\{(i,j) \in \mathbb{N}^2 \mid i \geq j\}$ in unseren Grundkörper und den Endomorphismus, der "jede Zeile eins nach unten rückt und die unterste Zeile zu Null mach". Der Hauptraum H zum Eigenwert Null besteht aus allen Funktionen, die nur auf endlich vielen Zeilen von Null verschieden sind. Betrachten wir den Vektor $v \in V$ mit

$$v(i,j) = \begin{cases} 1 & i = 2j; \\ 0 & \text{sonst}, \end{cases}$$

so ist sein Bild $\bar{v} \in V/H$ ein von Null verschiedener Vektor, der im Bild jeder Potenz unseres Endomorphismus liegt. In V selbst gibt es jedoch keinen derartigen von Null verschiedenen Vektor, folglich kann $H \subset V$ kein unter unserem Endomorphismus stabiles Komplement besitzen.

5.4 Jordan-Zerlegung

Satz 5.4.1 (**Jordan-Zerlegung**). *Sei V ein endlichdimensionaler Vektorraum über einem algebraisch abgeschlossenen Körper und $x \in \mathrm{End}\, V$ ein Endomorphismus von V. So gibt es genau eine Zerlegung $x = x_\mathrm{s} + x_\mathrm{n}$ mit x_s diagonalisierbar, x_n nilpotent und $x_\mathrm{s}x_\mathrm{n} = x_\mathrm{n}x_\mathrm{s}$.*

5.4.2. Der untere Index s bei x_s steht für "semisimple", die deutsche Übersetzung dafür ist "halbeinfach". Ein Endomorphismus a eines Vektorraums V über einem Körper k heißt ganz allgemein halbeinfach genau dann, wenn er über einem algebraischen Abschluß von k diagonalisierbar ist. In der Situation des Lemmas heißen x_s bzw. x_n der **halbeinfache** bzw. der **nilpotente Anteil** von x.

Beweis. Gegeben ein Endomorphismus x eines endlichdimensionalen Vektorraums über einem algebraisch abgeschlossenen Körper erklären wir einen Endomorphismus x_s durch die Vorschrift, daß er auf dem Hauptraum $\mathrm{Hau}(x;\lambda)$ von x zum Eigenwert λ jeweils durch die Multiplikation mit λ operieren soll. Dann ist x_s diagonalisierbar, und setzen wir $x_\mathrm{n} = x - x_\mathrm{s}$, so ist x_n nilpotent und x_s kommutiert mit x und dann auch mit x_n. Das zeigt die Existenz unserer Zerlegung. Ist $x = s + n$ eine weitere Zerlegung mit s diagonalisierbar, n nilpotent und $sn = ns$, so folgt zunächst $sx = xs$ und dann, da s die Haupträume von x stabilisieren muß, auch $sx_\mathrm{s} = x_\mathrm{s}s$. So erkennen wir, daß x, s, n, x_s und x_n paarweise kommutieren. Natürlich ist dann $x_\mathrm{n} - n$ nilpotent. Da s die Haupträume von x stabilisiert und da nach 2.8.13 auch die Restriktion von s auf besagte Haupträume diagonalisierbar ist, folgt aus der Definition von x_s, daß auch $x_\mathrm{s} - s$ diagonalisierbar sein muß. Aus $x_\mathrm{n} - n = s - x_\mathrm{s}$ folgt dann aber sofort, daß beide Seiten Null sind. Das zeigt die Eindeutigkeit unserer Zerlegung. □

Bemerkung 5.4.3. Hier lassen sich x_s und x_n sogar als Polynome in x ohne konstanten Term ausdrücken, d.h. es gibt $P, Q \in T\mathbb{C}[T]$ mit $x_\mathrm{s} = P(x)$ und $x_\mathrm{n} = Q(x)$. In der Tat, falls N so groß ist, daß gilt $\mathrm{Hau}(x;\lambda) = \ker(x-\lambda)^N$ für alle λ, so erhält man ein mögliches P aus dem chinesischen Restsatz **??** als simultane Lösung der Kongruenzen $P \equiv \lambda \pmod{(T-\lambda)^N}$ für alle Eigenwerte λ von x und für $\lambda = 0$, und ein mögliches Q ist dann $T - P(T)$. Ich mag die in der Literatur weit verbreitete Argumentation mit diesen Polynomen jedoch nicht besonders.

Satz 5.4.4 (Funktorialität der Jordan-Zerlegung). *Sei gegeben ein kommutatives Diagramm endlichdimensionaler komplexer Vektorräume der Gestalt*

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ x\downarrow & & \downarrow y \\ V & \xrightarrow{f} & W \end{array}$$

Sind $x = x_\mathrm{s} + x_\mathrm{n}$ und $y = y_\mathrm{s} + y_\mathrm{n}$ die Jordan-Zerlegungen von x und y, so kommutieren auch die Diagramme

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ x_{\mathrm{s}}\downarrow & & \downarrow y_{\mathrm{s}} \\ V & \xrightarrow{f} & W \end{array} \qquad \begin{array}{ccc} V & \xrightarrow{f} & W \\ x_{\mathrm{n}}\downarrow & & \downarrow y_{\mathrm{n}} \\ V & \xrightarrow{f} & W \end{array}$$

Beweis. Aus $fx = yf$ folgt wegen $f(\mathrm{Hau}(x;\lambda)) \subset \mathrm{Hau}(y;\lambda)$ nach der im vorhergehenden Beweis von 5.4.1 gegebenen Beschreibung der Jordan-Zerlegung unmittelbar erst $fx_{\mathrm{s}} = y_{\mathrm{s}}f$ und dann $fx_{\mathrm{n}} = y_{\mathrm{n}}f$. □

5.4.5. Stabilisiert speziell ein Endomorphismus eines endlichdimensionalen Vektorraums einen vorgegebenen Teilraum, so stabilisieren nach 5.4.4 auch sein halbeinfacher und sein nilpotenter Anteil besagten Teilraum.

5.4.6. Beide Sätze 5.4.1 und 5.4.4 gelten weiter und mit demselben Beweis auch noch, wenn man statt der Endlichdimensionalität der darin auftauchenden Vektorräume nur fordert, daß die fraglichen Endomorphismen x und y im Sinne von 5.3.16 lokal endlich sein sollen, und von x_{n} schwächer nur fordert, daß es lokal nilpotent sein soll.

Übung 5.4.7. Gegeben ein endlichdimensionaler komplexer Vektorraum V und ein Endomorphismus $x : V \to V$ haben wir stets

$$\operatorname{im} x \supset \operatorname{im} x_{\mathrm{s}}$$

Hinweis: Das Bild von x_{s} ist genau die Summe der Haupträume zu von Null verschiedenen Eigenwerten und das Bild von x umfaßt offensichtlich diese Summe. Alternativ erkennt man $\operatorname{im} x \supset \operatorname{im}(x_{\mathrm{s}}^N)$ für hinreichend großes N durch Entwicklung von $x_{\mathrm{s}}^N = (x - x_{\mathrm{n}})^N$ nach der binomischen Formel und Ausklammern von x, und die Behauptung folgt wegen $\operatorname{im} x_{\mathrm{s}} = \operatorname{im} x_{\mathrm{s}}^N$.

Übung 5.4.8. Jeder Endomorphismus der Ordnung zwei eines Vektorraums über einem Körper einer von zwei verschiedenen Charakteristik ist diagonalisierbar. Hinweis: Später zeigen wir das als **??**. Jeder Endomorphismus der Ordnung vier eines komplexen Vektorraums ist diagonalisierbar. Hinweis: Man zerlege zunächst in Eigenräume unter dem Quadrat unseres Endomorphismus. Allgemeiner werden Sie in 5.4.9 zeigen, daß jeder Endomorphismus endlicher Ordnung eines komplexen Vektorraums diagonalisierbar ist.

Übung 5.4.9. Sei k ein algebraisch abgeschlossener Körper der Charakteristik Null. Sei V ein k-Vektorraum und $\varphi : V \to V$ ein Endomorphismus "endlicher Ordnung", als da heißt, es gebe $n \geq 1$ mit $\varphi^n = \mathrm{id}$. So ist V die direkte Summe der Eigenräume von φ. Hinweis: Man behandle zunächst den endlichdimensionalen Fall mithilfe der Jordan-Zerlegung und beachte dabei, daß höhere Potenzen eines nilpotenten Endomorphismus stets größere Kerne haben müssen, solange nicht beide fraglichen Potenzen bereits Null sind. Fortgeschrittene erkennen einen Spezialfall des Satzes von Maschke **??**.

Definition 5.4.10. Ein Endomorphismus f eines Vektorraums heißt **unipotent** genau dann, wenn $(f - \mathrm{id})$ nilpotent ist.

Übung 5.4.11. Ein unipotenter Endomorphismus endlicher Ordnung eines Vektorraums über einem Körper der Charakteristik Null ist bereits die Identität.

Übung 5.4.12. Gegeben ein Endomorphismus eines endlichdimensionalen Vektorraums über einem algebraisch abgeschlossenen Körper besteht der Hauptraum des transponierten Endomorphismus des Dualraums genau aus den Linearformen, die auf allen Haupträumen zu anderen Eigenwerten des ursprünglichen Endomorphismus verschwinden.

5.5 Jordan'sche Normalform

Definition 5.5.1. Gegeben $r \geq 1$ definieren wir eine $(r \times r)$-Matrix $J(r)$, genannt der **nilpotente Jordan-Block der Größe** r, durch die Vorschrift $J(r)_{i,j} = 1$ für $j = i + 1$ und $J(r)_{i,j} = 0$ sonst. Insbesondere ist also $J(1)$ die (1×1)-Matrix mit dem einzigem Eintrag Null.

Satz 5.5.2 (Normalform nilpotenter Endomorphismen). *Gegeben ein nilpotenter Endomorphismus eines endlichdimensionalen Vektorraums gibt es stets eine Basis derart, daß die Matrix unseres Endomorphismus in dieser Basis blockdiagonal ist mit nilpotenten Jordanblöcken auf der Diagonalen, also von der Gestalt*

$$\mathrm{diag}(J(r_1), \ldots, J(r_n))$$

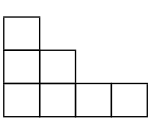
Die positiven natürlichen Zahlen $r_1, \ldots, r_n$ sind hierbei durch unseren nilpotenten Endomorphismus eindeutig bestimmt bis auf Reihenfolge.

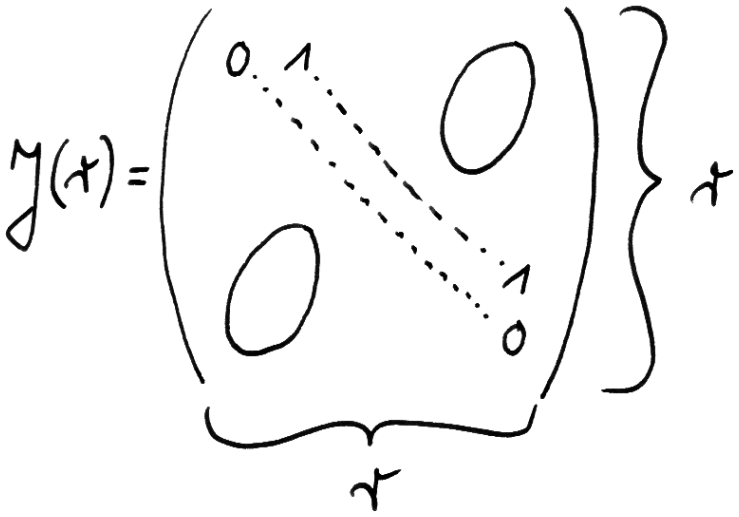
Beweis von 5.5.2. Wir beginnen mit dem Beweis der Eindeutigkeit und beginnen diesen mit einer Definition.

Definition 5.5.3. Unter einem **Youngdiagramm** verstehen wir eine endliche Teilmenge $T \subset \mathbb{N} \times \mathbb{N}$ mit der Eigenschaft

$$((i, j) \in T \text{ und } i' \leq i \text{ und } j' \leq j) \Rightarrow (i', j') \in T$$

Die Elemente von T nennen wir die "Kästchen" unseres Youngdiagramms und stellen uns ein Paar (i, j) vor als das Kästchen auf einem Rechenpapier, bei dem die Koordinaten der linken unteren Ecke gerade (i, j) sind. Zum Beispiel stellt das Bild





Der nilpotente Jordan-Block $J(r)$ der Größe r. Steht auf der Diagonalen statt der Nullen ein Skalar λ, so nennen wir die entsprechende Matrix einen **Jordan-Block der Größe r zum Eigenwert λ** und notieren diese Matrix

$$J(r;\lambda) = J(r) + \lambda I_r$$

das Youngdiagramm dar, das formal zu beschreiben wäre durch die Menge $\{(0,0),(0,1),(0,2),(1,0),(1,1),(2,0),(3,0)\}$. In der Praxis denkt man bei Youngdiagrammen meist an Bilder dieser Art.

Zu jedem Young-Diagramm T bilden wir ein Paar (V_T, N_T) bestehend aus einem Vektorraum mit einem nilpotentem Endomorphismus, indem wir die Kästchen von T als Basis von V_T nehmen und denjenigen Endomorphismus N_T von V_T betrachten, der jedes Kästchen um eins nach rechts schiebt bzw. es anulliert, falls es beim "Um-eins-nach-rechts-schieben aus dem Young-Diagramm herausrutscht". In Formeln ist also $V_T = kT$ der freie k-Vektorraum über T, und bezeichnen wir mit $v_{i,j}$ den zu $(i,j) \in T$ gehörenden Basisvektor, so wird $N_T : kT \to kT$ beschrieben durch die Vorschrift

$$N_T(v_{i,j}) = \begin{cases} v_{i+1,j} & \text{falls } (i+1,j) \in T; \\ 0 & \text{sonst.} \end{cases}$$

Gegeben zwei Paare (V, A) und (W, B) bestehend aus einem k-Vektorraum mit einem Endomorphismus schreiben wir

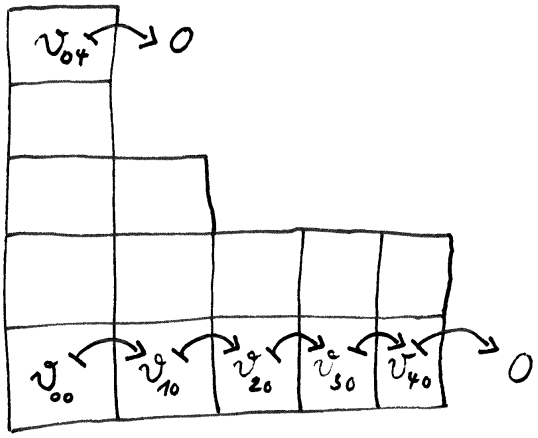
$$(V, A) \cong (W, B)$$

und nennen unsere Paare **isomorph** genau dann, wenn es einen Isomorphismus $\varphi : V \xrightarrow{\sim} W$ gibt mit $B \circ \varphi = \varphi \circ A$. In dieser Sprache ausgedrückt sagt unser Satz 5.5.2, daß für jedes Paar (V, N) bestehend aus einem endlichdimensionalen Vektorraum mit einem nilpotenten Endomorphismus genau ein Young-Diagramm T existiert mit $(V, N) \cong (V_T, N_T)$. Um hier die Eindeutigkeit zu zeigen, müssen wir nur die Implikation

$$(V_T, N_T) \cong (V_{T'}, N_{T'}) \quad \Rightarrow \quad T = T'$$

nachweisen. Offensichtlich besteht jedoch eine mögliche Basis von $(\operatorname{im} N_T^n)$ aus allen Kästchen von T, die in der $(n+1)$-ten Spalte unseres Young-Diagramms oder noch weiter rechts stehen, in Formeln aus allen $v_{i,j}$ mit $(i,j) \in T$ und $i \geq n$. Folglich ist $\dim(\operatorname{im} N_T^n / \operatorname{im} N_T^{n+1})$ gerade die Höhe der $(n+1)$-ten Spalte des Young-Diagramms T, für alle $n \geq 0$, und das zeigt die Eindeutigkeit. Die Existenz folgt unmittelbar aus Lemma 5.5.4, das wir gleich im Anschluß beweisen. □

Lemma 5.5.4. *Ist $N : V \to V$ ein nilpotenter Endomorphismus eines endlichdimensionalen Vektorraums V, so gibt es eine Basis B von V derart, daß $B \cup \{0\}$ stabil ist unter N und daß jedes Element von B unter N höchstens ein Urbild in B hat.*



Zum Beweis von 5.5.2

Bemerkung 5.5.5. Dieses Lemma gilt sogar ohne die Voraussetzung, daß V endlichdimensional ist. Um das zu zeigen, müssen wir nur die Induktion statt über die Dimension von V über die Nilpotenzordnung von N laufen lassen und 1.4.30 verwenden. Für einen lokal nilpotenten Endomorphismus N findet man jedoch im Allgemeinen keine "Jordan-Basis" wie in 5.5.4 mehr. Wir betrachten zum Beispiel den Raum V aller Abbildungen von der Menge $\{(i,j) \in \mathbb{N}^2 \mid i \geq j\}$ nach $\mathbb{R}$, die nur in endlich vielen Zeilen nicht identisch Null sind, und den Endomorphismus $N : V \to V$, der "jede Zeile um eins nach unten drückt und die nullte Zeile annulliert". Sicher hat V/NV eine abzählbare Basis, so daß es nur abzählbar viele "Jordan-Ketten" geben könnte. Die Existenz einer "Jordan-Basis" stünde also im Widerspruch dazu, daß V keine abzählbare Basis besitzt.

Beweis. Wir betrachten die Sequenz

$$\ker N \hookrightarrow V \twoheadrightarrow \operatorname{im} N$$

Mit Induktion über die Dimension von V dürfen wir annehmen, daß wir für das Bild von N eine derartige Basis bereits gefunden haben, sagen wir die Basis A. Jetzt ergänzen wir $\{a \in A \mid N(a) = 0\}$ durch irgendwelche $b_1, \ldots, b_s$ zu einer Basis des Kerns von N und wählen Urbilder $c_1, \ldots, c_r \in V$ für die Elemente von $A \backslash N(A)$ und behaupten, daß

$$B = A \cup \{b_1, \ldots, b_s\} \cup \{c_1, \ldots, c_r\}$$

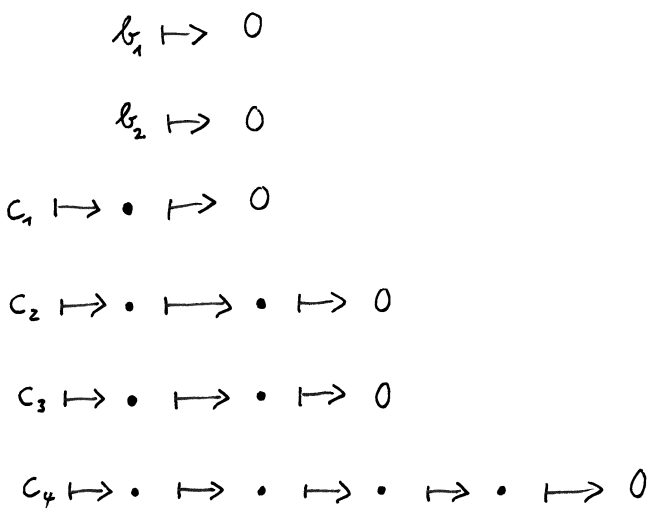
eine Basis von V ist mit den geforderten Eigenschaften. Nach Konstruktion ist $B \cup \{0\}$ stabil unter N und jedes Element von B hat unter N höchstens ein Urbild in B. Wir müssen also nur noch zeigen, daß B eine Basis von V ist. Dazu schreiben wir B als die Vereinigung der beiden Mengen

$$\begin{array}{lcl} \{a \in A \mid N(a) \neq 0\} & \cup & \{c_1, \ldots, c_r\} \\ \{a \in A \mid N(a) = 0\} & \cup & \{b_1, \ldots, b_s\} \end{array}$$

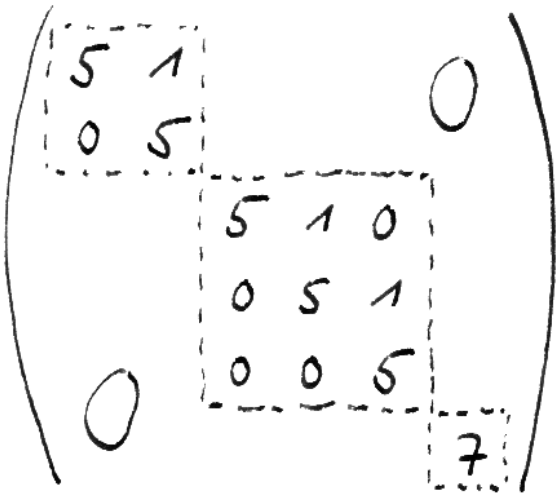
und bemerken, daß die erste ein System von Urbildern unter N für unsere Basis A von $(\operatorname{im} N)$ ist, wohingegen die zweite eine Basis von $(\ker N)$ ist. Damit ist unsere große Vereinigung eine Basis von V nach 1.6.11. □

Korollar 5.5.6 (Jordan'sche Normalform). *Gegeben ein Endomorphismus eines endlichdimensionalen Vektorraums über einem algebraisch abgeschlossenen Körper gibt es eine Basis unseres Vektorraums derart, daß die Matrix unseres Endomorphismus bezüglich dieser Basis blockdiagonal ist von der Gestalt*

$$\operatorname{diag}(J(r_1; \lambda_1), \ldots, J(r_t; \lambda_t))$$



Zum Beweis von 5.5.4. Die fetten Punkte stellen die Elemente der Basis A des Bildes im N dar. Die c_i zusammen mit den $a \in A$ mit $N(a) \neq 0$ bilden ein System von Urbildern unter N der Elemente von A.



Ein Matrix in Jordan'scher Normalform mit drei Jordanblöcken.

Die Jordan-Blöcke sind hierbei durch unseren Endomorphismus wohlbestimmt bis auf Reihenfolge.

Bemerkung 5.5.7. Das Korollar gilt mit demselben Beweis auch, wenn wir statt der algebraischen Abgeschlossenheit des Grundkörpers nur voraussetzen, daß das charakteristische Polynom unseres Endomorphismus über unserem Körper vollständig in Linearfaktoren zerfällt.

Beweis. Sei f unser Endomorphismus. Der Satz über die Hauptraumzerlegung 5.3.14 zeigt, daß wir ohne Beschränkung der Allgemeinheit annehmen dürfen, daß es einen Skalar λ gibt derart, daß $(f - \lambda\,\mathrm{id})$ nilpotent ist. Der Satz über die Normalform nilpotenter Endomorphismen 5.5.2 beendet dann den Beweis. □

6 Algebra und Symmetrie

6.1 Gruppenwirkungen

Definition 6.1.1. Eine **Operation** oder **Wirkung** einer Gruppe G auf einer Menge X ist eine Abbildung

$$\begin{array}{ccc} G \times X & \to & X \\ (g,x) & \mapsto & gx \end{array}$$

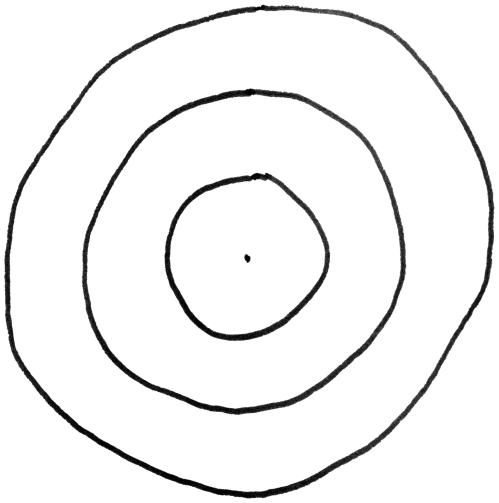
derart, daß gilt $g(hx) = (gh)x$ für alle $g, h \in G$, $x \in X$ und $ex = x$ für das neutrale Element $e \in G$ und alle $x \in X$. Ich ziehe die Bezeichnung als Operation vor, da der Begriff der "Wirkung" in der Physik in einer anderen Bedeutung verwendet wird. Eine Menge mit einer Operation einer Gruppe G nennt man eine G-**Menge**. Die Aussage "X ist eine G-Menge" schreiben wir in Formeln

$$G \curvearrowright X$$

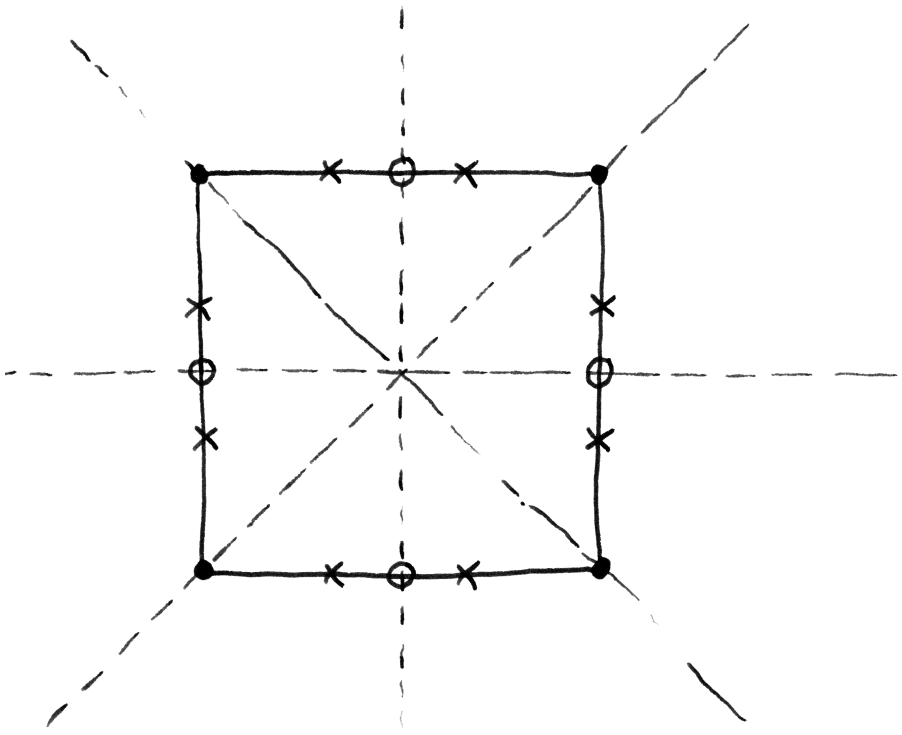
Bemerkung 6.1.2. In derselben Weise erklärt man allgemeiner auch den Begriff der Operation eines Monoids auf einer Menge. Allerdings ist der Begriff in dieser Allgemeinheit wesentlich weniger nützlich, da viele der im folgenden bewiesenen Aussagen wie die Zerlegung in Bahnen 6.1.10 oder die Darstellung von Bahnen als Quotienten 6.3.1 in dieser Allgemeinheit nicht mehr gelten.

Beispiele 6.1.3.

1. Ist X ein G-Menge, so ist auch die Potenzmenge $\mathcal{P}(X)$ eine G-Menge in natürlicher Weise.
2. Das Anwenden eines Isomorphismus definiert für jeden Vektorraum V eine Operation $\mathrm{GL}(V) \times V \to V$ von $\mathrm{GL}(V)$ auf V.
3. Jede Gruppe operiert vermittels ihrer Verknüpfung $G \times G \to G$ auf sich selbst.
4. Die symmetrische Gruppe $\mathcal{S}_n$ operiert in offensichtlicher Weise auf der Menge $\{1, 2, \ldots, n\}$.
5. Jede Gruppe G operiert auf jeder Menge X vermittels der **trivialen Operation** $gx = x \; \forall g \in G, x \in X$.
6. Ist G eine Gruppe und X eine G-Menge und $H \subset G$ eine Untergruppe, so ist X auch eine H-Menge in offensichtlicher Weise. Ist allgemeiner X eine G-Menge und $H \to G$ ein Gruppenhomomorphismus, so kann X in offensichtlicher Weise mit einer Operation von H versehen werden.



Einige Bahnen von S^1 auf $\mathbb{C}$



Einige Bahnen der Symmetriegruppe eines Quadrats

Übung 6.1.4. Gegeben ein Monoid G und eine Menge X induziert unsere Bijektion $\mathrm{Ens}(G \times X, X) \xrightarrow{\sim} \mathrm{Ens}(G, \mathrm{Ens}(X, X))$ aus I.2.2.23 eine Bijektion

$$\left\{ \begin{array}{c} \text{Operationen des Monoids } G \\ \text{auf der Menge } X \end{array} \right\} \xrightarrow{\sim} \left\{ \begin{array}{c} \text{Monoidhomomorphismen} \\ G \to \mathrm{Ens}(X) \end{array} \right\}$$

Ist G eine Gruppe, so erhalten wir insbesondere eine Bijektion

$$\left\{ \begin{array}{c} \text{Operationen der Gruppe } G \\ \text{auf der Menge } X \end{array} \right\} \xrightarrow{\sim} \left\{ \begin{array}{c} \text{Gruppenhomomorphismen} \\ G \to \mathrm{Ens}^\times(X) \end{array} \right\}$$

In gewisser Weise ist also eine Operation einer Gruppe G auf einer Menge X "dasselbe" wie ein Gruppenhomomorphismus $G \to \mathrm{Ens}^\times(X)$.

6.1.5. Ist ganz allgemein $X \times Y \to Z$ eine Abbildung, etwa $(x, y) \mapsto x \top y$, und sind $A \subset X$ und $B \subset Y$ Teilmengen, so notieren wir $(A \top B) \subset Z$ die Teilmenge

$$(A \top B) = \{x \top y \mid x \in A,\ y \in B\}$$

Wir haben derartige Notationen auch bereits oft verwendet, zum Beispiel, wenn wir das Erzeugnis eines Vektors in einem reellen Vektorraum als $\langle v \rangle = \mathbb{R}v$ schreiben, oder wenn wir das Erzeugnis von zwei Teilräumen U, W eines Vektorraums V als $U + W$ schreiben.

Definition 6.1.6. Sei X eine Menge mit einer Operation einer Gruppe G, also eine G-Menge.

1. Die Menge aller **Fixpunkte** von G notiert man

$$X^G = \{x \in X \mid gx = x\ \forall g \in G\}$$

   In vielen Situationen nennt man die Fixpunkte auch **Invarianten**.

2. Die **Standgruppe** oder **Isotropiegruppe** oder auch der **Fixator** oder **Stabilisator** eines Punktes $x \in X$ ist die Menge

$$G_x = \{g \in G \mid gx = x\}$$

   Sie ist eine Untergruppe von G. Ist allgemeiner $A \subset X$ eine Teilmenge, so unterscheiden wir zwischen dem **Stabilisator** $\{g \in G \mid gA = A\}$ und dem **Fixator** $\{g \in G \mid gx = x\ \forall x \in A\}$. Beide sind Untergruppen von G.

3. Eine G-Menge X heißt **frei** genau dann, wenn die Standgruppen aller ihrer Punkte trivial sind, in Formeln $(gx = x \text{ für ein } x \in X) \Rightarrow (g = e)$.

4. Für $A \subset X$, $H \subset G$ schreiben wir kurz HA für die Menge $HA = \{ha \mid h \in H, a \in A\}$. Für jede Teilmenge $A \subset X$ ist GA eine G-Menge in offensichtlicher Weise. Eine Teilmenge $Y \subset X$ heißt **G-stabil** genau dann, wenn gilt $GY = Y$, wenn also $Y \in \mathcal{P}(X)$ für die auf der Potenzmenge induzierte Wirkung ein Fixpunkt ist.

5. Sei $x \in X$. Die Menge

$$Gx = \{gx \mid g \in G\} \subset X$$

heißt die **Bahn** (englisch und französisch **orbit**) von x.

6. Eine Operation heißt **transitiv** und X heißt ein **homogener Raum** für G genau dann, wenn es ein $x \in X$ gibt mit $X = Gx$.

7. Eine Menge X mit einer freien transitiven Operation einer Gruppe G heißt ein **prinzipaler homogener Raum** für die Gruppe G oder auch kürzer ein **G-Torsor**.

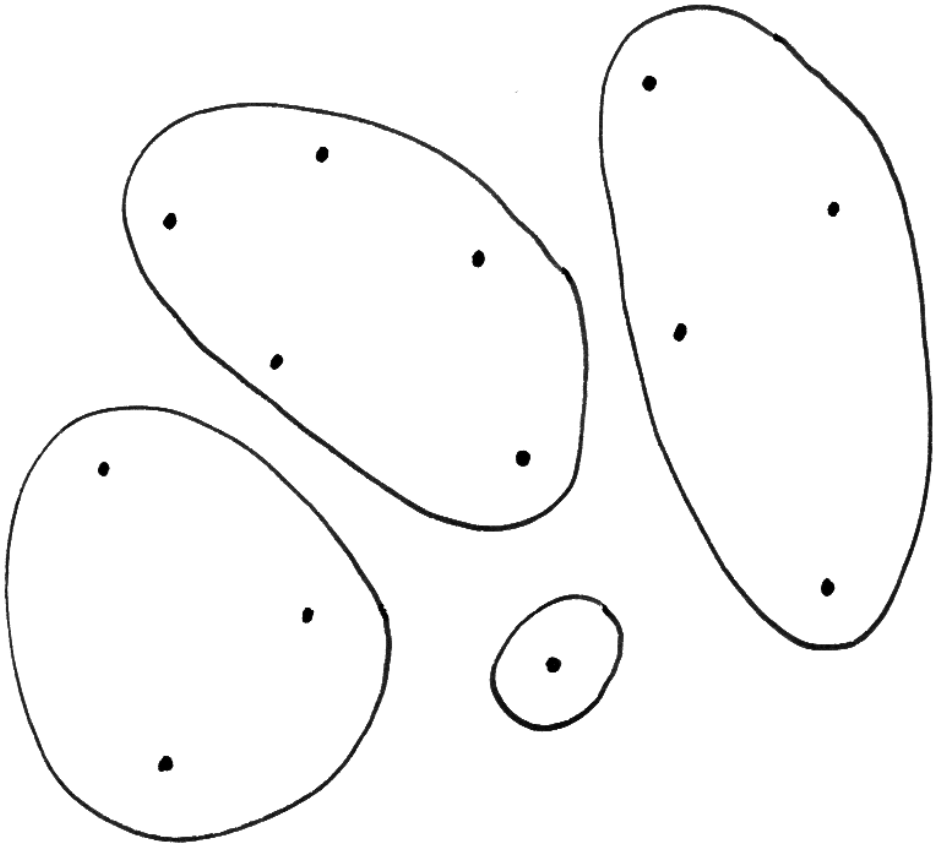
Beispiele 6.1.7. In jedem eindimensionalen Vektorraum über einem Körper k bilden die von Null verschiedenen Vektoren einen Torsor über der multiplikativen Gruppe $k^\times$ unseres Körpers. Jeder affine Raum ist ein Torsor über seinem Richtungsraum. Jede Menge mit genau zwei Elementen ist in natürlicher Weise ein $(\mathbb{Z}/2\mathbb{Z})$-Torsor. Jede Gruppe G kann in offensichtlicher Weise aufgefaßt werden als ein G-Torsor.

Bemerkung 6.1.8. Die Wirkung einer Gruppe auf der leeren Menge ist in unseren Konventionen nicht transitiv. Hier sind jedoch auch andere Konventionen gebräuchlich, zum Beispiel nennt Bourbaki die Wirkung einer Gruppe auf der leeren Menge durchaus transitiv. Noch mehr Terminologie zu Mengen mit Gruppenwirkung führen wir in **??** ein.

6.1.9. Eine **Partition einer Menge** X ist ein System $\mathcal{U} \subset \mathcal{P}(X)$ von paarweise disjunkten nichtleeren Teilmengen, deren Vereinigung ganz X ist.

Lemma 6.1.10 (Zerlegung in Bahnen). *Gegeben eine Menge mit Gruppenoperation bilden die Bahnen eine Partition unserer Menge.*

Beweis. Sei $G \curvearrowright X$ unsere Menge mit Gruppenoperation. Natürlich liegt jedes $x \in X$ in einer G-Bahn, nämlich in der G-Bahn Gx. Andererseits folgt aus $Gx \cap Gy \neq \emptyset$ schon $Gx = Gy$: In der Tat liefert $gx = hy$ wegen $Gg = G$ ja $Gx = Ggx = Ghy = Gy$. Die Bahnen sind also auch paarweise disjunkt. □

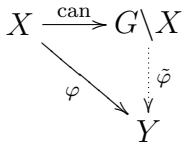


Eine Partition einer Menge mit dreizehn Elementen durch vier Teilmengen.

Definition 6.1.11. Gegeben eine Menge mit Gruppenoperation bezeichnet man das Mengensystem der Bahnen auch als den **Bahnenraum**. Ist $G \curvearrowright X$ unsere Menge mit Gruppenoperation, so ist der Bahnenraum also die Teilmenge $\{Gx \mid x \in X\} \subset \mathcal{P}(X)$ der Potenzmenge von X. Man notiert den Bahnenraum meist $G\backslash X$ oder auch X/G. Wir haben eine kanonische Surjektion $\mathrm{can} : X \twoheadrightarrow G\backslash X$, $x \mapsto Gx$, die jedem Element von X seine Bahn zuordnet.

Beispiel 6.1.12. Wir betrachten die Menge $X = \mathbb{C}$ der komplexen Zahlen mit der Operation von $G = S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ durch Multiplikation. Die Standgruppen sind $G_x = 1$ falls $x \neq 0$ und $G_0 = S^1$. Die Bahnen sind genau alle Kreise um den Nullpunkt mit Radius $r \geq 0$. Die Einbettung $\mathbb{R}_{\geq 0} \hookrightarrow \mathbb{C}$ induziert eine Bijektion mit dem Bahnenraum $\mathbb{R}_{\geq 0} \xrightarrow{\sim} (S^1\backslash\mathbb{C})$.

6.1.13 (**Universelle Eigenschaft des Bahnenraums**). Gegeben eine Menge mit Gruppenoperation $G \curvearrowright X$ und eine Abbildung in eine weitere Menge $\varphi : X \to Y$ mit der Eigenschaft $\varphi(gx) = \varphi(x)$ für alle $g \in G, x \in X$ existiert genau eine Abbildung $\tilde{\varphi} : G\backslash X \to Y$ mit $\tilde{\varphi} \circ \mathrm{can} = \varphi$, im Diagramm



In der Tat können und müssen wir $\tilde{\varphi}(Gx)$ als das einzige Element der Menge $\varphi(Gx)$ definieren. Man mag das auch als einen Spezialfall der universellen Eigenschaft des Raums der Äquivalenzklassen einer Äquivalenzrelation im Sinne von 2.4.3 verstehen.

Definition 6.1.14. Sei X eine Menge und G eine Gruppe. Eine **Rechtsoperation** von G auf X ist eine Abbildung

$$\begin{array}{ccc} X \times G & \to & X \\ (x, g) & \mapsto & xg \end{array}$$

derart, daß $x(gh) = (xg)h$ für alle $g, h \in G$, $x \in X$, und daß gilt $xe = x$ für das neutrale Element $e \in G$ und alle $x \in X$. Eine Menge mit einer Rechtsoperation einer Gruppe G nennt man auch eine G-**Rechtsmenge**. Eine freie und transitive G-Rechtsmenge nennt man einen G-**Rechtstorsor** oder auch kurz ein G-**Torsor** in der Hoffnung, daß der Leser aus dem Kontext erschließen kann, ob im jeweils vorliegenden Fall eine Menge mit freier und transitiver Rechts- oder mit freier und transitiver Linksoperation gemeint ist.

Bemerkung 6.1.15. Jede G-Rechtsmenge X wird zu einer G-Menge durch die Operation $gx = xg^{-1}$, die Begriffsbildung einer G-Rechtsmenge ist also in gewisser Weise obsolet. Sie dient im wesentlichen dem Zweck, in manchen Situationen suggestivere Notationen zu ermöglichen.

6.1.16. Gegeben zwei Gruppen G und H können wir auch ihr kartesisches Produkt $G \times H$ zu einer Gruppe machen, indem wir darauf die komponentenweise Verknüpfung $(g, h)(g', h') = (gg', hh')$ betrachten. Analog versehen wir das Produkt einer beliebigen Familie von Gruppen mit der Struktur einer Gruppe.

Übung 6.1.17. Sei k ein Körper. Man zeige, daß wir eine Operation der Gruppe $\mathrm{GL}(n;k) \times \mathrm{GL}(m;k)$ auf der Menge $\mathrm{M}(n \times m; k)$ erhalten durch die Vorschrift $(A, B)M = AMB^{-1}$. Man zeige weiter, daß die Bahnen unserer Operation genau die nichtleeren Fasern der durch den Rang gegebenen Abbildung $\mathrm{rk} : \mathrm{M}(n \times m; k) \to \mathbb{N}$ sind. Hinweis: Smith-Normalform 1.7.41.

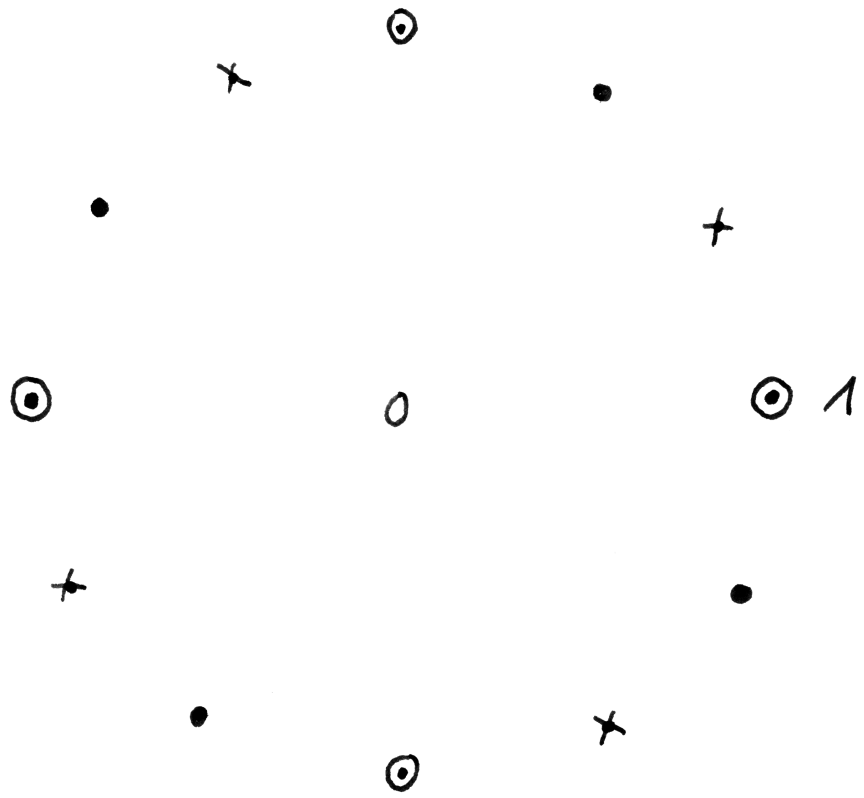
Übung 6.1.18. Sei k ein Körper. Man zeige, daß wir eine Operation der Gruppe $\mathrm{GL}(n;k)$ auf der Menge $\mathrm{M}(n \times n; k)$ erhalten durch die Vorschrift $A.M = AMA^{-1}$. Man zeige, wie für einen algebraisch abgeschlossenen Körper k die Theorie der Jordan'schen Normalform eine Bijektion liefert zwischen dem Bahnenraum zu dieser "Operation durch Konjugation" und der Menge aller endlichen Multimengen von Paaren aus $\mathbb{N}_{\geq 1} \times k$, deren erste Komponenten sich zu n aufaddieren.

6.2 Restklassen

6.2.1. Ist $(G, \perp)$ eine Menge mit Verknüpfung und sind $A, B \subset G$ Teilmengen, so schreiben wir $A \perp B = \{a \perp b \mid a \in A,\ b \in B\} \subset G$ und erhalten auf diese Weise eine Verknüpfung auf der Menge aller Teilmengen von G, der sogenannten Potenzmenge $\mathcal{P}(G)$. Ist unsere ursprüngliche Verknüpfung assoziativ, so auch die induzierte Verknüpfung auf der Potenzmenge. Wir kürzen in diesem Zusammenhang oft die einelementige Menge $\{a\}$ mit a ab, so daß also zum Beispiel $a \perp B$ als $\{a\} \perp B$ zu verstehen ist.

Definition 6.2.2. Ist G eine Gruppe, $H \subset G$ eine Untergruppe und $g \in G$ ein Element, so nennen wir die Menge gH die **Linksnebenklasse von** g **unter** H und die Menge Hg die **Rechtsnebenklasse von** g **unter** H. Diese Nebenklassen sind also Teilmengen von G. Ein Element einer Nebenklasse nennt man einen **Repräsentanten** der besagten Nebenklasse. Weiter betrachten wir in G die beiden Mengensysteme

$$\begin{aligned} G/H &= \{gH \mid g \in G\} \\ H\backslash G &= \{Hg \mid g \in G\} \end{aligned}$$



Die drei Nebenklassen der Gruppe der vierten Einheitswurzeln in der Gruppe der zwölften Einheitswurzeln. Da diese Gruppe kommutativ ist, fallen hier Rechtsnebenklassen und Linksnebenklassen zusammen.

aller Links- bzw. Rechtsnebenklassen von H in G. Die Elemente von G/H und von $H\backslash G$ sind also Teilmengen von G und G/H sowie $H\backslash G$ selbst sind dementsprechend Teilmengen der Potenzmenge $\mathcal{P}(G)$ von G.

6.2.3. Per definitionem sind die Rechtsnebenklassen von H in G genau die Bahnen der durch Multiplikation gegebenen Operation von H auf G. Insbesondere bilden sie also eine Partition von G. Analoges gilt für die Linksnebenklassen.

6.2.4. Ist G eine Gruppe und $H \subset G$ eine Untergruppe, so ist die Menge der Linksnebenklassen $X = G/H$ eine G-Menge in offensichtlicher Weise.

Beispiel 6.2.5. Im Fall $G = \mathbb{Z} \supset H = m\mathbb{Z}$ haben wir die Menge der Nebenklassen $\mathbb{Z}/m\mathbb{Z}$ bereits in 2.1.7 diskutiert und sogar selbst mit der Struktur einer Gruppe, ja sogar mit der Struktur eines Rings versehen. Im allgemeinen trägt G/H nur dann eine natürliche Gruppenstruktur, wenn wir an unsere Untergruppe H zusätzliche Forderungen stellen, vergleiche 6.4.

Satz 6.2.6 (**Lagrange**). *Gegeben eine endliche Gruppe teilt die Kardinalität jeder Untergruppe die Kardinalität der ganzen Gruppe. Ist G unsere endliche Gruppe und $H \subset G$ eine Untergruppe, so gilt genauer*

$$|G| = |H| \cdot |G/H| = |H| \cdot |H\backslash G|$$

Beweis. Jedes Element von G gehört zu genau einer Links- bzw. Rechtsnebenklasse unter H, und jede dieser Nebenklassen hat genau $|H|$ Elemente. □

Definition 6.2.7. Gegeben eine Gruppe G mit einer Untergruppe H heißt die Zahl $|G/H|$ der Restklassen der **Index** von H in G.

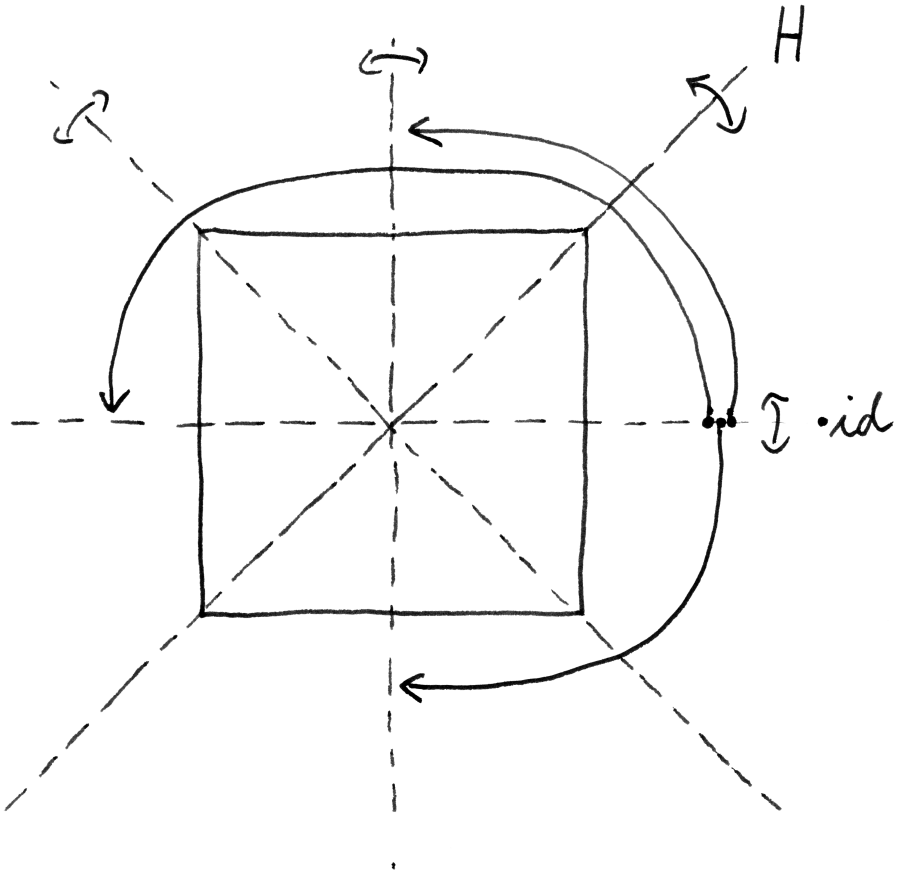
Übung 6.2.8. Seien $G \supset H$ eine Gruppe und eine Untergruppe. Man zeige, daß es eine Bijektion zwischen G/H und $H \setminus G$ gibt.

Übung 6.2.9. Haben zwei endliche Untergruppen einer Gruppe teilerfremde Kardinalitäten, so besteht ihr Schnitt nur aus dem neutralen Element.

6.3 Bahnformel

Lemma 6.3.1 (**Bahnen als Quotienten**). *Sei G eine Gruppe, X eine G-Menge und $x \in X$ ein Punkt. So induziert die Abbildung $G \to X$, $g \mapsto gx$ eine Bijektion*

$$G/G_x \xrightarrow{\sim} Gx$$



Die acht Symmetrien des Quadrats. Eine Rechtsnebenklasse der von der Spiegelung an der Nordost-Diagonalen erzeugten Untergruppe besteht aus den beiden Symmetrien des Quadrats, die die obere rechte Ecke in eine vorgegebene weitere Ecke überführen. Eine Linksnebenklasse besteht dahingegen aus den beiden Symmetrien des Quadrats, bei denen die obere rechte Ecke von einer vorgegebenen weiteren Ecke herkommt.

Beweis. Für jede G_x-Nebenklasse $N \subset G$ im Sinne von 6.2.2 besteht die Menge Nx nur aus einem Punkt, für $N = gG_x$ haben wir genauer $Nx = gG_xx = \{gx\}$. Die Abbildung im Lemma wird nun definiert durch die Bedingung, daß sie jeder Nebenklasse $N \in G/G_x$ das einzige Element von Nx zuordnet. Diese Abbildung ist offensichtlich surjektiv. Sie ist aber auch injektiv, denn aus $gG_xx = hG_xx$ folgt $gx = hx$, also $h^{-1}g \in G_x$, also $gG_x = hG_x$. □

6.3.2. Ist G eine endliche Gruppe und X eine G-Menge, so folgt mit dem vorhergehenden Lemma 6.3.1 aus 6.2.6 für alle $x \in X$ insbesondere die sogenannte **Bahnformel**

$$|G| = |G_x| \cdot |Gx|$$

Die Kardinalität jeder Bahn teilt also die Kardinalität der ganzen Gruppe, und die Kardinalität der Isotropiegruppen ist konstant auf den Bahnen. Genauer prüft man für beliebiges G die Formel $G_{gx} = gG_xg^{-1}$ für $g \in G$, $x \in X$.

Beispiel 6.3.3. Seien $k \leq n$ natürliche Zahlen. Auf der Menge X aller k-elementigen Teilmengen der Menge $\{1, 2, \ldots, n\}$ operiert die symmetrische Gruppe $\mathcal{S}_n$ transitiv. Die Isotropiegruppe des Punktes $x \in X$, der durch die k-elementige Teilmenge $\{1, 2, \ldots, k\}$ gegeben wird, ist isomorph zu $\mathcal{S}_k \times \mathcal{S}_{n-k}$. Die Bahnformel liefert folglich $|X| = n!/(k!(n-k)!)$ in Übereinstimmung mit unseren Erkenntnissen aus I.1.1.16. Ähnlich kann man auch die in I.2.2.26 diskutierten Formeln für die Multinomialkoeffizienten herleiten.

Beispiel 6.3.4. Wir können unsere Bahnformel auch umgekehrt anwenden, wenn wir zum Beispiel die Drehungen zählen wollen, die einen Würfel in sich überführen. Die Gruppe G dieser Drehungen operiert sicher transitiv auf der Menge E der acht Ecken des Würfels und die Isotropiegruppe jeder Ecke p hat drei Elemente. Wir folgern $|G| = |G_p| \cdot |E| = 3 \cdot 8 = 24$.

Übung 6.3.5. Sind Q, H Untergruppen einer Gruppe G, so induziert die Einbettung $Q \hookrightarrow G$ eine Bijektion $Q/Q \cap H \xrightarrow{\sim} QH/H$.

6.4 Normalteiler

Definition 6.4.1. Eine Untergruppe einer gegebenen Gruppe heißt ein **Normalteiler** genau dann, wenn die Rechtsnebenklassen unserer Untergruppe mit ihren Linksnebenklassen übereinstimmen. Ist G unsere Gruppe, so heißt also in Formeln eine Untergruppe $H \subset G$ ein Normalteiler genau dann, wenn gilt $gH = Hg \quad \forall g \in G$.

Beispiele 6.4.2. In einer kommutativen Gruppe ist jede Untergruppe ein Normalteiler. In der Gruppe $\mathcal{S}_3$ der Permutationen von 3 Elementen ist die Untergruppe $\mathcal{S}_2 \subset \mathcal{S}_3$ aller Permutationen, die die dritte Stelle festhalten, kein Normalteiler.

Übung 6.4.3. Der Kern eines Gruppenhomomorphismus ist stets ein Normalteiler. Allgemeiner ist das Urbild eines Normalteilers unter einem Gruppenhomomorphismus stets ein Normalteiler, und das Bild eines Normalteilers unter einem surjektiven Gruppenhomomorphismus ist wieder ein Normalteiler. Jede Untergruppe vom Index zwei ist ein Normalteiler.

Übung 6.4.4. Genau dann stimmen also für einen gegebenen homogenen Raum alle Isotropiegruppen überein, wenn er isomorph ist zum Quotienten der Gruppe nach einem Normalteiler. Wir sagen dann auch, der homogene Raum sei **normal**. Hinweis: 6.3.2.

Satz 6.4.5 (**Konstruktion der Restklassengruppe**). *Ist $H \subset G$ ein Normalteiler, so ist G/H abgeschlossen unter der induzierten Verknüpfung auf der Potenzmenge $\mathcal{P}(G)$ von G und wird damit eine Gruppe, genannt die* **Restklassengruppe** *oder der* **Quotient** *von G nach H.*

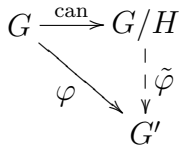
Beweis. Es gilt $(gH)(g_1H) = gg_1HH = gg_1H$, also ist unsere Menge stabil unter der Verknüpfung. Das Assoziativgesetz gilt eh, das neutrale Element ist H, und das Inverse zu gH ist $g^{-1}H$. □

Beispiel 6.4.6. Die Restklassengruppe $\mathbb{Z}/m\mathbb{Z}$ kennen wir bereits aus 2.1.7, wo wir darauf sogar noch eine Multiplikation erklärt hatten, die sie zu einem Ring macht. Sie hat genau m Elemente.

Satz 6.4.7 (**Universelle Eigenschaft der Restklassengruppe**). *Sei G eine Gruppe und $H \subset G$ ein Normalteiler.*

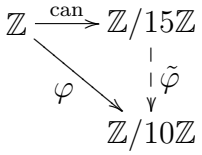
1. *Die Abbildung* $\mathrm{can} : G \to G/H$, $g \mapsto gH$ *ist ein Gruppenhomomorphismus mit Kern H.*
2. *Ist $\varphi : G \to G'$ ein Gruppenhomomorphismus mit $\varphi(H) = \{1\}$, so gibt es genau einen Gruppenhomomorphismus $\tilde{\varphi} : G/H \to G'$ mit $\varphi = \tilde{\varphi} \circ \mathrm{can}$.*

6.4.8. Der Übersichtlichkeit halber stelle ich die in diesem Satz auftauchenden Gruppen und Morphismen auch nocheinmal in einem Diagramm dar:



Man sagt auch, φ **faktorisiert über** die kanonische Abbildung can in den Quotienten.

Beispiel 6.4.9. Wir haben etwa



oder in Worten: Die Abbildung $\varphi = 2\,\mathrm{can} : \mathbb{Z} \to \mathbb{Z}/10\mathbb{Z}$, $n \mapsto (2n + 10\mathbb{Z})$ faktorisiert über $\mathbb{Z}/15\mathbb{Z}$ und induziert so einen Gruppenhomomorphismus $\mathbb{Z}/15\mathbb{Z} \to \mathbb{Z}/10\mathbb{Z}$.

Beweis. Die erste Aussage ist klar. Für die zweite Aussage beachten wir, daß unter der Annahme $\varphi(H) = \{1\}$ das Bild einer H-Nebenklasse $\varphi(gH) = \varphi(g)\varphi(H) = \{\varphi(g)\}$ nur aus einem einzigen Element besteht. Dies Element nennen wir $\tilde{\varphi}(gH)$, so daß also gilt $\tilde{\varphi}(gH) = \varphi(g)$ und $\varphi(gH) = \{\tilde{\varphi}(gH)\}$. Auf diese Weise erhalten wir das gesuchte $\tilde{\varphi}$. □

Satz 6.4.10 (**Isomorphiesatz**). *Jeder Homomorphismus $\varphi : G \to H$ von Gruppen induziert einen Isomorphismus $\tilde{\varphi} : G/\ker\varphi \xrightarrow{\sim} \operatorname{im}\varphi$.*

Beweis. Sicher ist unser $\tilde{\varphi}$ surjektiv. Es ist nach 2.2.12 aber auch injektiv, denn sein Kern besteht nur aus dem neutralen Element der Restklassengruppe. □

Korollar 6.4.11 (**Noether'scher Isomorphiesatz**). *Ist G eine Gruppe und sind $K \subset H \subset G$ zwei Normalteiler von G, so induziert die Komposition von kanonischen Abbildungen $G \twoheadrightarrow (G/K) \twoheadrightarrow (G/K)/(H/K)$ einen Isomorphismus*

$$G/H \xrightarrow{\sim} (G/K)/(H/K)$$

Beweis. Sicher ist unsere Komposition surjektiv. Unsere Aussage folgt also aus dem Isomorphiesatz 6.4.10, sobald wir zeigen, daß H der Kern unserer Komposition ist. Sicher ist H eine Teilmenge dieses Kerns. Liegt umgekehrt $g \in G$ im Kern unserer Komposition $G \twoheadrightarrow (G/K)/(H/K)$, so liegt die Nebenklasse gK im Kern von $(G/K) \twoheadrightarrow (G/K)/(H/K)$, als da heißt, es gibt $h \in H$ mit $gK = hK$, und daraus folgt sofort $g \in H$. □

6.4.12. Beim Arbeiten mit Restklassengruppen ermöglicht oft der Formalismus der "exakten Sequenzen" besonders transparente Formulierungen. Wir führen ihn deshalb im folgenden kurz ein.

Definition 6.4.13. 1. Eine Sequenz von Gruppen und Gruppenhomomorphismen $A' \xrightarrow{x} A \xrightarrow{y} A''$ heißt **exakt bei** A genau dann, wenn das Bild der ersten Abbildung zusammenfällt mit dem Kern der zweiten Abbildung, in Formeln $\operatorname{im} x = \ker y$.

2. Eine Sequenz von Gruppen $\ldots \to A_{i+1} \to A_i \to A_{i-1} \to \ldots$ heißt **exakt** genau dann, wenn sie exakt ist an jeder Stelle A_i.

Beispiel 6.4.14. Eine Sequenz der Gestalt $A \xrightarrow{r} B \xrightarrow{s} 1$ ist exakt genau dann, wenn r surjektiv ist. Eine Sequenz der Gestalt $1 \xrightarrow{r} B \xrightarrow{s} C$ ist exakt genau dann, wenn s injektiv ist.

Beispiel 6.4.15. Für jeden Homomorphismus $x : M \to N$ von abelschen Gruppen ist die Sequenz

$$1 \to (\ker x) \to M \to N \to (N/\operatorname{im} x) \to 1$$

exakt. Man nennt wegen dieser Symmetrie in dieser Situation den Quotienten nach dem Bild auch den **Kokern** unseres Morphismus von abelschen Gruppen und notiert ihn $\operatorname{cok} x := (N/\operatorname{im} x)$.

Übung 6.4.16. Gegeben ein kommutatives Diagramm von abelschen Gruppen mit exakten Zeilen

$$\begin{array}{ccccccc} M' & \longrightarrow & M & \longrightarrow & M'' & \longrightarrow & 0 \\ \downarrow{\scriptstyle a} & & \downarrow{\scriptstyle b} & & \vdots{\scriptstyle c} & & \downarrow \\ N' & \longrightarrow & N & \longrightarrow & N'' & \longrightarrow & 0 \end{array}$$

ohne den gestrichelten Pfeil existiert genau ein Gruppenhomomorphismus c wie durch den gestrichelten Pfeil angedeutet, der das mittlere Quadrat unseres Diagramms zum Kommutieren bringt. Sind a und b Isomorphismen, so auch c. Hinweis: Man beginne mit dem Fall $N' = M'$, $N = M$, $a = \mathrm{id}$, $b = \mathrm{id}$ und $M'' = \operatorname{cok} x$ für $x : M' \to M$.

Übung 6.4.17. Gegeben Sequenzen von Gruppen $A \xrightarrow{r} B \xrightarrow{s} C$ und $A' \xrightarrow{r'} B' \xrightarrow{s'} C'$ ist ihr **Produkt**

$$(A \times A') \xrightarrow{r \times r'} (B \times B') \xrightarrow{s \times s'} (C \times C')$$

exakt genau dann, wenn die beiden Ausgangssequenzen exakt sind. Analoges gilt sowohl für das Produkt als auch für die direkte Summe einer beliebigen Familie von Sequenzen. Diese Aussage enthält im Lichte von 6.4.16 insbesondere eine Präzisierung der Erwartung, daß das “Bilden von Produkten mit dem Bilden von Quotienten kommutieren sollte”.

6.5 Zyklische Gruppen

Definition 6.5.1. Eine Gruppe heißt **zyklisch** genau dann, wenn sie von einem einzigen Element erzeugt wird.

6.5.2. Zum Beispiel ist eine Gruppe G, deren Kardinalität eine Primzahl ist, notwendig zyklisch, da sie nach 6.2.6 außer $H = G$ und $H = 1$ keine weiteren Untergruppen haben kann. Für jede Gruppe G können wir die von einem Element $g \in G$ erzeugte Untergruppe beschreiben als

$$\langle g \rangle = \{g^n \mid n \in \mathbb{Z}\}$$

Definition 6.5.3. Sei g ein Element einer Gruppe G. Die **Ordnung**

$$\operatorname{ord} g$$

von g ist die kleinste natürliche Zahl $n \geq 1$ mit $g^n = 1_G$. Gibt es kein solches n, so setzen wir $\operatorname{ord} g = \infty$ und sagen, g habe **unendliche Ordnung**. Elemente der Ordnung 2 heißen auch **Involutionen**. Elemente, die ihre eigenen Inversen sind, nenne ich **selbstinvers**. Die Selbstinversen sind also genau die Involutionen mitsamt dem neutralen Element.

Lemma 6.5.4 (Struktur zyklischer Gruppen). *Ist G eine Gruppe und $g \in G$ ein Element, so stimmt die Ordnung von g überein mit der Kardinalität der von g erzeugten Untergruppe, in Formeln $\operatorname{ord} g = |\langle g \rangle|$. Genauer gilt:*

1. *Hat g unendliche Ordnung, so ist die Abbildung $\nu \mapsto g^\nu$ ein Isomorphismus $\mathbb{Z} \xrightarrow{\sim} \langle g \rangle$.*
2. *Hat g endliche Ordnung $\operatorname{ord} g = n$, so induziert $\nu \mapsto g^\nu$ einen Isomorphismus $\mathbb{Z}/n\mathbb{Z} \xrightarrow{\sim} \langle g \rangle$.*

Beweis. Wir betrachten den Gruppenhomomorphismus $\varphi : \mathbb{Z} \to G$, $\nu \mapsto g^\nu$. Nach 6.4.10 haben wir einen Isomorphismus $\mathbb{Z}/\ker\varphi \xrightarrow{\sim} \operatorname{im}\varphi = \langle g \rangle$. Nach 2.2.13 ist $\ker\varphi$ von der Form $\ker\varphi = n\mathbb{Z}$ für ein $n \in \mathbb{Z}$, $n \geq 0$, und dann gilt notwendig $n = \operatorname{ord} g$ für g von endlicher Ordnung bzw. $n = 0$ für g von unendlicher Ordnung. □

6.5.5. Motiviert durch dies Lemma nennt man die Kardinalität einer Gruppe auch oft die **Ordnung der Gruppe**. Wir haben mit unserem Lemma im Übrigen auch bewiesen, daß jede Gruppe mit genau 5 Elementen isomorph ist zu $\mathbb{Z}/5\mathbb{Z}$.

Korollar 6.5.6. *Bei einer endlichen Gruppe G teilt die Ordnung jedes Elements $g \in G$ die Ordnung der ganzen Gruppe, in Formeln*

$$g^{|G|} = 1$$

Beweis. Man wende den Satz von Lagrange 6.2.6 an auf die von unserem Element erzeugte Untergruppe. □

Korollar 6.5.7 (**Kleiner Fermat**). *Ist p eine Primzahl, so gilt für alle ganzen Zahlen $a \in \mathbb{Z}$ die Kongruenz*

$$a^p \equiv a \pmod{p}$$

Beweis. Die multiplikative Gruppe $(\mathbb{Z}/p\mathbb{Z})^\times$ des Körpers $\mathbb{Z}/p\mathbb{Z}$ hat genau $p-1$ Elemente, nach 6.5.6 gilt also $b^{p-1} = 1$ für alle $b \in (\mathbb{Z}/p\mathbb{Z})^\times$. Es folgt $b^p = b$ für alle $b \neq 0$, und für $b = 0$ gilt diese Gleichung eh. Mit $b = a + p\mathbb{Z}$ ergibt sich dann die Behauptung. □

Übung 6.5.8. Sei k ein endlicher Körper mit $|k| = q$ Elementen. Man zeige $a^q = a$ für alle $a \in k$. Man zeige weiter, daß der Kern unserer Surjektion $k[X_1, \ldots, X_n] \to \mathrm{Ens}(k^n, k)$ aus 2.3.29 genau aus denjenigen Polynomen besteht, die sich als Summe $P_1(X_1^q - X_1) + \ldots + P_n(X_n^q - X_n)$ der Produkte von irgendwelchen Polynomen $P_i \in k[X_1, \ldots, X_n]$ mit den Polynomen $(X_i^q - X_i)$ schreiben lassen. Hinweis: Unsere Summen von Produkten bilden einen Untervektorraum, zu dem der Untervektorraum aller Polynome, in denen kein X_i in der Potenz q oder höher vorkommt, komplementär ist.

Übung 6.5.9. Man zeige: Jede Untergruppe einer zyklischen Gruppe ist zyklisch. Genauer haben wir für beliebiges $m \in \mathbb{N}$ eine Bijektion

$$\begin{array}{ccc} \{\text{Teiler } d \in \mathbb{N} \text{ von } m\} & \xrightarrow{\sim} & \{\text{Untergruppen von } \mathbb{Z}/m\mathbb{Z}\} \\ d & \mapsto & d\mathbb{Z}/m\mathbb{Z} \end{array}$$

Man folgere, daß jede von der ganzen Gruppe verschiedene Untergruppe einer zyklischen Gruppe von Primzahlpotenzordnung $\mathbb{Z}/p^r\mathbb{Z}$ in der Untergruppe $p\mathbb{Z}/p^r\mathbb{Z} \subset \mathbb{Z}/p^r\mathbb{Z}$ enthalten sein muß. Hinweis: 2.2.13.

Übung 6.5.10. Man zeige, daß jede Untergruppe einer endlich erzeugten abelschen Gruppe endlich erzeugt ist, und daß man für die Untergruppe höchstens soviele Erzeuger benötigt wie für die ganze Gruppe. Hinweis: Induktion über die Zahl der Erzeuger. Als Basis mag man 6.5.9 nehmen. Dann bilde man geeignete Restklassengruppen. Vom höheren Standpunkt aus wird das in **??** nocheinmal bewiesen.

Übung 6.5.11. Sei m eine von Null verschiedene natürliche Zahl. Man zeige, daß die Vorschrift $\varphi \mapsto \varphi(\bar{1})$ für eine beliebige Gruppe G eine Bijektion liefert

$$\mathrm{Grp}(\mathbb{Z}/m\mathbb{Z}, G) \xrightarrow{\sim} \{g \in G \mid \text{Die Ordnung von } g \text{ ist endlich und teilt } m\}$$

Übung 6.5.12. Jede endlich erzeugte Untergruppe von $\mathbb{Q}$ ist zyklisch.

Übung 6.5.13. Man zeige, daß die additive Gruppe aller Gruppenhomomorphismen $\mathrm{Grp}(\mathbb{Z}/n\mathbb{Z}, \mathbb{Q}/\mathbb{Z})$ unter punktweiser Addition isomorph ist zu $\mathbb{Z}/n\mathbb{Z}$, für alle $n \geq 1$.

6.5.14. Gibt es natürliche Zahlen $n \in \mathbb{N}$, die

bei Division durch 6 Rest 4 lassen,

bei Division durch 13 Rest 2, und

bei Division durch 11 Rest 9?

Da $\langle 6, 13\rangle = \langle 13, 11\rangle = \langle 6, 11\rangle = \langle 1\rangle$ lautet die Antwort ja, wie man aus dem anschließenden Korollar 6.5.17 folgert.

Satz 6.5.15. *Ist $m = ab$ ein Produkt von zwei zueinander teilerfremden Faktoren, so liefert die offensichtliche Abbildung einen Isomorphismus*

$$\mathbb{Z}/m\mathbb{Z} \xrightarrow{\sim} \mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z}$$

6.5.16. Übung 6.6.13 zeigt, daß diese Gruppen im Fall nicht teilerfremder Faktoren auch nicht isomorph sind.

Beweis. Wir betrachten die Abbildung

$$\begin{array}{rccc} \varphi: & \mathbb{Z} & \to & \mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z} \\ & n & \mapsto & (n + a\mathbb{Z}, n + b\mathbb{Z}) \end{array}$$

Ihr Kern besteht aus allen $n \in \mathbb{Z}$, die durch a und b teilbar sind, also aus allen Vielfachen von m. Der Isomorphiesatz liefert mithin einen Isomorphismus $\mathbb{Z}/m\mathbb{Z} \xrightarrow{\sim} \operatorname{im}\varphi$, und daraus folgt hinwiederum $\operatorname{im}\varphi = \mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z}$, da unsere Untergruppe $\operatorname{im}\varphi$ selbst auch schon $m = ab$ Elemente hat. □

Korollar 6.5.17 (Chinesischer Restsatz). *Ist $m = q_1 \ldots q_s$ ein Produkt von paarweise teilerfremden ganzen Zahlen, so liefert die offensichtliche Abbildung einen Isomorphismus*

$$\mathbb{Z}/m\mathbb{Z} \xrightarrow{\sim} \mathbb{Z}/q_1\mathbb{Z} \times \ldots \times \mathbb{Z}/q_s\mathbb{Z}$$

Beweis. Übung. □

Übung 6.5.18. Man gebe alle Zahlen an, die bei bei Division durch 6 Rest 4 lassen, bei Division durch 13 Rest 2, und bei Division durch 11 Rest 9. Hinweis: Der euklidische Algorithmus liefert schon mal Lösungen, wenn ein Rest 1 ist und die anderen Null.

Übung 6.5.19. Gegeben x, y zwei Elemente endlicher Ordnung in einer kommutativen Gruppe teilt die Ordnung ihres Produkts das Produkt ihrer Ordnungen, in Formeln $\operatorname{ord}(xy) | (\operatorname{ord} x)(\operatorname{ord} y)$. Sind hier die Ordnungen von x und y teilerfremd, so gilt sogar $\operatorname{ord}(xy) = (\operatorname{ord} x)(\operatorname{ord} y)$. Hinweis: 6.5.15, 6.2.9, 2.2.5.

Übung 6.5.20. In jeder endlichen kommutativen Gruppe wird die maximal von einem Gruppenelement erreichte Ordnung geteilt von den Ordnungen aller Gruppenelemente. Hinweis: Bezeichnet $M \subset \mathbb{N}$ die Menge aller Ordnungen von Elementen unserer Gruppe, so enthält M mit jeder Zahl auch alle ihre Teiler. Weiter enthält M nach 6.5.19 mit je zwei teilerfremden Zahlen auch ihr Produkt.

Definition 6.5.21. Gegeben eine Gruppe G heißt die kleinste Zahl $e \geq 1$ mit $g^e = 1 \ \ \forall g \in G$ der **Exponent** unserer Gruppe. Gibt es kein solches e, so sagen wir, die Gruppe habe unendlichen Exponenten.

6.6 Endlich erzeugte abelsche Gruppen

6.6.1. Unter einer **Primzahlpotenz** verstehen wir im folgenden eine natürliche Zahl der Gestalt $q = p^r$ für p prim und $r \geq 1$. Gegeben eine Primzahl p verstehen wir unter einer **p-Potenz** dahingegen eine natürliche Zahl der Gestalt $q = p^r$ für p prim und $r \geq 0$. Man möge mir nachsehen, daß in dieser Terminologie nicht alle p-Potenzen Primzahlpotenzen sind. Die beiden folgenden Sätze geben zwei **Klassifikationen der endlich erzeugten abelschen Gruppen**.

Satz 6.6.2. *Gegeben eine endlich erzeugte abelsche Gruppe G gibt es genau eine endliche Folge von von 1 verschiedenen natürlichen Zahlen $a_1, \ldots, a_s \in \{0, 2, 3, \ldots\}$ mit $a_i | a_{i+1}$ für $1 \leq i < s$ derart, daß gilt*

$$G \cong \mathbb{Z}/a_1\mathbb{Z} \times \ldots \times \mathbb{Z}/a_s\mathbb{Z}$$

Satz 6.6.3. *Gegeben eine endlich erzeugte abelsche Gruppe G gibt es Primzahlpotenzen $q_1, \ldots, q_t$ und eine natürliche Zahl $r \in \mathbb{N}$ mit*

$$G \cong \mathbb{Z}/q_1\mathbb{Z} \times \ldots \times \mathbb{Z}/q_t\mathbb{Z} \times \mathbb{Z}^r$$

Die Zahl r wird durch G eindeutig festgelegt und heißt der **Rang** *von G. Die Primzahlpotenzen q_τ sind eindeutig bis auf Reihenfolge.*

Korollar 6.6.4. *Jede endliche abelsche Gruppe ist ein Produkt von zyklischen Gruppen von Primzahlpotenzordnung.*

6.6.5. Man beachte in beiden Sätzen, daß die Faktoren keineswegs eindeutig sind "als Untergruppen unserer abelschen Gruppe". Der Beweis der beiden Sätze wird uns bis zum Ende des Abschnitts beschäftigen. Eine erste wesentliche Zutat ist der gleich folgende Elementarteilersatz 6.6.8.

6.6.6. Ein Element endlicher Ordnung in einer Gruppe heißt ein **Torsionselement**. Eine Gruppe, in der alle Elemente außer dem neutralen Element unendliche Ordnung haben, heißt **torsionsfrei**. Zum Beispiel sind die abelschen Gruppen $\mathbb{Z}$, $\mathbb{Q}$ und $\mathbb{R}$ torsionsfrei. Jede endlich erzeugte torsionsfreie abelsche Gruppe ist nach unserer Klassifikation isomorph zu $\mathbb{Z}^r$ für geeignetes $r \in \mathbb{N}$.

6.6.7. Die Menge aller Torsionselemente ist in jeder abelschen Gruppe eine Untergruppe. Das Produkt aller endlichen Faktoren in jeder unserer beiden Darstellungen ist also eine wohldefinierte Untergruppe. Genauer ist sogar das Produkt aller Faktoren in der zweiten Zerlegung, deren Ordnungen Potenzen einer festen Primzahl p sind, eine wohlbestimmte Untergruppe unserer endlich erzeugten abelschen Gruppe, nämlich die Untergruppe aller Elemente von p-Potenzordnung.

Satz 6.6.8 (Elementarteilersatz). *1. Gegeben eine nicht notwendig quadratische Matrix A mit ganzzahligen Einträgen gibt es stets quadratische ganzzahlig invertierbare Matrizen mit ganzzahligen Einträgen P und Q derart, daß $B = PAQ$ eine Matrix mit Nullen außerhalb der Diagonalen ist, in der die Diagonaleinträge weiter vorn jeweils die Diagonaleinträge weiter hinten teilen, in Formeln $i \neq j \Rightarrow B_{i,j} = 0$ und $B_{i,i}|B_{i+1,i+1}\ \forall i$.*

2. Wir können durch geeignete Wahl von P und Q sogar zusätzlich erreichen, daß alle Diagonaleinträge nichtnegativ sind, und unter dieser Zusatzannahme werden besagte Diagonaleinträge durch die Matrix A bereits eindeutig festgelegt.

6.6.9. Den Beweis der analogen Aussage für Polynomringe dürfen Sie selbst als Übung 6.6.18 ausarbeiten, eine gemeinsame Verallgemeinerung für sogenannte "Hauptidealringe" wird in **??** dargestellt.

Beweis. Wir beginnen mit dem Nachweis der Existenz. Ist A die Nullmatrix, so ist nichts mehr zu zeigen. Sonst finden wir P, Q invertierbar derart, daß PAQ oben links einen positiven Eintrag hat, und zwar den kleinstmöglichen unter allen PAQ mit positivem Eintrag dort. Dann teilt dieser Eintrag notwendig alle anderen Einträge der ersten Spalte, da wir sonst durch Zeilenoperationen, genauer durch Subtraktion eines Vielfachen der ersten Zeile von einer anderen Zeile, Multiplikation einer Zeile mit -1 und Vertauschung zweier Zeilen einen noch kleineren positiven Eintrag oben links erzeugen könnten. Ebenso teilt unser Eintrag auch alle anderen Einträge in der ersten Zeile. Durch entsprechende Zeilen- und Spaltenoperationen können wir also zusätzlich die erste Zeile und Spalte bis auf den ersten Eintrag als genullt

annehmen. Teilt nun unser positiver Eintrag oben links nicht alle anderen Einträge unserer Matrix, sagen wir nicht $a_{i,j}$ für $i \neq 1 \neq j$, so könnten wir durch Addieren der ersten Zeile zur i-ten Zeile gefolgt von einer Subtraktion eines Vielfachen der ersten Spalte von von der j-ten Spalte einen noch kleineren positiven Eintrag an der Stelle (i, j) erzeugen, und ihn durch Zeilen- und Spaltenvertauschung in die linke obere Ecke bringen im Widerspruch zu unserer Annahme. Also teilt unser positiver Eintrag oben links alle anderen Einträge unserer Matrix und eine offensichtliche Induktion beendet den Beweis der Existenz. Um die Eindeutigkeit zu zeigen bemerken wir, daß sich für gegebenes $r \geq 1$ der größte gemeinsame Teiler G_r aller $(r \times r)$-Minoren unter Zeilen- und Spaltenoperationen nicht ändert. Folglich sind die $G_r = d_1 \ldots d_r$ wohlbestimmt durch A, und dasselbe gilt dann auch für die d_i. □

Beweis von 6.6.2. Wir notieren in diesem Beweis unsere abelsche Gruppe G additiv. Gegeben ein Erzeugendensystem $g_1, \ldots, g_n$ von G erklären wir in offensichtlicher Weise einen surjektiven Gruppenhomomorphismus

$$\mathbb{Z}^n \twoheadrightarrow G$$

mit $(a_1, \ldots, a_n) \mapsto a_1 g_1 + \ldots + a_n g_n$. Dessen Kern ist nach 6.5.10 eine endlich erzeugte abelsche Gruppe K, für die wir wieder einen surjektiven Gruppenhomomorphismus $\mathbb{Z}^m \twoheadrightarrow K$ finden. Mit der Komposition $\mathbb{Z}^m \twoheadrightarrow K \hookrightarrow \mathbb{Z}^n$ als erster Abbildung entsteht so eine exakte Sequenz von abelschen Gruppen

$$\mathbb{Z}^m \to \mathbb{Z}^n \to G \to 0$$

im Sinne von 6.4.13. Genau wie bei Vektorräumen überlegt man sich, daß die Gruppenhomomorphismen $\mathbb{Z}^m \to \mathbb{Z}^n$ genau die Multiplikationen von links mit ganzzahligen $(n \times m)$-Matrizen sind, falls Elemente aus $\mathbb{Z}^m$ bzw. $\mathbb{Z}^n$ als Spaltenvektoren aufgefasst werden. Weiter überlegt man sich, daß auch in dieser Situation die Verknüpfung von Homomorphismen der Multiplikation von Matrizen entspricht. Bezeichnet nun A die Matrix unserer Abbildung $\mathbb{Z}^m \to \mathbb{Z}^n$ und wählen wir P und Q wie im Elementarteilersatz, so ergibt sich ein kommutatives Diagramm von abelschen Gruppen

$$\begin{array}{ccc} \mathbb{Z}^m & \xrightarrow{A\circ} & \mathbb{Z}^n \\ {\scriptstyle Q\circ}\uparrow\wr & & \wr\downarrow{\scriptstyle P\circ} \\ \mathbb{Z}^m & \xrightarrow{D\circ} & \mathbb{Z}^n \end{array}$$

für eine nicht notwendig quadratische Diagonalmatrix D mit nichtnegativen Einträgen $d_1 | d_2 | \ldots | d_r$ für $r = \min(m, n)$. Bilden wir nun andererseits das Produkt der exakten Sequenzen $\mathbb{Z} \xrightarrow{d_i} \mathbb{Z} \to \mathbb{Z}/d_i\mathbb{Z} \to 0$ mit $m - r$ Kopien

der exakten Sequenzen $\mathbb{Z} \to 0 \to 0 \to 0$ im Fall $m > n$ bzw. $n - r$ Kopien der exakten Sequenzen $0 \to \mathbb{Z} \xrightarrow{\mathrm{id}} \mathbb{Z} \to 0$ im Fall $n > m$, so erhalten wir mit 6.4.17 die untere Horizontale in einem kommutativen Diagramm mit exakten Zeilen

$$\begin{array}{ccccccc}
\mathbb{Z}^m & \xrightarrow{A\circ} & \mathbb{Z}^n & \longrightarrow & G & \longrightarrow & 0 \\
{\scriptstyle Q\circ}\uparrow\wr & & {\scriptstyle P\circ}\downarrow\wr & & & & \downarrow \\
\mathbb{Z}^m & \xrightarrow{D\circ} & \mathbb{Z}^n & \longrightarrow & (\mathbb{Z}/d_1\mathbb{Z} \times \ldots \times \mathbb{Z}/d_r\mathbb{Z} \times \mathbb{Z}^{n-r}) & \longrightarrow & 0
\end{array}$$

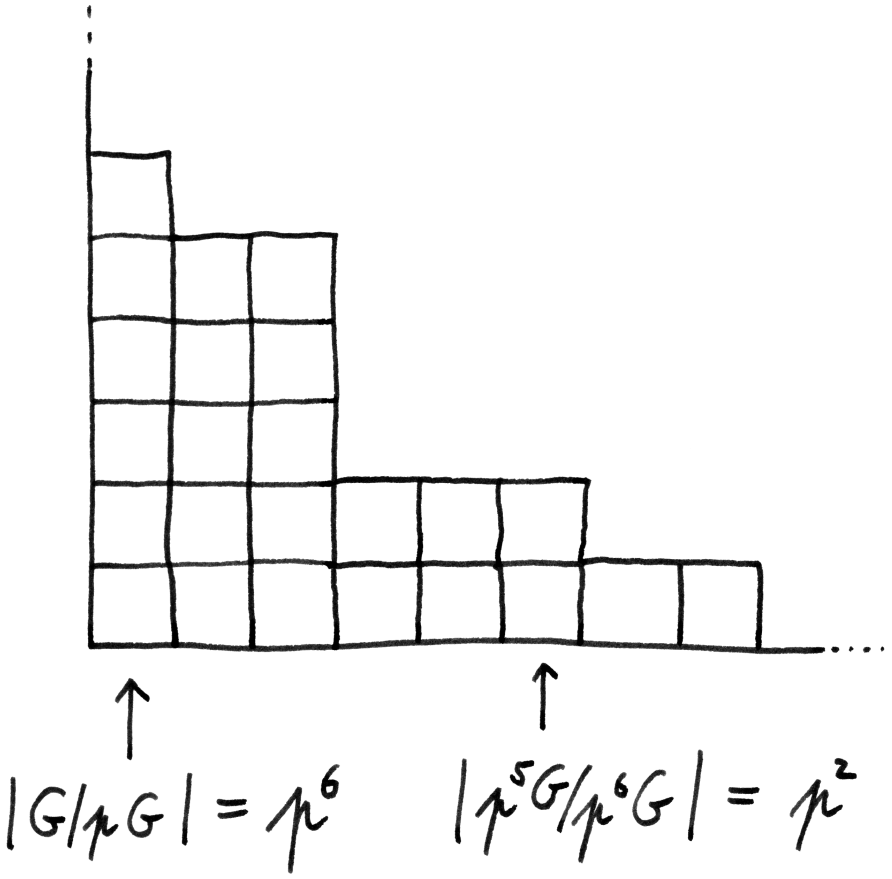
6.4.16 liefert damit einen Isomorphismus $G \cong \mathbb{Z}/d_1\mathbb{Z} \times \ldots \times \mathbb{Z}/d_r\mathbb{Z} \times \mathbb{Z}^{n-r}$. Lassen wir von unserer Folge $d_1|d_2|\ldots|d_r$ alle Einsen vorne weg und ergänzen am Ende $(n-r)$ Nullen, so erhalten wir eine Folge $a_1|\ldots|a_s$ wie im Satz 6.6.2 gefordert, und die Existenz ist gezeigt. Um die Eindeutigkeit zu zeigen bemerken wir, daß für jede endlich erzeugte abelsche Gruppe G und jede Primzahl p und alle $n \in \mathbb{N}$ der Quotient $p^nG/p^{n+1}G$ nach 2.2.32 und 6.5.10 in eindeutiger Weise ein endlichdimensionaler Vektorraum über $\mathbb{F}_p$ ist. Wir notieren seine Dimension

$$D_p^n(G) \in \mathbb{N}$$

Alternativ mag man $D_p^n(G)$ auch als die eindeutig bestimmte natürliche Zahl $D \in \mathbb{N}$ mit $|p^nG/p^{n+1}G| = p^D$ charakterisieren. Aus 6.4.17 folgert man unmittelbar $D_p^n(G \times H) = D_p^n(G) + D_p^n(H)$ für je zwei endlich erzeugte abelsche Gruppen G und H. Für zyklische Gruppen $G \cong \mathbb{Z}/a\mathbb{Z}$ behaupten wir weiter

$$D_p^n(\mathbb{Z}/a\mathbb{Z}) \quad = \quad \begin{cases} 1 & p^{n+1}|a; \\ 0 & \text{sonst.} \end{cases}$$

In der Tat ist das klar für $a = p^m$, mit 2.2.37 erkennen wir es für a teilerfremd zu p, und mit 6.5.15 folgt es im allgemeinen. Gegeben eine endliche abelsche Gruppe G und eine Primzahl p bilden wir nun ein Youngdiagramm $Y_p(G)$ im Sinne von 5.5.3, indem wir jeweils $D_p^n(G)$ Kästchen in der $(n+1)$-ten Spalte übereinander türmen. Ist G nun isomorph zu $\mathbb{Z}/a_1\mathbb{Z} \times \ldots \times \mathbb{Z}/a_s\mathbb{Z}$ mit $a_1 \geq 2$ und $a_1|a_2|\ldots|a_s$, so sind die Zeilenlängen unseres Diagramms von unten angefangen das größte z mit p^z teilt a_s, das größte z mit p^z teilt a_{s-1}, und so weiter. Wenden wir diese Erkenntnis an auf alle Primzahlen p, so folgt bereits die im Satz behauptete Eindeutigkeit für endliche abelsche Gruppen. Ist unsere abelsche Gruppe nur endlich erzeugt, so können wir das vorstehende Argument dahingehend modifizieren, daß wir unseren Diagrammen auch unendliche Zeilen erlauben. Dann ist eben die Länge der untersten Zeile das Supremum über alle $\nu \in \mathbb{N}$ mit $p^\nu|a_s$ und so weiter, als da heißt, es gibt für eine und jede Primzahl ebensoviele unendliche Zeilen, wie Nullen in der Kette $a_1|a_2|\ldots|a_s$, wie Faktoren $\mathbb{Z}$. □



Das Youngdiagramm $Y_p(G)$ der abelschen p-Gruppe
$G = (\mathbb{Z}/p\mathbb{Z}) \times (\mathbb{Z}/p^3\mathbb{Z})^3 \times (\mathbb{Z}/p^6\mathbb{Z}) \times (\mathbb{Z}/p^8\mathbb{Z})$.

Beweis von 6.6.3. Die Existenz folgt aus 6.6.2 mit dem Chinesischen Restsatz 6.5.17. Die Eindeutigkeit erkennt man, indem man sich überlegt, daß verschiedene Folgen $a_1|a_2|\ldots|a_s$ auch zu verschiedenen Produkten wie in 6.6.3 führen. Genauer kann man a_1 beschreiben als das Produkt der jeweils höchsten Primzahlpotenzen für alle vorkommenden Primzahlen, a_2 als das Produkt der jeweils zweithöchsten und so weiter, bis am Ende die Zahl der Nullen gerade die Zahl der Faktoren $\mathbb{Z}$ in der Zerlegung 6.6.3 sein muß. □

Übung 6.6.10. Sind $a, b \in \mathbb{Z}$ teilerfremd, in Formeln $\langle a, b\rangle = \langle 1\rangle$, so läßt sich das Element $(a, b) \in \mathbb{Z}^2$ ergänzen zu einem Erzeugendensystem von $\mathbb{Z}^2$. Man formuliere und zeige auch die analoge Aussage für $\mathbb{Z}^n$.

Übung 6.6.11. Der Rang ist einer endlich erzeugten abelschen Gruppe kann beschrieben werden als die Dimension des $\mathbb{Q}$-Vektorraums $\mathrm{Grp}(G, \mathbb{Q})$ aller Gruppenhomomorphismen von G nach $\mathbb{Q}$.

Übung 6.6.12. Man gebe ein dreielementiges bezüglich Inklusion minimales Erzeugendensystem der Gruppe $\mathbb{Z}$ an.

Übung 6.6.13. Gegeben $a, b \in \mathbb{N}_{\geq 1}$ gibt es einen Gruppenisomorphismus $\mathbb{Z}/ab\mathbb{Z} \cong \mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z}$ genau dann, wenn a und b teilerfremd sind.

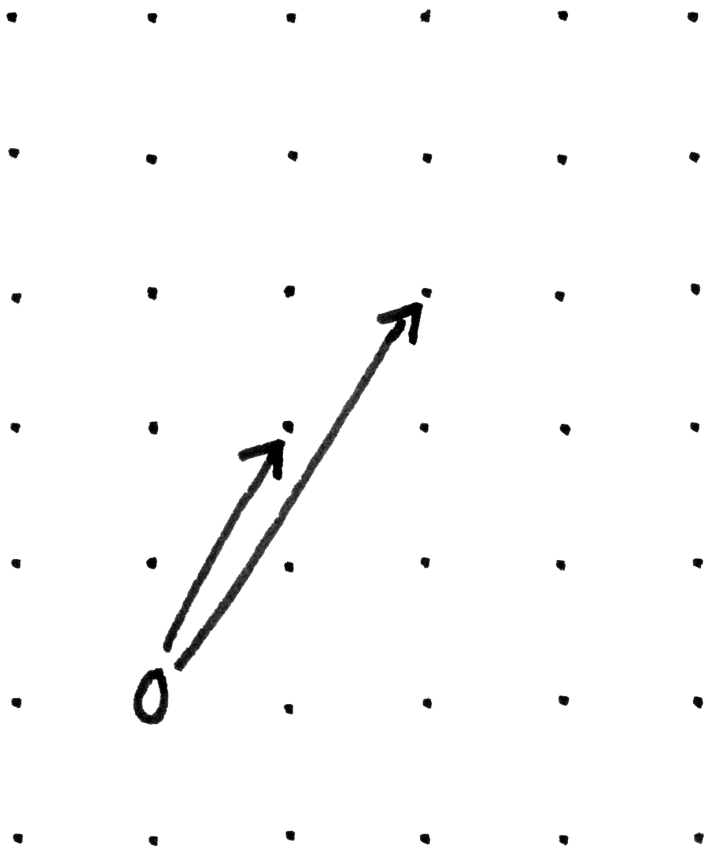
Übung 6.6.14. Man zeige, daß es für jede zyklische Gruppe G gerader Ordnung genau ein Element der Ordnung zwei und genau einen surjektiven Gruppenhomomorphismus in die zweielementige Gruppe gibt.

Übung 6.6.15. Man berechne die Elementarteiler der Matrix

$$\begin{pmatrix} 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 \\ 5 & 5 & 5 & 5 \end{pmatrix}$$

Beispiel 6.6.16. Gegeben eine abelsche Gruppe M bilden die Elemente endlicher Ordnung stets eine Untergruppe $M_{\mathrm{tor}} \subset M$ und der Quotient M/M_{tor} ist torsionsfrei. Allerdings gibt es im Gegensatz zum Fall endlich erzeugter abelscher Gruppen im allgemeinen keinen Gruppenisomorphismus zwischen M und $M_{\mathrm{tor}} \times (M/M_{\mathrm{tor}})$. Betrachten wir etwa in der Gruppe $M = \prod_{n=0}^{\infty} \mathbb{Z}/p^n\mathbb{Z}$ das Element $v = (\overline{p^0}, 0, \overline{p^1}, 0, \overline{p^2}, 0, \ldots)$, so ist $\overline{v} \in M/M_{\mathrm{tor}}$ nicht Null und für alle $i \geq 0$ gibt es $w = w_i \in M/M_{\mathrm{tor}}$ mit $p^i w = v$. Das einzige Element von M, das in dieser Weise "durch alle p-Potenzen teilbar ist", ist jedoch die Null, folglich existiert kein Gruppenisomorphismus zwischen M und $M_{\mathrm{tor}} \times (M/M_{\mathrm{tor}})$. Dies Beispiel ist im übrigen eine Variation von 5.3.18.

Übung 6.6.17. Man finde ein Repräsentantensystem für die Bahnen unter der offensichtlichen Wirkung von $\mathrm{GL}(n;\mathbb{Z}) \times \mathrm{GL}(m;\mathbb{Z})$ auf $\mathrm{M}(n \times m;\mathbb{Q})$. Hinweis: 6.6.8.



Ein Erzeugendensystem von $\mathbb{Z}^2$

Übung 6.6.18 (**Smith-Zerlegung**). Gegeben eine nicht notwendig quadratische Matrix A mit Einträgen in einem Polynomring $k[X]$ zeige man: (1) Es gibt quadratische im Matrizenring über $k[X]$ invertierbare Matrizen mit polynomialen Einträgen P und Q derart, daß $B = PAQ$ eine Matrix mit Nullen außerhalb der Diagonalen ist, in der die Diagonaleinträge weiter vorn jeweils die Diagonaleinträge weiter hinten teilen, in Formeln $i \neq j \Rightarrow B_{i,j} = 0$ und $B_{i,i}|B_{i+1,i+1}\ \forall i$; (2) Wir können durch geeignete Wahl von P und Q sogar zusätzlich erreichen, daß alle von Null verschiedenen Diagonaleinträge normiert sind, und unter dieser Zusatzannahme werden besagte Diagonaleinträge durch die Matrix A bereits eindeutig festgelegt. Hinweis: Vielleicht wäre es eine gute Idee, gleich hier den Zusammenhang mit der Jordan'schen Normalform aufzuzeigen. Eigentlich sollte diese Übung die Polynomdivision mit Rest abwarten.

6.7 Konjugationsklassen

Definition 6.7.1. Ist G eine Gruppe und $x \in G$ ein Element, so ist die Abbildung

$$\begin{array}{rccc} (\operatorname{int} x): & G & \to & G \\ & g & \mapsto & xgx^{-1} \end{array}$$

ein Isomorphismus der Gruppe G mit sich selber. Er heißt die **Konjugation mit** x. Ganz allgemein nennt man einen Isomorphismus einer Gruppe mit sich selber auch einen **Automorphismus** der Gruppe. Die Automorphismen einer Gruppe G bilden selber eine Gruppe mit der Verknüpfung von Abbildungen als Verknüpfung. Sie heißt die **Automorphismengruppe** von G und wir notieren sie $\mathrm{Grp}^\times(G)$. Diejenigen Automorphismen einer Gruppe, die sich als Konjugation mit einem geeigneten Gruppenelement schreiben lassen, heißen **innere Automorphismen** und auf englisch **interior automorphisms**, daher die Notation int. Sicher gilt $(\operatorname{int} x) \circ (\operatorname{int} y) = \operatorname{int}(xy)$, folglich ist $x \mapsto \operatorname{int} x$ ein Gruppenhomomorphismus $\operatorname{int}: G \to \mathrm{Grp}^\times(G)$ und insbesondere eine Operation der Gruppe G auf der Menge G, die **Operation durch Konjugation**

$$\begin{array}{rcl} G \times G & \to & G \\ (x, y) & \mapsto & (\operatorname{int} x)(y) = xyx^{-1} \end{array}$$

Die Bahnen unter dieser Operation heißen die **Konjugationsklassen** unserer Gruppe.

Beispiele 6.7.2. Die Konjugationsklassen in einer kommutativen Gruppe sind einelementig. Die Theorie der Jordan'schen Normalform beschreibt die Konjugationsklassen in $\mathrm{GL}(n;\mathbb{C})$, vergleichen 6.1.18.

Übung 6.7.3. Sei A eine zyklische Gruppe der Ordnung $n \in \mathbb{N}$. So gibt es genau einen Ringisomorphismus $\mathbb{Z}/n\mathbb{Z} \xrightarrow{\sim} \operatorname{End} A$, und dieser Ringisomorphismus induziert einen Isomorphismus zwischen der Einheitengruppe $(\mathbb{Z}/n\mathbb{Z})^\times$ und der Automorphismengruppe von A.

Übung 6.7.4. Man gebe jeweils ein Repräsentantensystem an für die Konjugationsklassen der Gruppe der Isometrien der affinen euklidischen Ebene $\mathbb{R}^2$ und der Untergruppe ihrer orientierungserhaltenden Isometrien. Hinweis: 3.4.7.

6.8 Endliche Untergruppen der Drehgruppe

Definition 6.8.1. Sei $A \subset \mathbb{R}^3$ eine Teilmenge. Unter einer **Symmetrie** von A verstehen wir ein Element der orthogonalen Gruppe $g \in \mathrm{O}(3)$ mit $gA = A$. Unter einer **Drehsymmetrie** von A verstehen wir eine Drehung $g \in \mathrm{SO}(3)$ mit $gA = A$.

Satz 6.8.2 (Klassifikation der endlichen Drehgruppen). *Jede endliche Untergruppe der Gruppe* $\mathrm{SO}(3)$ *aller Drehungen des Raums ist genau eine der folgenden Gruppen:*

1. *Eine* **zyklische Gruppe** C_k *mit* $k \geq 1$ *Elementen, bestehend aus allen Drehungen zu einer festen Drehachse um Winkel der Gestalt* $2\pi n/k$. *Der Fall* $k = 1$ *deckt hier den Fall der trivialen Gruppe ab, die nur aus der Identität besteht.*

2. *Eine* **Diedergruppe** D_k *mit* $2k$ *Elementen für* $k \geq 2$. *Im Fall* $k > 2$ *ist das die Gruppe aller Drehsymmetrien eines ebenen gleichseitigen* k*-Ecks, aufgefaßt als räumliche Figur. Im Fall* $k = 2$ *ist es die Gruppe aller derjenigen Drehungen, die von einem Paar orthogonaler Geraden jede in sich überführen.*

3. *Eine* **Tetraedergruppe** T *aller 12 Drehsymmetrien eines Tetraeders.*

4. *Eine* **Würfelgruppe** W *aller 24 Drehsymmetrien eines Würfels.*

5. *Eine* **Ikosaedergruppe** I *aller 60 Drehsymmetrien eines Ikosaeders.*

6.8.3. Will man diesen Satz einem Laien erklären, der mit dem Gruppenbegriff nicht vertraut ist, so mag man nach 2.2.4 auch einfacher von endlichen Mengen von Drehungen reden, die mit je zwei Drehungen stets auch deren Hintereinanderausführung enthalten. Vom mathematischen Standpunkt aus mag man das Resultat als eine Klassifikation aller Konjugationsklassen von endlichen Untergruppen der Drehgruppe ansehen.

6.8.4. Das Evozieren der platonischen Körper stellt insofern einen Stilbruch dar, als wir uns zumindest implizit darauf verständigt hatten, alle unsere Überlegungen ausschließlich im Rahmen der Mengenlehre durchzuführen. Ein möglicher **Würfel** ist schnell beschrieben, man mag als Ecken die acht Vektoren $(\pm 1, \pm 1, \pm 1)$ nehmen. Die Ecken eines **Tetraeders** erhält man, wenn man nur die vier Ecken dieses Würfels nimmt, bei denen das Produkt der Koordinaten Eins ist. Den **Ikosaeder** besprechen wir in 6.8.13 noch ausführlich. Zu den fünf sogenannten "platonischen Körpern" rechnet man außer diesen dreien noch den **Oktaeder** und den **Dodekaeder**. Die Eckenmenge eines Oktaeders bilden etwa die drei Vektoren der Standardbasis des $\mathbb{R}^3$ mitsamt ihren Negativen. Die Eckenmenge eines Dodekaeders mag man anschaulich als die Menge der "Flächenmitten eines Ikosaeders" beschreiben und formal als die Menge der "Pole der Polordnung drei" im Sinne des gleich folgenden Beweises im Fall der Symmetriegruppe eines Ikosaeders. Die Bezeichnungen Tetraeder, Oktaeder, Dodekaeder und Ikosaeder für die platonischen Körper außer dem Würfel kommen von den griechischen Worten für die Anzahlen 4, 8, 12 und 20 ihrer Flächen her. Man findet für den Würfel wegen seiner 6 Flächen manchmal auch die Bezeichnung "Hexaeder".

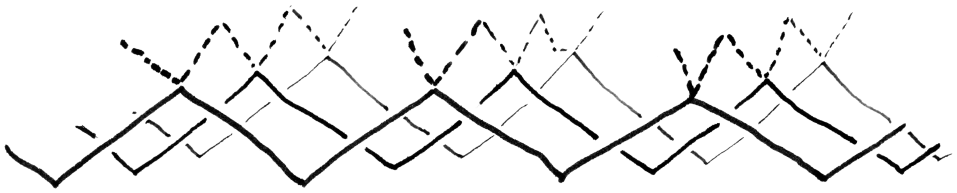
6.8.5. Unser Satz 6.8.2 ist ein möglicher Ausgangspunkt der Kristallographie: Unter einem **n-dimensionalen Kristall** verstehen wir hier eine Teilmenge K eines n-dimensionalen affinen reellen euklidischen Raums E, etwa die Menge der Orte der Atome eines Kristallgitters, mit der Eigenschaft, daß (1) die Translationen aus ihrer Symmetriegruppe den Richtungsraum aufspannen und daß es (2) eine positive untere Schranke gibt für die Längen aller von Null verschiedenen Translationen aus besagter Symmetriegruppe. Die zweite Eigenschaft schließt etwa den Fall aus, daß unsere Teilmenge einfach der ganze besagte euklidische Raum ist. Unter der **Punktgruppe** P eines Kristalls verstehen wir die Untergruppe $P \subset \mathrm{O}(\vec{E})$ aller linearen Anteile von Symmetrien unseres Kristalls, unter seiner **Drehgruppe** $D \subset \mathrm{SO}(\vec{E})$ die Menge aller orientierungserhaltenden Elemente der Punktgruppe. Man zeigt, daß die Punktgruppe eines Kristalls stets endlich sein muß, und daß als Drehgruppen von räumlichen, als da heißt dreidimensionalen Kristallen nur die Gruppen C_k und D_k mit $k \in \{1, 2, 3, 4, 6\}$ sowie die Tetraedergruppe und die Würfelgruppe auftreten können. Die Einteilung nach Drehgruppen entspricht in etwa, aber leider nicht ganz, der in der Kristallographie gebräuchlichen Einteilung in die sieben **Kristallsysteme**. Genauer entsprechen dem "kubischen System" die Würfelgruppe und die Tetraedergruppe, dem "tetragonalen System" die Drehgruppen C_4 und D_4, dem "hexagonalen System" die Drehgruppen C_6 und D_6, dem "trigonalen System" die Drehgruppen C_3 und D_3, aber das "orthorhombische", "monokline" und "trikline

System" lassen sich erst anhand ihrer Punktgruppen unterscheiden. Auch in den übrigen Fällen liefert die Punktgruppe eine feinere Klassifikation, für sie gibt es 32 Möglichkeiten, nach denen die Kristalle in die sogenannten **Kristallklassen** eingeteilt werden. Die eigentliche Klassifikation beschreibt alle als Symmetriegruppen von räumlichen Kristallen möglichen Bewegungsgruppen des Anschauungsraums bis auf Konjugation mit affinen, nicht notwendig euklidischen Automorphismen. Es gibt hierfür 230 Möglichkeiten. Das **achtzehnte Hilbert'sche Problem** fragte unter anderem danach, ob es analog in jeder Dimension nur endlich viele Möglichkeiten für wesentlich verschiedene Kristalle gibt. Bieberbach konnte dafür einen Beweis geben.

Übung 6.8.6 (**Das Kristallgitter des Diamants**). Seien $v_1, \ldots, v_4$ Richtungsvektoren des dreidimensionalen Anschauungsraums, die vom Schwerpunkt eines Tetraeders zu seinen vier Ecken zeigen. Wir betrachten alle Linearkombinationen $\sum_{i=1}^4 n_i v_i$ mit $\sum_{i=1}^4 n_i \in \{0, 1\}$ und behaupten, daß diese Linearkombinationen gerade die Punkte beschreiben, an denen in einem Diamant die Kohlenstoff-Atome sitzen. In der Tat sind unsere Linearkombinationen paarweise verschieden, die einzige Relation $v_1 + v_2 + v_3 + v_4 = 0$ unserer Vektoren führt aufgrund unserer Einschränkungen nicht zu Mehrdeutigkeiten, und sie lassen sich auch beschreiben als die Elemente des von den Richtungsvektoren $v_1 - v_2, v_1 - v_3$ und $v_1 - v_4$ erzeugten Gitters mitsamt dem um v_1 verschobenen Gitter. Jeder Punkt hat vier nächste Nachbarn, der Nullpunkt etwa $v_1, \ldots, v_4$, und zu diesen ist er gebunden im Diamantkristall. Anschaulich mag man sich eine Lage von parallelen horizontalen Zick-Zack-Linien denken, die Zick-Zacks darin nach oben und unten, dann eine weitere horizontale Lage senkrecht dazu, bei denen die Tiefpunkte immer gerade die Hochpunkte der Lage darunter berühren, und so weiter, und schließlich an jedem dieser Berührungspunkte ein Kohlenstoffatom.

6.8.7. Eine Würfelgruppe kann auch als die Gruppe aller Drehsymmetrien desjenigen Oktaeders aufgefaßt werden, dessen Ecken die Mittelpunkte der Flächen des Würfels sind. Ähnlich kann eine Ikosaedergruppe auch als Gruppe aller Drehsymmetrien eines Dodekaeders aufgefaßt werden. Die Kantenmitten eines Tetraeders bilden die Ecken eines Oktaeders, so erhält man eine Einbettung der Tetraedergruppe in die Würfelgruppe.

6.8.8. Die Diedergruppe D_2 ist offensichtlich isomorph zur Klein'schen Vierergruppe $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. Sie kann vielleicht übersichtlicher auch beschrieben werden als die Gruppe aller Drehungen, die von einem Tripel paarweise orthogonaler Geraden jede in sich überführen. Neben der Identität liegen darin also die Drehungen um 180° um jede dieser drei Geraden. Die Tetraedergruppe kann man in die symmetrische Gruppe $\mathcal{S}_4$ einbetten vermittels ihrer Operation auf den Ecken des Tetraeders. Wir erhalten so einen Isomorphismus der



Versuch einer graphischen Darstellung der räumlichen Struktur des Diamantkristalls. Die durchgezogenen und gestrichelten Linien sind nur der Transparenz halber verschiedenartig gezeichnet und bedeuten die Bindungen zwischen den Kohlenstoffatomen, die jeweils an den Ecken der Zick-Zack-Linien sitzen. Die hier gezeichnete Struktur gilt es nun übereinanderzuschichten, so daß sich jeweils die Ecken treffen.

Tetraedergruppe mit der alternierenden Gruppe A_4 aller geraden Permutationen von vier Elementen. Die Würfelgruppe operiert auf der Menge der vier räumlichen Diagonalen des Würfels und wir erhalten so einen Isomorphismus $W \cong \mathcal{S}_4$. Die Ikosaedergruppe operiert auf der Menge der fünf eingeschriebenen Würfel eines Dodekaeders, von denen einer in nebenstehendem Bild schematisch dargestellt ist. Mit etwas Geduld kann man direkt einsehen, daß diese Operation einen Isomorphismus der Ikosaedergruppe I mit der alternierenden Gruppe A_5 aller geraden Permutationen von 5 Elementen liefert. In **??** werden wir erklären, wie man das auch mit weniger Geduld aber mehr Gruppentheorie einsehen kann, und in **??** werden wir zusätzlich einen Isomorphismus dieser Gruppe mit der Gruppe $\mathrm{SL}(2;\mathbb{F}_5)/\{\pm \mathrm{id}\}$ herleiten.

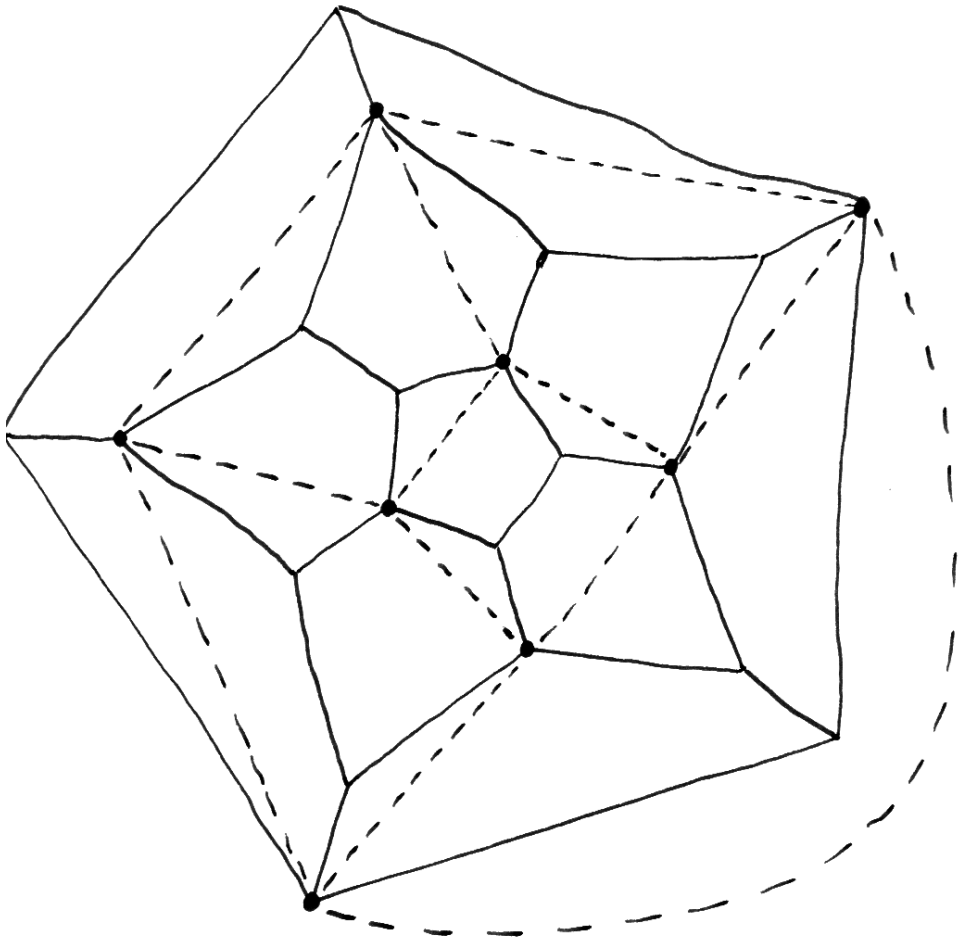
Übung 6.8.9. Man konstruiere einen surjektiven Gruppenhomomorphismus $\mathcal{S}_4 \twoheadrightarrow \mathcal{S}_3$. Hinweis: Geometrisch mag man sich die $\mathcal{S}_4$ nach 6.8.8 als die Gruppe der Drehsymmetrien eines Würfels denken und den fraglichen Gruppenhomomorphismus konstruieren vermittels der Operation dieser Gruppe auf der Menge der drei Mittelsenkrechten auf den Flächen des Würfels.

Beweis von Satz 6.8.2. Sei $G \subset \mathrm{SO}(3)$ eine endliche Untergruppe. Für jedes vom neutralen Element verschiedene Element $g \in G\backslash 1$ unserer Gruppe definieren wir seine "Pole" als die beiden Schnittpunkte seiner Drehachse mit der Einheitssphäre. Sei P die Menge aller Pole von Elementen aus $G\backslash 1$. Natürlich ist P eine endliche Menge und G operiert auf P. Wir zählen nun die Menge M aller Paare (g,p) mit $g \in G\backslash 1$ und p einem Pol von g auf zwei Weisen: Einmal gehört jedes von 1 verschiedene Gruppenelement $g \in G\backslash 1$ zu genau zwei Polen, also haben wir $|M| = 2(|G|-1)$. Andererseits gehört jeder Pol $p \in P$ zu genau $|G_p|-1$ von 1 verschiedenen Gruppenelementen, also haben wir $|M| = \sum_{p\in P}(|G_p|-1)$. Zusammen erhalten wir

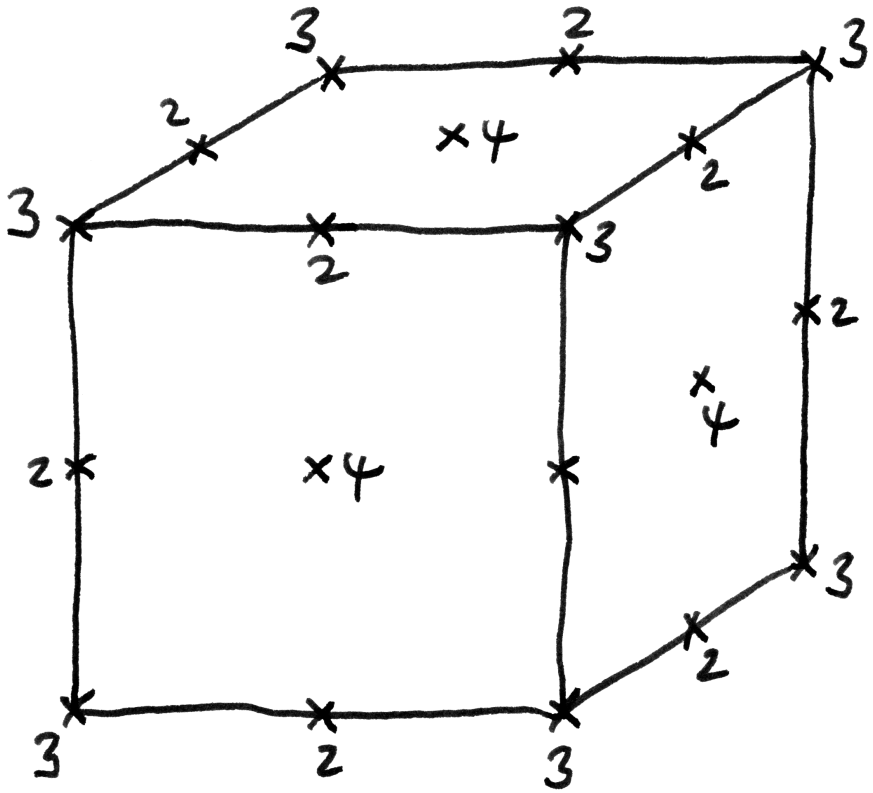
$$2(|G|-1) = \sum_{p\in P}(|G_p|-1)$$

Sei nun $P = P_1 \cup \ldots \cup P_r$ die Bahnzerlegung von P und seien $p_i \in P_i$ fest gewählt. Die Isotropiegruppe von p_i habe sagen wir $n_i \geq 2$ Elemente. Die zugehörige Bahn hat dann $|P_i| = |G|/n_i$ Elemente und alle Isotropiegruppen zu Polen aus P_i haben n_i Elemente. Die Kardinalität der Isotropiegruppe eines Pols nennen wir auch kürzer die **Polordnung**. In dieser Terminologie ist also n_i die Polordnung des Pols p_i. Fassen wir also die Pole jeder Bahn in unserer Summe zu einem Summanden zusammen, so können wir in unserer Gleichung die rechte Seite umformen zu $\sum_{i=1}^{r}(|G|/n_i)(n_i-1)$ und es ergibt sich die Gleichung

$$2 - \frac{2}{|G|} = \sum_{i=1}^{r}\left(1 - \frac{1}{n_i}\right)$$



Einer der fünf eingeschriebenen Würfel eines Dodekaeders, mit gestrichelt eingezeichneten Kanten.



Die “von vorne sichtbaren” Pole der Würfelgruppe mit den Kardinalitäten der jeweiligen Isotropiegruppen

Jeder Summand auf der rechten Seite ist mindestens 1/2, der Ausdruck links ist aber kleiner als 2. Es kommen also nur bis zu drei Bahnen von Polen in Betracht. Wir machen nun eine Fallunterscheidung nach der Zahl r der Bahnen von Polen.

Fall 0: Es gibt überhaupt keine Pole. In diesem Fall besteht G nur aus dem neutralen Element, und wir haben die Gruppe C_1 vor uns.

Fall 1: Ganz P ist eine Bahn. Das ist unmöglich, denn es muß gelten $|G| \geq 2$ wenn es überhaupt Pole geben soll, und damit hätten wir $2 - \frac{2}{|G|} \geq 1 > 1 - \frac{1}{n_1}$ im Widerspruch zu unserer Gleichung.

Fall 2: Es gibt genau zwei Bahnen P_1 und P_2 in P. Wir haben dann

$$\frac{2}{|G|} = \frac{1}{n_1} + \frac{1}{n_2}$$

Da n_1 und n_2 Teiler sind von $|G|$, haben wir $n_i \leq |G|$ und damit notwendig $n_1 = n_2 = |G|$. Alle Pole werden also von der Gruppe festgehalten, es gibt folglich nur zwei Pole, die sich notwendig gegenüberliegen müssen. Damit sind wir im Fall der zyklischen Gruppen C_k mit $k > 1$.

Fall 3: Es gibt genau drei Bahnen P_1, P_2 und P_3 in P, wir haben also

$$\frac{2}{|G|} = \frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3} - 1$$

Wir dürfen annehmen $n_1 \leq n_2 \leq n_3$. Sicher gilt dann $n_1 = 2$, sonst wäre die rechte Seite ≤ 0. Haben wir auch $n_2 = 2$, so kann n_3 beliebige Werte annehmen und wir haben $|G| = 2n_3$. Die Bahn P_3 besteht dann aus zwei Polen, die sich notwendig gegenüberliegen, und die von den Gruppenelementen zu den anderen Polen vertauscht werden. Die Gruppe wird damit eine Diedergruppe. Sicher sind $(2, 4, 4)$ und $(2, 3, 6)$ unmöglich für (n_1, n_2, n_3), da ja gilt $\frac{1}{2} + \frac{1}{4} + \frac{1}{4} - 1 = 0 = \frac{1}{2} + \frac{1}{3} + \frac{1}{6}$, also bleiben nur die Fälle $(2, 3, 3), (2, 3, 4)$ und $(2, 3, 5)$, und man berechnet leicht die zugehörigen Gruppenordnungen zu $12, 24$, und 60. Den Stand unseres Beweises bis hierher können wir so zusammenfassen: Wir haben eine Abbildung konstruiert—man mag sie die Bahnpolordnungsabbildung nennen—die jeder endlichen Untergruppe der Drehgruppe eine endliche Multimenge natürlicher Zahlen zuordnet, und haben gezeigt, daß in ihrem Bild höchstens die folgenden Multimengen liegen:

$$\emptyset, \{k, k\}_{k\geq 2}, \{2, 2, k\}_{k\geq 2}, \{2, 3, 3\}, \{2, 3, 4\} \text{ und } \{2, 3, 5\}.$$

Wir müssen nun noch zeigen, daß (1) die angegebenen Multimengen genau das Bild unserer Bahnpolordnungsabbildung bilden, und daß (2) je zwei Drehgruppen aus derselben Faser zueinander konjugiert sind. Wenn wir das alles

gezeigt haben, so folgt, daß die Bahnpolordungsabbildung eine Bijektion

$$\left\{ \begin{array}{c} \text{endliche Untergruppen} \\ \text{der Drehgruppe SO(3),} \\ \text{bis auf Konjugation} \end{array} \right\} \xrightarrow{\sim} \left\{ \begin{array}{c} \emptyset, \{k,k\}_{k\geq 2}, \{2,2,k\}_{k\geq 2}, \\ \{2,3,3\}, \{2,3,4\}, \{2,3,5\} \end{array} \right\}$$

definiert, und zusammen mit der beim Beweis erzeugten Anschauung zeigt das unseren Satz. Die Existenz endlicher Untergruppen der Drehgruppe mit derartigen Polbahnen und Polordnungen scheint mir anschaulich klar. Zum Beispiel hat die Würfelgruppe drei Polbahnen, als da sind: Eine Bahn aus den 8 Ecken zur Polordnung 3; eine Bahn aus den auf Länge Eins normierten 12 Mittelpunkten der Kanten, zur Polordnung 2; und eine Bahn aus den auf Länge Eins normierten 6 Mittelpunkten der Flächen, zur Polordnung 4. Diese Anschauung läßt sich auch leicht zu einem formalen Beweis präzisieren in allen Fällen mit Ausnahme des Ikosaeder-Falls $(2,3,5)$. In diesem Fall folgt die Existenz formal korrekt aus 6.8.13. Daß je zwei zyklische Gruppen derselben endlichen Ordnung und je zwei Diedergruppen derselben endlichen Ordnung in der Drehgruppe zueinander konjugiert sind, scheint mir offensichtlich. Die folgenden beiden Lemmata 6.8.10 und 6.8.11 zeigen, daß auch je zwei Gruppen zu gegebenem Typ $(2,3,n)$ in der Drehgruppe zueinander konjugiert sind. Damit vervollständigen sie den Beweis unseres Satzes. □

Lemma 6.8.10. *1. Jede endliche Untergruppe einer Drehgruppe von einem der beiden Typen $(2,3,4)$ oder $(2,3,5)$ ist maximal unter allen endlichen Untergruppen der Drehgruppe.*

2. Eine endliche Drehgruppe von einem der Typen $(2,3,3)$, $(2,3,4)$ oder $(2,3,5)$ kann beschrieben werden als die Symmetriegruppe jeder ihrer beiden kleineren Bahnen von Polen.

Beweis. Nach unseren bisherigen Erkenntnissen kommen bei endlichen Drehgruppen für die Paare (Ordnung eines Pols, Kardinalität seiner Bahn) nur die Paare $(n,1)$, $(n,2)$, $(2,n)$, $(3,4)$, $(3,8)$, $(3,20)$, $(4,6)$ und $(5,12)$ in Frage. Für jeden Pol müssen sich bei Übergang zu einer echt größeren Gruppe nach der Bahnformel entweder seine Polordnung oder die Kardinalität seiner Bahn oder beide vervielfachen. Das ist aber bei $(4,6)$ und $(5,12)$ unmöglich und wir erhalten die erste Behauptung. In den drei Fällen der zweiten Behauptung enthält weiter jede Bahn von Polen mindestens drei Punkte, also auch zwei verschiedene nicht gegenüberliegende Punkte. Folglich operiert sogar die Symmetriegruppe der Bahn P_i treu auf P_i und ist insbesondere endlich. Nun muss P_i auch unter der fraglichen Symmetriegruppe eine Bahn von Polen sein. Wenn diese Symmetriegruppe größer sein will, muss sie also

an diesen Polen größere Polordnungen haben. Wieder ist das unmöglich bei $(3,4)$, $(3,8)$, $(3,20)$, $(4,6)$ und $(5,12)$. □

Lemma 6.8.11. *Sind zwei endliche Drehgruppen vom selben Typ* $(2,3,n)$ *mit* $3 \leq n \leq 5$ *gegeben und sind* P_3 *und* $\tilde{P}_3$ *jeweils zugehörige Polbahnen kleinstmöglicher Kardinalität, so gibt es eine Drehung, die* P_3 *in* $\tilde{P}_3$ *überführt.*

Beweis. Gegeben eine Bahn von Polen P_i betrachten wir ganz allgemein die Operation von G auf $P_i \times P_i$ und beachten, daß aus geometrischen Gründen die Isotropiegruppe eines Paars (p,q) mit $p \neq \pm q$ trivial sein muß. Nach dieser Vorüberlegung betrachten wir die drei Fälle der Reihe nach.

Im Fall $(2,3,3)$ haben wir $|P_3| = 4$ und $|P_3 \times P_3| = 16$. Folglich gibt es in $P_3 \times P_3$ ein Paar mit trivialer Isotropiegruppe, das also eine 12-elementige Bahn hat, die notwendig aus allen (p,q) mit $p \neq q$ bestehen muß. Je zwei verschiedene Punkte aus P_3 haben also denselben Abstand. Ich hoffe, daß damit sowohl die Aussage des Lemmas im Fall $n = 3$ klar wird als auch, daß die Punkte aus P_3 die Ecken eines Tetraeders bilden.

Im Fall $(2,3,4)$ haben wir $|P_3| = 6$ und $|P_3 \times P_3| = 36$. Folglich gibt es in $P_3 \times P_3$ ein Paar mit trivialer Isotropiegruppe, das also eine 24-elementige Bahn hat, die notwendig aus allen (p,q) mit $p \neq \pm q$ bestehen muß. Die anderen Bahnen müssen aus Paaren mit nichttrivialer Isotropiegruppe bestehen, und da die Bahn der sechs Paare der Gestalt (p,p) noch nicht genug Elemente liefert, muß auch noch eine Bahn von der Gestalt $(p,-p)$ vorkommen. Wir sehen so einerseits, daß P_3 stabil ist unter Punktspiegelung am Ursprung, und andererseits, daß je zwei voneinander verschiedene Pole aus P_3, die sich nicht gegenüberliegen, denselben Abstand haben. So erkennen wir hoffentlich sowohl die Aussage des Lemmas im Fall $n = 4$ als auch, daß die Elemente von P_3 die Ecken eines Oktaeders bilden müssen.

Im Fall $(2,3,5)$ haben wir $|P_3| = 12$ und $|P_3 \times P_3| = 144$. Wieder haben wir an Bahnen in $|P_3 \times P_3|$ die zwölfelementige Bahn aller Paare (p,p), möglicherweise eine eine zwölfelementige Bahn aller Paare $(p,-p)$, und daneben nur Bahnen mit 60 Elementen. Es folgt, daß $P_3 \times P_3$ in vier Bahnen zerfällt, und zwar die Bahn der Paare gleicher Pole, die Bahn der Paare von sich gegenüberliegenden Polen, und zwei weitere Bahnen von Polpaaren. Nehmen wir irgendeinen Pol $p \in P_3$, so bilden die Bilder von jedem Pol $q \in P_3$ mit $q \neq \pm p$ unter den Drehungen aus unserer Gruppe mit Fixpunkt p ein regelmäßiges Fünfeck, und für zwei verschiedene Ecken eines Fünfecks gibt es zwei Möglichkeiten für ihren Abstand, deren Verhältnis übrigends nach **??** oder elementargeometrischen Überlegungen gerade der goldene Schnitt ist. Unsere beiden 60-elementigen Bahnen müssen sich also im Abstand zwischen den

Polen eines Paars unterscheiden. Zu jedem Pol aus P_3 gibt es damit außer dem Pol selbst und dem gegenüberliegenden Pol noch 5 "nahe" Pole und 5 "weite" Pole. Nun bilden zwei sich gegenüberliegende Pole aus P_3 mit jedem weiteren Pol ein Dreieck, das nach dem Satz des Thales bei diesem weiteren Pol einen rechten Winkel hat, wobei dieser Pol notwendig zu einem von unseren beiden sich gegenüberliegenden Polen nah sein muß und zum anderen weit, da ja zu jedem unserer sich gegenüberliegenden Pole von den zehn verbleibenden Polen fünf nah und fünf weit sein müssen. Da unser Dreieck eine Hypothenuse der Länge 2 hat, wird dadurch der Abstand zwischen nahen Polen und der zwischen weiten Polen bereits vollständig beschrieben und hängt insbesondere nicht von unserer Gruppe ab. Damit erkennen wir, daß im Fall $(2,3,5)$ die Bahn P_3 bestehen muß aus (1) zwei gegenüberliegenden Punkten N und $S = -N$ sowie (2) zwei regelmäßigen Fünfecken der fünf zu N nahen Pole und der fünf zu S nahen Pole mit jeweils von der speziellen Gruppe unabhängigem Abstand der Ecken dieser Fünfecke zu den jeweiligen Polen. Jede Ecke des "nördlichen" Fünfecks muß aber auch einer Ecke des "südlichen" Fünfecks gegenüberliegen. Unser Lemma folgt unmittelbar. □

Der Rest dieses Abschnitts war es in der Vorlesung 2008/09 nicht dran.

Übung 6.8.12. Die Multiplikation definiert einen Isomorphismus zwischen der Gruppe aller Symmetrien aus O(3) eines Ikosaeders und dem Produkt der Gruppe seiner Drehsymmetrien mit der zweielementigen Gruppe, die von der Punktspiegelung am Ursprung erzeugt wird. Insbesondere ist die "nichtorientierte Ikosaedergruppe" keineswegs isomorph zur symmetrischen Gruppe $\mathcal{S}_5$.

Lemma 6.8.13 (Existenz des Ikosaeders). *Es gibt in der Einheitssphäre eine zwölfelementige Teilmenge, die stabil ist unter den Drehungen mit den Winkeln $\pm 2\pi/5$ um die Ursprungsgeraden durch alle ihre Punkte, und je zwei derartige Teilmengen lassen sich durch eine Drehung ineinander überführen.*

6.8.14. Anschaulich mag man sich eine derartige Teilmenge der Einheitssphäre als die Menge der Ecken eines Isosaeders denken, was hoffentlich im gleich folgenden Beweis noch deutlicher werden wird.

Beweis. Wir betrachten die Menge $\mathcal{D} \subset \mathcal{P}(S^2)$ aller gleichseitigen Dreiecke mit Ecken auf der Einheitssphäre, die nicht in einer Ebene mit dem Ursprung liegen, also formal

$$\mathcal{D} = \left\{ \{a,b,c\} \;\middle|\; \begin{array}{l} a,b,c \in \mathbb{R}^3, \\ \|a\| = \|b\| = \|c\| = 1, \\ \|a-b\| = \|b-c\| = \|c-a\|, \\ \langle a,b,c\rangle_{\mathbb{R}} = \mathbb{R}^3. \end{array} \right\}$$

Gegeben ein Dreieck $\Delta \in \mathcal{D}$ und eine Ecke $a \in \Delta$ definieren wir das **umgeklappte Dreieck** $\Delta^a \in \mathcal{D}$ als das eindeutig bestimmte Dreieck $\Delta^a \in \mathcal{D}$ mit $\Delta \cap \Delta^a = \{b, c\}$. Haben wir $\Delta^a = \{\alpha, b, c\}$, so gilt dann natürlich auch umgekehrt $(\Delta^a)^\alpha = \Delta$. Definieren wir zu einem Dreieck $\Delta \in \mathcal{D}$ die Menge $\mathcal{D}(\Delta)$ als die kleinste Teilmenge $\mathcal{D}(\Delta) \subset \mathcal{D}$, die Δ enthält und die stabil ist unter dem Umklappen von Dreiecken, so gilt offensichtlich $\mathcal{D}(\Delta) = \mathcal{D}(\Delta')$ für alle $\Delta' \in \mathcal{D}(\Delta)$. Ist $r \in \mathrm{O}(3)$ orthogonal, so gilt offensichtlich

$$\{ra, rb, rc\}^{ra} = r(\{a, b, c\}^a)$$

für jedes Dreieck $\{a, b, c\} \in \mathcal{D}$ und insbesondere $r(\mathcal{D}(\Delta)) = \mathcal{D}(r\Delta)$. Haben wir nun zusätzlich $|(r\Delta) \cap \Delta| \geq 2$, so folgt $r\Delta \in \mathcal{D}(\Delta)$ und damit $\mathcal{D}(r\Delta) = \mathcal{D}(\Delta)$. Haben wir genauer $\Delta = \{a, b, c\}$ und $r\Delta \cap \Delta = \{b, c\}$, so folgt unmittelbar $r\Delta = \Delta^a$.

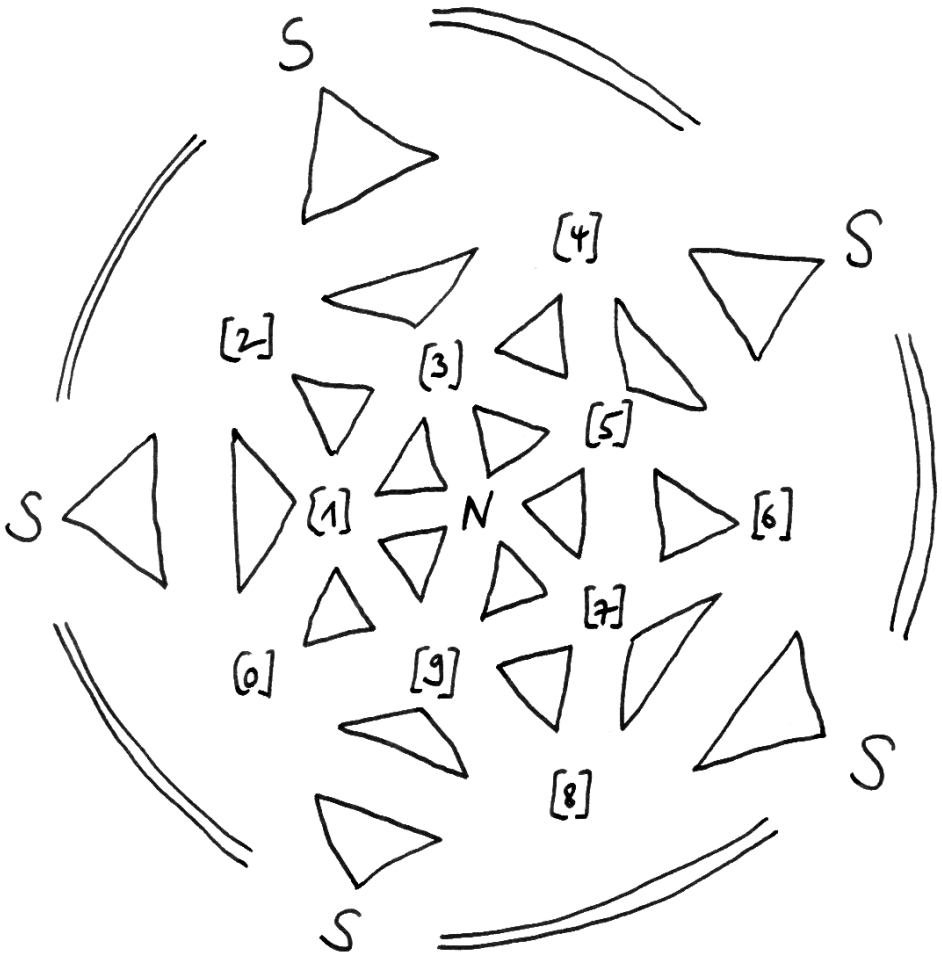
Nach diesen Vorüberlegungen gehen wir nun aus von einem regelmäßigen Fünfeck, dessen Ecken wir der Reihe nach mit $[1], [3], [5], [7], [9]$ bezeichnen, bilden darauf die Pyramide mit Spitze N und aufsteigenden Kanten von derselben Länge wie die Kanten des Fünfecks, und schrumpfen oder strecken diese Pyramide so, daß wir sie als "Polkappe" in die Einheitssphäre legen können. Dann gehen die fünf gleichseitigen Dreiecke dieser Polkappe durch Umklappen auseinander hervor, genauer haben wir $\{[n], [n+2], N\}^{[n]} = \{[n+4], [n+2], N\}$ für $n \in \mathbb{Z}/10\mathbb{Z}$ ungerade, wie der innerste Teil des nebenstehenden Bildes verdeutlicht. Bezeichne $\mathcal{D}^* \subset \mathcal{D}$ die kleinste unter Umklappen stabile Menge von Dreiecken, die diese fünf Dreiecke umfaßt. Wir zeigen im folgenden, daß $\mathcal{D}^*$ endlich ist und daß die Menge aller Ecken von Dreiecken aus $\mathcal{D}^*$ genau zwölf Elemente hat. Das ist dann offensichtlich eine zwölfelementige Teilmenge der Einheitssphäre von der im Lemma gesuchten Art. An die Arbeit! Zunächst definieren wir Ecken $[n] \in S^2$ für $n \in \mathbb{Z}/10\mathbb{Z}$ gerade durch

$$\{N, [n-1], [n+1]\}^N = \{[n], [n-1], [n+1]\}$$

oder bildlich gesprochen durch das "vom Nordpol weg nach unten Klappen" unserer fünf Dreiecke mit Ecke bei N. Dann zeigen wir, daß für $n \in \mathbb{Z}/10\mathbb{Z}$ gerade gilt

$$\{[n], [n-1], [n+1]\}^{[n-1]} = \{n, [n+1], [n+2]\}$$

oder bildlich gesprochen: Wenn wir zwei benachbarte Dreiecke unserer Ursprungspyramide nach unten klappen, so haben die beiden Bilder des Nordpols genau die Kantenlänge aller unserer Dreiecke als Abstand. Ist $\{N, a, b\}$


S
S
S
S
S
[1]
[2]
[3]
[4]
[5]
[6]
[7]
[8]
[9]
[0]
N

ein Dreieck unserer Ausgangspyramide, so gibt es ja genau eine Drehung $r \in \mathrm{SO}(3)$ mit $r(N) = N$ und $r(a) = b$ und $r^5 = \mathrm{id}$. Natürlich gilt dann auch

$$r\{N, a, b\} = \{N, a, b\}^a,$$

denn die linke Seite ist ein von $\{N, a, b\}$ verschiedenes Dreieck, das $\{N, a, b\}$ in $\{N, b\}$ trifft. Wir folgern, daß es für jedes Dreieck $\Delta = \{p, a, b\} \in \mathcal{D}^*$ genau ein $r \in \mathrm{SO}(3)$ gibt mit $r(p) = p$ und $r(a) = b$ und $r^5 = \mathrm{id}$, und daß für dieses r gilt $r\{p, a, b\} = \{p, a, b\}^a$. Wir betrachten nun diese Drehung r für $\Delta = \{[1], [3], N\}$, also $r \in \mathrm{SO}(3)$ mit $r : [1] \mapsto [1], [3] \mapsto N$. Ich behaupte

$$\begin{array}{ccccccc}
\{[1],[2],[3]\} & \stackrel{r}{\longmapsto} & \{[1],[3],N\} & \stackrel{r}{\longmapsto} & \{[1],N,[9]\} & \stackrel{r}{\longmapsto} & \{[1],[9],[0]\} \\
\uparrow r & & & & & & \downarrow r \\
\{[1],\beta,[2]\} & = & & = & & = & \{[1],[0],\alpha\}
\end{array}$$

mit a priori unbekannten $\alpha, \beta \in S^2$. In der Tat ist jedes Anwenden der Drehung r ein Umklappen an einer Kante, die $[1]$ enthält, und r^2 kann keines unserer Dreiecke auf sich selber werfen. Aus r^5 folgt nun aber $\beta = [0]$ und $[\alpha] = 2$ und $\{[1], [0], [2]\} \in \mathcal{D}^*$. Analog können wir statt für $[1]$ auch für die anderen ungeraden Ecken $[n]$ argumentieren und erhalten so alle im Bild dargestellten Dreiecke mit den ebenfalls dort dargestellten Umklappbeziehungen, mit Ausnahme der Dreiecke mit der als S bezeichneten Ecke. Dazu betrachten wir nun die Drehung $r \in \mathrm{SO}(3)$ mit $r : [2] \mapsto [2], [4] \mapsto [3]$ und $r^5 = \mathrm{id}$ und folgern

$$\begin{array}{ccccc}
\{[2],[4],[3]\} & \stackrel{r}{\longmapsto} & \{[2],[3],[1]\} & \stackrel{r}{\longmapsto} & \{[2],[1],[0]\} \\
\uparrow r & & & & \downarrow r \\
\{[2],\beta,[4]\} & \longleftarrow & \underset{r}{\longleftarrow} & \longleftarrow & \{[2],[0],\alpha\}
\end{array}$$

mit a priori unbestimmten $\alpha, \beta \in S^2$, von denen dann aber folgt, daß $\alpha = \beta$ gelten muß und daß für diese Ecke $S \in S^2$ gilt

$$\begin{array}{rcl}
\{[2],[1],[0]\}^{[1]} & = & \{[2],S,[0]\} \\
\{[2],[3],[4]\}^{[3]} & = & \{[2],S,[4]\}
\end{array}$$

Arbeiten wir statt mit $[2]$ mit den anderen "geraden" Ecken, so erkennen wir

$$\{[n-1],[n],[n+1]\}^{[n]} = \{[n-1],S,[n+1]\}$$

für alle geraden $n \in \mathbb{Z}/10\mathbb{Z}$. Damit haben wir genau die im Bild dargestellten Dreiecke von $\mathcal{D}^*$ mitsamt ihren in ebendiesem Bild dargestellten Umklappbeziehungen erhalten. Da diese 20 Dreiecke ein unter Umklappen stabiles System bilden, haben wir bereits ganz $\mathcal{D}^*$ gefunden und insbesondere gilt $|\mathcal{D}^*| = 20 < \infty$. □

6.8.15. Die obigen Überlegungen kann man dahingehend zusammenfassen, daß gegeben ein gleichseitiges Dreieck $\Delta = \{a, b, c\}$, für das es eine Drehung r um die Ursprungsgerade durch a gibt mit $r^5 = \mathrm{id}$ und $r : b \mapsto c$, die Menge $\mathcal{D}(\Delta)$ der daraus durch Umklappen entstehenden Dreiecke endlich ist. Die hier geforderte Eigenschaft hat sicher jedes Dreieck, das anschaulich gesprochen "Fläche eines Ikosaeders" ist. Es gibt aber auch noch andere gleichseitige Dreiecke mit dieser Eigenschaft, nämlich diejenigen gleichseitigen Dreiecke, die anschaulich gesprochen die "Diagonale unseres Ausgangsfünfecks" als Seitenlänge haben.

Bemerkung 6.8.16. Mit welchen platonischen Körpern kann man den Raum füllen? Ich vermute, das geht nur mit Würfeln: Die anderen sollten als Winkel zwischen an einer Kante angrenzenden Flächen nie einen Winkel der Gestalt $2\pi/n$ haben.

Bemerkung 6.8.17. Vielleicht ist es vernünftig, platonische Körper zu definieren über die Mengen ihrer Ecken, die man wohl wie folgt charakterisieren kann: Man definiere für eine endliche Teilmenge E des Raums ihre **Abständezahl** $A(E)$ als die Zahl der möglichen von Null verschiedenen verschiedenen Abstände zwischen ihren Elementen. Eine endliche Teilmenge E einer Sphäre heißt nun Tetraeder bei $|E| = 4$, $A(E) = 1$, Würfel bei $|E| = 8$, $A(E) = 3$, Oktaeder bei $|E| = 6$, $A(E) = 2$, Ikosaeder bei $|E| = 12$, $A(E) = 3$, Dodekaeder bei $|E| = 20$, $A(E) = 4$. Stimmt das eigentlich? Möglicherweise sollte man bei allen außer dem Tetraeder noch fordern, daß E stabil ist unter Punktspiegelung am Ursprung.

6.9 Skalarprodukte zu Drehgruppen

Das war in der Vorlesung 2008/09 nicht dran.

6.9.1. In diesem Abschnitt holen wir den Rest des Beweises von Satz 3.1.5 über den Zusammenhang zwischen Bewegungsgruppen und Skalarprodukten nach. Ich erinnere daran, daß wir in 3.1.2 eine Bewegungsgruppe eines dreidimensionalen reellen affinen Raums E definiert hatten als eine alle Translationen umfassende Untergruppe $B \subset \mathrm{Aut}\, E$ der Automorphismengruppe unseres affinen Raums derart, daß es für je zwei Paare (H, L) von Teilmengen von E bestehend aus einer Halbebene und einer Halbgerade auf ihrem Rand genau einen Automorphismus aus B gibt, der sie ineinander überführt. Die Elemente der Isotropiegruppe $B_p \subset B$ eines Punktes $p \in E$ nennen wir **Drehungen um den Punkt** p. Da unsere Bewegungsgruppe nach Annahme alle Translationen enthält, liefert das Bilden des linearen Anteils einen Isomorphismus der Isotropiegruppe B_p jedes Punktes $p \in E$ mit derselben

Gruppe $D \subset \mathrm{GL}(\vec{E})$ von Automorphismen des Richtungsraums. Die Elemente von D nennen wir **Drehungen im Richtungsraum**. Nach unseren Annahmen bilden die linearen Anteile der Elemente einer Bewegungsgruppe im Sinne der gleich folgenden Definition eine Drehgruppe.

Definition 6.9.2. Unter einem **Strahl** L in einem reellen Vektorraum V verstehen wir eine Teilmenge $L \subset V$ mit der Eigenschaft, daß es in V einen Vektor $v \neq 0$ gibt mit $L = \mathbb{R}_{\geq 0} v$. Unter einer **Drehgruppe** in einem dreidimensionalen reellen Vektorraum verstehen wir eine Untergruppe seiner Automorphismengruppe mit der Eigenschaft, daß es für je zwei Paare von Teilmengen unseres Vektorraums bestehend aus einer linearen Halbebene und einem Strahl auf ihrem Rand genau ein Element unserer Untergruppe gibt, die das eine Paar in das andere überführt. Die Elemente einer solchen Drehgruppe bezeichnen wir dann auch als **Drehungen**.

Satz 6.9.3 (Drehgruppen und Skalarprodukte). *Gegeben ein dreidimensionaler reeller Vektorraum V und ein ausgezeichneter Vektor $m \in V \backslash 0$ liefert die Abbildung $b \mapsto \mathrm{SO}(V; b)$ eine Bijektion*

$$\{\textit{Skalarprodukte } b \textit{ auf } V \textit{ mit } b(m,m) = 1\} \quad \tilde{\to} \quad \{\textit{Drehgruppen } D \subset \mathrm{GL}(V)\}$$

6.9.4. Aus diesem Satz folgen unmittelbar die noch unbewiesenen Behauptungen von Satz 3.1.5 über den Zusammenhang zwischen Bewegungsgruppen und Skalarprodukten.

Beweis. Den Nachweis, daß für jedes Skalarprodukt b auf V die Gruppe $\mathrm{SO}(V; b) = \{d \in \mathrm{GL}(V) \mid b(dv, dw) = b(v, w)\ \forall v, w \in V \text{ und } \det d = 1\}$ in der Tat eine Drehgruppe ist, überlasse ich dem Leser und beginne gleich mit der Konstruktion der Umkehrabbildung. Sei also V ein dreidimensionaler reeller Vektorraum und $D \subset \mathrm{GL}(V)$ eine Drehgruppe im Sinne unserer Definition 6.9.2. Gegeben $v, w \in V$ vereinbaren wir die Sprechweise, w **stehe drehsenkrecht auf** v und schreiben

$$w \vdash v$$

genau dann, wenn es eine Drehung $r \in D$ gibt mit $r(w) = -w$ und $r(v) = v$. Aus $w \vdash v$ folgt leicht $dw \vdash dv$ für jede Drehung d und $\lambda w \vdash \mu v$ für alle $\mu, \lambda \in \mathbb{R}$. Des weiteren steht nur der Nullvektor drehsenkrecht auf sich selbst. Gegeben linear unabhängige Vektoren $v, w \in V$ vereinbaren wir nun für das dadurch bestimmt Paar aus einer Halbebene nebst einem Strahl auf ihrem Rand speziell für diesen Beweis die Notation

$$[v, w] = (\mathbb{R}v + \mathbb{R}_{\geq 0} w, \mathbb{R}_{\geq 0} v)$$

Unsere Definition einer Drehgruppe besagt in dieser Notation, daß es für je zwei Paare (v, w) und (v', w') von linear unabhängigen Vektoren genau ein Element r unserer Drehgruppe gibt mit $r : [v, w] \mapsto [v', w']$. Als nächstes zeigen wir die Symmetrie der Relation des Drehsenkrechtstehens.

Lemma 6.9.5. *Es gilt* $w \vdash v \Rightarrow v \vdash w$.

Beweis. Zunächst zeigen wir das für v, w linear unabhängig. Gilt $w \vdash v$, so gibt es ja per definitionem eine Drehung r mit $rw = -w$ und $rv = v$. Das muß natürlich die Drehung r sein mit $r : [v, w] \mapsto [v, -w]$. Betrachten wir zusätzlich die Drehung s mit $s : [v, w] \mapsto [-v, w]$, so folgt $s^2 = \mathrm{id}$ und weiter $sr = rs$, da beide Abbildungen die Eigenschaft $[v, w] \mapsto [-v, -w]$ haben. Daraus folgt erst $sv = -v$ und dann $sw = w$ durch explizite Rechnung oder konzeptioneller, da s die Eigenräume von r im Erzeugnis $\mathbb{R}v + \mathbb{R}w$ stabilisiert. Das liefert dann $v \vdash w$ wie behauptet. Gilt $w \vdash v$ für linear abhängige Vektoren, so muß mindestens einer der Nullvektor sein. Im Fall $v = 0$ ist $0 \vdash w$ offensichtlich, bereits die Identität hält dann w fest und bildet v auf sein Negatives ab. Es reicht also, wenn wir $v \vdash 0$ zeigen für alle $v \neq 0$. Unter dieser Annahme gibt es jedoch für $u \notin \mathbb{R}v$ eine Drehung s mit $s : [v, u] \mapsto [-v, u]$. Wegen $s^2 : [v, u] \mapsto [v, u]$ gilt $s^2 = \mathrm{id}$ und daraus folgt $s(v) = -v$ und damit haben wir in der Tat $v \vdash 0$. □

Lemma 6.9.6. *1. Die auf allen Vektoren einer Ebene drehsenkrecht stehenden Vektoren bilden eine Gerade.*

2. Die auf allen Vektoren einer Gerade drehsenkrecht stehenden Vektoren bilden eine Ebene.

3. Für jeden von Null verschiedenen Vektor $n \neq 0$ gibt es genau eine Drehung r_n mit $r_n n = n$ und $u \vdash n \Leftrightarrow r_n u = -u$.

Beweis. Gegeben $G \subset P \subset V$ eine Gerade in einer Ebene gibt es genau eine Drehung, die die Gerade G punktweise festhält und die beiden zugehörigen Halbebenen von P vertauscht: Schreiben wir etwa $G = \mathbb{R}v$ und $P = \mathbb{R}v + \mathbb{R}w$, so kann unsere Drehung charakterisiert werden durch $[v, w] \mapsto [v, -w]$. Es folgt, daß die Menge der auf allen Vektoren aus G drehsenkrecht stehenden Vektoren von P eine Gerade G' ist, eben der (-1)-Eigenraum dieser Drehung in P, und nach 6.9.5 ist die Menge der auf allen Vektoren aus G drehsenkrecht stehenden Vektoren von P dann wieder unsere ursprüngliche Gerade G. Gegeben linear unabhängige Vektoren v, w mit $v \vdash w$ hat die Drehung d mit $d : [v, w] \mapsto [w, -v]$ folglich die Eigenschaft $d(w) \in \mathbb{R}v$ und es ergibt sich sofort $d : [w, -v] \mapsto [-v, -w]$, also $d^4 = \mathrm{id}$. Wir erkennen $d^2 v = -v$, $d^2 w = -w$ und folglich $d^2 u = -u$ für alle $u \in P$. Andererseits

haben wir $d^2 \neq -\mathrm{id}$, etwa da die Determinante eines Quadrats nie negativ sein kann, folglich hat d^2 einen von Null verschiedenen Fixvektor n und es folgt $P = \{u \in V \mid n \vdash u\}$. Wir erkennen so, daß die auf einer vorgegebenen Ebene drehsenkrechten Vektoren stets eine Gerade bilden, und daß es zu einem von Null verschiedenen Vektor $n \neq 0$ stets genau eine Drehung r_n gibt mit $r_n(n) = n$ und $u \vdash n \Rightarrow r_n(u) = -u$. Daß die auf allen Vektoren einer Gerade drehsenkrecht stehenden Vektoren eine Ebene bilden, folgt daraus dann unmittelbar. □

Übung 6.9.7. Gegeben ein Vektor $n \neq 0$ gilt für jede Drehung d die Identität $r_{dn} = d \circ r_n \circ d^{-1}$ und für jeden von Null verschiedenen Skalar $\lambda \in \mathbb{R}^\times$ haben wir $r_{\lambda n} = r_n$.

Lemma 6.9.8. *Gegeben zwei linear unabhängige Vektoren v, w gilt für die Drehung r mit $r : [v, w] \mapsto [w, v]$ die Identität $r^2 = \mathrm{id}$ und es gibt $\lambda > 0$ mit $rv = \lambda w$ und $r\lambda w = v$.*

Beweis. Die Restriktion von r auf die Ebene $\mathbb{R}v + \mathbb{R}w$ hat negative Determinante, da ihre Matrix in der Basis v, w oben links eine Null hat und in der Nebendiagonalen positive Einträge. Damit hat unsere Matrix zwei verschiedene reelle Eigenwerte und r^2 hat zwei positive reelle Eigenwerte, etwa mit Eigenvektoren n und m, und wegen $r^2 : [n, m] \mapsto [n, m]$ folgt $r^2 = \mathrm{id}$. Der Rest des Lemmas folgt leicht. □

Lemma 6.9.9. *Bildet eine Drehung einen Strahl bijektiv auf sich selber ab, so hält sie ihn bereits punktweise fest.*

6.9.10. Dies Lemma formalisiert die Erfahrungstatsache, daß eine Achse beim Drehen ihre Länge nicht ändert, und es mag lächerlich wirken, das beweisen zu wollen. In der Tat hätten wir diese Aussage auch als zusätzliche Bedingung zu unserer Definition des Anschauungsraums und zur Definition des Begriffs einer Drehgruppe hinzunehmen können. Daß ich das nicht getan habe, hat rein ästhetische Gründe: Wir können so eine größere Wegstrecke mit reiner Logik zurücklegen.

Beweis. Es gilt für $u \neq 0$ und jede Drehung $d \in D$ zu zeigen

$$d(\mathbb{R}_{\geq 0}u) = \mathbb{R}_{\geq 0}u \;\Rightarrow\; du = u$$

Dazu wählen wir $v \neq 0$ mit $v \vdash u$. Gilt $dv \in \mathbb{R}v$, so folgt $d^2v \in \mathbb{R}_{>0}v$ und damit $d^2 : [u, v] \mapsto [u, v]$ und so $d^2 = \mathrm{id}$ und dann $du = u$. Sonst spannen v und dv die zu u drehsenkrechte Ebene auf. Nach 6.9.8 gibt es $\lambda > 0$ und eine Drehung r, die λv mit dv vertauscht. Deren Quadrat ist die Identität, woraus leicht folgt $ru = -u$. Für die Verknüpfung rr_v gilt dann $\lambda v \mapsto dv$

und $u \mapsto u$, woraus folgt $rr_v : [u, v] \mapsto [u, dv]$, also $rr_v = d$ und damit dann $du = u$ wie gewünscht. □

Lemma 6.9.11. *Jede Bahn einer Drehgruppe trifft jeden Strahl in genau einem Punkt.*

Beweis. Daß jede Bahn jeden Strahl in höchstens einem Punkt trifft, folgt sofort aus 6.9.9. Daß jede Bahn jeden Strahl in mindestens einem Punkt trifft, folgt unmittelbar aus unserer Definition einer Drehgruppe. □

6.9.12. Wie in 3.1.7 erklären wir die Drehnorm eines Vektors v als diejenige nichtnegative reelle Zahl $\|v\| = \lambda$, für die es eine Drehung d gibt mit $d(v) = \lambda m$. Das vorhergehende Lemma 6.9.11 zeigt, daß es genau ein $\lambda \geq 0$ mit dieser Eigenschaft gibt.

Nach diesen Vorbereitungen konstruieren wir nun unser Skalarprodukt. Gegeben $v \neq 0$ und $w \notin \mathbb{R}v$ gilt für unser r_v aus 6.9.6.1 sicher $r_v w = \alpha v - \gamma w$ mit $\gamma \geq 0$. Wegen $r_v^2 w = \alpha v - \alpha\gamma v + \gamma^2 w = w$ folgt $\gamma = 1$. Es gibt folglich für alle $w \in V$ genau eine reelle Zahl $\alpha_v(w)$ mit der Eigenschaft

$$r_v w + w = \alpha_v(w) v$$

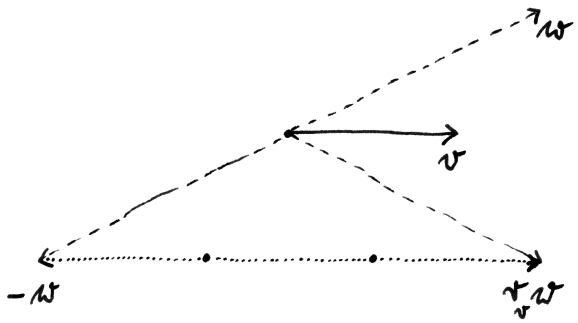
Man erkennt unschwer, daß α_v eine Linearform auf V ist. Wir können α_v auch charakterisieren als die eindeutig bestimmte Linearform, die auf v den Wert Zwei annimmt und auf allen zu v drehsenkrechten Vektoren den Wert Null. Unsere Definitionen liefern für jede weitere Drehung d die Identität $\alpha_{dv} \circ d = \alpha_v$ alias $\alpha_{dv} = \alpha_v \circ d^{-1}$ und für jeden von Null verschiedenen Skalar $\lambda \in \mathbb{R}^\times$ die Identität $\alpha_{\lambda v} = \lambda^{-1}\alpha_v$. Werden zwei von Null verschiedene Vektoren v, w durch eine Drehung untereinander vertauscht, so gilt für unsere Ausdrücke weiter die Identität $\alpha_v(w) = \alpha_w(v)$. In der Tat, aus $rv = w$ und $rw = v$ folgt $rr_v = r_w r$ und aus der von der Mitte ausgehend zu entwickelnden Gleichungskette

$$\alpha_w(v)w - v = r_w(v) = r_w rw = rr_v w = r(\alpha_v(w)v - w) = \alpha_v(w)w - v$$

ergibt sich die Behauptung. Nun wählen wir die durch unseren ausgezeichneten Vektor m gegebene Drehnorm und erklären die Abbildung $b = b_m : V \times V \to \mathbb{R}$ durch die Vorschrift

$$b(v, w) = \begin{cases} \|v\|^2 \alpha_v(w)/2 & v \neq 0; \\ 0 & v = 0. \end{cases}$$

Offensichtlich gilt $\|v\|^2 = b(v, v)$ und $w \mapsto b(v, w)$ ist linear für alle v. Schließlich beachten wir, daß für je zwei von Null verschiedene Vektoren $v, w \in V$ die



Diese Abbildung illustriert die Definition von $\alpha_v(w)$. Im hier dargestellten Fall hätten wir etwa $\alpha_v(w) = 3$ und für das Skalarprodukt b_l mit $v \in l$ hätten wir $b_l(v, w) = 3/2$.

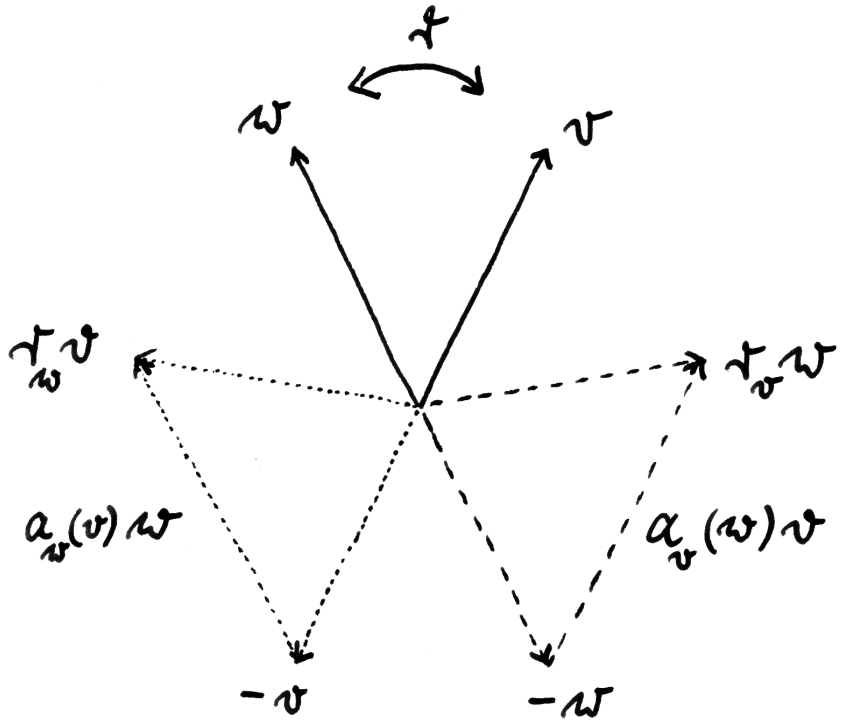


Illustration der Identität $\alpha_v(w) = \alpha_w(v)$ unter der Annahme, daß es eine Drehung r gibt, die v und w vertauscht.

Vektoren $\|v\|^{-1}v$ und $\|w\|^{-1}w$ durch eine Drehung untereinander vertauscht werden. Nach dem Vorhergehenden folgt $\|v\|\|w\|^{-1}\alpha_v(w) = \|w\|\|v\|^{-1}\alpha_w(v)$ alias $\|v\|^2\alpha_v(w) = \|w\|^2\alpha_w(v)$ und damit $b(v,w) = b(w,v)$ erst für je zwei von Null verschiedene Vektoren, aber dann auch sofort für alle $v, w \in V$. In der Terminologie aus 3.1.4 ist also b ein Skalarprodukt auf V und wir haben wie versprochen eine Abbildung in die Gegenrichtung konstruiert. Daß unsere beiden Abbildungen in der Tat zueinander invers sind, mag der Leser selbst prüfen. □

7 Universelle Konstruktionen

7.1 Quotientenvektorräume

Satz 7.1.1 (**Quotientenvektorraum**). *Sei k ein Körper. Gegeben $V \supset U$ ein k-Vektorraum mit einem Teilraum existiert auf der Restklassengruppe V/U genau eine Struktur als k-Vektorraum $k \times V/U \to V/U$ derart, daß die kanonische Projektion*

$$\mathrm{can} : V \twoheadrightarrow V/U$$

eine k-lineare Abbildung wird. Mit dieser Vektorraumstruktur heißt V/U der **Quotient von V nach U**.

Beweis. Wir betrachten die Abbildung

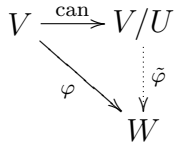
$$\begin{array}{rcl} k \times \mathcal{P}(V) & \to & \mathcal{P}(V) \\ (\lambda\ ,\ A) & \mapsto & \lambda.A = \lambda A + U \end{array}$$

Für $A = v + U$ finden wir $\lambda A + U = \lambda v + U$, so daß unsere Abbildung eine Abbildung $k \times V/U \to V/U$ induziert, die die Eigenschaft $\overline{\lambda v} = \lambda.\bar{v}$ hat für alle $\lambda \in k$, $v \in V$. Damit folgt sofort, daß unsere Abbildung $k \times V/U \to V/U$ auf der abelschen Gruppe V/U eine Struktur als k-Vektorraum definiert, und daß die Projektion $V \twoheadrightarrow V/U$ für diese Struktur k-linear ist. Umgekehrt ist auch klar, daß das die einzige Struktur als k-Vektorraum auf der abelschen Gruppe V/U ist, für die die Projektion $V \twoheadrightarrow V/U$ eine k-lineare Abbildung sein kann. □

Satz 7.1.2 (**Universelle Eigenschaft des Quotientenvektorraums**). *Sei k ein Körper. Seien $V \supset U$ ein k-Vektorraum mit einem Untervektorraum und sei* $\mathrm{can} : V \twoheadrightarrow V/U$ *die kanonische Projektion. So haben wir* $\ker(\mathrm{can}) = U$ *und für jeden weiteren Vektorraum W liefert das Vorschalten der kanonischen Projektion eine Bijektion*

$$\mathrm{Hom}_k(V/U, W) \xrightarrow{\circ\,\mathrm{can}}_{\sim} \{\varphi \in \mathrm{Hom}_k(V, W) \mid \varphi(U) = 0\}$$

7.1.3. Mit dem Diagramm



können wir die Aussage des Satzes auch dahingehend formulieren, daß jede lineare Abbildung $\varphi : V \to W$ mit $\varphi(U) = 0$ auf genau eine Weise k-linear "über die kanonische Projektion $\mathrm{can} : V \twoheadrightarrow V/U$ faktorisiert". Genau genommen hat also gar nicht der Quotientenvektorraum die universelle Eigenschaft, sondern der Homomorphismus can in den Quotientenvektorraum.

Beweis. Es muß nur gezeigt werden, daß der nach der universellen Eigenschaft der Restklassengruppe 6.4.7 wohldefinierte Gruppenhomomorphismus $\tilde{\varphi}$ in unserer Situation auch k-linear ist. Das folgt jedoch aus $\tilde{\varphi}(\lambda\bar{v}) = \tilde{\varphi}(\overline{\lambda v}) = \varphi(\lambda v) = \lambda\varphi(v) = \lambda\tilde{\varphi}(\bar{v})$. □

7.1.4. Jeder Vektorraumhomomorphismus $f : V \to W$ induziert einen Vektorraumisomorphismus $V/\ker f \xrightarrow{\sim} \operatorname{im} f$: Das folgt unmittelbar aus der entsprechenden Aussage für Gruppen 6.4.10. Gegeben Vektorräume $V \supset W \supset U$ induziert die Komposition von kanonischen Abbildungen $V \twoheadrightarrow V/U \twoheadrightarrow (V/U)/(W/U)$ einen Vektorraumisomorphismus $V/W \xrightarrow{\sim} (V/U)/(W/U)$: Das folgt unmittelbar aus dem Noether'schen Isomorphiesatz 6.4.11.

Definition 7.1.5. Gegeben $V \subset U$ ein Vektorraum mit einem Untervektorraum heißt $\dim(V/U)$ auch die **Kodimension** von U in V.

7.1.6. Ist V endlichdimensional, so haben wir nach 7.1.2 und der Dimensionsformel 1.6.10 die Identität $\dim(V/U) = \dim(V) - \dim(U)$, aber es gibt auch in unendlichdimensionalen Räumen durchaus Teilräume endlicher Kodimension. Eine Teilmenge eines Vektorraums ist eine lineare Hyperebene im Sinne von 1.3.20 genau dann, wenn sie ein Untervektorraum der Kodimension Eins ist.

Übung 7.1.7. Gegeben eine bilineare Abbildung $b : V \times W \to L$ und Untervektorräume $A \subset V$ und $B \subset W$ mit $b(A \times W) = 0 = b(V \times B)$ gibt es genau eine bilineare Abbildung $\bar{b} : (V/A) \times (W/B) \to L$ derart, daß das folgende Diagramm kommutiert:

$$\begin{array}{ccc} V \times W & \xrightarrow{b} & L \\ {\scriptstyle \mathrm{can} \times \mathrm{can}}\downarrow & & \| \\ V/A \times W/B & \xrightarrow{\bar{b}} & L \end{array}$$

Bemerkung 7.1.8. Um die Beziehung des Quotientenraums zu anderen Konstruktionen wie etwa dem Dualraum zu diskutieren, ist die Sprache der exakten Sequenzen aus 6.4.13 und besonders der kurzen exakten Sequenzen, wie wir sie gleich einführen werden, gut geeignet.

Definition 7.1.9. Eine Sequenz von Gruppen $A' \to A \to A''$ heißt eine **kurze exakte Sequenz** genau dann, wenn sie exakt ist in der Mitte und außerdem die erste Abbildung injektiv ist und die zweite surjektiv. Gleichbedeutend ist die Forderung, daß die in trivialer Weise erweiterte Sequenz $1 \to A' \to A \to A'' \to 1$ an jeder Stelle exakt ist. Wir notieren kurze exakte Sequenzen meist $A' \hookrightarrow A \twoheadrightarrow A''$.

Beispiel 7.1.10. Für jeden Normalteiler $N \subset G$ ist $N \hookrightarrow G \twoheadrightarrow G/N$ eine kurze exakte Sequenz von Gruppen. Für jeden Untervektorraum $U \subset V$ ist speziell $U \hookrightarrow V \twoheadrightarrow V/U$ eine kurze exakte Sequenz von Vektorräumen.

Beispiel 7.1.11. Für jeden surjektiven Gruppenhomomorphismus $x : G \twoheadrightarrow G''$ ist $\ker x \hookrightarrow G \twoheadrightarrow G''$ eine kurze exakte Sequenz von Gruppen. Für jede surjektive lineare Abbildung $U \subset V \twoheadrightarrow W$ ist speziell $\ker x \hookrightarrow V \twoheadrightarrow W$ eine kurze exakte Sequenz von Vektorräumen.

7.1.12. Die Dimensionsformel kann in dieser Terminologie auch dahingehend formuliert werden, daß für jede kurze exakte Sequenz $V' \hookrightarrow V \twoheadrightarrow V''$ von Vektorräumen gilt

$$\dim V = \dim V' + \dim V''$$

Definition 7.1.13. Gegeben Sequenzen $A \xrightarrow{r} B \xrightarrow{s} C$ und $A' \xrightarrow{r'} B' \xrightarrow{s'} C'$ verstehen wir unter einem **Homomophismus von Sequenzen** ein Tripel (f, g, h) von Homomophismen derart, daß das folgende Diagramm kommutiert:

$$\begin{array}{ccccc} A & \xrightarrow{r} & B & \xrightarrow{s} & C \\ \downarrow f & & \downarrow g & & \downarrow h \\ A' & \xrightarrow{r'} & B' & \xrightarrow{s'} & C' \end{array}$$

Solch ein Morphismus heißt ein **Isomorphismus von Sequenzen** genau dann, wenn alle drei vertikalen Abbildungen f, g und h Isomorphismen sind.

7.1.14. Offensichtlich ist mit einer exakten Sequenz auch jede dazu isomorphe Sequenz exakt. Für jede kurze exakte Sequenz von Gruppen $A' \hookrightarrow A \twoheadrightarrow A''$ ist das Bild $N \subset A$ von A' ein Normalteiler und wir erhalten einen Isomorphismus von kurzen exakten Sequenzen

$$\begin{array}{ccccc} N & \hookrightarrow & A & \twoheadrightarrow & A/N \\ \wr\downarrow f & & \| & & \wr\downarrow h \\ A' & \hookrightarrow & A & \twoheadrightarrow & A'' \end{array}$$

indem wir für f die Inverse der von der Einbettung $A' \hookrightarrow A$ induzierten Bijektion $A' \xrightarrow{\sim} N$ nehmen und für h die von der universellen Eigenschaft des Quotienten 6.4.7 induzierte Abbildung. Arbeiten wir speziell mit Vektorräumen, so finden wir mit 1.5.18 in A einen zu N komplementären Teilraum U und nach 1.6.9 induziert die kanonische Abbildung einen Isomorphismus $U \xrightarrow{\sim} A/N$. In diesem Fall erhalten wir also zusätzlich einen Isomorphismus von kurzen exakten Sequenzen

$$\begin{array}{ccccc} N & \hookrightarrow & N \oplus U & \twoheadrightarrow & U \\ \| & & \wr\downarrow g & & \wr\downarrow h \\ N & \hookrightarrow & A & \twoheadrightarrow & A/N \end{array}$$

wobei implizit zu verstehen ist, daß die Morphismen der oberen Horizontale schlicht die kanonische Injektion und Projektion sein sollen. Im Fall von Gruppen, selbst im Fall von abelschen Gruppen, liegen die Verhältnisse komplizierter, wie etwa der Fall der kurzen exakten Sequenz $\mathbb{Z} \hookrightarrow \mathbb{Z} \twoheadrightarrow \mathbb{Z}/2\mathbb{Z}$ zeigt, mit der Multiplikation mit Zwei als erster Abbildung.

7.1.15. Gegeben eine kurze exakte Sequenz von Vektorräumen ist auch die duale Sequenz eine kurze exakte Sequenz. Das ist in der Tat offensichtlich im Fall einer kurzen exakten Sequenz der Gestalt $N \hookrightarrow N \oplus U \twoheadrightarrow U$ und folgt dann mit 7.1.14 im allgemeinen. Speziell ist die Transponierte einer Injektion eine Surjektion und die Transponierte einer Surjektion eine Injektion. Wir verallgemeinern diese Argumente nun noch auf den Fall beliebiger exakter Sequenzen von Vektorräumen.

Proposition 7.1.16. *Jede exakte drei-Term-Sequenz von Vektorräumen ist isomorph zu einer direkten Summe von vier exakten Sequenzen der folgenden vier Typen:*

$$\begin{array}{ccccc} U & \to & 0 & \to & 0 \\ V & \xrightarrow{\mathrm{id}} & V & \to & 0 \\ 0 & \to & W & \xrightarrow{\mathrm{id}} & W \\ 0 & \to & 0 & \to & X \end{array}$$

Beweis. Sei $A \xrightarrow{r} B \xrightarrow{s} C$ unsere Sequenz. Nach 1.5.18 besitzt $(\operatorname{im} r) = (\ker s)$ ein Komplement $W \subset B$, und nach 1.6.9 induziert s einen Isomorphismus $W \xrightarrow{\sim} (\operatorname{im} s)$. Nach 1.5.18 besitzt auch $(\ker r)$ ein Komplement $V \subset A$ und nach 1.6.9 induziert r einen Isomorphismus $V \xrightarrow{\sim} (\operatorname{im} r)$. Wählen wir nun noch ein Komplement $X \subset C$ von $(\operatorname{im} s)$, so induziert die Einbettung folglich einen Isomorphismus zwischen unserer ursprünglichen Sequenz und der direkten Summe der vier Sequenzen

$$\begin{array}{ccccc} (\ker r) & \to & 0 & \to & 0 \\ V & \xrightarrow{\sim} & (\operatorname{im} r) & \to & 0 \\ 0 & \to & W & \xrightarrow{\sim} & (\operatorname{im} s) \\ 0 & \to & 0 & \to & X \end{array}$$

Die Proposition folgt unmittelbar. □

Korollar 7.1.17. *Gegeben eine exakte Sequenz $U \xrightarrow{r} V \xrightarrow{s} W$ von Vektorräumen erhalten wir für jeden weiteren Vektorraum L exakte induzierte Sequenzen*

$$\begin{array}{ccccc} \operatorname{Hom}(W,L) & \xrightarrow{\circ s} & \operatorname{Hom}(V,L) & \xrightarrow{\circ r} & \operatorname{Hom}(U,L) \\ \operatorname{Hom}(L,U) & \xrightarrow{r\circ} & \operatorname{Hom}(L,V) & \xrightarrow{s\circ} & \operatorname{Hom}(L,W) \end{array}$$

Beweis. Das folgt unmittelbar aus der vorhergehenden Proposition 7.1.16 zusammen mit der Erkenntnis, daß das Bilden des Homomorphismenraums "mit endlichen direkten Summen vertauscht". □

Bemerkung 7.1.18. Nehmen wir in diesem Korollar als L den Grundkörper, so folgt insbesondere, daß jede exakte drei-Term-Sequenz beim Dualisieren wieder eine exakte Sequenz liefert.

Übung 7.1.19 (**Additivität der Spur**). Gegeben ein kommutatives Diagramm von endlichdimensionalen Vektorräumen mit zweimal derselben kurzen exakten Zeile

$$\begin{array}{ccccc} V' & \hookrightarrow & V & \twoheadrightarrow & V'' \\ f'\downarrow & & f\downarrow & & f''\downarrow \\ V' & \hookrightarrow & V & \twoheadrightarrow & V'' \end{array}$$

gilt für die Spuren der Vertikalen die Identität $\operatorname{tr}(f) = \operatorname{tr}(f') + \operatorname{tr}(f'')$. Allgemeiner hat im Fall beliebiger Vektorräume der Homomorphismus f endlichen Rang genau dann, wenn f' und f'' endlichen Rang haben, und unter dieser Voraussetzung gilt für die Spuren der Vertikalen wieder $\operatorname{tr}(f) = \operatorname{tr}(f') + \operatorname{tr}(f'')$.

7.2 Tensorprodukte von Vektorräumen

7.2.1. Gegeben eine Menge X und ein Körper k hatten wir in 1.3.17 den freien Vektorraum kX über X eingeführt durch die Vorschrift

$$kX = \left\{ f : X \to k \;\middle|\; \begin{array}{l} f \text{ nimmt nur an endlich vielen Stellen} \\ x \in X \text{ einen Wert ungleich Null an} \end{array} \right\}$$

Wir hatten weiter in 1.5.13 die kanonische Abbildung $\operatorname{can} : X \to kX$ eingeführt, die jedem $x \in X$ die Funktion zuordnet, die Eins ist an der Stelle x und Null sonst. Oft kürzen wir im folgenden $\operatorname{can}(x)$ schlicht mit x ab. Schließlich hatten wir in 1.5.14 die universelle Eigenschaft freier Vektorräume diskutiert, nach der es für jeden k-Vektorraum V und jede Abbildung $f : X \to V$ genau eine k-lineare Abbildung $\tilde{f} : kX \to V$ gibt mit $\tilde{f} \circ \operatorname{can} = f$. In anderen Worten liefert also für jeden k-Vektorraum V und jede Menge X das Vorschalten der kanonischen Einbettung eine Bijektion

$$\operatorname{Hom}_k(kX, V) \xrightarrow[\sim]{\circ\,\operatorname{can}} \operatorname{Ens}(X, V)$$

Definition 7.2.2. Sind V und W zwei Vektorräume über einem Körper k, so definieren wir einen weiteren k-Vektorraum $V \otimes W = V \otimes_k W$, das

Tensorprodukt von V und W, mitsamt einer k-bilinearen Abbildung

$$\begin{array}{rccc} \tau: & V \times W & \to & V \otimes W \\ & (v,w) & \mapsto & v \otimes w \end{array}$$

wie folgt: Wir betrachten die Menge $V \times W$, darüber den freien k-Vektorraum $k(V \times W)$, und darin den Untervektorraum $U \subset k(V \times W)$, der erzeugt wird von allen Ausdrücken

$$\begin{array}{l} \mathrm{can}(v+v',w) - \mathrm{can}(v,w) - \mathrm{can}(v',w) \\ \mathrm{can}(\lambda v,w) - \lambda\,\mathrm{can}(v,w) \\ \mathrm{can}(v,w+w') - \mathrm{can}(v,w) - \mathrm{can}(v,w') \\ \mathrm{can}(v,\lambda w) - \lambda\,\mathrm{can}(v,w) \end{array}$$

für $v, v' \in V$, $w, w' \in W$ und $\lambda \in k$. Dann definieren wir unser Tensorprodukt als den Quotientenvektorraum

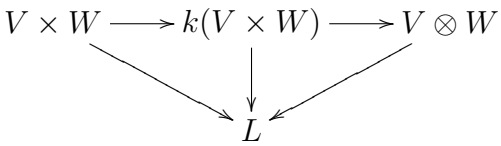
$$V \otimes W = k(V \times W)/U$$

und erklären $v \otimes w$ als die Nebenklasse von $\mathrm{can}(v,w)$. Die Bilinearität von $(v,w) \mapsto v \otimes w$ folgt sofort aus der Definition des herausgeteilten Untervektorraums U.

Satz 7.2.3 (Universelle Eigenschaft des Tensorprodukts). *Gegeben ein Körper k, Vektorräume V, W, L über k und eine k-bilineare Abbildung $b: V \times W \to L$ existiert genau eine k-lineare Abbildung $\hat{b}: V \otimes W \to L$ mit $b(v,w) = \hat{b}(v \otimes w) \quad \forall v \in V, w \in W$. In anderen Worten liefert also das Vorschalten der kanonischen Abbildung eine Bijektion*

$$\mathrm{Hom}_k(V \otimes W, L) \xrightarrow{\circ\tau}_{\sim} \mathrm{Hom}_k^{(2)}(V \times W, L)$$

Beweis. Wir arbeiten mit dem Diagramm



Für jede Abbildung $b: V \times W \to L$ gibt es nach 7.2.1 genau eine k-lineare Abbildung $\tilde{b}: k(V \times W) \to L$ mit $\tilde{b} \circ \mathrm{can} = b$. Ist b bilinear, so gilt offensichtlich $\tilde{b}(U) = 0$, also gibt es $\hat{b}: V \otimes W \to L$ linear mit $\hat{b}(v \otimes w) = b(v,w)$. Diese Abbildung $\hat{b}$ ist eindeutig bestimmt durch b, da die $v \otimes w$ ja das Tensorprodukt als Vektorraum erzeugen. □

7.2.4. Das Tensorprodukt wird durch seine universelle Eigenschaft bereits bis auf eindeutigen Isomorphismus festgelegt. Seien genauer gegeben ein Körper k, Vektorräume V, W über k und eine k-bilineare Abbildung $c : V \times W \to T$ in einen weiteren k-Vektorraum T mit der Eigenschaft, daß für jede k-bilineare Abbildung $b : V \times W \to L$ in einen k-Vektorraum L genau eine k-lineare Abbildung $\tilde{b} : T \to L$ existiert mit $b(v, w) = \tilde{b}(c(v, w)) \quad \forall v \in V, w \in W$. So ist die Abbildung $\hat{c}$ ein Isomorphismus

$$\hat{c} : V \otimes W \xrightarrow{\sim} T$$

Um das einzusehen, betrachten wir das nebenstehende Diagramm. Notieren wir etwa $\tau : V \times W \to V \otimes W$ unsere kanonische bilineare Abbildung $(v, w) \mapsto v \otimes w$, so ist $\tilde{\tau}$ invers zu $\hat{c}$. In der Tat gilt $\tilde{\tau} \circ \hat{c} \circ \tau = \tilde{\tau} \circ c = \tau$ und aus der universellen Eigenschaft von τ folgt, daß es nur eine lineare Abbildung $f = \hat{\tau} : V \otimes W \to V \otimes W$ geben darf mit $f \circ \tau = \tau$. Da nun die Identität auf dem Tensorprodukt eine mögliche derartige lineare Abbildung ist, folgt schon einmal $\mathrm{id} = \tilde{\tau} \circ \hat{c}$. Weiter gilt $\hat{c} \circ \tilde{\tau} \circ c = \hat{c} \circ \tau = c$ und aus der universellen Eigenschaft von c folgt, daß es nur eine lineare Abbildung $g = \tilde{c} : T \to T$ geben darf mit $g \circ c = c$. Da nun die Identität auf T eine mögliche derartige lineare Abbildung ist, folgt auch $\tilde{c} = \mathrm{id} = \hat{c} \circ \tilde{\tau}$.

7.2.5. Da es uns eigentlich nur auf die universelle Eigenschaft ankommt, kann man natürlich auch andere Konstruktionen versuchen. Ist V endlichdimensional, so hat auch der Raum $T_1 = \mathrm{Hom}_k(V^*, W)$ mit der bilinearen Abbildung $c_1 : V \times W \to T_1$, $(v, w) \mapsto (f \mapsto f(v)w)$ die geforderte universelle Eigenschaft, und ist darüber hinaus auch W endlichdimensional, so gilt dasselbe für das Paar (T_2, c_2) mit $T_2 = \mathrm{Hom}^{(2)}(V^* \times W^*, k)$ und $c_2 : (v, w) \mapsto ((f, g) \mapsto f(v)g(w))$. In vielen Quellen wählt man diese Konstruktion zur Definition des Tensorprodukts, vermutlich mit dem Ziel, Quotientenvektorräume zu vermeiden. Das geschieht dann aber doch um den Preis einer eingeschränkten Allgemeinheit. Daß es durchaus sinnvoll und nützlich sein kann, auch zwei unendlichdimensionale Vektorräume tensorieren zu können, mögen die Übungen 7.2.15 und 7.2.16 illustrieren.

7.2.6. Ein in der Literatur auch oft beschrittener Zugang, der sogar Tensorprodukte unendlichdimensionaler Räume liefert, geht so: Man wählt Basen $A \subset V$ und $B \subset W$, setzt $T_3 = k(A \times B)$ und erklärt die kanonische bilineare Abbildung $c_3 : V \times W \to T_3$ als die eindeutig bestimmte bilineare Abbildung mit $(a, b) \mapsto (a, b)$ für alle $a \in A$, $b \in B$.Im zweiten Beweis von 7.2.9 führen wir aus, warum auch diese Konstruktion eine bilineare Abbildung mit der das Tensorprodukt charakterisierenden universellen Eigenschaft liefert. Ich mag diesen Zugang nicht aus den folgenden zwei Gründen: Erstens hängt sie von

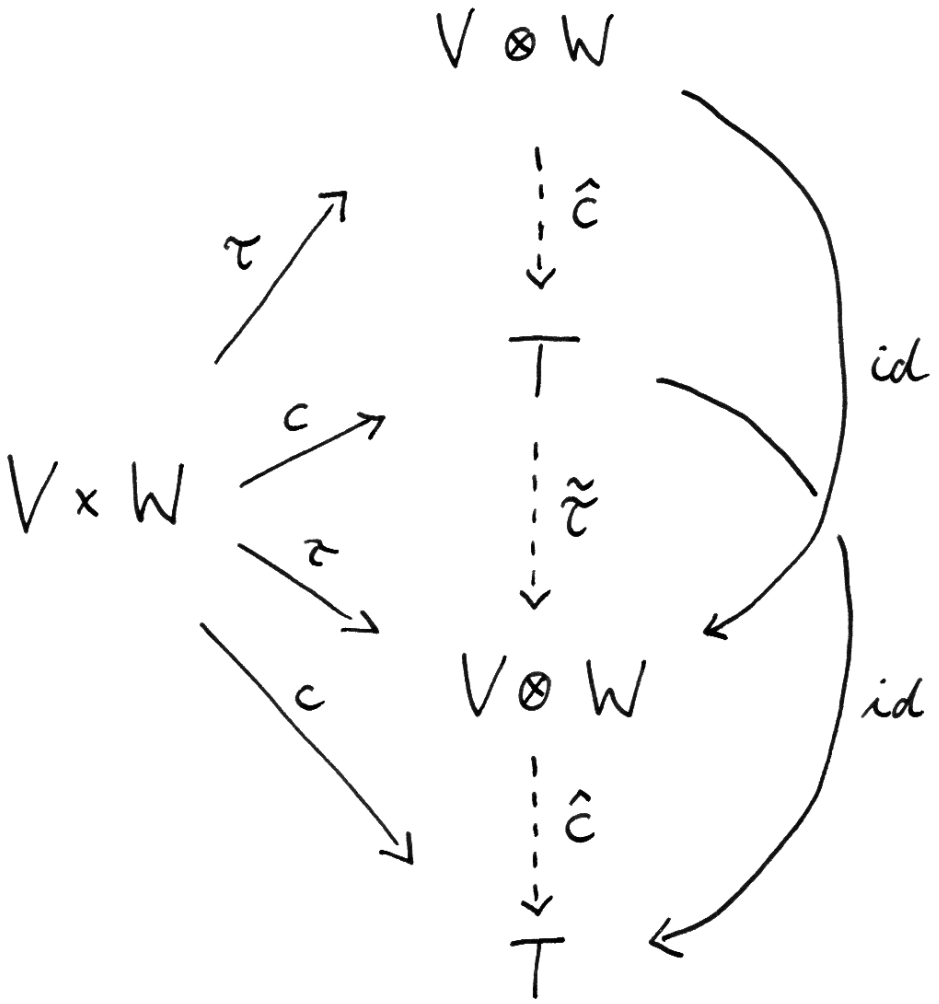


Illustration zum Beweis der Festlegbarkeit des Tensorprodukts bis auf eindeutigen Isomorphismus durch seine universelle Eigenschaft.

den gewählten Basen ab, so daß wir nicht eigentlich einen Vektorraum sondern vielmehr eine durch die möglichen Wahlen von Basen indizierte Familie von paarweise in kanonischer und verträglicher Weise isomorphen Vektorräumen erhalten. Und zweitens läßt sich diese Konstruktion im Gegensatz zu der hier in diesem Text gewählten Konstruktion nicht unmittelbar zu einer Konstruktion von Tensorprodukten über Ringen **??** verallgemeinern.

7.2.7. Keineswegs jedes Element eines Tensorprodukts ist von der Form $v \otimes w$, die Elemente dieser Gestalt erzeugen jedoch das Tensorprodukt als Vektorraum und sogar als abelsche Gruppe. Besitzt weiter ein Element eines Tensorprodukts eine Darstellung in der Gestalt $v \otimes w$, so besitzt es meist sogar viele verschiedene Darstellungen dieser Gestalt. Geben wir eine Abbildung von einem Tensorprodukt in einen Vektorraum L an durch eine Vorschrift der Gestalt $v \otimes w \mapsto b(v, w)$, so ist der Leser implizit gefordert, die Bilinearität der Abbildung $b : V \times W \to L$ zu prüfen, und gemeint ist die durch die universelle Eigenschaft definierte Abbildung $\hat{b} : V \otimes W \to L$.

Definition 7.2.8. Sind $f : V \to V'$ und $g : W \to W'$ lineare Abbildungen, so definieren wir eine lineare Abbildung $f \otimes g : V \otimes W \to V' \otimes W'$ durch die Vorschrift $(f \otimes g)(v \otimes w) = f(v) \otimes g(w)$.

Lemma 7.2.9 (Basen von Tensorprodukten). *Ist $v_1, \ldots, v_n$ eine Basis von V und $w_1, \ldots w_m$ eine Basis von W, so bilden die $v_i \otimes w_j$ eine Basis von $V \otimes W$. Insbesondere gilt im endlichdimensionalen Fall*

$$\dim_k(V \otimes W) = (\dim_k V)(\dim_k W)$$

7.2.10. Dieselbe Aussage gilt mit demselben Beweis auch für nicht notwendig endliche Basen.

7.2.11. Ein Spezialfall von Tensorprodukten ist im Übrigen das "Rechnen mit Einheiten": Ist zum Beispiel L der eindimensionale "$\mathbb{R}$-Vektorraum aller Längen", so kann $L \otimes L$ interpretiert werden als der eindimensionale "$\mathbb{R}$-Vektorraum aller Flächen", und ist m eine Basis von L, so ist $m \otimes m$ diejenige Basis von $L \otimes L$, die man üblicherweise m^2 notiert.

Erster Beweis. Nur die lineare Unabhänigkeit der $v_i \otimes w_j$ ist nicht auf Anhieb klar. Aber sind $v_1^*, \ldots, v_n^* \in V^*$ und $w_1^*, \ldots, w_m^* \in W^*$ die Vektoren der dualen Basen und betrachten wir die bilinearen Abbildungen

$$\begin{array}{rccc} b_{ij} : & V \times W & \to & k \\ & (v, w) & \mapsto & v_i^*(v) w_j^*(w) \end{array}$$

so ist $\hat{b}_{ij} : V \otimes W \to k$ eine lineare Abbildung, die $v_i \otimes w_j$ auf 1 abbildet und alle anderen $v_\nu \otimes w_\mu$ auf Null. □

Zweiter Beweis. Wir führen den zweiten Beweis gleich für den Fall beliebiger Dimension. Sind $A \subset V$ und $B \subset W$ Basen, so liefern nach 1.5.24 und 7.2.1 die Einschränkung bzw. das Vorschalten der kanonischen Einbettung Bijektionen

$$\mathrm{Hom}_k^{(2)}(V \times W, L) \xrightarrow{\sim} \mathrm{Ens}(A \times B, L) \xleftarrow{\sim} \mathrm{Hom}_k(k(A \times B), L)$$

Bezeichnet $c : V \times W \to k(A \times B)$ die bilineare Abbildung, die unter diesen Isomorphismen der Identität auf $k(A \times B)$ entspricht, so kann man unschwer einsehen, daß c auch die von einem Tensorprodukt geforderte universelle Eigenschaft hat. Dann muß jedoch nach 7.2.4 die durch diese universelle Eigenschaft gegebene Abbildung einen Isomorphismus $\tilde{\tau} : k(A \times B) \xrightarrow{\sim} V \otimes W$ liefern, und man prüft leicht, daß dieser Isomorphismus $(a, b) \in k(A \times B)$ auf $a \otimes b \in V \otimes W$ abbildet. Aus 1.4.11 folgt dann unmittelbar, daß die $a \otimes b$ eine Basis von $V \otimes W$ bilden. □

7.2.12. Gegeben Vektorräume V, W mit angeordneten Basen $\mathcal{A} = (v_1, \ldots v_n)$ und $\mathcal{B} = (w_1, \ldots, w_n)$ bilden wir in $V \otimes W$ die angeordnete Basis

$$\begin{aligned} \mathcal{A} \otimes \mathcal{B} \quad = \quad & (v_1 \otimes w_1, v_1 \otimes w_2, \ldots, v_1 \otimes w_m, \\ & v_2 \otimes w_1, v_2 \otimes w_2, \ldots, v_2 \otimes w_m, \\ & \qquad \ldots \quad \ldots \quad \ldots \\ & v_n \otimes w_1, v_n \otimes w_2, \ldots, v_n \otimes w_m) \end{aligned}$$

Gegeben zusätzlich weitere Vektorräume V', W' mit angeordneten Basen $\mathcal{A}' = (v'_1, \ldots, v'_{n'})$ und $\mathcal{B}' = (w'_1, \ldots, w'_{m'})$ und lineare Abbildungen $f : V \to V'$ und $g : W \to W'$ können wir die Matrix ${}_{\mathcal{A}' \otimes \mathcal{B}'}[f \otimes g]_{\mathcal{A} \otimes \mathcal{B}}$ von $f \otimes g$ in den Basen $\mathcal{A} \otimes \mathcal{B}$ und $\mathcal{A}' \otimes \mathcal{B}'$ wie folgt durch die Matrizen $A = {}_{\mathcal{A}'}[f]_{\mathcal{A}}$ und $B = {}_{\mathcal{B}'}[g]_{\mathcal{B}}$ ausdrücken: Haben wir etwa $A = (a_{ij})$, so wird

$${}_{\mathcal{A}' \otimes \mathcal{B}'}[f \otimes g]_{\mathcal{A} \otimes \mathcal{B}} \quad = \quad \begin{pmatrix} a_{11}B & \ldots & a_{1n}B \\ \vdots & & \vdots \\ a_{n'1}B & \ldots & a_{n'n}B \end{pmatrix}$$

Auf der rechten Seite ist hier die Matrix geblockt geschrieben, es handelt sich ja eigentlich um eine $(n'm' \times nm)$-Matrix. Sie heißt auch das **Kronecker-Produkt** der Matrizen A und B und wird $A \otimes B$ notiert, so daß wir unsere Identität oben also auch schreiben können in der Form

$${}_{\mathcal{A}' \otimes \mathcal{B}'}[f \otimes g]_{\mathcal{A} \otimes \mathcal{B}} = {}_{\mathcal{A}'}[f]_{\mathcal{A}} \otimes {}_{\mathcal{B}'}[g]_{\mathcal{B}}$$

Um diese Identität zu prüfen beginnen wir mit

$$f(v_i) = \sum_j a_{ji} v'_j \quad \text{und} \quad g(w_k) = \sum_l b_{lk} w'_l$$

und folgern

$$(f \otimes g)(w_i \otimes w_k) = \sum_{j,l} a_{ji} b_{lk} v'_j \otimes w'_l$$

Die Einträge der Matrix von $f \otimes g$ sind also alle Produkte von einem Eintrag der Matrix von f mit einem Eintrag der Matrix von g. Daß diese Einträge dann auch noch an den oben beschriebenen Stellen der Matrix von $f \otimes g$ stehen, mag sich der Leser am einfachsten selbst überlegen.

Proposition 7.2.13. *Gegeben Vektorräume U, V, W erhalten wir einen Isomorphismus*

$$\mathrm{Hom}(U, \mathrm{Hom}(V, W)) \xrightarrow{\sim} \mathrm{Hom}(U \otimes V, W)$$

durch die Vorschrift $f \mapsto \tilde{f}$ mit $\tilde{f}(u \otimes v) = (f(u))(v)$.

Beweis. Beide Seiten sind in offensichtlicher und mit der angegebenen Abbildung verträglicher Weise in Bijektion zur Menge $\mathrm{Hom}^{(2)}(U \times V, W)$ aller bilinearen Abbildungen $U \times V \to W$. □

Übung 7.2.14. Für jeden k-Vektorraum V definiert das Auswerten oder lateinisierend "Evaluieren" eine lineare Abbildung $\mathrm{ev} : V^* \otimes V \to k$, $\xi \otimes v \mapsto \xi(v)$. Man zeige, daß sie unter dem Isomorphismus aus 7.2.13 der Identität auf V^* entspricht.

Übung 7.2.15. Gegeben ein Körper k induziert die Multiplikation einen Isomorphismus $k[X] \otimes_k k[Y] \xrightarrow{\sim} k[X, Y]$.

Übung 7.2.16. Die Multiplikation induziert einen Isomorphismus von reellen Vektorräumen $\mathbb{R} \otimes_{\mathbb{Q}} \mathbb{Q}[X] \xrightarrow{\sim} \mathbb{R}[X]$. Analoges gilt, wenn man $\mathbb{R} \supset \mathbb{Q}$ ersetzt durch ein beliebiges Paar $K \supset k$ bestehend aus einem Körper mit einem Teilkörper.

Definition 7.2.17. Induktiv bilden wir auch längere Tensorprodukte durch die Vorschrift $V_1 \otimes \ldots \otimes V_{n-1} \otimes V_n = (V_1 \otimes \ldots \otimes V_{n-1}) \otimes V_n$ und definieren darin die Vektoren $v_1 \otimes \ldots \otimes v_n = (v_1 \otimes \ldots \otimes v_{n-1}) \otimes v_n$.

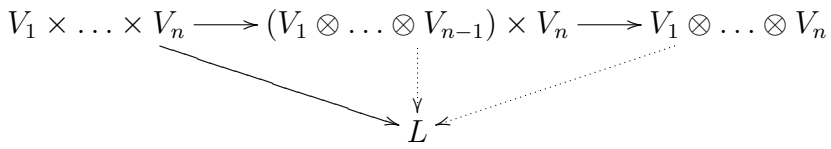
Proposition 7.2.18. *1. Gegeben Vektorräume $V_1, \ldots, V_n$ erhalten wir durch das Tensorieren von Vektoren eine multilineare Abbildung*

$$\begin{array}{rcl} V_1 \times \ldots \times V_n & \to & V_1 \otimes \ldots \otimes V_n \\ (v_1, \ldots, v_n) & \mapsto & v_1 \otimes \ldots \otimes v_n \end{array}$$

2. Ist L ein weiterer k-Vektorraum und $F : V_1 \times \ldots \times V_n \to L$ eine beliebige multilineare Abbildung, so gibt es genau eine lineare Abbildung $\hat{F} : V_1 \otimes \ldots \otimes V_n \to L$ mit $F(v_1, \ldots, v_n) = \hat{F}(v_1 \otimes \ldots \otimes v_n)$ für alle $v_1 \in V_1, \ldots, v_n \in V_n$.

7.2.19. Damit diese Proposition auch im Fall $n = 0$ gilt, vereinbaren wir für das Tensorprodukt mit überhaupt keinem Faktor, daß damit schlicht der Grundkörper gemeint sein soll, und als multilineare Abbildung des kartesischen Produkts von Vektorräumen mit überhaupt keinem Faktor, das ja nach unseren Konventionen "die" einpunktige Menge ist, in sein Tensorprodukt vereinbaren wir die Abbildung, die diesen einzigen Punkt auf $1 \in k$ wirft.

Beweis. Mit Induktion über n. Wir argumentieren mit dem Diagramm



mit hoffentlich offensichtlichen horizontalen Morphismen. Wir zeigen nur die Existenz von $\hat{F}$ und überlassen den Nachweis der anderen Behauptungen dem Leser. Für jedes feste $v_n \in V_n$ ist die Abbildung $(v_1, \ldots, v_{n-1}) \mapsto F(v_1, \ldots, v_n)$ multilinear und induziert so nach Induktionsannahme eine lineare Abbildung $V_1 \otimes \ldots \otimes V_{n-1} \to L$. Das liefert die mittlere Vertikale. Man überzeugt sich nun leicht, daß die mittlere Vertikale auch in $v_n \in V_n$ linear sein muß. Als bilineare Abbildung faktorisiert sie also über $V_1 \otimes \ldots \otimes V_n$. □

Proposition 7.2.20. *1. Sind V, W Vektorräume, so ist die lineare Abbildung $V \otimes W \to W \otimes V$ gegeben durch $v \otimes w \mapsto w \otimes v$ ein Isomorphismus.*

2. Sind $V_1, \ldots, V_p, V_{p+1}, \ldots, V_n$ Vektorräume, so ist die lineare Abbildung $(V_1 \otimes \ldots \otimes V_p) \otimes (V_{p+1} \otimes \ldots \otimes V_n) \to V_1 \otimes \ldots \otimes V_p \otimes V_{p+1} \otimes \ldots \otimes V_n$ mit $(v_1 \otimes \ldots \otimes v_p) \otimes (v_{p+1} \otimes \ldots \otimes v_n) \mapsto v_1 \otimes \ldots \otimes v_p \otimes v_{p+1} \otimes \ldots \otimes v_n$ ein Isomorphismus.

3. Die Abbildung $\lambda \otimes v \mapsto \lambda v$ liefert einen Isomorphismus

$$k \otimes V \xrightarrow{\sim} V$$

vom Tensorprodukt des Grundkörpers mit einem Vektorraum in besagten Vektorraum selber. Analoges gilt für den zweiten Tensorfaktor.

4. Gegeben Vektorräume W, V, V' liefert $w \otimes (v, v') \mapsto (w \otimes v, w \otimes v')$ einen Isomorphismus

$$W \otimes (V \oplus V') \xrightarrow{\sim} (W \otimes V) \oplus (W \otimes V')$$

Analoges gilt auch für den ersten Tensorfaktor.

Beweis. Man kann in allen diesen Fällen leicht einsehen, daß die betrachteten Abbildungen eine geeignete Basis des Ausgangsraums bijektiv mit einer Basis des Zielraums identifizieren. Die Details seien dem Leser überlassen. □

7.2.21. In gewisser Weise liefert das die Kommutativität, Assoziativität und das neutrale Element für das Tensorprodukt, nebst der Distributivität mit direkten Summen. In größerer Abstraktion formalisiert das der Formalismus der "Tensorkategorie", wie er in **??** diskutiert wird.

Übung 7.2.22. Gegeben ein Vektorraum V und eine Familie von Vektorräumen $(W_i)_{i\in I}$ und liefert die kanonische Abbildung stets einen Isomorphismus

$$V \otimes \left(\bigoplus W_i\right) \xrightarrow{\sim} \bigoplus (V \otimes W_i)$$

Analoges gilt für den anderen Tensorfaktor.

Übung 7.2.23. Gegeben Vektorräume U, V, W kommutiert das Diagramm

$$\begin{array}{ccc} U^* \otimes V \otimes V^* \otimes W & \hookrightarrow & \operatorname{Hom}(U,V) \otimes \operatorname{Hom}(V,W) \\ \downarrow & & \downarrow \\ U^* \otimes W & \hookrightarrow & \operatorname{Hom}(U,W) \end{array}$$

mit den von den kanonischen Injektionen 7.3.1 induzierten Horizontalen, dem Verknüpfen von linearen Abbildungen als rechter Vertikale, und in der linken Vertikalen der Verknüpfung

$$U^* \otimes V \otimes V^* \otimes W \to U^* \otimes k \otimes W \xrightarrow{\sim} U^* \otimes W$$

der vom Auswerten $\mathrm{id} \otimes \mathrm{ev} \otimes \mathrm{id}$ induzierten Abbildung gefolgt vom Isomorphismus $U^* \otimes k \otimes W \xrightarrow{\sim} U^* \otimes W$, $f \otimes \lambda \otimes w \mapsto \lambda(f \otimes w)$. Die obige Verknüpfung bezeichnet man in dieser und ähnlichen Situationen auch als die **Verjüngung von Tensoren**.

Übung 7.2.24. Gegeben Körper $k \subset K$ und ein k-Vektorraum V wird $V_K = K \otimes_k V$ in offensichtlicher Weise ein K-Vektorraum. Man sagt, er entstehe aus V durch **Erweiterung der Skalare**. Die "kanonische" k-lineare Abbildung $\mathrm{can} : V \to V_K, v \mapsto 1 \otimes v$ hat dann die universelle Eigenschaft, daß für jeden K-Vektorraum W das Vorschalten von can eine Bijektion

$$\operatorname{Hom}_K(V_K, W) \xrightarrow{\sim} \operatorname{Hom}_k(V, W)$$

liefert. Weiter ist das Bild unter can jeder k-Basis von V eine K-Basis von V_K, und gegeben ein weiterer k-Vektorraum W induziert die k-lineare Abbildung $V \otimes_k W \to V_K \otimes_K W_K$, $v \otimes w \mapsto \mathrm{can}(v) \otimes \mathrm{can}(w)$ einen Isomorphismus

$$(V \otimes_k W)_K \xrightarrow{\sim} V_K \otimes_K W_K$$

Ist V oder K endlichdimensional über k, so liefert auch die k-lineare Abbildung $\mathrm{Hom}_k(V,W) \to \mathrm{Hom}_K(V_K, W_K)$ gegeben durch $f \mapsto \mathrm{id} \otimes f$ einen Isomorphismus

$$(\mathrm{Hom}_k(V,W))_K \xrightarrow{\sim} \mathrm{Hom}_K(V_K, W_K)$$

und insbesondere "vertauscht unter dieser Endlichkeitsannahme das Erweitern der Skalare mit dem Bilden des Dualraums". Im Spezialfall $\mathbb{R} \subset \mathbb{C}$ bezeichnet man $V_{\mathbb{C}}$ als die **Komplexifizierung** von V. Ein alternativer vom Tensorprodukt unabhängiger Zugang zur Komplexifizierung wird in **??** erklärt.

7.3 Kanonische Injektionen bei Tensorprodukten

Satz 7.3.1. *Für beliebig vorgegebene k-Vektorräume V, W liefert die Vorschrift $f \otimes w \mapsto (v \mapsto f(v)w)$ eine Injektion*

$$\mathrm{can} : V^* \otimes W \hookrightarrow \mathrm{Hom}(V,W)$$

Sind V oder W endlichdimensional, so ist diese Injektion ein Isomorphismus. Im allgemeinen besteht ihr Bild genau aus allen Homomorphismen endlichen Ranges.

7.3.2. Ist $v_1, \ldots, v_n$ eine Basis von V und $v_1^*, \ldots, v_n^* \in V^*$ die duale Basis von V^*, so können wir im Satz die inverse Abbildung angeben durch die Vorschrift $f \mapsto v_1^* \otimes f(v_1) + \ldots + v_n^* \otimes f(v_n)$. Ist zusätzlich $w_1, \ldots, w_m$ eine Basis von W, so wird $v_i^* \otimes w_j$ abgebildet auf diejenige lineare Abbildung, deren Matrix in Bezug auf die gegebenen Basen die Basismatrix E_{ji} aus 1.7.13 ist. Im Fall endlichdimensionaler Räume kann der Satz also leicht mit Basen überprüft werden. Der Beweis gilt den unendlichdimensionalen Fällen.

Beweis. Daß das Bild unserer kanonischen Abbildung enthalten ist in der Menge aller Abbildungen endlichen Ranges scheint mir offensichtlich. Den Nachweis, daß auch jede Abbildung endlichen Ranges im Bild liegt, überlasse ich dem Leser und zeige nur die Injektivität. Es reicht zu zeigen, daß für $f_1, \ldots, f_n \in V^*$ und $w_1, \ldots, w_m \in W$ jeweils linear unabhängig die Bilder der $f_i \otimes w_s$ eine linear unabhängige Familie in $\mathrm{Hom}(V,W)$ bilden. Um das zu zeigen suche man $v_1, \ldots, v_n \in V$ mit $f_i(v_j) = \delta_{ij}$, etwa mithilfe von 1.8.6, und $g_1, \ldots, g_m \in W^*$ mit $g_t(w_s) = \delta_{ts}$. Gegeben eine verschwindende Linearkombination im Hom-Raum $0 = \sum c_{is}\,\mathrm{can}(f_i \otimes w_s)$ folgt dann

$$0 = g_t\left(\left(\sum c_{is}\,\mathrm{can}(f_i \otimes w_s)\right)(v_j)\right) = c_{jt} \quad \forall j,t \qquad \square$$

Übung 7.3.3. Gegeben ein endlichdimensionaler k-Vektorraum V kommutiert das Diagramm

$$\begin{array}{ccccc} \sum_{i=1}^{n} v_i \otimes v_i^* & \in & V \otimes V^* & \xrightarrow{\text{ev}} & k \\ \downarrow & & \downarrow\wr & & \| \\ \mathrm{id} & \in & \mathrm{End}\, V & \xrightarrow{\text{tr}} & k \end{array}$$

wo $v_1, \ldots, v_n$ eine beliebige Basis von V bedeuten möge und $v_1^*, \ldots, v_n^*$ die duale Basis von V^* meint und die mittlere Vertikale die in 7.3.1 erklärte Injektion und ev das Auswerten im Sinne von 7.2.14. Hat V unendliche Dimension, so kommutiert das rechte Quadrat immer noch, wenn wir unten links nur Endomorphismen endlichen Ranges betrachten und ihre Spur wie in 1.7.38 nehmen. Allerdings ist dann unser Tensorausdruck nicht mehr sinnvoll definiert und die Identität gehört auch nicht mehr zu den Endomorphismen endlichen Ranges.

Korollar 7.3.4 (Tensorprodukt und Dualität). *Gegeben Vektorräume V, W liefert die Abbildung $f \otimes g \mapsto (v \otimes w \mapsto f(v)g(w))$ eine Injektion*

$$V^* \otimes W^* \hookrightarrow (V \otimes W)^*$$

vom Tensorprodukt der Dualräume in den Dualraum des Tensorprodukts. Sie ist ein Isomorphismus genau dann, wenn einer unserer beiden Räume endlichdimensional ist.

Bemerkung 7.3.5. Im Fall endlichdimensionaler Räume kann das leicht mit Basen überprüft werden. Der Beweis gilt den unendlichdimensionalen Fällen.

Beweis. Wir argumentieren mit dem Diagramm

$$\begin{array}{ccc} (V \otimes W)^* & = & \mathrm{Hom}(V \otimes W, k) \\ & & \uparrow\wr \\ & & \mathrm{Hom}(V, \mathrm{Hom}(W, k)) \\ & & \| \\ V^* \otimes W^* & \hookrightarrow & \mathrm{Hom}(V, W^*) \end{array}$$

Der vertikale Isomorphismus kommt aus 7.2.13, die horizontale Injektion aus 7.3.1. Daß deren Komposition genau die im Korollar beschriebene Abbildung ist, mag der Leser selbst prüfen. □

Übung 7.3.6. (Hinweis: 7.3.4.) Gegeben Mengen X, Y und ein beliebiger Körper k liefert die offensichtliche Abbildung eine Injektion

$$\mathrm{Ens}(X, k) \otimes \mathrm{Ens}(Y, k) \hookrightarrow \mathrm{Ens}(X \times Y, k)$$

Übung 7.3.7. Gegeben ein Vektorraum V und eine Familie von Vektorräumen $(W_i)_{i\in I}$ und liefert die kanonische Abbildung stets eine Injektion

$$V \otimes \left(\prod W_i\right) \hookrightarrow \prod (V \otimes W_i)$$

die jedoch im allgemeinen kein Isomorphismus ist. Genauer ist sie nur ein Isomorphismus, falls entweder V endlichdimensional ist oder nur für endlich viele i der zugehörige Vektorraum W_i von Null verschieden ist. Hinweis: Man folgere aus 7.3.1 die Injektivität der Komposition $V \otimes W \to V^{**} \otimes W \to \mathrm{Hom}(V^*, W)$ und bette beide Seiten verträglich ein in $\prod \mathrm{Hom}(V^*, W_i) \cong \mathrm{Hom}(V^*, \prod W_i)$.

Übung 7.3.8. Gegeben Vektorräume V, V', W, W' liefert das Tensorieren von Abbildungen eine Injektion

$$\mathrm{Hom}(V, V') \otimes \mathrm{Hom}(W, W') \hookrightarrow \mathrm{Hom}(V \otimes W, V' \otimes W')$$

Hinweis: Man mag V und W als direkte Summe eindimensionaler Räume schreiben und 7.3.7 anwenden. Alternative: Man mag ohne Beschränkung der Allgemeinheit annehmen, daß unsere Vektorräume frei sind über den Mengen X, X', Y, Y', und kann dann unter Verwendung von 7.3.6 beide Seiten in verträglicher Weise einbetten in den Raum $\mathrm{Ens}(X \times X' \times Y \times Y', k)$ aller Abbildungen von besagtem Produkt in den Grundkörper k.

7.4 Alternierende Tensoren und Determinante

7.4.1. Gegeben ein Körper k und ein k-Vektorraum V und eine natürliche Zahl $r \in \mathbb{N}$ vereinbaren wir

$$V^{\otimes r} = \underbrace{V \otimes \ldots \otimes V}_{r \text{ Faktoren}}$$

und verstehen $V^{\otimes 0} = k$ im Sinne der vorhergehenden Bemerkung 7.2.19. Gegeben eine lineare Abbildung $f : V \to W$ verwenden wir weiter die Abkürzung $(f \otimes \ldots \otimes f) = f^{\otimes r} : V^{\otimes r} \to W^{\otimes r}$. Gegeben $v \in V$ schreiben wir kurz $(v \otimes \ldots \otimes v) = v^{\otimes r}$ für das Bild von $(v, \ldots, v)$ unter der kanonischen multilinearen Abbildung $V^r \to V^{\otimes r}$ und verstehen insbesondere $v^{\otimes 0} = 1$.

7.4.2. Ich erinnere daran, daß wir in 2.7.14 die Determinante

$$\det : \mathrm{M}(n \times n; k) \to k$$

charakterisiert hatten als die eindeutig bestimmte multilineare alternierende Funktion der Spaltenvektoren, die der Einheitsmatrix die Eins zuordnet.

Gegeben ein beliebiger k-Vektorraum V und $r \geq 0$ setzen wir nun

$$\mathrm{Alt}^r(V) = \{f : \underbrace{V \times \ldots \times V}_{r \text{ Faktoren}} \to k \mid f \text{ ist multilinear und alternierend}\}$$

Im Spezialfall $r = 0$ ist das leere Produkt als einpunktige Menge zu verstehen und $\mathrm{Alt}^0(V)$ als die Menge aller Abbildungen von dieser einpunktigen Menge nach k, so daß das Auswerten einen kanonischen Isomorphismus $\mathrm{Alt}^0(V) \xrightarrow{\sim} k$ definiert. Wir fassen diesen Isomorphismus in unserer Notation hinfort als Gleichheit $\mathrm{Alt}^0(V) = k$ auf. Bezeichnet $J_r \subset V^{\otimes r}$ das Erzeugnis aller Tensoren mit zwei gleichen Einträgen, so liefert das Vorschalten der Verknüpfung $V \times \ldots \times V \to V^{\otimes r} \to V^{\otimes r}/J_r$ der kanonischen multilinearen Abbildung mit der kanonischen Abbildung auf den Quotienten aufgrund universellen Eigenschaften Isomorphismen

$$\mathrm{Alt}^r(V) \xleftarrow{\sim} \{g \in \mathrm{Hom}(V^{\otimes r}, k) \mid g(J_r) = 0\} \xleftarrow{\sim} \mathrm{Hom}(V^{\otimes r}/J_r, k)$$

Der Quotient $V^{\otimes r}/J_r$ heißt die **r-te äußere Potenz von** V und wird für gewöhnlich

$$V^{\otimes r}/J_r = \bigwedge^r V$$

notiert. Mit dieser Notation haben wir also einen kanonischen Isomorphismus $(\bigwedge^r V)^* \xrightarrow{\sim} \mathrm{Alt}^r(V)$ erhalten. Im Extremfall $r = 0$ verstehen wir hier wieder $V^{\otimes 0} = \bigwedge^0 V = k$ und unsere Aussage behält ihre Gültigkeit.

Übung 7.4.3. Das Erzeugnis $J_r \subset V^{\otimes r}$ aller Tensoren mit zwei gleichen Einträgen fällt zusammen mit dem Erzeugnis $J'_r \subset V^{\otimes r}$ aller Tensoren mit zwei benachbarten gleichen Einträgen.

Übung 7.4.4. Gegeben ein Vektorraum V definiert jede Permutation $\sigma \in \mathcal{S}_r$ einen Endomorphismus $[\sigma] : V^{\otimes r} \xrightarrow{\sim} V^{\otimes r}$ durch das "Permutieren der Tensorfaktoren", in Formeln $[\sigma] : v_1 \otimes \ldots \otimes v_r \mapsto v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(r)}$. Man zeige, daß diese Endomorphismen eine Rechtsoperation definieren, in Formeln $[\sigma] \circ [\tau] = [\tau \circ \sigma]$. Unter **alternierenden Tensoren** versteht man diejenigen Elemente von $V^{\otimes r}$, die beim Vertauschen von zwei Tensorfaktoren ihr Vorzeichen wechseln, in Formeln die Elemente des Teilraums

$$(V^{\otimes r})^{\mathrm{sgn}} = \{t \in V^{\otimes r} \mid [\sigma](t) = (\mathrm{sgn}\,\sigma)t \; \forall \sigma \in \mathcal{S}_r\}$$

Man zeige, daß im Fall eines Grundkörpers der Charakteristik Null die kanonische Projektion $V^{\otimes r} \twoheadrightarrow \bigwedge^r V$ einen Isomorphismus $(V^{\otimes r})^{\mathrm{sgn}} \xrightarrow{\sim} \bigwedge^r V$ induziert. Man zeige weiter, daß im Fall eines Grundkörpers der Charakteristik Null der **Alternator** $\mathrm{alt} : V^{\otimes r} \to (V^{\otimes r})^{\mathrm{sgn}}$ gegeben durch

$$t \mapsto \frac{1}{r!} \sum_{\sigma \in \mathcal{S}_r} \mathrm{sgn}(\sigma)[\sigma](t)$$

eine Projektion im Sinne von 1.6.4 ist.

7.4.5. Gegeben $r, s \geq 0$ gibt es genau eine Abbildung

$$\bigwedge^r V \times \bigwedge^s V \;\to\; \bigwedge^{r+s} V$$

derart, daß mit dem Zusammentensorieren von Tensoren in der oberen Horizontale und besagter Abbildung in der unteren Horizontale das Diagramm

$$\begin{array}{ccc} V^{\otimes r} \times V^{\otimes s} & \longrightarrow & V^{\otimes (r+s)} \\ \downarrow & & \downarrow \\ \bigwedge^r V \times \bigwedge^s V & \longrightarrow & \bigwedge^{r+s} V \end{array}$$

kommutiert. Man folgert das unschwer aus Übung 7.1.7, da die obere Horizontale unseres Diagramms sowohl $J_r \times V^{\otimes s}$ als auch $V^{\otimes r} \times J_s$ auf Teilmengen von J_{r+s} abbildet. Unsere so konstruierte Abbildung $\bigwedge^r V \times \bigwedge^s V \to \bigwedge^{r+s} V$ ist offensichtlich bilinear. Sie wird notiert als $(\omega, \eta) \mapsto \omega \wedge \eta$ und heißt das **Dachprodukt**, englisch **wedge-product**, französisch **produit extérieur**. Aus der Konstruktion ergibt sich unmittelbar seine Assoziativität, in $\bigwedge^{r+s+t} V$ gilt also

$$(\omega \wedge \eta) \wedge \xi = \omega \wedge (\eta \wedge \xi)$$

für alle $\omega \in \bigwedge^r V$, $\eta \in \bigwedge^s V$ und $\xi \in \bigwedge^t V$. Die direkte Summe

$$\bigwedge V = \bigoplus_{r \geq 0} \bigwedge^r V$$

wird mit dem komponentenweisen Dachprodukt insbesondere ein Ring mit Eins-Element $1 \in k = \bigwedge^0 V$.

7.4.6. Ganz allgemein bezeichnet man einen k-Vektorraum A mit einer bilinearen Verknüpfung $A \times A \to A$ als eine **k-Algebra** und versteht unter einem **Algebrenhomomorphismus** in eine weitere k-Algebra eine k-lineare Abbildung, die mit den jeweiligen Verknüpfungen verträglich ist. Ist die Verknüpfung einer Algebra assoziativ, so spricht man von einer **assoziativen Algebra**. Gibt es für diese Verknüpfung ein neutrales Element, so spricht man von einer **unitären Algebra**. Eine Algebra ist also genau dann assoziativ und unitär, wenn die zugrundeliegende Menge mit der Vektorraum-Addition als Addition und der bilinearen Verknüpfung als Multiplikation ein Ring ist. Ich schlage deshalb vor, derartige Algebren **Ringalgebren** zu nennen. Unter einem **Homomorphismus von Ringalgebren** verstehen wir dann einen

Algebrenhomomorphismus, der auch ein Ringhomomorphismus ist. Wir können diese Abbildungen sowohl charakterisieren als Algebrenhomomorphismen, die das neutrale Element auf das neutrale Element werfen, als auch als über dem Grundkörper lineare Ringhomomorphismen. Wir vereinbaren für die Menge der Ringalgebrenhomomorphismen von einer k-Ringalgebra A in eine k-Ringalgebra B die Notation $\mathrm{Ralg}_k(A, B)$.

7.4.7. In dieser Terminologie ist $\bigwedge V$ eine Ringalgebra. Sie heißt die **äußere Algebra** des Vektorraums V. Die offensichtliche Identifikation $V \xrightarrow{\sim} \bigwedge^1 V$ notieren wir kurzerhand $v \mapsto v$ und behandeln sie auch sprachlich als Gleichheit. Gegeben $v \in V$ gilt in $\bigwedge^2 V$ wegen $v \otimes v \in J_2$ natürlich $v \wedge v = 0$, und mit 2.7.10 folgt

$$v \wedge w = -w \wedge v \quad \forall v, w \in V$$

Definition 7.4.8. Sei k ein Körper, V ein k-Vektorraum, $(v_i)_{i\in I}$ eine Basis von V und $\leq$ eine Anordnung von I. Gegeben $J \subset I$ mit $|J| = r < \infty$ erklären wir dann ein Element $v_J \in \bigwedge^r V$ als das Dachprodukt

$$v_J = v_{i_1} \wedge \ldots \wedge v_{i_r}$$

für $i_1 < i_2 < \ldots < i_r$ die der Größe nach geordneten Elemente von J. Im Extremfall $r = 0$ vereinbaren wir $v_\emptyset = 1 \in k = \bigwedge^0 V$.

Proposition 7.4.9 (Basen der äußeren Potenzen). *Sei k ein Körper, V ein k-Vektorraum und $(v_i)_{i\in I}$ eine Basis von V. Sei eine Anordnung auf I gewählt. Gegeben $r \geq 0$ bilden dann die v_J mit $J \subset I$ und $|J| = r$ eine Basis der r-ten äußeren Potenz $\bigwedge^r V$.*

Beweis. Alle Tensoren $v_{i_1} \otimes \ldots \otimes v_{i_r}$ für $i_1, \ldots, i_r \in I$ beliebig erzeugen nach 7.2.9 die r-te Tensorpotenz $V^{\otimes r}$. Alle Dachprodukte $v_{i_1} \wedge \ldots \wedge v_{i_r}$ erzeugen folglich die r-te äußere Potenz $\bigwedge^r V$. Beim Umordnen derartiger Dachprodukte ändert sich höchstens das Vorzeichen, und kommt ein Vektor doppelt vor, ist das fragliche Dachprodukt eh null. Folglich erzeugen die Dachprodukte $v_{i_1} \wedge \ldots \wedge v_{i_r}$ mit $i_1 < \ldots < i_r$ unsere r-te äußere Potenz, und es bleibt nur, ihre lineare Unabhängigkeit nachzuweisen. Dazu betrachten wir für $f_1, \ldots, f_r \in V^*$ beliebig die lineare Abbildung

$$\begin{array}{rccc} \mathrm{alt}(f_1, \ldots, f_r): & V^{\otimes r} & \to & k \\ & w_1 \otimes \ldots \otimes w_r & \mapsto & \det(f_i(w_j)) \end{array}$$

Sie verschwindet offensichtlich auf J_r und induziert folglich eine lineare Abbildung $\bigwedge^r V \to k$. Betrachten wir nun die Koordinatenfunktionen $v_i^* \in V^*$ zu unserer Basis $(v_i)_{i\in I}$ und bezeichnen für jedes r-elementige $J \subset I$ bestehend aus den der Größe nach geordneten Elementen $i_1 < \ldots < i_r$ mit

$\operatorname{alt}(v_{i_1}^*, \ldots, v_{i_r}^*) = v_J^* : \bigwedge^r V \to k$ die zu $v_{i_1}^*, \ldots, v_{i_r}^*$ gehörige Linearform, so folgt aus den Eigenschaften der Determinante für je zwei r-elementige Teilmengen $J, K \subset I$ unmittelbar

$$v_J^*(v_K) = \begin{cases} 1 & J = K; \\ 0 & \text{sonst;} \end{cases}$$

Das impliziert die lineare Unabhängigkeit der v_K. □

Proposition 7.4.10 (Äußere Potenzen und Dualisieren). *Sei k ein Körper, V ein k-Vektorraum und $r \geq 0$. So existiert genau eine bilineare Abbildung*

$$\bigwedge^r (V^*) \times \bigwedge^r V \to k$$

mit $((f_1 \wedge \ldots \wedge f_r), (w_1 \wedge \ldots \wedge w_r)) \mapsto \det(f_i(w_j))$, und im Fall $\dim V < \infty$ induziert diese Abbildung einen Isomorphismus

$$\bigwedge^r (V^*) \xrightarrow{\sim} \left(\bigwedge^r V \right)^*$$

7.4.11. Im Zweifelsfall interpretieren wir $\bigwedge^r V^*$ im folgenden als $\bigwedge^r (V^*)$.

Beweis. Das wurde im wesentlichen bereits im Laufe des Beweises der vorhergehenden Proposition 7.4.9 gezeigt. Die Details bleiben dem Leser überlassen. □

7.4.12. Man beachte, daß sich im Fall eines endlichdimensionalen Vektorraums V mithilfe dieser Proposition unser Isomorphismus $(\bigwedge^r V)^* \xrightarrow{\sim} \operatorname{Alt}^r(V)$ aus 7.4.2 zu einem Isomorphismus $\bigwedge^r (V^*) \xrightarrow{\sim} \operatorname{Alt}^r(V)$ verlängern läßt. Mit den durch diese Isomorphismen gegebenen Vertikalen und dem Dachprodukt in der oberen Horizontalen und dem Dachprodukt, wie wir es im Rahmen des Stokes'schen Satzes in **??** direkt einführen, in der unteren Horizontalen kommutiert dann das Diagramm

$$\begin{array}{ccc} \bigwedge^r V^* \times \bigwedge^s V^* & \longrightarrow & \bigwedge^{r+s} V^* \\ \wr\downarrow & & \wr\downarrow \\ \operatorname{Alt}^r(V) \times \operatorname{Alt}^s(V) & \longrightarrow & \operatorname{Alt}^{r+s}(V) \end{array}$$

Die hier gegebene Konstruktion des Dachprodukts benötigt zwar den größeren begrifflichen Aufwand, scheint mir aber sehr viel durchsichtiger als die im Rahmen des Beweises von **??** gegebene direkte Konstruktion.

7.4.13. Aus 7.4.9 folgt für einen Vektorraum V endlicher Dimension $\dim V = d < \infty$ sofort

$$\dim \bigwedge^r V = \binom{d}{r}$$

und insbesondere $\dim \bigwedge^d V = 1$ und $\bigwedge^r V = 0$ für $r > d$. Man kürzt deshalb im endlichdimensionalen Fall oft $\bigwedge^{\dim V} V = \bigwedge^{\max} V$ ab.

Beispiel 7.4.14. Eine Basis von $\bigwedge^2 \mathbb{R}^4$ besteht etwa aus den sechs Vektoren $\mathrm{e}_1 \wedge \mathrm{e}_2, \mathrm{e}_1 \wedge \mathrm{e}_3, \mathrm{e}_1 \wedge \mathrm{e}_4, \mathrm{e}_2 \wedge \mathrm{e}_3, \mathrm{e}_2 \wedge \mathrm{e}_4$ und $\mathrm{e}_3 \wedge \mathrm{e}_4$.

7.4.15 (**Maximale äußere Potenz und Determinante**). Jede lineare Abbildung $f : V \to W$ induziert lineare Abbildungen $f^{\otimes r} : V^{\otimes r} \to W^{\otimes r}$ und durch Übergang zu den Quotienten lineare Abbildungen $\bigwedge^r f : \bigwedge^r V \to \bigwedge^r W$, die in ihrer Gesamtheit einen Ringhomomorphismus

$$\bigwedge f : \bigwedge V \to \bigwedge W$$

liefern. Natürlich gilt auch $\bigwedge(f \circ g) = (\bigwedge f) \circ (\bigwedge g)$ und $\bigwedge(\mathrm{id}) = \mathrm{id}$. Ist speziell $f : V \to V$ ein Endomorphismus eines endlichdimensionalen Vektorraums, so ist $\bigwedge^{\max} f : \bigwedge^{\max} V \to \bigwedge^{\max} V$ ein Endomorphismus eines eindimensionalen Vektorraums alias ein Skalar. Wir zeigen nun, daß dieser Skalar genau die Determinante von f ist, in Formeln

$$\bigwedge^{\max} f = \det f$$

Sei dazu $v_1, \ldots, v_n$ eine Basis von V. Dann ist $v_1 \wedge \ldots \wedge v_n$ nach 7.4.9 eine Basis von $\bigwedge^n V$. Haben wir $f(v_i) = \sum a_{ji} v_j$, so folgt

$$\begin{aligned}
(\textstyle\bigwedge f)(v_1 \wedge \ldots \wedge v_n) &= f(v_1) \wedge \ldots \wedge f(v_n) \\
&= (\textstyle\sum a_{j1} v_j) \wedge \ldots \wedge (\sum a_{ln} v_l) \\
&= \textstyle\sum_{\sigma:\{1,\ldots,n\}\to\{1,\ldots,n\}} a_{\sigma(1)1} v_{\sigma(1)} \wedge \ldots \wedge a_{\sigma(n)n} v_{\sigma(n)} \\
&= \textstyle\sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \ldots a_{\sigma(n)n} v_1 \wedge \ldots \wedge v_n \\
&= (\det f) v_1 \wedge \ldots \wedge v_n
\end{aligned}$$

Die Multiplikationsregel für Determinanten folgt mit diesen Erkenntnissen unmittelbar aus der Relation $\bigwedge^{\max}(f \circ g) = (\bigwedge^{\max} f) \circ (\bigwedge^{\max} g)$. Daß die Determinante eines Endomorphismus $f : V \to V$ verschwindet, falls dieser nicht vollen Rang hat, kann man in diesem Formalismus auch wie folgt einsehen: Man schreibt f als Verknüpfung $V \twoheadrightarrow \operatorname{im} f \hookrightarrow V$, und unter der Annahme $d = \dim V > \dim(\operatorname{im} f)$ folgt $\bigwedge^d(\operatorname{im} f) = 0$, womit dann auch die Komposition $\bigwedge^d V \to \bigwedge^d(\operatorname{im} f) \to \bigwedge^d V$ die Nullabbildung sein muß.

Übung 7.4.16. Gegeben eine $(n \times m)$-Matrix A und eine $(m \times n)$-Matrix B kann man die Determinante der $(n \times n)$-Matrix AB bestimmen wie folgt: Für jede n-elementige Teilmenge $I \subset \{1, \ldots m\}$ mit Elementen $i_1 < \ldots < i_n$ möge A_I gerade aus den Spalten von A der Indizes $i_1, \ldots, i_n$ bestehen und B^I aus den Zeilen von B der Indizes $i_1, \ldots, i_n$. So gilt

$$\det(AB) = \sum_{|I|=n} (\det A_I)(\det B^I)$$

7.4.17. Für einen k-Vektorraum V endlicher Dimension $\dim V = n$ liefert das Dachprodukt nichtausgeartete Paarungen $\bigwedge^d V \times \bigwedge^{n-d} V \to \bigwedge^n V$, denn wir haben $v_I \wedge v_J = \pm v_1 \wedge \ldots \wedge v_n$ falls I das Komplement von J ist und Null sonst. Jeder Isomorphismus $\bigwedge^n V \xrightarrow{\sim} k$ definiert also insbesondere einen Isomorphismus $\bigwedge^{n-1} V \cong V^*$.

Übung 7.4.18. Ist $V' \hookrightarrow V \twoheadrightarrow V''$ eine kurze exakte Sequenz endlichdimensionaler Vektorräume, so induziert mit der Notation $d = \dim V''$ das Dachprodukt $\bigwedge^{\max} V' \otimes \bigwedge^d V \xrightarrow{\sim} \bigwedge^{\max} V$ einen Isomorphismus, den sogenannten **kanonischen Isomorphismus**

$$\bigwedge^{\max} V' \otimes \bigwedge^{\max} V'' \xrightarrow{\sim} \bigwedge^{\max} V$$

7.5 Das kanonische Skalarprodukt

Das kam in der Vorlesung 2008/2009 nicht vor.

Definition 7.5.1. Sei V ein reeller Vektorraum und L ein eindimensionaler orientierter reeller Vektorraum. Unter einem **Skalarprodukt auf V mit Einheiten** L verstehen wir eine symmetrische bilineare Abbildung

$$s : V \times V \to L^{\otimes 2}$$

mit $v \neq 0 \Rightarrow s(v, v) > 0$ für diejenige Orientierung auf $L^{\otimes 2}$, die charakterisiert wird durch die Eigenschaft $a \otimes a \in L^{\otimes 2}_{>0}$ für alle $a \in L \backslash 0$. Die Orientierung auf L setzen wir hier nur deshalb voraus, damit wir eine Wurzelabbildung $\surd : L^{\otimes 2}_{\geq 0} \to L_{\geq 0}$ erklären können als das Inverse des Quadrierens $L_{\geq 0} \xrightarrow{\sim} L^{\otimes 2}_{\geq 0}$, $a \mapsto a \otimes a$, so daß wir die **Länge** eines Vektors erklären können als

$$\|v\|_s = \sqrt{s(v, v)} \in L_{\geq 0}$$

7.5.2. Gegeben ein dreidimensionaler reeller Vektorraum mit ausgezeichneter Drehgruppe im Sinne von 6.9.2 will ich nun in vollständiger kanonischer Weise ein Skalarprodukt mit Einheiten konstruieren. Dazu müssen wir aus

diesen Daten zuerst einmal einen orientierten eindimensionalen Vektorraum konstruieren, den wir den "Vektorraum der Längen" nennen werden. Das braucht bereits einige Vorbereitungen.

Definition 7.5.3. Gegeben eine Gruppe G, eine G-Menge X und eine G-Rechtsmenge Y definieren wir ihr **balanciertes Produkt**

$$Y \times_G X$$

als den Quotienten des kartesischen Produkts $Y \times X$ nach der Äquivalenzrelation $(yg, x) \sim (y, gx) \quad \forall y \in Y, x \in X$ und $g \in G$, alias den Bahnenraum von $Y \times X$ unter der durch die Vorschrift $g.(y, x) = (yg^{-1}, gx)$ gegebenen G-Operation. Die Äquivalenzklasse alias Bahn von $(y, x) \in Y \times X$ notieren wir $[y, x] \in Y \times_G X$.

Übung 7.5.4. Seien k ein Körper, V ein k-Vektorraum und G eine Gruppe. Seien weiter $\rho : G \to \mathrm{GL}(V)$ ein Gruppenhomomorphismus und Y ein G-Torsor. So gibt es auf dem balancierten Produkt

$$Y \times_G V = Y \times_G^{\rho} V$$

genau eine Struktur als k-Vektorraum derart, daß für alle $y \in Y$ die Abbildung $v \mapsto [y, v]$ einen Vektorraumisomorphismus $V \xrightarrow{\sim} Y \times_G V$ liefert.

Definition 7.5.5. Sei V ein dreidimensionaler reeller Vektorraum V mit einer ausgezeichneten Drehgruppe $D \subset \mathrm{GL}(V)$ im Sinne von 6.9.2. Eine Bahn $l \subset V \backslash 0$ unserer Drehgruppe D nennen wir eine **positive Länge**. Diese positiven Längen bilden in natürlicher Weise einen $\mathbb{R}_{>0}$-Torsor im Sinne von 6.1.6.7. Den zugehörigen orientierten eindimensionalen **Vektorraum der Längen** erklären wir als das balancierte Produkt

$$L_D = L = L_{>0} \times_{\mathbb{R}_{>0}} \mathbb{R}$$

versehen mit derjenigen Orientierung, für die alle Vektoren $[l, \alpha]$ mit $l \in L_{>0}$ und $\alpha > 0$ positiv orientierte Basen sind. Die Injektion $L_{>0} \hookrightarrow L$, $l \mapsto [l, 1]$ notieren wir nicht extra, sondern fassen sie im weiteren als die Einbettung einer Teilmenge auf, deren Bild im übrigen genau die positiven Vektoren unseres orientierten eindimensionalen Vektorraums sind, so daß wir mit dieser Notation keine Zweideutigkeiten erzeugen. Wir nennen L_D die **Längengerade** unserer Drehgruppe.

7.5.6. Gegeben ein dreidimensionaler reeller Vektorraum V mit ausgezeichneter Drehgruppe $D \subset \mathrm{GL}(V)$ gibt es genau ein Skalarprodukt auf V mit

Einheiten in der zugehörigen Längengerade L, in Formeln genau eine positiv definite symmetrische bilineare Abbildung $s : V \times V \to L^{\otimes 2}$ derart, daß gilt $s(v,v) = Dv \otimes Dv$ für alle $v \neq 0$, wo $Dv \in L_{>0} \subset L$ die Bahn von v unter der Drehgruppe D meint. Hier folgt die Eindeutigkeit von s aus der Polarisierungsidentität und die Existenz erhält man, indem man das Ende des Beweises von 6.9.3 geeignet variiert. Ich schlage vor, diese Abbildung

$$s : V \times V \to L^{\otimes 2}$$

das **kanonische Skalarprodukt** unserer Drehgruppe zu nennen, da sie ein Skalarprodukt mit Einheiten im Sinne unserer Definition 7.5.1 ist, das nur von der ausgezeichneten Drehgruppe und sonst von keinerlei Wahlen abhängt.

7.5.7. Ich fasse nun nocheinmal unser mathematisches Modell des Anschauungsraums zusammen: Wir modellieren den Anschauungsraum, wie in 1.9.5 erklärt, als einen dreidimensionalen reellen affinen Raum

$$\mathbb{E}$$

zusammen mit einer ausgezeichneten Bewegungsgruppe im Sinne von 3.1.2. Die Geraden entsprechen unseren Sichtlinien. Die ausgezeichneten Bewegungen entsprechen den anschaulichen Bewegungen, wie in 3.1.3 ausgeführt wird. Im Richtungsraum des Anschaungsraums erhalten wir dann eine ausgezeichnete Drehgruppe im Sinne von 6.9.2, die hinwiederum wie in 7.5.5 und 7.5.6 erklärt eine ausgezeichnete Längengerade liefert, die wir in diesem Fall mit

$$\mathbb{L}$$

bezeichnen, nebst einem Skalarprodukt $\langle\ ,\ \rangle$ mit Einheiten in dieser Längengerade. Zumindest in Europa wird diese Längengerade meist vermittels des in der französischen Revolution gewählten **Meters** $m \in \mathbb{L}_{>0}$ mit $\mathbb{R}$ identifiziert. Das kanonische Skalarprodukt auf dem Richtungsraum des Anschauungsraums nimmt also Werte in einem eindimensionalen reellen Vektorraum an, in dem das "Quadratmeter" eine Basis ist. Man notiert es meist abkürzend m^2 statt $m^{\otimes 2}$, wie es eigentlich unserer Vereinbarung 7.4.1 entsprechen würde.

Übung 7.5.8. Wir erinnern 7.5.4. Sei k ein Körper. Wir betrachten für alle $r \in \mathbb{Z}$ den Gruppenhomomorphismus $k^\times \to \mathrm{GL}(k) = k^\times$ gegeben durch $\lambda \mapsto \lambda^r$. So erhalten wir im Fall $r \geq 0$ für jeden eindimensionalen Vektorraum V einen Isomorphismus

$$\begin{array}{ccc} (V\backslash 0) \times_{k^\times} k & \xrightarrow{\sim} & V^{\otimes r} \\ [v,\lambda] & \mapsto & \lambda(v^{\otimes r}) \end{array}$$

und im Fall $r \leq 0$ für jeden eindimensionalen Vektorraum V einen Isomorphismus

$$\begin{array}{ccc} (V\backslash 0) \times_{k^\times} k & \xrightarrow{\sim} & (V^*)^{\otimes(-r)} \\ [v, \lambda] & \mapsto & \lambda((v^*)^{\otimes(-r)}) \end{array}$$

mit v^* definiert durch $v^*(v) = 1$. In diesem Sinne werden wir von nun an negative Tensorpotenzen eines eindimensionalen Vektorraums als die entsprechenden positiven Potenzen des Dualraums interpretieren.

Übung 7.5.9. Wir erinnern 7.5.4. Seien V, W zwei k-Vektorräume derselben Dimension. Wir betrachten die Gruppe $G = \mathrm{GL}(W)$ und den Gruppenhomomorphismus $G \to \mathrm{GL}(\bigwedge^r W)$ gegeben durch $g \mapsto \bigwedge^r g$. Sei $Y = \mathrm{Hom}^\times(W, V)$ der G-Torsor aller Isomorphismen $W \xrightarrow{\sim} V$. So erhalten wir einen Isomorphismus

$$\begin{array}{ccc} Y \times_G \bigwedge^r W & \xrightarrow{\sim} & \bigwedge^r V \\ [f \quad , \quad w] & \mapsto & (\bigwedge^r f)(w) \end{array}$$

Kapitel IV

Typische Prüfungsfragen

1 Lineare Algebra

1. Was ist ein Körper? Wie leitet man die Regel für das Addieren von Brüchen aus den Körperaxiomen ab?

2. Was ist eine Basis eines Vektorraums? Könnte es passieren, daß ich in demselben Vektorraum eine Basis mit 13 Elementen und eine mit 17 Elementen finde? Was ist die Dimension eines Vektorraums? Hat jeder Vektorraum eine Basis? Was ist überhaupt ein Vektorraum? Wie leitet man $0v = 0$ aus den Vektorraumaxiomen ab?

3. Warum ist jedes unverkürzbare Erzeugendensystem eine Basis? Warum ist jede unverlängerbare linear unabhängige Teilmenge eine Basis?

4. Wie versieht man die Menge der Homomorphismen von einem Vektorraum zu einem anderen mit der Struktur eines Vektorraums? Wie berechnet man die Dimension eines derartigen Raums von Homomorphismen?

5. Was ist die Matrix einer ebenen Drehung um 45°? Was ist ihre Determinante? Ihre Eigenwerte?

6. Geben Sie eine (3×3)-Matrix vom Rang ... ohne Nullen an. Was ist deren Determinante? Was ist die Lösungsmenge des zugehörigen Gleichungssystems? Was sind die Eigenwerte?

7. Was ist die Determinante einer Matrix? Wie rechnet man sie aus? Warum hat die transponierte Matrix dieselbe Determinante? Warum ist jede Matrix mit von Null verschiedener Determinante invertierbar?

8. Besitzt jede Matrix einen Eigenwert? Ist jede Matrix diagonalisierbar? Beispiel? Gegenbeispiel? Ist jede relle symmetrische Matrix diagonalisierbar? Beweis?

9. Was ist ein Eigenwert einer linearen Abbildung? Welche Eigenwerte hat das Ableiten, aufgefaßt als lineare Abbildung vom Raum der beliebig oft differenzierbaren reellen Funktionen auf der reellen Zahlengeraden $\mathcal{C}^\infty_{\mathbb{R}}(\mathbb{R})$ in sich selbst? Welche Eigenwerte hat das Ableiten aufgefaßt als lineare Abildung vom Raum der Polynome in sich selbst? Was sind die Eigenräume? Und wenn man Koeffizienten in einem Körper der Charakteristik ... nimmt?

10. Wieviele Untervektorräume hat ein ...-dimensionaler Vektorraum über dem Körper mit ... Elementen?

11. Wieviele angeordnete Basen hat ein . . .-dimensionaler Vektorraum über dem Körper mit . . . Elementen?

12. Nimmt die quadratische Form . . . positive und negative Werte an? Wie findet man so etwas im allgemeinen heraus?

13. Berechnen Sie die inverse Matrix zu . . .

14. Was versteht man unter dem Rang einer Matrix? Warum stimmen Zeilenrang und Spaltenrang stets überein?

15. Wie hängen die Eigenwerte einer invertierbaren Matrix zusammen mit den Eigenwerten ihrer Inversen? Wie hängt die Jordan'sche Normalform einer invertierbaren Matrix zusammen mit der Jordan'schen Normalform ihrer Inversen?

2 Algebra

1. Gibt es eine Gruppe mit . . . Elementen? Gibt es eine abelsche Gruppe mit . . . Elementen? Wie konstruiert man überhaupt so eine Restklassengruppe? Was ist das Inverse zu . . . in $\mathbb{Z}/a\mathbb{Z}$? Wieviele paarweise nicht isomorphe abelsche Gruppen gibt es mit . . . Elementen? Welche?

2. Wieviele Gruppenhomomorphismen gibt es von $\mathbb{Z}/4\mathbb{Z}$ nach $\mathbb{Z}/6\mathbb{Z}$?

3. Wieviele Elemente hat $\mathrm{GL}(3;\mathbb{F}_7)$? Wie größ ist die 7-Sylow darin? Können Sie eine 7-Sylow angeben?

4. Hat jedes Polynom eine Nullstelle? Kann man den Grundkörper so vergrößern, daß es eine kriegt? Wie geht das?

5. Ist das Polynom . . . irreduzibel? Was ist ein irreduzibles Polynom? Inwieweit ist die Zerlegung eines Polynoms in irreduzible Faktoren eindeutig? Warum?

6. Wieviele Nullstellen kann das Polynom . . . höchstens haben? Warum? Gibt es zu vorgegebenen Nullstellen stets ein Polynom, das genau diese Nullstellen hat? Warum-warum nicht?

7. Gibt es einen Körper mit . . . Elementen? Wie zeigt man das? Wann ist $\mathbb{Z}/a\mathbb{Z}$ ein Körper? Warum ist $\mathbb{Z}/10\mathbb{Z}$ kein Körper? Wie rechnet man in diesem Ring? Besitzt . . . darin ein multiplikatives Inverses? Und

zwar welches? Welche abelsche Gruppe erhält man als Einheitengruppe? Welche abelsche Gruppe ist die multiplikative Gruppe des Körpers mit ... Elementen? Welche Kardinalität kann ein endlicher Körper haben? Warum? Sind je zwei Körper mit ... Elementen isomorph? Warum?

8. Was ist die Automorphismengruppe des Körpers mit ... Elementen? Wie zeigt man das?

9. Ist das regelmäßige ...-Eck konstruierbar mit Zirkel und Lineal? Warum oder warum nicht? Welche regelmäßigen n-Ecke sind eigentlich konstruierbar? Warum-warum nicht?

3 Analysis

1. Wie bestimmt man die Ableitung des Arcustangens? Was ist überhaupt die Ableitung? Was bedeutet darin das Symbol $\varinjlim$? Wie entwickelt man arctg in eine Potenzreihe? Warum ist diese Rechnung erlaubt?...

2. Was ist $\lim_{x\to\infty} \frac{x+\mathrm{e}^x}{\log x+\mathrm{e}^x}$? Was bedeutet $\lim_{x\to\infty} g(x) = b$? Wie ist die Exponentialfunktion definiert? Kennen Sie andere Funktionen, die ihre eigene Ableitung sind? Sind das alle? Warum?

3. Berechnen Sie den Schwerpunkt eines Kuchenstücks: Stellen Sie es auf die Spitze und integrieren die Höhe y über das entsprechende Gebiet. Wie lautet allgemein die Formel zur Transformation von Mehrfachintegralen auf krummlinige Koordinaten? Was ist die Beziehung zur Substitutionsregel?

4. Finden Sie eine Stammfunktion für den Arcustangens, $\int \arctan$; wie ist überhaupt das Integral definiert? Warum kann es mittels Stammfunktionen berechnet werden?

5. Was bedeutet $\lim_{n\to\infty} a_n = a$? Schreiben Sie es mit den zugehörigen „für alle“ und „es gibt“ einmal auf. Wie folgt $\lim_{n\to\infty} 1/n = 0$?

6. Wie ist die Exponentialfunktion definiert? Warum konvergiert diese Reihe? Wie zeigt man das Quotientenkriterium? Das Majorantenkriterium?

7. (Falls es dran war) Kennen Sie eine Funktion, die ihre eigene dritte Ableitung ist, $f''' = f$? Können Sie alle derartigen Funktionen $f\colon \mathbb{R} \to \mathbb{C}$ angeben? Können Sie alle derartigen Funktionen $f\colon \mathbb{R} \to \mathbb{R}$ angeben?

8. Wie ist das Integral $\int_a^b f(x)\,\mathrm{d}x$ für $f\colon [a,b] \to \mathbb{R}$ stetig definiert? Welche Probleme können für f unstetig auftreten? Was bedeutet gleichmäßig stetig?

9. Wie ist der Logarithmus definiert? Warum wird jede positive reelle Zahl als Wert der Exponentialfunktion angenommen? Was ist die Ableitung des Logarithmus? Seine Potenzreihenentwicklung? Das Integral? Die Potenzreihenentwicklung um den Punkt $p = 5$? Der Konvergenzradius daselbst?

10. Was ist das höherdimensionale Analogon der Ableitung? Wie hängt das totale Differential mit den partiellen Ableitungen zusammen? Wie lautet in dieser Allgemeinheit die Kettenregel? Wie zeigt man sie?

11. Was ist das Lebesgue-Maß? Wie ist das Lebesgue-Integral definiert? Was ist seine Beziehung zu absoluter Konvergenz?

Literaturverzeichnis

[Maz08] Barry Mazur, *Mathematical platonism and its opposites*, Newsletter of the EMS **68** (2008), 19–21.

Index

E_{ij} Basismatrizen, 107
$X - Y$ Differenz von Mengen, 26
$X \backslash Y$ Differenz von Mengen, 26
$X \times Y$ kartesisches Produkt, 28
$X \cap Y$ Schnitt, 26
$X \cup Y$ Vereinigung, 26
Y^X bei Mengen, 35
$\Leftarrow$
    folgt aus, 42
$\Leftrightarrow$ gleichbedeutend, 42
$\Rightarrow$
    impliziert, 42
$\bigoplus$
    Summe von Vektorräumen, 222
$\circ$
    Matrixprodukt, 102
    Verknüpfung von Abbildungen, 36
$\coprod$ disjunkte Vereinigung, 222
$\emptyset$ leere Menge, 25
$\forall$ für alle, 42
$\oplus$ direkte Summe
    von Untervektorräumen, 96
    von Vektorräumen, 82
$\overrightarrow{AB}$ Richtungsvektor, 119
$:=$ definiert durch, 11
$\perp$ Orthogonalität, 176
$\perp$ steht senkrecht auf, 176
$\prod$
    Produkt von Mengen, 222
    Produkt von Vektorräumen, 222
$\prod$ Produkt
    von Zahlen, 13
$\sharp$ Kardinalität, 26
$\subset$ Teilmenge, 25
$\subseteq$ Teilmenge, 25
$\subsetneq$ echte Teilmenge, 25
$\sum$ Summe
    von Zahlen, 11
$\vec{u} + e$, 119
$\{\ \}$ Multimenge, 40
$k((X))$ formale Laurentreihen, 145
$k(X)$ rationale Funktionen, 150
$k[[X]]$ formale Potenzreihen, 145
$n!$ Fakultät, 13
r-te äußere Potenz von V, 301
$|\ |$
    Kardinalität, 26
$\mapsto$ wird abgebildet auf, 35
$\rightarrow$ Abbildung, 33

Abb, 35
Abbildung, 33
    einwertige, 35
    identische Abbildung, 35
    inverse Abbildung, 39
    konstante, 35
    Projektionsabbildung, 77
    Umkehrabbildung, 39
abelsch
    Gruppe, 50
abgeschlossen
    algebraisch, 144
    unter Verknüpfung, 49
Abspalten von Linearfaktoren, 143
Abständezahl, 278
Addition
    in Ring, 129

adjungiert, 205
    Matrix, 163
Äquivalenzklasse, 148
Äquivalenzrelation, 148
äußere Algebra, 303
affin
    Abbildung, 121
    Raum, 119
    Raum, über Vektorraum, 121
affiner Teilraum
    von affinem Raum, 122
    von Vektorraum, 99
Algebra, 302
algebraisch
    abgeschlossen, Körper, 144
Algebrenhomomorphismus, 302
allgemeine lineare Gruppe, 93, 106
Alternator, 301
alternierend, 158, 159
    Tensor, 301
alternierende Gruppe, 152
assoziativ, 46
aufgespannt
    Untervektorraum, 83
Ausartungsraum, 216
ausgeartet
    Paarung, 216
Automorphismengruppe
    eines Vektorraums, 93
Automorphismus
    einer Gruppe, 263
    eines Vektorraums, 92
    von affinem Raum, 121

Bahn, 242
Bahnenraum, 244
Bahnformel, 249
balanciertes Produkt, 307
baryzentrische Koordinaten, 128
Basis, 85
    angeordnete, 85
    indizierte, 85
    orientierte, 192
    von Vektorraum, 85
Basisexistenzsatz, 87
Basismatrix, 107
Basiswechselmatrix, 111
Bessel'sche Ungleichung, 179
Bewegungen, 170
Bewegungsgruppe, 170
Bidualraum, 117
Bijektion, 38
bijektiv
    Abbildung, 38
Bild, 33, 35
    von Gruppenhomomorphismus, 135
Bildmenge, 35
bilinear, 97
Bilinearform, 173
Binomialkoeffizienten, 14
binomische Formel, 15
Bruchzahlen, 25

$\subset$ Teilmenge, 25
$\subseteq$ Teilmenge, 25
$\subsetneq$ echte Teilmenge, 25
card, 26
Catalan-Zahl, 48
Cauchy-Schwarz'sche Ungleichung, 177
Cayley-Hamilton, 167
char
    Charakteristik, 140
Charakteristik
    eines Rings, 140
charakteristisches Polynom, 166
    von Endomorphismus, 166
Chinesischer Restsatz, 255
cok Kokern, 252
corps, 55

Dachprodukt, 302
darstellende Matrix, 101, 110

Definition, 11
Definitionsbereich, 33, 151
Determinante, 154
    von Endomorphismus, 162
diag($\lambda_1, \ldots, \lambda_n$) Diagonalmatrix, 108
Diagonale, 78
diagonalisierbar
    Endomorphismus, 167
    Matrix, 169
Diagonalmatrix, 108
Diedergruppe, 264
Differenz
    von Mengen, 26
Dimension
    von affinem Raum, 119
    von Vektorraum, 90
Dimensionsformel, 99
direkte Summe
    von Untervektorräumen, 96
    von Vektorräumen, 82, 222
disjunkt, 25
disjunkte Vereinigung, 222
Distributivgesetz, 55, 129
Dodekaeder, 265
Drehgruppe, 265, 279
Drehnorm, 172
drehsenkrecht zu $\vec{v}$, 172
Drehsymmetrie, 264
Drehung, 172, 279
    um Punkt, 278
Dreiecksungleichung, 179
duale, 115
duale Basis, 116
Dualraum, 113

$\in, \notin$, 25
$\exists$ es existiert ein, 42
$\mathbb{E}$, 170
$\mathbb{E}$ der Anschauungsraum, 120
$\exists!$ es existiert genau ein, 42
Ebene
    affine, 122
echt
    Teilmenge, 25
Eig, 225
Eigenraum, 225
Eigenvektor, 165
Eigenwert, 165
Einbettung
    einer Teilmenge, 38
Einheit, 133
Einheitsmatrix, 101
Eins
    in Ring, 129
Eins-Element
    in Ring, 129
Einschränkung, 38
Einsetzen in Polynome, 142
einwertige Abbildung, 35
Element, 24
Elementarmatrix, 108
Elementarteilersatz
    über dem Grundring $\mathbb{Z}$, 257
endlich erzeugbar, 83
endlich erzeugt, 83
endliche Primkörper, 138
Endomorphismenring, 130
Endomorphismus
    von abelscher Gruppe, 130
    von Vektorräumen, 92
Ens, 35
$\mathrm{Ens}^{\times}$, 52
ensemble, 35
Erweiterung der Skalare, 297
Erzeugendensystem, 83
    von affinem Raum, 123
Erzeugnis, 83
erzeugt
    affiner Teilraum, 123
    Untervektorraum, 83
euklidisch
    Norm, 176

reeller affiner Raum, 189
Vektorraum, 176
Euler'sche Winkel, 187
ev Auswertungsabbildung, 295
exakt, 252
Sequenz, 251
Exponent, 256

Faktoren, 13
Fakultät, 13
Familie, 79
Faser
einer Abbildung, 36
Fehlstand, 151
Fibonacci-Folge, 17
field, 55
Fitting-Zerlegung
von Vektorräumen, 228
Fixator, 241
Fixpunkt, 100
von Gruppenwirkung, 241
Form
quadratische, 218
Frac, 150
fraction field, 150
frei
Gruppenwirkung, 241
Vektorraum, 84
Frobenius-Abbildung, 140
Fundamentalmatrix, 209
Funktion
rationale, 150
Umkehrfunktion, 39

ganze Zahlen, 25
Gauß-Algorithmus, 71
general linear group, 93, 106
gerade
Permutation, 152, 154
Zahl, 131
Gerade
affine, 122
GL(V) allgemeine lineare Gruppe, 93
GL($n; K$) allgemeine lineare Gruppe, 106
Goldbach-Vermutung, 138
goldener Schnitt, 19
Grad eines Polynoms, 142
Gram-Schmidt, 187, 188
Graph
einer Abbildung, 33
größter gemeinsamer Teiler, 135
Grp Gruppenhomomorphismen, 59
Gruppe, 50
Gruppe der Einheiten, 133
Gruppenhomomorphismus, 58

halbeinfacher Anteil
eines Endomorphismus, 230
Hau, Hauptraum, 225
Hauptraum, 225
Hauptvektor, 225
hermitesch, 175, 206
Hilbert'sche Probleme
Nummer 18, 266
Hilbertraum
endlichdimensionaler, 176
homogen
lineares Gleichungssystem, 71
homogener Raum, 242
homogenisieren
lineares Gleichungssystem, 71
Homomophismus
von Sequenzen, 287
Homomorphismus, 58
von Monoiden, 59
von Ringalgebren, 302
von Vektorräumen, 92
Hurwitz-Kriterium, 212
Hyperebene
affine, 122
lineare, 84

id, 35
idempotent
  in Rng, 130
  lineare Abbildung, 99
Identität, 35
Ikosaeder, 265
Ikosaedergruppe, 264
im, 98
image, 135
in, Morphismus in Koprodukt, 223
in_i
  Injektionen bei Summen, 93
indefinit, 211
Index
  von Untergruppe, 247
Induktion
  Induktionsannahme, 10
  Induktionsbasis, 10
  Induktionsschritt, 10
  Induktionsvoraussetzung, 10
  vollständige, 10
Injektion, 36
  kanonische, 93
injektiv
  Abbildung, 36
Inklusion, 38
innerer Automorphismus, 263
Integritätsbereich, 132
interior automorphisms, 263
Invariante
  von Gruppenwirkung, 241
invers, 49
  Matrix, 106
invertierbar, 50, 106, 133
Involution, 253
Isometrie, 189
isometrisch
  Isomorphismus, 189
isomorph
  Vektorräume, 92
Isomorphiesatz, 251
  Noether'scher, 251
Isomorphismus
  isometrischer, 189
  von affinen Räumen, 121
  von Gruppen, 58
  von Sequenzen, 287
  von Vektorräumen, 92
Isotropiegruppe, 241
Iwasawa-Zerlegung
  für $\mathrm{GL}(n;\mathbb{C})$, 189
  für $\mathrm{GL}(n;\mathbb{R})$, 188

Jägerzaunformel, 156
Jordan'sche Normalform, 236
Jordan-Block, 238
  nilpotenter, 232
Jordan-Zerlegung, 229

kanonisch
  Injektion, 93
kanonisches Skalarprodukt, 308
Kardinalität, 26
kartesisches Produkt, 28, 222
ker
  von linearer Abbildung, 98
Kern
  von Gruppenhomomorphismus, 135
  von linearer Abbildung, 98
Klassifikation
  abelsche Gruppen, 256
Kleiner Fermat, 254
Kodimension
  eines Untervektorraums, 286
Koeffizientenmatrix, 73
  erweiterte, 73
Körper, 55
Körperhomomorphismus, 59
Körperisomorphismus, 59
Kokern, 252
kommutativ
  Verknüpfung, 46

kommutativer Ring, 129
kommutieren, 142
Komplement, 26
    orthogonales, 180
komplementär
    Untervektorräume, 96
Komplexifizierung, 298
kongruent modulo, 130
Konjugation, 263
konjugiert
    Vektorraum, komplexer, 205
konstant
    Abbildung, 35
Koordinaten, 116
Koordinatenfunktionen, 116
Kreuzprodukt
    abstraktes, 201
    auf dem $\mathbb{R}^3$, 199
Kring
    kommutativer Ring, 129
Kristall, 265
Kristallklasse, 266
Kristallsystem, 265
Kronecker-Produkt, 294
Kroneckerdelta, 101
Kürzen in Ringen, 132
kurze exakte Sequenz, 286

Länge
    eines Vektors, 176
    in Einheiten, 306
    positive, 307
    von Permutation, 151
Lagrange
    Satz von, 247
Laufindex, 11
Laurentreihe
    formale, 145
leeren Familie, 79
Leibniz-Formel, 154
Leitkoeffizient, 143
Lemma, 48
linear
    Abbildung, 92
linear abhängig
    Familie, 85
    Teilmenge, 84
linear unabhängig
    Familie, 85
    Teilmenge, 84
lineare Anteil, 121
linearen Gleichungssystem, 71
Linearform, 113
Linearkombination, 83
Linksinverses, 96
Linksnebenklasse, 245
Lösungsmenge, 71
lokal
    nilpotent, 225
lokal endlich, 228
Lorentz-Metrik, 214

Mächtigkeit, 26
Matrix, 73
    quadratische, 73
Matrixmultiplikation, 102
Menge, 24
    G-Menge, 239
    leere Menge, 25
    Potenzmenge, 26
    Teilmenge, 25
Meter, 308
min, 46
Monoid, 49
Monoidhomomorphismus, 59
Morphismus
    von Monoiden, 59
multilinear, 159
Multimenge, 40
Multinomialkoeffizient, 39
Multiplikation
    in Ring, 129

$\mathbb{N}$ natürliche Zahlen, 25
$\mathbb{N}_0$, 25
natürliche Zahlen, 25
Nebenklasse, 245
negativ
    Vektor, 193
negativ definit, 211
negativ semidefinit, 211
Negatives, 52
neutrales Element, 49
nichtausgeartet
    Paarung, 216
nichtnegativ
    Vektor, 193
nilpotent
    Endomorphismus, 113
    in Rng, 130
    lokal, 225
nilpotenter Anteil
    eines Endomorphismus, 230
Noether'scher Isomorphiesatz, 251
Norm, 173
normal
    Endomorphismus, 226
    homogener Raum, 250
    Vektor, 176
Normalteiler, 249
normiert
    Polynom, 143
Nullring, 129
Nullstelle, 142
Nullteiler, 132
nullteilerfrei, 132
Nullvektor, 80
Nullvektorraum, 81

$\bigoplus$
    Summe von Vektorräumen, 222
$\mathrm{O}(V)$ orthogonale Automorphismen, 182
$\mathrm{O}(n)$ orthogonale Matrizen, 183
$\mathrm{U}(V)$ unitäre Automorphismen, 182
Oktaeder, 265
Operation
    einer Gruppe, 239
    triviale, 239
orbit, 242
$\operatorname{ord} g$ Ordnung von g, 253
Ordnung
    einer Gruppe, 253
    einer Nullstelle, 144
    von Gruppenelement, 253
orientierten Winkel, 196
Orientierung
    von Vektorraum, 192
orientierungserhaltend
    lineare Abbildung, 192
orientierungsumkehrend
    lineare Abbildung, 192
orthogonal, 176, 180
    Komplement, 180
    Matrix, 182
    Teilräume, 180
Orthogonalbasis, 215
    für Bilinearform, 212
orthogonale Projektion, 177
Orthogonalraum, 180
Orthonormalbasis, 177
Orthonormalsystem, 177

$\mathcal{P}(X)$ Potenzmenge, 26
$\prod$
    Produkt von Mengen, 222
    Produkt von Vektorräumen, 222
Paarung
    bilineare, 216
    nichtausgeartete, 216
parallel
    affine Teilräume, 122
Parallelogrammregel, 173
Partition
    einer Menge, 242

Pascal'sches Dreieck, 16
Permutation, 52
Polarisierungsidentität, 181
Polordnung, 268
Polstelle, 151
Polynomring, 141
positiv
    Vektor, 193
positiv definit, 211
    Bilinearform, 173
    hermitesche Matrix, 213
positiv orientiert
    Vektor, 193
positiv semidefinit, 211
Potenz
    p-Potenz, 256
    Primzahlpotenz, 256
Potenzmenge, 26, 78
Potenzreihe
    formale, 145
pr, Projektion aus Produkt, 223
pr_X
    Projektion, 35
pr_i
    Projektion, 77
Prä-Hilbertraum, 176
prim
    Restklasse, 141
Primfaktorzerlegung, 137
Primkörper, 140
Primzahl, 136
Primzahlpotenz, 256
Primzahlzwillinge, 138
prinzipaler homogener Raum, 242
produit extérieur, 302
Produkt
    von Mengen, 222
    von Sequenzen, 252
    von Vektorräumen, 222
Projektion
    bei zwei Mengen, 35
    längs Teilraum, 99
    von kartesischem Produkt, 77
Punkt, 25
Punktgruppe, 265
Pythagoras, Satz von, 176

$\mathbb{Q}$ rationale Zahlen, 25
quadratisch
    Matrix, 73, 106
quadratische Form, 201, 218
Quersumme, 131
Quot, 149
Quotient
    von Gruppe, 250
Quotientenvektorraum, 285

Radikal, 216
Raleigh-Quotient, 207
Rang
    einer abelschen Gruppe, 256
    einer Bilinearform, 216
    einer linearen Abbildung, 109
    einer Matrix, 109
rank, 109
rationale Funktion, 150
rationale Zahlen, 25
Raum, 25
    affiner, 119
Rechtsinverses, 96
Rechtsmenge, 244
Rechtsnebenklasse, 245
Rechtsoperation, 244
Rechtstorsor, 244
reeller Vektorraum, 23
Repräsentant, 148, 245
Repräsentantensystem, 148
Reskalierung
    von Translationen, 119
Restklasse, 130
    prime, 141
Restklassengruppe, 250

Restklassenring, 130
Richtungsraum, 119
Richtungsvektor, 119
Ring, 129
Ringalgebra, 302
Ringhomomorphismus, 132
Rng, 129
Rnghomomorphismen, 133

Schnitt
  von Mengenfamilie, 79
  von Mengensystem, 78
  zweier Mengen, 26
Schwerpunkt, 128
Sekunde, 194
selbstadjungiert, 206
selbstinvers, 253
Sequenz
  kurze exakte, 286
Sesquilinearform, 175
Sieb des Eratosthenes, 137
Signatur, 218
Signum, 154
Signum einer Permutation, 152
Skalar, 80
Skalarprodukt
  auf komplexem Vektorraum, 175
  auf reellem Vektorraum, 171
  kanonisches, 308
Skalarprodukt auf V mit Einheiten L, 306
skalarproduktsenkrecht, 173
Smith-Normalform, 108, 113
Smith-Zerlegung, 263
$\mathrm{SO}(V)$ spezielle orthogonale Automorphismen, 184
$\mathrm{SO}(n)$ spezielle orthogonale Matrizen, 184
Spaltenindex, 73
Spaltenrang, 109
Spaltenvektor, 101
Spann, 83
Spat, 200
Spatprodukt, 199
Spur
  einer Matrix, 112
  eines Endomorphismus
    von endlichdimensionalem Raum, 112
    von endlichem Rang, 112
stabil
  Teilmenge unter Abbildung, 224
  unter Gruppe, 242
Stabilisator, 241
Standardbasis, 86
Standardorientierung, 192
Standardskalarprodukt, 175
Standgruppe, 241
Strahl, 279
Streichmatrix, 162
$\mathrm{SU}(V)$ spezielle unitäre Automorphismen, 184
$\mathrm{SU}(n)$ spezielle unitäre Matrizen, 184
Summanden, 11
Summe
  von Untervektorräumen, 224
  von Vektorräumen, 222
Surjektion, 36
surjektiv
  Abbildung, 36
Sylvester
  Trägheitssatz, 217
Symmetrie, 264
  für Relation, 148
symmetrisch
  bilineare Abbildung, 158
  Bilinearform, 173
  Matrix, 203
symplektische Form, 220
symplektischer Vektorraum, 220
System von Teilmengen, 78

$\mathbb{T}$, 120
Teilen in Polynomringen, 143
Teiler, 132
teilerfremd, 135
Teilmenge, 25
    echte, 25
Teilraum, 82
teilt, 132, 135
Tensorprodukt
    über Körper, 290
Tetraeder, 265
Tetraedergruppe, 264
torsionsfrei
    Gruppe, 257
Torsor
    von links, 242
    von rechts, 244
tr Spur alias “trace”, 112
trace
    einer Matrix, 112
Trägheitssatz
    Sylvester’scher, 217
transitiv
    Gruppenwirkung, 242
Translation, 119
transponiert
    Matrix, 104
transponierte, 115
trivial
    Operation, 239
Tupel, 77

U(n) unitäre Matrizen, 183
umgeklappte Dreieck, 275
Umkehrfunktion, 39
ungerade
    Permutation, 152, 154
    Zahl, 131
unipotent, 232
unitär, 180
    Matrix, 182
unitärer Raum, 176
Universelle Eigenschaft
    des Quotientenraums, 285
    des Raums der Äquivalenzklassen, 149
Untergruppe, 134
    erzeugt von Teilmenge, 134
    triviale, 134
Untervektorraum, 82
Urbild
    von Menge, 36

van-de-Ven-Diagramme, 27
van-der-Monde-Determinante, 161
Vektor
    Element eines Vektorraums, 80
Vektorraum, 79
    komplex konjugierter, 205
Vereinigung, 26
    von Mengenfamilie, 79
    von Mengensystem, 78
Verjüngung von Tensoren, 297
Verknüpfung
    auf einer Menge, 44
    von Abbildungen, 36
Verknüpfungstabelle, 45
Vielfachheit
    einer Nullstelle, 144
voll
    Rang, 109

wedge-product, 302
Wert, 33
Wertebereich, 33
Wilson
    Satz von, 141
Winkel, 191
Wirkung
    einer Gruppe, 239
Würfel, 265
Würfelgruppe, 264

Wurzel
    von Polynom, 142

×
    kartesisches Produkt, 28

Young-Diagramm, 232

$\mathbb{Z}$ ganze Zahlen, 25
Zahl
    ganze, 25
    gerade, 131
    natürliche, 25
    rationale, 25
    ungerade, 131
Zeilenindex, 73
Zeilenrang, 109
Zeilenstufenform, 71
Zeilenvektor, 104
Zeiteinheit
    nichtrelativistische, 194
Zeitpunkt, 120
Zeitspanne, 194
Zerlegung in Linearfaktoren, 144
zyklisch
    Anordnung, 40
    Gruppe, 252