5 Matrizen

5.1 Vorbemerkung.

In der weiterführenden Übung zum letzten Kapitel haben wir den nun anstehenden Schritt im Wesentlichen vorweggenommen. Fassen wir zusammen: wir betrachten lineare Abbildungen $L : \mathbb{R}^n \to \mathbb{R}^m$ zwischen den Vektorräumen $\mathbb{R}^n$ und $\mathbb{R}^m$. Sowohl im Definitions- als auch im Zielraum von L wählen wir die entsprechende Basis aus Einheitsvektoren $\mathbf{e}_i$. Es ist L wegen seiner Linearität bereits durch den Satz der $L\mathbf{e}_i$ vollständig gegeben, und die zu L gehörende Matrix ist nichts anderes als das Zahlenschema, das sich gemäß gewisser Schreib-Konventionen ergibt, wenn wir für jedes der $\mathbf{e}_i \in \mathbb{R}^n$ die sich ergebenden $L\mathbf{e}_i \in \mathbb{R}^m$ in die Basis der $\mathbf{e}_i \in \mathbb{R}^m$ entwickeln. Der erste Abschnitt dieses Kapitels führt dieses einfache Vorhaben explizit aus.

Hat man den Begriff der Matrix in dieser Weise aber einmal definiert, beginnen Matrizen als eigenständige mathematische Objekte zu leben. Zum Beispiel können wir (unter bestimmten Voraussetzungen) Matrizen addieren oder multiplizieren. Davon wird in den Folgeabschnitten die Rede sein. Im Zusammenhang mit der Frage, wann einer gegebenen Matrix eine multiplikativ inverse Matrix zugeordnet werden kann, betritt die sogenannte *Determinante* einer Matrix die Bühne — dafür gibt es dann aber ein eigenes Kapitel; ebenso ein eigenes Kapitel verdient das Thema der Lösung linearer Gleichungssysteme mit Hilfe der Matrix- und Determinantenbetrachtungen.

5.2 Matrizen.

Das Setting für die Definition des Begriffs einer *Matrix* ist, wie in der Vorbemerkung skizziert, das folgende: gegeben seien die beiden Vektorräume $\mathbb{R}^n$, ausgestattet mit der Basis aus Einheitsvektoren $(\mathbf{e}_1, \dots, \mathbf{e}_n)$, und $\mathbb{R}^m$ versehen mit $(\mathbf{e}_1, \dots, \mathbf{e}_m)$. Und außerdem natürlich eine lineare Abbildung $L : \mathbb{R}^n \to \mathbb{R}^m$. Wir wollen die zu L gehörende Matrix definieren.

Man halte sich vor Augen, dass es in $\mathbb{R}^n$ für *jeden* Vektor $\mathbf{v}$ eine *eindeutige* Darstellung $\mathbf{v} = \sum_{i=1}^n v_i\mathbf{e}_i$ gibt (das ist die zentrale Eigenschaft einer Basis) und entsprechend in $\mathbb{R}^m$ für jedes $\mathbf{w}$ eine eindeutige Darstellung $\mathbf{w} = \sum_{i=j}^m w_j\mathbf{e}_j$; wir verzichten darauf, die n Einheitsvektoren in $\mathbb{R}^n$ anders zu bezeichnen als die m Einheitsvektoren in $\mathbb{R}^m$. Somit lässt sich $L\mathbf{e}_1$ als Linearkombination der $\mathbb{R}^m$-Basis schreiben: $L\mathbf{e}_1 = \sum_{j=1}^m (L\mathbf{e}_1)_j \mathbf{e}_j$, wobei die m Zahlen $(L\mathbf{e}_1)_j$ gerade die Komponenten des Spaltenvektors $L\mathbf{e}_1 \in \mathbb{R}^m$ sind. Das gleiche Manöver klappt natürlich auch für die anderen $L\mathbf{e}_i$, $i = 2 \ldots n$, so dass wir insgesamt haben:

$$L\mathbf{e}_i = \sum_{j=1}^{m} (L\mathbf{e}_i)_j \, \mathbf{e}_j \text{ für } i = 1 \ldots n \ .$$

Es ist reine Konvention, aber eine praktische – weil mnemotechnisch vorteilhafte – Konvention, nun die Gesamtheit der Zahlen $L_{ji} := (L\mathbf{e}_i)_j$ zu definieren und als *Matrix zur linearen Abbildung* L zu bezeichnen. Man bemerke die Reihenfolge der auftretenden Indizes! und wir werden gleich sehen, inwiefern diese Festlegung das Merken der Rechenoperationen erleichtert. Wieviel dieser Zahlen L_{ji} gibt es? Nun, i läuft von 1 bis n und j von 1 bis m. Das ergibt insgesamt $n \cdot m$ Zahlen L_{ji}. Ebenso eine zwar im Kern willkürliche, aber aus der Praxis gesehen vorteilhafte Konvention ist es, diese $n \cdot m$ Zahlen wie folgt zu arrangieren:

$$\begin{pmatrix} L_{11} & L_{12} & \cdots & L_{1n} \\ L_{21} & L_{22} & \cdots & L_{2n} \\ \vdots & \vdots & & \vdots \\ L_{m1} & L_{m2} & \cdots & L_{mn} \end{pmatrix}$$

Also: der Index j, der von 1 bis m, also über die Dimension des *Ziel*-Vektorraums läuft, wird als *Zeilen*index gelesen. Der Index i, der von 1 bis n, also über die Dimension des *Definitions*-Vektorraums läuft, bildet den *Spalten*index.[1] Das Gesamtschema heißt, die Leserin oder der Leser werden es bereits erraten haben, *Matrix* der linearen Abbildung L.

Hier ist noch eine weitere Sicht auf die Welt dieser Schreibkonventionen. Wir hatten ja vereinbart, Vektoren stets in Spaltenform zu schreiben und, wo aus Platzgründen dies nicht opportun erscheint, mit einem hochgestellten “T” anzudeuten, dass zum Beispiel der zeilensparend geschriebene Vektor $(5, -1, 3)^\mathsf{T}$ eben *doch* als Spaltenvektor gelesen werden will. Wir können uns eine Matrix von etwa 4

[1] Übrigens: was tun, wenn einer oder beide der Indizes an L_{ij} mehrstellig werden? Was soll L_{125} heißen: der Eintrag in der 12. Zeile und der 5. Spalte oder doch der in der 1. Zeile und der 25. Spalte? Erstaunlicherweise bleibt die Frage nach einer diesbezüglichen Notationskonvention unbeantwortet – weil es in der Praxis praktisch nie vorkommt. Von Hand bearbeitet kein Mensch derartig große Matrizen, und in numerischen rechnergestützten Anwendungen stellt sich das Problem ebenfalls nicht. In der Mathematik trifft man sowieso nur auf abstrakt gegebene Matrizen, so dass dort als Indizes nur die 1, vielleicht noch die 2 und im übrigen Buchstaben oder mit Buchstaben indizierte Buchstaben auftreten. Wenn sich also jemand aus welchen Gründen auch immer doch vor der Notwendigkeit gestellt sieht, klar anzugeben, was mit L_{125} gemeint ist, muss er oder sie kreativ werden. Man könnte ja $L_{12\ 5}$ schreiben oder so etwas.

Spalten und 3 Zeilen dadurch entstanden denken, dass wir jede der 4 Spalten als Vektor betrachten und die 4 Vektoren nebeneinander setzen. Haben wir konkret die 4 Vektoren $\mathbf{v}_1 = \begin{pmatrix} (\mathbf{v}_1)_1 \\ (\mathbf{v}_1)_2 \\ (\mathbf{v}_1)_3 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} (\mathbf{v}_2)_1 \\ (\mathbf{v}_2)_2 \\ (\mathbf{v}_2)_3 \end{pmatrix}, \ldots, \mathbf{v}_4 = \begin{pmatrix} (\mathbf{v}_4)_1 \\ (\mathbf{v}_4)_2 \\ (\mathbf{v}_1)_4 \end{pmatrix}$ gegeben, dann ergeben diese entsprechend zusammengestellt die Matrix $\begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4 \end{pmatrix} = \begin{pmatrix} (\mathbf{v}_1)_1 & (\mathbf{v}_2)_1 & \cdots & (\mathbf{v}_4)_1 \\ (\mathbf{v}_1)_2 & (\mathbf{v}_2)_2 & \cdots & (\mathbf{v}_4)_2 \\ (\mathbf{v}_1)_3 & (\mathbf{v}_2)_3 & \cdots & (\mathbf{v}_4)_3 \end{pmatrix}$. Hier sehen wir wieder die zunächst so eigenartig erscheinende "Umstellung" der Matrixindizes aus der Definition $L_{ji} := (L\mathbf{e}_i)_j$, denn die Matrix der L_{ji} ist in dem gerade geschilderten Sinne ja nichts als $\begin{pmatrix} L\mathbf{e}_1 & L\mathbf{e}_2 & \cdots & L\mathbf{e}_n \end{pmatrix}$, wo der "innere" Index i in $(L\mathbf{e}_i)_j$ gerade als *Spalten*index fungiert und der "äußere" Index j über die Dimension m dieser Spaltenvektoren läuft. Man kann daher formulieren, die Spalten einer Matrix seien die Bilder der Einheitsvektoren des Definitions-Vektorraums unter L. Diese Formulierung ist sehr prägnant und hilfreich und wir werden sie im nächsten Abschnitt als eine Art Merksatz herausstellen.

Wer pedantisch ist – und wer wäre das nicht? – unterscheidet strikt zwischen einer linearen Abbildung L und ihrer Matrix. Letztere wird häufig als (L_{ij}) bezeichnet oder auch als (l_{ij}), die Menge aller Matrizen mit rellen Komponenten, m Zeilen und n Spalten als $M(m \times n, \mathbb{R})$ oder als $\mathbb{R}^{m \times n}$ oder hie und da auch als $\mathbb{R}^{m,n}$. Nun haben wir bis jetzt lediglich *definiert*, was die Matrix einer linearen Abbildung sein soll. Als eine Art von komponentenorientierter Repräsentation einer linearen Abbildung muss eine Matrix aber zumindest folgendes leisten: sie muss das konkrete Berechnen des Bildvektors unter L für jeden beliebigen Vektor aus dem Definitions-Vektorraum möglich machen. Natürlich schafft sie das. Sie ist ja entstanden aus der sukzessiven Anwendung von L auf die Einheitsvektor-Basis des Definitions-Vektorraums, und der Linearität von L wegen brauchen wir mehr nicht zu wissen. Die explizite Ausführung dieses Gedankens bildet den Inhalt des folgenden kurzen Abschnitts.

5.3 Anwendung von Matrizen auf Vektoren.

Gegeben sei eine lineare Abbildung $L : \mathbb{R}^n \to \mathbb{R}^m$ und ihre Matrix $(L_{ji}) \in M(m \times n, \mathbb{R})$. Wir wenden L auf Vektoren $\mathbf{v} = (v_1, v_2, \ldots, v_n)^\mathsf{T}$ an und möchten dazu die Rechenregel herleiten, die uns die m Komponenten von $L\mathbf{v}$ liefert, wenn wir die n Komponenten von $\mathbf{v}$ und die $m \cdot n$ Komponenten von (L_{ji}) kennen.

Um diese Rechenregel herzuleiten, benötigen wir lediglich die Definition $L_{ji} := (L\mathbf{e}_i)_j$ und sonst nichts. Denn wenn wir L auf irgend einen Vektor $\mathbf{v}$ aus $\mathbb{R}^n$ anwenden, der die Komponentendarstellung $\mathbf{v} = (v_1, \ldots, v_n)^\mathsf{T} = \sum_i v_i\mathbf{e}_i$ hat, resultiert wegen der Linearität von L ja einfach $L\mathbf{v} = L(\sum_i v_i\mathbf{e}_i) = \sum_i v_i\, L\mathbf{e}_i$. Wenn wir also von dem Vektor $L\mathbf{v} \in \mathbb{R}^m$ die j-te Komponente wissen wollen, brauchen wir die j-te Komponente von $L\mathbf{e}_i \in \mathbb{R}^m$ (die v_i sind ja nur *Zahlen*), und die ist eben gerade $(L\mathbf{e}_i)_j = L_{ji}$. Eingesetzt folgt $(L\mathbf{v})_j = \sum_i v_i L_{ji}$, und das ist das gewünschte Ergebnis[2]:

Sei $L : \mathbb{R}^n \to \mathbb{R}^m$ linear mit Matrix $(L_{ij}) \in M(m \times n, \mathbb{R})$ sowie $\mathbf{v} \in \mathbb{R}^n$, dann ist

$$(L\mathbf{v})_i = \textstyle\sum_{j=1}^n L_{ij}v_j$$

die i-te Komponente von $L\mathbf{v}$, wenn v_j die j-te Komponente von $\mathbf{v}$ ist.

Direkt als Zahlenschema notiert liest sich das so:

$$\begin{pmatrix} L_{11}v_1 + L_{12}v_2 + \cdots + L_{1n}v_n \\ L_{21}v_1 + L_{22}v_2 + \cdots + L_{2n}v_n \\ \cdots \\ L_{m1}v_1 + L_{m2}v_2 + \cdots + L_{mn}v_n \end{pmatrix} = \begin{pmatrix} L_{11} & L_{12} & \cdots & L_{1n} \\ L_{21} & L_{22} & \cdots & L_{2n} \\ \vdots & \vdots & & \vdots \\ L_{m1} & L_{m2} & \cdots & L_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \cdots \\ v_n \end{pmatrix}$$

Man legt den $\mathbf{v}$-Vektor quer über die *erste* Zeile der Matrix (und die Längen passen, sonst stimmt etwas nicht!), multipliziert die dann übereinander liegenden Zahlen und addiert die Podukte; das ergibt die *erste* Komponente des Ergebnisvektors. Das gleiche macht man dann mit der zweiten Zeile, mit der dritten Zeile und so weiter, um vom Ergebnisvektor die zweite Komponente zu erhalten, die dritte Komponente und so weiter.

In den Übungen finden sich Vorschläge zum Ausprobieren und Einüben dieser elementaren, aber zentralen Rechenoperation. Als erstes überzeuge man sich aber davon, dass das Anwenden einer beliebigen Matrix auf den *ersten* Einheitsvektor $\mathbf{e}_1 = (1, 0, 0, \ldots, 0)^\mathsf{T}$ gerade die *erste* Spalte der Matrix sozusagen “herauspräpariert”, die Anwendung auf den *zweiten* Einheitsvektor $\mathbf{e}_2 = (0, 1, 0, \ldots, 0)^\mathsf{T}$ die *zweite* Spalte und so weiter. Sich dies klarzumachen und dann auch gut zu merken ist sehr nützlich und hilft bei vielen späteren Überlegungen rasch weiter. Wir schließen uns Klaus Jänich an, der in seinem Lehrbuch über Lineare Algebra einen Merkspruch aus dem

[2] wir tauschen j und i durch Umbenennen für leichtere Eingängigkeit.

genannten Sachverhalt macht und vorschlägt, diesen zu memorieren: “Die Spalten / Sind Bilder / Der Einheits- / Vektoren.” Zwar reimt sich das nicht, das ist ja aber bei moderner Lyrik häufig so. Der Spruch hat jedoch einen gewissen Sprachrhythmus, und dies erleichtert das Merken. Also: auswendiglernen!

5.4 Rechnen mit Matrizen.

Was können wir mit Matrizen alles machen? Diese Frage verweisen wir zurück auf die Frage, was man mit linearen Abbildungen machen kann. Zum einen kann man zwei lineare Abbildungen punktweise *addieren*, und das Ergebnis ist wieder eine lineare Abbildung: seien K und L beides Abbildungen von $\mathbb{R}^n$ nach $\mathbb{R}^m$ und setzen wir $(K + L)(\mathbf{v}) := K\mathbf{v} + L\mathbf{v}$ für jedes $\mathbf{v} \in \mathbb{R}^n$, dann ist – bitte nachrechnen! – das so festgelegte $K + L$ eine lineare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$. Welche Matrix $((K + L)_{ji})$ entspricht dieser Abbildung? Den Leser wird es nicht wundern: für jedes Indexpaar j, i ist einfach $(K + L)_{ji} = K_{ji} + L_{ji}$, oder in Worten ausgedrückt: man addiert komponentenweise die beiden (gleichdimensionierten!) Matrizen zu K und zu L zur Summenmatrix von $K + L$. Außerdem gilt, dass mit K auch $\lambda \cdot K$ wieder linear ist für jede reelle Zahl λ – auch hier wird λK aus K punktweise definiert –, und die Matrix zu λK ergibt sich aus der zu K dadurch, dass jeder Eintrag der letzteren mit λ malgenommen wird. Das ist alles eher mäßig interessant, und wer’s nicht glaubt, möge es in der entsprechenden Übung explizit nachrechnen. Erwähnenswert ist vielleicht aber immerhin, dass durch diese Operationen die Menge $M(m \times n, \mathbb{R})$ selbst wieder zu einem Vektorraum wird, denn wir haben nun Addition und Äußere Multiplikation dieser “Vektoren” (aka Matrizen) axiomenkonform erklärt. Letztlich ist in diesem Sinne eine Matrix nichts anderes als ein Vektor, dessen Komponenten ein bisschen anders auf dem Rechenpapier arrangiert sind, woraus auch sofort folgt, dass dieser Lineare Raum die Dimension $m \cdot n$ hat.

Außerdem ist aber die *Hintereinanderausführung* zweier linearer Abbildungen wieder eine lineare Abbildung, und *das* ist der eigentlich interessante Punkt, den wir in diesem Abschnitt auf den Matrizenkalkül überspielen. Angenommen also, vor uns liegen zwei lineare Abbildungen $K : \mathbb{R}^n \to \mathbb{R}^m$ und $L : \mathbb{R}^m \to \mathbb{R}^s$. Diese beiden können wir dann hintereinander ausführen im Sinne von $L \circ K$. Das heißt, wir wenden erst K auf einen n-dimensionalen Vektor an, und dann L auf das m-dimensionale Ergebnis der Anwendung von K. Schließlich erhalten wir einen s-dimensionalen Vektor.

Sowohl K als auch L sind Matrizen zugeordnet, K eine $m \times n$-Matrix (K_{ij}) und L eine $s \times m$-Matrix (L_{kl}). Aber zu der zusammengesetzten Abbildung $L \circ K$ gehört natürlich ebenfalls eine Matrix! und zwar eine $s \times n$-dimensionale, denn $L \circ K$ bildet ja von $\mathbb{R}^n$ nach $\mathbb{R}^s$ ab. Die Frage ist: lässt sich die Matrix $((L \circ K)_{pq})$ aus den Matrizen (K_{ij}) und (L_{kl}) irgendwie berechnen? Natürlich lässt sie sich das. Und wie?

Hier müssen wir wieder einmal mit Sorgfalt eine eigentlich einfache Rechnung so durchführen, dass der Überblick nicht verlorengeht. Wir betrachten drei lineare Abbildungen und die zugehörigen Matrizen:

$L : \mathbb{R}^m \to \mathbb{R}^s$ mit $L(\mathbf{e}_k) = \sum_{i=1}^{s} L_{ik}\mathbf{e}_i$
$K : \mathbb{R}^n \to \mathbb{R}^m$ mit $K(\mathbf{e}_j) = \sum_{k=1}^{m} K_{kj}\mathbf{e}_k$
$(L \circ K) : \mathbb{R}^n \to \mathbb{R}^s$ mit $(L \circ K)(\mathbf{e}_j) = \sum_{i=1}^{s}(L \circ K)_{ij}\mathbf{e}_i$.

Sodann stellen wir die folgende Gleichungskette auf: $\sum_{i=1}^{s}(L \circ K)_{ij}\mathbf{e}_i = (L \circ K)(\mathbf{e}_j) = L(K(\mathbf{e}_j)) = L(\sum_{k=1}^{m} K_{kj}\mathbf{e}_k) = \sum_{k=1}^{m} K_{kj}L(\mathbf{e}_k) = \sum_{k=1}^{m} K_{kj}\sum_{i=1}^{s} L_{ik}\mathbf{e}_i = \sum_{i=1}^{s}\big(\sum_{k=1}^{m} L_{ik}K_{kj}\big)\mathbf{e}_i$.

Und schließlich brauchen wir nur noch den ersten mit dem letzten Ausdruck zu vergleichen, um die gewünschte Beziehung zu erhalten:

$$(L \circ K)_{ij} = \sum_{k=1}^{m} L_{ik}K_{kj}$$

Auch hier wieder sollte man ein paar konkrete Matrixmultiplikationen durchrechnen, um Übung zu bekommen und das Prinzip durchdacht zu haben. Das Übungsblatt macht in dieser Hinsicht ein paar Vorschläge.

Im Zusammenhang mit quadratischen Matrizen – die eben wegen ihrer Übereinstimmung der Zeilen- mit der Spaltenanzahl in jeder Reihenfolge aufeinander multipliziert werden können – ist die sogenannte *Einheitsmatrix* $E = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$ interessant. Man versäume es nicht, die Produkte $E \cdot L$ und $L \cdot E$ für irgend eine Matrix L passender Dimension explizit auszurechnen, um sich davon zu überzeugen, dass E bezüglich der Matrixmultiplikation wie eine “1” agiert: $L \cdot E = L$

und $E \cdot L = L$. Wer dieser Vorlesung auch in ihren abstrakteren Teilen aufmerksam gefolgt ist, wird sich sofort fragen, ob die Menge der Matrizen $M(n \times n, \mathbb{R})$ mit der Matrixmultiplikation denn vielleicht eine Gruppe sein könnte mit E als Neutralem Element dieser Verknüpfung? Leider nein[3], und der Grund dafür ist, dass nicht jede quadratische Matrix L ein *Inverses* besitzt, also eine Matrix L^{-1} mit der Eigenschaft $L^{-1} \cdot L = L \cdot L^{-1} = E$; müsste sie aber, wenn $(M(n \times n, \mathbb{R}), \cdot)$ eine Gruppe sein sollte. Zum Thema "Invertieren von Matrizen" folgt noch weiter unten ein Abschnitt.

Eine andere Operation mit Matrizen, die häufig vorkommt, ist das *Transponieren* einer gegebenen Matrix. Sei dazu $(L_{ij}) \in M(m \times n, \mathbb{R})$ eine Matrix mit m Zeilen und n Spalten, dann ist $(L_{ij}^{\mathsf{T}}) \in M(n \times m, \mathbb{R})$ die dazu transponierte Matrix mit n Zeilen und m Spalten, wenn $L_{ij}^{\mathsf{T}} = L_{ji}$ ist; man beachte die Reihenfolge der Indizes. In Worten: die Transposition vertauscht Zeilen und Spalten. Die Transponierte hat als Zeilen das, was bei der ursprünglichen Matrix die Spalten waren, und umgekehrt. In diesem Sinne können wir auch das hochgestellte "T" interpretieren, das wir immer benutzt haben, um einen als Zeile geschriebenen Vektor zu einem Spaltenvektor zu machen: der Zeilenvektor kann identifiziert werden mit einer Matrix aus $M(1 \times n, \mathbb{R})$ und wird durch das Transponieren deshalb zu einer Matrix aus $M(n \times 1, \mathbb{R})$, die als Spaltenvektor interpretiert werden kann. Wir werden auf das Transponieren insbesondere im Zusammenhang mit dem Skalarprodukt stoßen.— Übrigens heißt eine quadratische Matrix L mit der Eigenschaft $L = L^{\mathsf{T}}$ aus naheliegenden Gründen *symmetrisch* und eine mit $L = -L^{\mathsf{T}}$ *antisymmetrisch* oder auch *schiefsymmetrisch*.

5.5 Eigenschaften der Matrixmultiplikation.

Es sind vor allem die beiden prominentesten Rechengesetze, die wir auf Gültigkeit im Falle der Matrixmultiplikation überprüfen müssen: das Assoziativgesetz und das Kommutativgesetz. In den Übungen wird der Leser aufgefordert, an Hand eines selbstausgedachten Beispiels für drei Matrizen K, L und M nachzurechnen, dass $(K \cdot L) \cdot M = K \cdot (L \cdot M)$ ist. Es ist aber auch nicht schwierig, dies mittels der allgemeinen Formel für die Matrixmultiplikation $(L \cdot K)_{ij} = \sum_{k=1}^{m} L_{ik} K_{kj}$ explizit zu beweisen. Unterm Strich bedeutet diese gute Nachricht, dass beim Bilden von Matrixprodukten aus mehreren Faktoren die Beklammerung gleichgültig ist (und mithin auch weggelassen werden kann).

[3]... aber zusammen mit der Matrixaddition ist $(M(n \times n, \mathbb{R}), +, \cdot)$ ein *Ring*.

Weniger eingängig sieht es aus, was das Kommutativgesetz betrifft, also die Frage nach der Reihenfolge der Faktoren bei einer Matrixmultiplikation. Einerseits sollte man sich klarmachen, dass man nicht notwendig $L \cdot K$ überhaupt bilden kann, selbst wenn $K \cdot L$ existiert. Aber selbst *wenn* beide beteiligten Matrizen quadratisch sind und von gleicher Dimension, ist im Allgemeinen keineswegs $K \cdot L = L \cdot K$. Das Kommutativgesetz gilt zwar für die Matrix*addition*, für die Matrix*multiplikation* aber *nicht*. Auch hier fordert die genannte Übungsaufgabe zur Bildung eines Beispiels auf.

Im Übrigen lässt sich wiederum mittels der zitierten allgemeinen Formel der Matrixmultiplikation quasi durch Hinschauen erkennen, dass man für Matrizen auch das *Distributivgesetz* einsetzen kann: $K \cdot (L + M) = K \cdot L + K \cdot M$ mit der üblichen Konvention "Punktrechnung vor Strichrechnung" auf der rechten Seite. So können wir insgesamt mit Matrizen praktisch so rechnen wie mit Zahlen, wenn wir nur stets beachten, dass wir die Reihenfolge der Faktoren in einer Matrixproduktbildung nicht ändern dürfen.

5.6 Die Inverse einer Matrix.

Im Abschnitt "Rechnen mit Matrizen" haben wir die Einheitsmatrix $E \equiv (E_{ij})$ kennengelernt (Einsen auf der Hauptdiagonalen, Nullen sonst überall) und gesehen, dass das Produkt irgendeiner quadratischen Matrix (L_{ij}) mit (E_{ij}) wieder (L_{ij}) ergibt, (E_{ij}) mithin also als eine Art Neutrales Element der Matrixmultiplikation fungiert. Für (L_{ij}) wünschen wir uns dann eine *inverse Matrix* (L_{ij}^{-1}) mit der Eigenschaft $(L_{ij}^{-1}) \cdot (L_{ij}) = (L_{ij}) \cdot (L_{ij}^{-1}) = (E_{ij})$ entsprechend der Inversenbildung beispielsweise bei der Multiplikation gewöhnlicher Zahlen[4]. In einer Übungsaufgabe zu diesem Kapitel sehen wir, wie man – wenigstens im einfachen Fall einer 2×2-Matrix – so eine Inverse ausrechnen kann, ohne allerdings auf eine allgemein anwendbare Strategie abzuheben; das wird dann das Thema eines späteren Kapitels werden. In eben dieser Übung finden wir aber auch ein Beispiel für eine 2×2-Matrix, die gar kein Inverses *hat*.

Was *ist* im abstrakten Sinn aber eigentlich das Inverse einer Matrix, wann existiert sie, was *ist* die Einheitsmatrix?

Die Einheitsmatrix fungiert als Neutrales Element der Matrixmultiplikation, wie wir gesehen

[4]... aber wir erinnern uns daran, dass nicht jede Zahl ein multiplikativ Inverses besitzt: die 0 nämlich nicht. Das hat im vorliegenden Zusammenhang eine Entsprechung, wie wir sehen werden.

haben: $(L_{ij}) \cdot (E_{ij}) = (E_{ij}) \cdot (L_{ij}) = (L_{ij})$ für jede quadratische Matrix (L_{ij}). Die Matrixmultiplikation ist nun aber nichts anderes als die Matrix-Version der Hintereinanderausführung der entsprechenden linearen Abbildungen - so haben wir sie weiter oben im Abschnitt **??** ja gerade konstruiert. Die Abbildung, die nichts tut und daher in einer Hintereinanderausführung keine Wirkung hat, ist natürlich die Identitätsabbildung Id; wir haben für eine lineare Abbildung $L : \mathrm{V} \to \mathrm{V}$ bekanntermaßen $L \circ \mathrm{Id} = \mathrm{Id} \circ L = L$ für $\mathrm{Id} : \mathrm{V} \to \mathrm{V}; \mathbf{v} \mapsto \mathbf{v}$, das genaue Analogon zur obigen Matrixgleichung. Die Einheitsmatrix ist demnach diejenige Matrix, die der Identität zugeordnet ist[5].

Damit wird deutlich, was das Inverse einer Matrix eigentlich ist und wann es existiert. Denn eine "Hintereinanderausführungs-Gleichung" für lineare Abbildungen der Art $L \circ L^{-1} = L^{-1} \circ L = \mathrm{Id}$ gilt dann und nur dann, wenn L invertierbar den Vektorraum V isomorph in sich selber abbildet (siehe Abschnitt "Isomorphismen" im Kapitel "Lineare Abbildungen"). Die zugeordnete Matrixgleichung ist nun gerade $(L_{ij}^{-1}) \cdot (L_{ij}) = (L_{ij}) \cdot (L_{ij}^{-1}) = (E_{ij})$. Folglich gilt:

> Eine Matrix hat genau dann eine Inverse, wenn die zugehörige lineare Abbildung bijektiv, also invertierbar respektive ein Isomorphismus ist. Die Inverse der Matrix ist in diesem Fall die der Umkehrabbildung zugeordnete Matrix.

Wenn übrigens K und L beides Matrizen sind[6], die ein Inverses besitzen, dann trifft das auch auf ihr Matrixprodukt $K \cdot L$ zu. Denn wenn wir $L^{-1} \cdot K^{-1}$ – und diese Größe existiert ja, weil K und L beide ein Inverses haben – auf $K \cdot L$ anwenden, bekommen wir $(L^{-1} \cdot K^{-1}) \cdot (K \cdot L) =$ (Assoziativität!) $L^{-1} \cdot (K^{-1} \cdot K) \cdot L = L^{-1} \cdot E \cdot L = L^{-1} \cdot L = E$. Richtig gelesen, heißt das aber nichts anderes, als dass das Inverse zu KL existiert und dann $(KL)^{-1} = L^{-1}K^{-1}$ ist. Die Reihenfolge der Matrizen kehrt sich also um.

Wir wollen die Frage nach der Bijektivität beziehungsweise Invertierbarkeit linearer Abbildungen – und damit das Problem der Existenz einer Inversen zu einer Matrix – noch etwas genauer betrachten. Die Dimensionsformel

$$\dim(\mathrm{Kern}\, L) + \dim(\mathrm{Bild}\, L) = \dim(\mathrm{V})$$

[5] Unser Spruch "Die Spalten / ..." und so weiter zeigt uns dies ebenfalls: die Spalten der Einheitsmatrix sind ja unverändert wieder die Einheitsvektoren.

[6] Aus praktischen Gründen benutzen wir im Folgenden einen Buchstaben wie L oder K zum Bezeichnen einer Matrix und gleichzeitig ihrer linearen Abbildung.

ist dabei, wie wir bereits andeuteten im entsprechenden Abschnitt des Kapitels über lineare Abbildungen, der Schlüssel zu dieser Frage. Was kann sie uns sagen über die Invertierbarkeit einer linearen Abbildung $L : \mathrm{V} \to \mathrm{W}; \mathbf{v} \mapsto L\mathbf{v}$ respektive der zugehörigen Matrix? Zunächst ist klar: weil L ja *bijektiv* sein muss, um invertierbar zu sein, können höchstens diejenigen linearen Abbildungen invertierbar sein, die von einem n-dimensionalen Vektorraum V in einen Vektorraum *gleicher Dimension* abbilden. Wäre der Zielraum W von höherer Dimension als $\dim(\mathrm{V}) = n$, dann könnte L nicht surjektiv sein, weil Bild L gemäß Dimensionsformel höchstens von Dimension n ist und damit echt kleiner wäre als W. Hätte der Zielraum W (und damit auch Bild $L \subseteq \mathrm{W}$) hingegen eine kleinere Dimension als n, müsste nach der Dimensionsformel $\dim \mathrm{Kern}\, L > 0$ sein. Es gäbe also verschiedene Vektoren in V, die von L auf den gleichen Vektor in W abgebildet würden – L wäre dann nicht injektiv. So bleibt für ein bijektives L nur die Möglichkeit $\dim\mathrm{V} = \dim\mathrm{W}$ übrig. Mit anderen Worten: *höchstens quadratische Matrizen sind invertierbar*, alle anderen ganz sicher nicht.

Die Bedingung, quadratisch zu sein, ist zwar *notwendig* für Invertierbarkeit, aber nicht *hinreichend*. Ein Blick auf die Dimensionsformel liefert uns aber gleich zwei hinreichende Bedingungen für Umkehrbarkeit: $\dim \mathrm{Kern}\, L = 0$ (wir erinnern uns: $\dim \mathrm{Kern}\, L = 0 \Leftrightarrow L$ ist injektiv) und – der Formel wegen gleichbedeutend damit – $\dim \mathrm{Bild}\, L = \dim\mathrm{V}$ (wir erinnern uns nochmal: $\dim \mathrm{Bild}\, L = \dim\mathrm{V} \Leftrightarrow L$ ist surjektiv). Wenn also eine lineare Abbildung zwischen zwei *gleich*dimensionalen Vektorräumen injektiv ist, dann ist sie automatisch gleich schon bijektiv. Die Linearität von L wie weiter oben schon explizit auf Linearkombinationen von Vektoren in V ausnutzend, folgt (das wissen wir aus dem Abschnitt über Isomorphismen im vorigen Kapitel): wenn $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ irgendeine Basis von V darstellt und $L : \mathrm{V} \to \mathrm{W}; \mathbf{v} \mapsto L\mathbf{v}$ ein Isomorphismus ist, dann ist $(L\mathbf{v}_1, \ldots, L\mathbf{v}_n)$ eine Basis von W – *und umgekehrt*. Damit haben wir eine direkte Möglichkeit, eine gegebene (quadratische!) Matrix daraufhin zu prüfen, ob ihre lineare Abbildung ein Isomorphismus, sie selbst also invertierbar ist. Wir wählen als Basis die Einheitsvektoren $\mathbf{e}_1 = (1, 0, \ldots, 0)^\mathsf{T}, \ldots, \mathbf{e}_n = (0, \ldots, 0, 1)^\mathsf{T}$, erinnern uns an den Spruch "die Spalten / Sind Bilder / Der Einheits / Vektoren" und wissen damit, dass die gegebene Matrix genau dann invertierbar ist, wenn ihre *Spalten* eine Basis von W bilden. Da es sich um n Spalten handelt und W n-dimensional ist, läuft das darauf hinaus, einen *Test auf lineare Unabhängigkeit der Matrixspalten* durchzuführen. Wie man von einer Anzahl Vektoren aber konkret überprüft, ob sie linear unabhängig sind, haben wir uns in einigen Übungsaufgaben jeweils *ad hoc* angeschaut; im nächsten Kapitel werden wir eine Variante für dieses Testverfahren kennenlernen, das den Vorteil hat, die inverse Matrix gleich mitzuliefern, *wenn* die Matrix denn invertierbar ist. Zusammenfassend und zum Merken:

> Eine quadratische Matrix ist genau dann invertierbar, wenn ihre Spalten linear unabhängig sind.

5.7 Permutationen.

And now for something completely different. Die sogenannten *Permutationen* sind eine eigene Klasse von Abbildungen extrem einfacher Art. Trotz ihrer Einfachheit kann man für sie eine verblüffende Menge wichtiger Struktureigenschaften ans Licht heben – was wir nicht tun wollen, denn mit der Linearen Algebra haben sie eigentlich nicht wirklich etwas zu tun. Warum wir hier überhaupt und wenigstens kurz über Permutationen sprechen wollen, hat seinen Grund darin, dass wir sie in der "Weiterführenden Übung" zu diesem Kapitel brauchen werden. Wir beschränken uns aber auf das Allernotwendigste.

Zunächst einmal ist eine Permutation nichts anderes als Bijektionen von $\{1, 2, \ldots, n\}$ nach $\{1, 2, \ldots, n\}$, also quasi Umordnungen dieser n Zahlen. Konkretes Beispiel: $p : \mathbb{N}_3 \to \mathbb{N}_3; 1 \mapsto 2, 2 \mapsto 3, 3 \mapsto 1$. Weil diese Dinger viel gebraucht werden[7], gibt es eine praktischere Schreibweise dafür: $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$. Und eigentlich ist die obere Reihe dieser Schreibweise auch noch obsolet, denn dort steht ja immer das gleiche, nämlich $(1, 2, 3)$. Daher geben wir eine Permutation $p : \mathbb{N}_n \to \mathbb{N}_n$ in der Form $(p(1), p(2), \ldots, p(n))$ an[8] und lesen die Permutation $(2, 3, 1)$ so: die 1 wird auf die 2 abgebildet, die 2 auf die 3 und die 3 auf die 1. Für gegebenes n gibt es, wie man durch Hinschauen und Abzählen herausfindet, nur endlich viele Permutationen, nämlich $n!$ viele ($n! := n \cdot (n-1) \cdot (n-2) \cdot \cdots \cdot 2 \cdot 1$, gesprochen "$n$ Fakultät"). Wie bei Funktionen üblich können wir zwei Permutationen hintereinander ausführen. So ist zum Beispiel $(1, 3, 2) \circ (2, 3, 1) = (3, 2, 1)$ [9]. Weil es sich ja um Bijektionen handelt, gibt es zu jeder Permutation auch eine Inverse, so dass die Hintereinanderausführung der beiden die Identische Abbildung $\mathrm{Id} = (1, 2, 3)$ ergibt. Übrigens: wer findet die Inverse zu $(2, 3, 1)$?[10].

[7] ganz prominent etwa im Differentialformenkalkül

[8] Aus dem Zusammenhang sollte jeweils klar sein, dass es sich dann *nicht* um ein Tupel handelt, sondern um die Angabe einer Abbildung.

[9] Leserichtung der Definition der Hintereinanderausführung von Funktionen beachten!

[10] Spoiler: es ist $(3, 1, 2)$

Die ausgebaute Gruppentheorie der Permutationen skippen wir hier einfach mal. Folgendes ist aber wichtig, und wir werden es brauchen: man unterscheidet *gerade* von *ungeraden* Permutationen. Was bedeuten diese Festlegungen? Zunächst wird jeder Permutation p ihre *Inversionszahl* $o(p)$ zugeordnet. Das ist diejenige Anzahl von Paaren $(p(i), p(j))$, bei denen zwar $i < j$, aber doch $p(i) > p(j)$ ist. Zum Beispiel hat unser $(2, 3, 1)$ die Inversionen $(2, 1)$ und $(3, 1)$. Man setzt dann $\operatorname{sign}(p) := (-1)^{o(p)}$. Wenn dieses Signum den Wert 1 hat, die Anzahl der Inversionen von p also gerade ist, heißt p *gerade Permutation*, andernfalls *ungerade Permutation*. Die dahinterstehende Idee ist, dass sich eine Permutation p stets als gerade bzw. ungerade Hintereinanderausführung von $o(p)$ *Transpositionen* darstellen lässt, das sind Permutationen, die nur genau zwei Ziffern vertauschen. Dazu abschließend ein Beispiel: nehmen wir $f = (5, 6, 1, 3, 4, 2)$. Zählen wir die Inversionen von f, kommen wir auf $4+4+1+1 = 10$, es ist also $\operatorname{sign}(f) := (-1)^{10} = 1$ und f somit gerade. Man macht sich leicht klar: besitzt eine beliebige Permutation p überhaupt keine Inversion der Form p_i, p_{i+1}, wo also zwei nebeneinanderstehende Funktionswerte in der "falschen" Reihenfolge sind, dann muss bereits $p = \mathrm{Id}$ sein. Unser f hat zwei solcher Inversionen, nämlich $(6, 1)$ und $(4, 2)$. Schnappen wir uns z.B. die erste und bilden die zugehörige Transposition $t_{(6,1)} = (6, 2, 3, 4, 5, 1)$, dann rechnen wir für $f' = t_{(6,1)} \circ f$ (Achtung, rechte Funktion zuerst anwenden!) die Werte $f' = (6, 2, 3, 4, 5, 1) \circ (5, 6, 1, 3, 4, 2) = (5, 1, 6, 3, 4, 2)$ aus. Der Witz ist nun: beim Auszählen der Inversionen von f' erhalten wir $4 + 3 + 1 + 1 = 9$, das heißt eine weniger als bei f. Wir könnten nun so fortfahren und wiederum aus f' eine Inversion benachbarter Werte heraussuchen und so weiter. Im Endeffekt hätten wir f zerlegt in eine Kette von Transpositionen, und zwar in eine Kette mit *geradzahlig* vielen Elementen, weil ja die Anzahl der Inversionen nach jeder Transposition um 1 sinkt und anfänglich $o(f) = 10$ war. Auf diesen elementaren, aber weit führenden Eigenschaften der Permutationen beruht ihre Bedeutung in der Algebra und ihren Anwendungen.

5.8 Basiswahlen und Basiswechsel.

Was machen wir übrigens, wenn wir lineare Abbildungen zwischen irgendwelchen endlichdimensionalen Vektorräumen und nicht zwischen $\mathbb{R}^n$ und $\mathbb{R}^m$ betrachten wollen? Kann man da auch irgendwie mit Matrizen rechnen? Sicher kann man das! wir müssen uns nur an die ursprüngliche Konstruktion einer Matrix erinnern: eine lineare Abbildung ist bereits dadurch definiert, dass wir sagen, wie sie auf die Basisvektoren des Ausgangsvektorraums wirkt, und die zugehörige Matrix ist nichts anderes als das Zahlenschema, mit dem wir zusammenfassen, wie sich die Bilder dieser

Basisvektoren unter der linearen Abbildung aus den Basisvektoren des Zielraums linearkombinieren. Mit anderen Worten: eine Matrix ist mindestens in verdeckter Form immer im Spiel, wenn für Definitions- und Zielmenge der betrachteten linearen Abbildung Basen gewählt sind. Dabei ist es wichtig, sich klarzumachen, dass die konkreten Einträge dieses Zahlenschemas explizit von allen beiden Basiswahlen abhängig sind.

Das bedeutet zweierlei. Einmal können wir, wenn wir es mit abtrakten endlichdimensionalen Vektorräumen zu tun haben, über konkrete Basiswahlen jeder linearen Abbildung eine ebenso konkrete Matrix zuweisen und damit dann rechnen[11]. Zum anderen sind wir in der Lage, beim Wechsel zwischen verschiedenen Basen der beteiligten Vektorräume die sich auf die ursprünglich gegebenen Basen beziehende Matrix auf die neuen Basen umzurechnen. Weil auch der Basiswechsel selber durch eine Matrix vermittelt wird, ist das nicht schwierig. In diesem Abschnitt gehen wir diese Basiswechsel einmal explizit durch.

Was eine Basis ist, wissen wir aus dem Kapitel "Basen und Dimensionen". In einer etwas anderen Sichtweise – die aber im Folgenden die hilfreichere ist – können wir eine gegebene Basis $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$ des n-dimensionalen Vektorraums V auch als lineare Abbildung lesen, nämlich als eben die, die jedem n-Tupel $(\lambda_1, \ldots, \lambda_n)$ die zugehörige Linearkombination der Basisvektoren $\mathbf{v} = \sum_i \lambda_i \mathbf{b}_i$ zuweist. In der Weiterführenden Übung zum Kapitel "Lineare Räume" haben wir diese lineare Abbildung jeweils benutzt, ohne sie allerdings mit einem expliziten Buchstaben zu benennen; wir haben dort einfach so etwas geschrieben wie $(-7, 2, 3)_{\mathbf{b}}^{\mathsf{T}}$. Und wir haben etwas unpräzise $(-7, 2, 3)_{\mathbf{b}}^{\mathsf{T}} \equiv (-7) \cdot \mathbf{b}_1 + 2 \cdot \mathbf{b}_2 + 3 \cdot \mathbf{b}_1$ identifiziert, was aber strenggenommen nicht so richtig passt, denn auf der linken Seite steht ein Zahlentripel, auf der rechten ein Vektor[12]. In diesem Abschnitt hier geben wir nun dieser Abbildung einen Namen und eine Bezeichnung: wir sagen jetzt $\varphi_{\mathbf{b}}(-7, 2, 3)^{\mathsf{T}} \equiv (-7) \cdot \mathbf{b}_1 + 2 \cdot \mathbf{b}_2 + 3 \cdot \mathbf{b}_3$ und nennen $\varphi_{\mathbf{b}} : \mathbb{R}^n \to \mathrm{V}$ den *kanonischen Basisisomorphismus* zur Basis $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$. Der kanonische Basisisomorphismus zu einer Basiswahl des Vektorraums V linearkombiniert in umkehrbar eindeutiger Weise aus jedem Zahlentupel, das ihm vor die Füße geworfen wird, den "zugehörigen" Vektor aus V mittels der

[11] Die Physiker haben dieses Vorgehen sogar für *nicht* endlichdimensionale Vektorräume praktiziert. Die sich dann ergebenden Matrizen haben dementsprechend auch nicht mehr nur endlich viele Zeilen und Spalten. Die Physiker, in mathematischen Dingen stets einigermaßen abgebrüht, haben mit dieser "HEISENBERGschen Matrizenmechanik" trotzdem umzugehen gelernt. Später hat sich gezeigt, dass diese unendlichdimensionale Matrizenmechanik äquivalent ist zu der *Schrödinger*schen Darstellung der Quantenmechanik.

[12] immerhin sei zu unseren Gunsten erwähnt, dass dieser Vektor aus $\mathbb{R}^3$ ist, also ebenfalls ein Zahlentripel; im vorliegenden Abschnitt wollen wir aber V immer als abstrakten Vektorraum sehen – was die Sache einfacher macht, weil klarer.

Basisvektoren. Wir sagen “zugehörig”, denn wenn die Basis $(\mathbf{b}_i)$ einmal gewählt ist[13], dann steht automatisch $\varphi_{\mathbf{b}}$ ohne weiteres Zutun vor uns[14]. Wenn man will, kann man also $\varphi_{\mathbf{b}}$ als eine Art “Labelling” der gewählten Basis lesen: $\varphi_{\mathbf{b}}(1,0,0)^{\mathsf{T}} = \mathbf{b}_1$, $\varphi_{\mathbf{b}}(0,1,0)^{\mathsf{T}} = \mathbf{b}_2$ und $\varphi_{\mathbf{b}}(0,0,1)^{\mathsf{T}} = \mathbf{b}_3$.

Dies bedeutet im Umkehrschluss natürlich: haben wir *zwei* verschiedene Basen $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$ und $(\mathbf{c}_1, \ldots, \mathbf{c}_n)$ gewählt, werden wir sofort mit *zwei* verschiedenen kanonischen Basisisomorphismen $\varphi_{\mathbf{b}}$ und $\varphi_{\mathbf{c}}$ versorgt – das ist in der Übung 8 zu diesem Kapitel einmal explizit durchgerechnet, wobei dort $\mathrm{V} = \mathbb{R}^3$ ist und eine der beiden Basen gerade die Standardnormalbasis $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$, so dass wir dort $\varphi_{\mathbf{e}}$ betrachten; und halten wir uns vor Augen, was $\varphi_{\mathbf{e}}$ eigentlich tut: jedem Tripel $(v_1, v_2, v_3)^{\mathsf{T}}$ wird das Tripel $\varphi_{\mathbf{e}}((v_1, v_2, v_3)^{\mathsf{T}}) = v_1 \cdot (1,0,0)^{\mathsf{T}} + v_2 \cdot (0,1,0)^{\mathsf{T}} + v_3 \cdot (0,0,1)^{\mathsf{T}} = (v_1, v_2, v_3)^{\mathsf{T}}$ zugewiesen: $\varphi_{\mathbf{e}}$ macht, wenn wir einen *Vektor* aus $\mathbb{R}^3$ mit seinem *Komponenten-Tripel* identifizieren, einfach gar nichts – es ist die Identität auf $\mathbb{R}^3$, eine mögliche Ursache für die eine oder andere Verwirrung. Deshalb ist es übersichtlicher, im Folgenden und insbesondere in der genannten Übung diese Identifizierung zu unterdrücken und stets explizit zu unterscheiden zwischen einem abstrakten Vektor $\mathbf{v} \in \mathrm{V} = \mathbb{R}^3$ und seinem Komponententripel $\varphi_{\mathbf{e}}^{-1}(\mathbf{v})$.

$$\begin{tikzcd} & \mathrm{V} & \\ \mathbb{R}^n \arrow[ur, "\varphi_{\mathbf{b}}"] \arrow[rr, "T_{\mathbf{cb}}"] & & \mathbb{R}^n \arrow[ul, "\varphi_{\mathbf{c}}"'] \end{tikzcd}$$

Das obige Diagramm verdeutlicht die Situation: einmal haben wir den abstrakten n-dimensionalen Vektorraum V, zum anderen zwei verschiedene kanonische Basisisomorphismen $\varphi_{\mathbf{e}}$ und $\varphi_{\mathbf{b}}$, die die ihnen vorgelegten Zahlentupel mittels zweier verschiedener Basen zu Vektoren aus V linearkombinieren. Da die Übersetzungen von Tupeln in Vektoren durch $\varphi_{\mathbf{b}}$ beziehungsweise $\varphi_{\mathbf{c}}$ bijektiv sind[15], das heißt $\varphi_{\mathbf{b}}^{-1}$ und $\varphi_{\mathbf{c}}^{-1}$ existieren, können wir die Hintereinanderausführung $\varphi_{\mathbf{c}}^{-1} \circ \varphi_{\mathbf{b}}$ zusammensetzen. Sie bildet – siehe Diagramm – offfenbar $\mathbb{R}^n$ wieder nach $\mathbb{R}^n$ ab. Und was tut $\varphi_{\mathbf{c}}^{-1} \circ \varphi_{\mathbf{b}} : \mathbb{R}^n \to \mathbb{R}^n$ explizit? Es nimmt ein Zahlentupel entgegen, linearkombiniert daraus einen Vektor $\mathbf{v} \in \mathrm{V}$ bezüglich der Basis $(\mathbf{b}_i)$ und bestimmt dann die Komponenten des so erhaltenen Vektors bezüglich der *anderen* Basis $(\mathbf{c}_i)$. Mit anderen Worten, diese Hintereinanderausführung bewerkstelligt einen *Basiswechsel* von $(\mathbf{b}_i)$ nach $(\mathbf{c}_i)$: aus einem Zahlentupel, das als Darstellung eines Vektors $\mathbf{v}$ in der Basis $(\mathbf{b}_i)$ zu verstehen ist, macht $\varphi_{\mathbf{c}}^{-1} \circ \varphi_{\mathbf{b}}$ das Zahlentupel für *denselben* Vektor $\mathbf{v}$ in der *anderen* Basis $(\mathbf{c}_i)$. Wir schreiben deshalb auch $T_{\mathbf{cb}} := \varphi_{\mathbf{c}}^{-1} \circ \varphi_{\mathbf{b}}$.

[13] *einschließlich* der Reihenfolge der ausgesuchten Basisvektoren!

[14] dieses “ohne weiteres Zutun” meinen die Mathematiker, wenn sie etwas “kanonisch” nennen – ein häufig verwendeter, aber eigentlich kein streng mathematisch definierter Begriff.

[15] ... wobei der Grund dafür natürlich gerade in den Basiseigenschaften zu suchen ist.

Als bijektive lineare Abbildung $T_{\mathbf{cb}} : \mathbb{R}^n \to \mathbb{R}^n$ wird $T_{\mathbf{cb}}$ wie üblich durch eine Matrix dargestellt; die Übung 8 zu diesem Kapitel ist nichts anderes als die Aufgabe, die Matrix $(T_{\mathbf{cb}ij})$ für zwei konkret gegebene Basen des $\mathbb{R}^3$ einmal explizit auszurechnen. Wenn man dies allgemein hinschreibt, liest sich der Basiswechsel ziemlich kompliziert, obwohl er im Grunde einfach ist.

Führen wir einen Basiswechsel dennoch einmal allgemein und explizit hier durch. Der abstrakte Vektor $\mathbf{v}$ habe einmal die Komponenten $(\mathbf{v}_1^{\mathbf{b}}, \ldots, \mathbf{v}_n^{\mathbf{b}})$ in der Basis $(\mathbf{b}_i)$ sowie $(\mathbf{v}_1^{\mathbf{c}}, \ldots, \mathbf{v}_n^{\mathbf{c}})$ in der Basis $(\mathbf{c}_i)$. Als abkürzende Schreibweise setzen wir $\mathbf{e}_i = (0, \ldots, \overset{i}{1}, \ldots, 0)^{\mathsf{T}}$, will heißen, $\mathbf{e}_i$ ist das *Zahlentupel* mit Nullen und der 1 an i-ter Stelle (aus $\mathbb{R}^n$; *und nicht etwa ein abstrakter Vektor aus* V). Es ist eben nur umständlich, immer $(0, 1, 0, \ldots, 0)^{\mathsf{T}}$ und so weiter zu schreiben. Mit dieser Schreibweise haben wir abkürzend $(\mathbf{v}_1^{\mathbf{b}}, \ldots, \mathbf{v}_n^{\mathbf{b}}) = \sum_j \mathbf{v}_j^{\mathbf{b}} \mathbf{e}_j =$ und $(\mathbf{v}_1^{\mathbf{c}}, \ldots, \mathbf{v}_n^{\mathbf{c}}) = \sum_j \mathbf{v}_j^{\mathbf{c}} \mathbf{e}_j$. Das sind natürlich *keine* Basisentwicklungen! Die beiden Tupeldarstellungen sind im Allgemeinen *nicht* gleich. Vielmehr suchen wir ja gerade nach der aus $T_{\mathbf{cb}}$ gebildeten Matrix, die die eine Darstellung in die andere überführt.

Nun ist ja $\sum_j \mathbf{v}_j^{\mathbf{c}} \mathbf{e}_j = \varphi_{\mathbf{c}}^{-1}(\mathbf{v})$ (denn definitionsgemäß ist $\varphi_{\mathbf{c}}(\mathbf{e}_j) = \mathbf{c}_j$, und nun wende man $\varphi_{\mathbf{c}}$ auf beide Seiten an!), so dass wir also schreiben können $\sum_j \mathbf{v}_j^{\mathbf{c}} \mathbf{e}_j = \varphi_{\mathbf{c}}^{-1}(\mathbf{v}) = (\varphi_{\mathbf{c}}^{-1} \circ \mathrm{Id})(\mathbf{v}) = (\varphi_{\mathbf{c}}^{-1} \circ (\varphi_{\mathbf{b}} \circ \varphi_{\mathbf{b}}^{-1}))(\mathbf{v}) = (\varphi_{\mathbf{c}}^{-1} \circ \varphi_{\mathbf{b}}) \circ \varphi_{\mathbf{b}}^{-1}(\mathbf{v})$ (genau *hier* findet jetzt der Basiswechsel statt!) $= T_{\mathbf{cb}} \circ \varphi_{\mathbf{b}}^{-1}(\mathbf{v}) = T_{\mathbf{cb}}(\sum_i \mathbf{v}_i^{\mathbf{b}} \mathbf{e}_i) = \sum_i \mathbf{v}_i^{\mathbf{b}} T_{\mathbf{cb}} \mathbf{e}_i$, und nach unserer allgemeinen Matrixdefinition ("Die Spalten / Sind Bilder / ...") auf Seite **??** ist dies weiter ... $= \sum_i \sum_j \mathbf{v}_i^{\mathbf{b}} (T_{\mathbf{cb}})_{ji} \mathbf{e}_j$. Vergleichen wir ganz links mit ganz rechts und benennen wir der Lesbarkeit halber $j \leftrightarrow i$ um: $\mathbf{v}_i^{\mathbf{c}} = \sum_j (T_{\mathbf{cb}})_{ij} \mathbf{v}_j^{\mathbf{b}}$. Anders formuliert: die Matrix, die wir auf die $\mathbf{b}$-Komponenten eines abstrakten Vektors $\mathbf{v}$ anwenden müssen, um seine $\mathbf{c}$-Komponenten zu erhalten, ist $(T_{\mathbf{cb}})_{ij}$, deren i-te Spalte $T_{\mathbf{cb}} \mathbf{e}_i$ wegen $T_{\mathbf{cb}} \mathbf{e}_i = (\varphi_{\mathbf{c}}^{-1} \circ \varphi_{\mathbf{b}})(\mathbf{e}_i) = \varphi_{\mathbf{c}}^{-1}(\mathbf{b}_i)$ ist – gerade dasjenige n-Zahlentupel, das der $\mathbf{c}$-Basisisomorphismus in den Basisvektor $\mathbf{b}_i$ abbildet. Wenn wir also, zum Beispiel so etwa wie in der Übung 8, einen Basiswechsel durchzuführen hätten, in dem $\mathbf{b}_2 = -\mathbf{c}_1 + 2 \cdot \mathbf{c}_2$ wäre, dann ergäbe sich für die 2. Spalte unserer Basiswechselmatrix $\varphi_{\mathbf{c}}^{-1}(\mathbf{b}_2) = \varphi_{\mathbf{c}}^{-1}(-\mathbf{c}_1 + 2\mathbf{c}_2) = \varphi_{\mathbf{c}}^{-1}(-\mathbf{c}_1) + 2 \cdot \varphi_{\mathbf{c}}^{-1}(\mathbf{c}_2) = -\mathbf{e}_1 + 2\mathbf{e}_2$, bei drei Dimensionen mithin $(-1, 2, 0)^{\mathsf{T}}$. Damit diese Prozedur als so einfach erkannt wird, wie sie im Grunde ist, empfiehlt es sich – mittels Übung 8 und gegebenenfalls darüber hinaus mit selbsterfundenen Basiswechseln –, die eine oder andere konkrete Realisierung durchzurechnen.

Wir beschließen diesen Abschnitt und das ganze Kapitel mit der – ebenfalls in der genannten Übung exemplifizierten – Operation, eine *Matrix* von einer Basis in die andere umzurechnen. Das folgende Diagramm stellt die Situation dar:

$$
\begin{tikzcd}
\mathbb{R}^n \arrow[rrr, "L^{\mathbf{pb}}"] \arrow[dr, "\varphi_{\mathbf{b}}"] \arrow[dd, "T_{\mathbf{cb}}=\varphi_{\mathbf{c}}^{-1}\circ\varphi_{\mathbf{b}}"'] & & & \mathbb{R}^m \arrow[dl, "\varphi_{\mathbf{p}}"'] \arrow[dd, "\varphi_{\mathbf{q}}^{-1}\circ\varphi_{\mathbf{p}}=T_{\mathbf{qp}}"] \\
& \mathbf{V} \arrow[r, "\boldsymbol{L}"] & \mathbf{W} & \\
\mathbb{R}^n \arrow[rrr, "L^{\mathbf{qc}}"'] \arrow[ur, "\varphi_{\mathbf{c}}"'] & & & \mathbb{R}^m \arrow[ul, "\varphi_{\mathbf{q}}"]
\end{tikzcd}
$$

Wir finden das obige “Dreieck des Basiswechsels” zweimal darin, denn die lineare Abbildung L, um die es geht, bildet einen abstrakten n-dimensionalen Vektorraum V in einen abstrakten m-dimensionalen Vektorraum W ab, und für *beide* Vektorräume wollen wir je zwei alternative Basen anbieten. Für V seien es die beiden Basiswahlen $\varphi_{\mathbf{b}} : \mathbb{R}^n \to \mathrm{V}$ und $\varphi_{\mathbf{c}} : \mathbb{R}^n \to \mathrm{V}$, für W die beiden Basiswahlen $\varphi_{\mathbf{p}} : \mathbb{R}^m \to \mathrm{W}$ und $\varphi_{\mathbf{q}} : \mathbb{R}^m \to \mathrm{W}$. Dementsprechend sind wir mit den beiden Basiswechselmatrizen $T_{\mathbf{cb}} : \mathbb{R}^n \to \mathbb{R}^n$ und $T_{\mathbf{qp}} : \mathbb{R}^m \to \mathbb{R}^m$ versorgt, die in Form des von oben bekannten Dreiecks eingezeichnet sind. Zwischen $\mathbb{R}^n$ und $\mathbb{R}^m$ bilden die beiden Matrizen $L^{\mathbf{pb}}$ und $L^{\mathbf{qc}}$ ab, *beides* Matizen zur linearen Abbildung $L : \mathrm{V} \to \mathrm{W}$, die eine aber bezüglich der Basiswahlen $\varphi_{\mathbf{b}}$ im Definitions-Vektorraum und $\varphi_{\mathbf{p}}$ im Ziel-Vektorraum, die andere bezüglich $\varphi_{\mathbf{c}}$ im Definitions- und $\varphi_{\mathbf{q}}$ im Ziel-Vektorraum. Die für all dies verantwortliche lineare Abildung L, sämtliche Basiswahlen, Basiswechsel und Basen-bezogene Matrixdarstellungen von L sind in diesem einen Diagramm zusammengestellt.

Gerade mit diesem Diagramm aber ist die Umrechnungsvorschrift von $L^{\mathbf{qc}}$ nach $L^{\mathbf{pb}}$ quasi durch Nachfahren mit dem Finger herleitbar, wenn man nur die Reihenfolgen beachtet und daran denkt, dass $T_{\mathbf{qp}}^{-1} = (\varphi_{\mathbf{q}}^{-1} \circ \varphi_{\mathbf{p}})^{-1} = \varphi_{\mathbf{p}}^{-1} \circ (\varphi_{\mathbf{q}}{}^{-1})^{-1} = \varphi_{\mathbf{p}}^{-1} \circ \varphi_{\mathbf{q}} = T_{\mathbf{pq}}$ ist (die Basiswechsel sind Isomorphismen); es ist nämlich $L^{\mathbf{pb}} = T_{\mathbf{qp}}{}^{-1} \circ L^{\mathbf{qc}} \circ T_{\mathbf{cb}}$ oder

$$L^{\mathbf{pb}} = T_{\mathbf{pq}} \circ L^{\mathbf{qc}} \circ T_{\mathbf{cb}} \ .$$

Letztlich war’s das auch schon. In den Übungen zu diesem Kapitel findet sich eine praktische Anwendung für eine konkret gegebene Matrix – *Matrix*, nicht abstrakt gegebene lineare Abbildung, denn das ist in der Praxis der am häufigsten auftretende Fall: in irgendwelchen Koordinaten, meistens in der Standardnormalbasis eines $\mathbb{R}^n$, ist eine quadratische Matrix vorgegeben, es ist also in unserem Diagramm $\mathbb{R}^n = \mathbb{R}^m$ und $\varphi_{\mathbf{c}} = \varphi_{\mathbf{q}}$ die Standardnormalbasis, und $L^{\mathbf{qc}} \equiv L^{\mathbf{e}}$ ist explizit gegeben. Nun ist die Frage, wie die Matrix $L^{\mathbf{b}}$ zur selben linearen Abbildung aussieht, bezogen auf eine andere Basis $\varphi_{\mathbf{b}} = \varphi_{\mathbf{p}}$. Häufig ist es so, dass diese andere Basis gerade so gewählt werden soll, dass die in dieser Basis formulierte Matrix besonders einfach aussieht, zum Beispiel Diagonalgestalt hat. Wir werden diesen Fall im Kapitel über Hauptachsentransformationen näher beleuchten. Manchmal hingegen wird man dazu gedrängt, eine neue Basis aus einer alten zu konstruieren, wobei die neue Basis einem gegebenen Problem angepasst ist, zum Beispiel einem Unterraum des Phasenraums eines physikalischen Systems, den dieses einiger Erhaltungssätze wegen nicht verlassen kann. In diesem Fall geht es dann darum, eine in der alten Basis gegebene Matrix in die neue, angepasste Basis umzurechnen.