Kapitel S

Tensorprodukt

Was ist das Tensorprodukt? Da stellen wir uns ganz dumm.
Jeder weiß, dass man Vektoren untereinander zwar addieren kann,
aber allgemein nicht multiplizieren. Wir tun es einfach trotzdem!
Das universelle Ergebnis ist das Tensorprodukt.

frei nach Prof. Günter Harder

Inhalt dieses Kapitels S

Tensorprodukte: grundlegend & allgemein = schwierig?

Wir sind auf der Zielgeraden dieser Vorlesung zur Linearen Algebra. Zum krönenden Abschluss behandeln wir das Tensorprodukt, zunächst über einem Körper (§S1), das vereinfacht, dann über beliebigen Ringen (§S2). Die Grundidee ist im Eingangszitat treffend zusammengefasst:

Wir multiplizieren Vektoren, naiv-mutig und so allgemein wie möglich. (Ich zitiere hier frei aus meinem Gedächtnis, wie Prof. Günter Harder in seinen Bonner Algebra-Vorlesungen das Tensorprodukt resümierte.) Etwas präziser gesagt, das Tensorprodukt realisiert folgendes Ziel:

Wir betrachten bilineare oder allgemein multilineare Abbildungen; mit dem Tensorprodukt können wir diese „universell linearisieren".

Wie alle neuen Werkzeuge kostet das Tensorprodukt anfangs Mühe, doch wie immer lohnt sich auch hier die langfristige Investition. Um auch kurzfristig schon erste schöne Anwendungen zu zeigen, beginne ich über einem Körper mit dem „Tensorprodukt für Eilige."

Zudem wage ich den Blick über den Tellerrand und diskutiere Anwendungen in der Quantenmechanik. Das zeigt zweierlei:

1. Die mathematischen Grundlagen sind vergleichsweise leicht.
2. Die Anwendungen sind erstaunlich vielseitig und komplex.

Schon die ersten Anwendungen sind spektakulär. Wer diese dennoch nicht wünscht oder braucht, kann die physikalische Einkleidung auch ignorieren und sich auf den mathematischen Kern konzentrieren. Abfragen werde ich es nicht, anbieten will ich es auf jeden Fall.

Erfahrungsgemäß ist es nämlich hilfreich, zu mathematischen Begriffen auch geometrische Anschauung und physikalische Intuition zu schulen. Dies dient der naturwissenschaftlich-technischen Allgemeinbildung und nützt auch direkt dem mathematischen Verständnis beim Lernen.

Mathematik (gr. μαθηματική τέχνη) ist die ‘Kunst des Erkennens’.

Woher kommen Tensoren?

Wie so viele fundamentale Begriffe wurden **Tensoren** ursprünglich in der Physik eingeführt und erst später mathematisch präzisiert. Das ist kein Makel, sondern Beleg ihrer tiefliegenden Bedeutung.

In der theoretischen Physik werden „Tensoren" seit langem verwendet, da sie zur effizienten Modellbildung nützen und in Rechnungen helfen. Dabei werden diese „Tensoren" zunächst ganz pragmatisch durch ihre **Rechenregeln** eingeführt. . . der Erfolg heiligt bekanntlich die Mittel.

Diesen Zugang „für Eilige" möchte ich unten als erstes vorstellen. Das entspricht recht genau der ideengeschichtlichen Entwicklung, und auch heute noch antwortet es aus didaktisch-physikalischer Sicht sehr direkt und auch motivierend auf die Frage: Was sind Tensoren?

Wahre Begebenheit: Der Einstieg dieses Kapitels entstand auf die sorgenvolle Frage meines Sohnes, was er mit Tensoren anfangen soll. . . Wir wollen und können Tensoren anpacken und sofort damit rechnen!

Wo werden Tensoren genutzt?

Aus **mathematischer Sicht** dauerte es hingegen etwas länger, bis es Mathematiker/innen schließlich gelang, für diese „Tensoren“ eine befriedigende theoretische Grundlage zu finden und zu erklären.

Seither werden sie auch in der Mathematik recht vielseitig genutzt und angewendet, insbesondere in der Algebra und der Differentialgeometrie, zudem in den Ingenieurwissenschaften (etwa der Mechanik, Elastizität, Strömungslehre, Elektrodynamik, ...) und der theoretischen Physik, von der Quantenmechanik bis zur Relativitätstheorie.

Dabei werden Tensoren meist zu Tensorfeldern verallgemeinert, die abkürzend-schludrig meist ebenfalls Tensoren genannt werden.

Ein Tensorfeld ordnet jedem Punkt des Raums einen Tensor zu. Viele naturwissenschaftliche Modelle und physikalische Theorien nutzen Tensorfelder, allen voran die allgemeine Relativitätstheorie, die damit die gekrümmte Geometrie der Raumzeit beschreibt. Das zugehörige mathematische Teilgebiet heißt Tensoranalysis.

Tensoren als Größen mit Indizes

Tensoren sind ein **Hilfsmittel** zur Beschreibung eines dahinterliegenden Objekts, das uns eigentlich interessiert, etwa eine lineare Abbildung oder eine Bilinearform uvm. Das Ergebnis unserer Rechnungen soll dabei invariant sein, also dasselbe in jedem Koordinatensystem.

0. Stufe: Wir beginnen viele unserer Rechnungen mit einem Skalar a: Dies ist einfach ein Element des Grundkörpers K, etwa $\mathbb{R}$ oder $\mathbb{C}$.

1. Stufe: Eine Folge $b = (b_1, \dots, b_k)$ solcher Skalare fassen wir zu einem neuen Objekt zusammen und rechnen damit als vektorielle Größe. Solch ein Vektor $b = (b_i)_i$ hat dabei einen einzigen Index i.

2. Stufe: Eine Folge von Vektoren $c = (c_{ij})_{ij}$ hat demnach zwei Indizes. Warum sagen wir nicht einfach „Matrix“? Tensoren sind allgemeiner!

3. Stufe: Ein Tensor $d = (d_{ijk})_{ijk}$ kann drei (oder mehr) Indizes haben, wie etwa das Levi-Civita-Symbol $\varepsilon_{ijk} = \operatorname{sign}(i, j, k)$ für $i, j, k \in \{1, 2, 3\}$. Die anfängliche Analogie mit Matrizen bricht daher schnell zusammen. Wir untersuchen Tensoren als eigene, universelle Konstruktion.

Wie erklären wir Tensoren mathematisch?

Tensoren sind ein vielseitiges und nützliches Werkzeug der Mathematik. Es lohnt sich daher, Zeit in diesen grundlegenden Begriff zu investieren, gerade zum Übergang von „koordinatengebunden“ zu „koordinatenfrei“. Das ist sehr präzise, doch abstrakt, und daher nicht allgemein beliebt.

Zur Motivation beginne ich mit meiner kurzen Vorschau „für Eilige“ und einer ersten spektakulären Anwendung in der Quantenmechanik: dem sogenannten No-Cloning-Theorem. Soviel Physik muss sein!

Zum Einstieg betone ich die **universelle Basiseigenschaft** S1B, die den Rechnungen zugrundeliegt und sofortigen Zugang ermöglicht.

Anschließend beschäftigen wir uns sorgsam und geduldig mit der mathematischen Grundlegung: Der richtige Ausgangspunkt ist hier die **universelle Abbildungseigenschaft** (S2E). Anschließend erarbeiten wir ihre segensreichen Konsequenzen wie die Basiseigenschaft S2G.

Dabei gehe ich den langen Weg und beginne so allgemein wie möglich: Wir legen das Thema breit an und spezialisieren dann schrittweise.

Erinnerung: das kartesische Produkt als direkte Summe

Satz S1A: das kartesische Produkt als direkte Summe

Gegeben seien zwei Vektorräume U und V über einem Körper K.

(1) Zu U, V haben wir das kartesische Produkt als **direkte Summe**:

$$\begin{aligned} &U \times V := \{\, (u, v) \mid u \in U,\ v \in V \,\} = U' \oplus V' \\ &\text{mit} \quad (i_1, p_1) : U \cong U' := U \times \{0\} : u \leftrightarrows (u, 0) \\ &\text{und} \quad (i_2, p_2) : V \cong V' := \{0\} \times V : v \leftrightarrows (0, v) \end{aligned}$$

(2) Zu je zwei gegebenen Basen $(u_i)_{i \in I}$ von U und $(v_j)_{j \in J}$ von V ist ihre Aneinanderhängung $(w_k)_{k \in I \sqcup J}$ eine Basis des Raums $U \times V$, wobei $I \cap J = \emptyset$ und $w_i = (u_i, 0)$ für $i \in I$ und $w_j = (0, v_j)$ für $j \in J$.

(3) Insbesondere addieren sich die Dimensionen:

$$\dim(U' \oplus V') = \dim(U') + \dim(V')$$

🙂 Damit können wir konkret rechnen, denn alles liegt explizit vor!

Erinnerung: das kartesische Produkt als direkte Summe

Dieser Satz dient als freundliche Erinnerung und Zusammenfassung von zwei universell nützlichen Konstruktionen: Produkt vs Summe.

Aus mengentheoretischer Sicht konstruieren wir den Raum $U \times V$ als kartesisches Produkt mit koordinatenweiser Addition und Skalierung. Aus Sicht der linearen Räume betrachten wir dies als direkte Summe. genauer $U \times V = U' \oplus V'$ mit den Kopien U' und V' von U und V.

Allgemein unterscheiden wir das Produkt $\prod_{i \in I} V_i$
und die (externe / interne) direkte Summe $\bigoplus_{i \in I} V_i$.
Allgemein gilt hierbei die Inklusion $\bigoplus_{i \in I} V_i \le \prod_{i \in I} V_i$,
und Gleichheit gilt für jede endliche Indexmenge I.

Die Konstruktion einer Basis durch Aneinanderhängen zeigt:
Die Dimension der Summe ist die Summe der Dimensionen.
Mit dieser Konstruktion können wir Vektorräume addieren.
Als nächstes wollen wir Vektorräume auch multiplizieren.

Vorschau: das Tensorprodukt für Eilige

Definition S1B: das Tensorprodukt und seine Basiseigenschaft

Gegeben seien drei Vektorräume U, V und W über einem Körper K.

(1) Ein **Produkt** von U und V nach W ist eine K–bilineare Abbildung

$$\otimes \;:\; U \times V \to W \;:\; (u, v) \mapsto w = u \otimes v.$$

Das heißt ausgeschrieben für alle $u, u' \in U$ und $v, v' \in V$ und $\lambda \in K$:

$$\begin{aligned} (u + u') \otimes v &= (u \otimes v) + (u' \otimes v), & \qquad (\lambda u) \otimes v &= \lambda(u \otimes v), \\ u \otimes (v + v') &= (u \otimes v) + (u \otimes v'), & \qquad u \otimes (v\lambda) &= (u \otimes v)\lambda. \end{aligned}$$

(2) Wir nennen $\otimes$ ein **Tensorprodukt** von U und V in den **Tensorraum** $W = U \otimes V$, falls zudem folgende **universelle Basiseigenschaft** gilt:

Zu je zwei gegebenen Basen $(u_i)_{i\in I}$ von U und $(v_j)_{j\in J}$ von V ist die Produktfamilie $(u_i \otimes v_j)_{(i,j)\in I\times J}$ eine Basis des Zielraums W.

(3) Insbesondere multiplizieren sich die Dimensionen:

$$\dim(U \otimes V) = \dim(U) \cdot \dim(V)$$

☺ Damit können wir konkret rechnen, denn wir haben eine Basis!

Diese Definition über die universelle Basiseigenschaft ist die schnellste, doch nicht unbedingt die beste. Wir optimieren und verallgemeinern sie später zur universellen Abbildungseigenschaft (Definition S2E).

Über einem Körper sind beide Eigenschaften äquivalent (Satz S2G), daher beginne ich hier mit dem einfacheren Spezialfall.

☺ Satz S1A(1) konstruiert explizit den Vektorraum $U \times V = U' \oplus V'$. Daraus ziehen wir Folgerungen für (2) Basen und (3) Dimensionen.

Definition S1B erklärt die geforderten Eigenschaften (1) und (2). Wir beweisen später die Existenz (etwa durch Matrizen S1C oder Polynome S1D, oder ganz allgemein basisfrei als Quotienten S2I) und die Eindeutigkeit (bis auf eindeutige Isomorphie, Satz S2H).

☺ Definition S1B enthält vorab schon genug Information, um sofort damit rechnen zu können, daher der Slogan „Tensorprodukt für Eilige“.

Wie rechnen wir mit Tensoren?

Aufgabe: (1) Nennen Sie eine Basis des Tensorraums $\mathbb{R}^2 \otimes \mathbb{R}^2$.
(2) Sind darin die Tensoren $(0,1) \otimes (1,0)$ und $(1,0) \otimes (0,1)$ gleich?

Lösung: (1) Wir wählen die Standardbasis (e_1, e_2) für $U = V = \mathbb{R}^2$.
Für den Tensorraum $W = U \otimes V$ erhalten wir daraus die Produktbasis

$$\big(\, e_1 \otimes e_1,\; e_1 \otimes e_2,\; e_2 \otimes e_1,\; e_2 \otimes e_2 \,\big).$$

(2) Damit können wir die Tensoren $e_2 \otimes e_1$ und $e_1 \otimes e_2$ vergleichen:

$$\begin{aligned}
e_2 \otimes e_1 &= 0 \cdot (e_1 \otimes e_1) + 0 \cdot (e_1 \otimes e_2) + 1 \cdot (e_2 \otimes e_1) + 0 \cdot (e_2 \otimes e_2) \\
e_1 \otimes e_2 &= 0 \cdot (e_1 \otimes e_1) + 1 \cdot (e_1 \otimes e_2) + 0 \cdot (e_2 \otimes e_1) + 0 \cdot (e_2 \otimes e_2)
\end{aligned}$$

Die Tensoren $e_2 \otimes e_1$ und $e_1 \otimes e_2$ in $W = U \otimes V$ sind verschieden!

In diesem schönen Beispiel stört nur der numerische Zufall $2 \cdot 2 = 2 + 2$.
Erinnerung: Beim Tensorprodukt multiplizieren sich die Dimensionen.
Das Tensorprodukt $\mathbb{R}^3 \otimes \mathbb{R}^5$ etwa hat die Dimension $3 \cdot 5 = 15$,
die Summe $\mathbb{R}^3 \oplus \mathbb{R}^5$ hingegen hat die Dimension $3 + 5 = 8$.

Wie rechnen wir mit Tensoren?

☺ Definition S1B begründet das folgende Vergleichsverfahren:

Algo S1B: Vergleich von zwei Tensoren in $U \otimes V$

Eingabe: zwei Elemente w_1, w_2 des Tensorraums $U \otimes V$
Ausgabe: „gleich“ falls $w_1 = w_2$ oder „ungleich“ falls $w_1 \neq w_2$

1: Schreibe w_1 und w_2 in der gewählten Produktbasis von $U \otimes V$
2: Vergleiche diese Linearkombinationen koeffizientenweise

Mit diesem einfachen Trick haben wir die vorige Aufgabe gelöst, und genau so gelingt es für jedes Tensorprodukt über einem Körper.

Dies betont die Bedeutung der universellen Basiseigenschaft S1B: So können wir alle Elemente eindeutig darstellen und vergleichen.

Zum effektiven Rechnen gehört wie immer als allererster Schritt, je zwei Elemente auf Un/Gleichheit prüfen zu können.

Definition S1B: einfache Tensoren

(4) Für Tensorprodukte vereinbaren wir folgende Sprechweisen.
Jedes Element $w \in U \otimes V$ des Tensorraums nennen wir einen **Tensor**.

Hat $w \in U \otimes V$ die besonders einfache Form $w = u \otimes v$ mit $u \in U$ und $v \in V$, so nennen wir den Tensor w **einfach** oder **elementar** oder **rein**.

Anders gesagt, die Menge der **einfachen Tensoren** ist das Bild der Produktabbildung $\otimes : U \times V \to U \otimes V$. Diese ist i.A. nicht surjektiv!

Der Tensorraum entsteht aus allen **Summen** einfacher Tensoren; solche Summen sind im Allgemeinen keine einfachen Tensoren.

Der Tensorraum erlaubt immer eine **Basis** aus einfachen Tensoren, zum Beispiel jede Produktbasis. Es gibt viele weitere Möglichkeiten.

Übung: Finden Sie in $\mathbb{R}^2 \otimes \mathbb{R}^2$ eine Basis aus vier Tensoren, von denen *keiner* einfach ist. (Die nächste Aufgabe hilft dabei.)

Wie rechnen wir mit Tensoren?

Aufgabe: Im Tensorraum $W = \mathbb{R}^2 \otimes \mathbb{R}^2$ betrachten wir den allgemeinen Tensor $w = a \cdot (e_1 \otimes e_1) + b \cdot (e_1 \otimes e_2) + c \cdot (e_2 \otimes e_1) + d \cdot (e_2 \otimes e_2)$. Genau dann ist w einfach, also $w = u \otimes v$, wenn $ad - bc = 0$ gilt.

Lösung: (1) Die Implikation „$\Rightarrow$" ist klar, denn wir finden

$$\begin{aligned} w = u \otimes v &= (u_1 e_1 + u_2 e_2) \otimes (v_1 e_1 + v_2 e_2) \\ &= u_1 v_1 (e_1 \otimes e_1) + u_1 v_2 (e_1 \otimes e_2) + u_2 v_1 (e_2 \otimes e_1) + u_2 v_2 (e_2 \otimes e_2). \end{aligned}$$

Für diesen Tensor w gilt somit $ad - bc = u_1 v_1 \cdot u_2 v_2 - u_1 v_2 \cdot u_2 v_1 = 0$.

(2) Die Umkehrung „$\Leftarrow$" ist interessanter. Der Fall $w = 0$ ist dabei klar. Sei also $w \neq 0$, etwa $a \neq 0$. Für $u = a e_1 + c e_2$ und $v = 1 e_1 + (b/a) e_2$ gilt dann $u \otimes v = w$. Die Fälle $b \neq 0$, $c \neq 0$, $d \neq 0$ sind analog.

Folgerung: Gilt $\dim U \geq 2$ und $\dim V \geq 2$, so ist die Produktabbildung $\otimes : U \times V \to U \otimes V$ nicht surjektiv, d.h. nicht jeder Tensor ist einfach.

Beispiel S1c: Matrixmodell des Tensorprodukts

Wir betrachten $U = K^{p\times 1}$ und $V = K^{1\times q}$ sowie $W = K^{p\times q}$ mit

$$\cdot \;:\; K^{p\times 1} \times K^{1\times q} \to K^{p\times q} \;:\; (u,v) \mapsto u \cdot v.$$

Dieses Matrixprodukt bedeutet ausgeschrieben für $(p,q) = (3,4)$:

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \cdot \begin{bmatrix} v_1 & v_2 & v_3 & v_4 \end{bmatrix} = \begin{bmatrix} u_1v_1 & u_1v_2 & u_1v_3 & u_1v_4 \\ u_2v_1 & u_2v_2 & u_2v_3 & u_2v_4 \\ u_3v_1 & u_3v_2 & u_3v_3 & u_3v_4 \end{bmatrix}$$

Es erfüllt die geforderten Eigenschaften eines Tensorprodukts:

1. Die Produktabbildung $\cdot$ ist bilinear über dem Grundkörper K.
2. Aus den Basen $(e_1,\dots,e_p)$ von $K^{p\times 1}$ und $(e_1^\intercal,\dots,e_q^\intercal)$ von $K^{1\times q}$ erhalten wir die Produktbasis $(e_{i,j} = e_i \cdot e_j^\intercal)_{i=1,\dots,p}^{j=1,\dots,q}$ von $K^{p\times q}$.

So unterscheiden wir $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right] \cdot \left[\begin{smallmatrix}1 & 0\end{smallmatrix}\right] = \left[\begin{smallmatrix}0 & 0\\1 & 0\end{smallmatrix}\right]$ und $\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right] \cdot \left[\begin{smallmatrix}0 & 1\end{smallmatrix}\right] = \left[\begin{smallmatrix}0 & 1\\0 & 0\end{smallmatrix}\right]$ (S105).

Notation: Für dieses **dyadische Produkt** nutzt man üblicherweise die bequeme Schreibweise $\otimes : K^p \times K^q \to K^{p\times q} : (u_i)_i \otimes (v_j)_j = (u_i v_j)_{i,j}$.
Dies verallgemeinern wir später zum Kronecker Produkt (S2P);
hier sehen wir den Spezialfall $\otimes : K^{p\times 1} \times K^{1\times q} \to K^{p\times q}$.

Aufgabe: Einfache Tensoren sind genau die Matrizen vom Rang ≤ 1.

Lösung: „$\Rightarrow$“: Für $u \in K^{p\times 1}$ und $v \in K^{1\times q}$ hat $w = uv$ den Rang ≤ 1, denn alle Spalten der Matrix w sind Vielfache des Spaltenvektors u.

Ebenso: Alle Zeilen von w sind Vielfache des Zeilenvektors v.
Wir erinnern uns: Es gilt Zeilenrang gleich Spaltenrang. (K2J)

„$\Leftarrow$“: Hat $w \in K^{p\times q}$ den Rang 1, so enthält w einen Spaltenvektor $u \in K^{p\times 1} \smallsetminus \{0\}$, und alle anderen Spalten sind Vielfache von u.

Also existiert ein Zeilenvektor $v \in K^{1\times q} \smallsetminus \{0\}$, sodass $w = uv$ gilt.
Hat w den Rang 0, so gilt $w = 0$, also $w = uv$ mit $u = 0$ oder $v = 0$.

Matrizen als vertrautes Beispiel

Aufgabe: Hat die Matrixmultiplikation $\cdot : K^{p\times r} \times K^{r\times q} \to K^{p\times q}$ allgemein für $r \in \mathbb{N}$ die Eigenschaften eines Tensorprodukts?

Lösung: Nein. Das Produkt $\cdot : K^{p\times r} \times K^{r\times q} \to K^{p\times q}$ ist zwar bilinear, aber für $r \ge 2$ fehlt ihm die geforderte universelle Basiseigenschaft:

Zu Basen $(u_i)_{i\in I}$ von $U = K^{p\times r}$ und $(v_j)_{j\in J}$ von $V = K^{r\times q}$ ist die Produktfamilie $(u_i \otimes v_j)_{(i,j)\in I\times J}$ keine Basis des Raums $W = K^{p\times q}$.

Wir sehen dies bereits an den Dimensionen, denn die erwartete Multiplikativität $pr \cdot rq \stackrel{!}{=} pq$ ist wegen $p, q \ge 1$ nur für $r = 1$ möglich.

🙂 Für $r = 1$ erhalten wir gerade das obige Beispiel S1C. Für $r \ge 2$ erhalten wir Produkte, die nicht die universelle Basiseigenschaft haben. Für unsere Repertoire an Gegen/Beispielen ist das ebenso nützlich. Das Kronecker–Produkt S2P erweist sich hier als Tensorprodukt.

Matrizen als vertrautes Beispiel

Aufgabe: Zeigen Sie, dass für je zwei endlich-dimensionale Räume U und V über K ein Tensorprodukt $\otimes : U \times V \to U \otimes V$ existiert.

Lösung: Nach Wahl von Basen haben wir $U \cong K^p$ und $V \cong K^q$. Für die Modellräume K^p und K^q haben wir das obige Tensorprodukt

$$\otimes \;:\; K^p \times K^q \to K^{p \times q} \;:\; (u_i)_i \otimes (v_j)_j = (u_i v_j)_{i,j}.$$

Die Komposition $U \times V \xrightarrow{\sim} K^p \times K^q \to K^{p \times q}$ ist bilinear und hat die universelle Basiseigenschaft, ist also ein Tensorprodukt von U und V.

🙂 Diese Konstruktion ist zwar korrekt, doch irgendwie „quick and dirty“. Die Wahl einer Basis ist immer ein Akt der Willkür, ganz besonders hier. Zudem gelingt dieser Trick zunächst nur in endlicher Dimension.

🙂 Wir werden in Satz S2ı eine allgemeine und basisfreie Konstruktion ausführen, statt „quick and dirty“ ist sie „long-lasting and elegant“.

Beispiel S1D: Polynommodell des Tensorprodukts

Wir betrachten $U = K[X]$ und $V = K[Y]$ und $W = K[X,Y]$ mit

$$\cdot \;:\; K[X] \times K[Y] \to K[X,Y] \;:\; (u,v) \mapsto u \cdot v.$$

Dieses Polynomprodukt bedeutet ausgeschrieben:

$$\big[\textstyle\sum_{i\in\mathbb{N}} u_i X^i\big] \cdot \big[\textstyle\sum_{j\in\mathbb{N}} v_j Y^j\big] \;=\; \textstyle\sum_{(i,j)\in\mathbb{N}^2} u_i v_j X^i Y^j$$

Es erfüllt die geforderten Eigenschaften eines Tensorprodukts:

1. Die Produktabbildung $\cdot$ ist bilinear über dem Grundkörper K.
2. Aus den Monombasen $(X^i)_{i\in\mathbb{N}}$ von $K[X]$ und $(Y^j)_{j\in\mathbb{N}}$ von $K[Y]$ erhalten wir die Produktbasis $(X^i Y^j)_{(i,j)\in\mathbb{N}^2}$ von $K[X,Y]$.

Dasselbe gilt endlich-dimensional für $U = K[X]_{<p}$ und $V = K[Y]_{<q}$ mit $W = \langle\, X^i Y^j \mid i < p,\, j < q \,\rangle_K^!$ und entspricht dem obigen Matrixmodell.

So interpretieren wir $(0,1) \otimes (1,0)$ als $(0X^0 + 1X^1) \cdot (1Y^0 + 0Y^1) = X$, hingegen $(1,0) \otimes (0,1)$ als $(1X^0 + 0X^1) \cdot (0Y^0 + 1Y^1) = Y$ (S105).

☺ Die Definition S1B über die universelle Basiseigenschaft funktioniert problemlos auch für unendlich-dimensionale Räume, wie hier zu sehen.

Aufgabe: Hat die Polynommultiplikation $\cdot : K[X] \times K[X] \to K[X]$ in einer Variablen ebenfalls die Eigenschaften eines Tensorprodukts?

Lösung: Nein. Das Produkt $\cdot : K[X] \times K[X] \to K[X]$ ist zwar bilinear, aber ihm fehlt leider die geforderte universelle Basiseigenschaft:

⚠ Zu den Basen $(X^i)_{i \in \mathbb{N}}$ von $U = K[X]$ und $(X^j)_{j \in \mathbb{N}}$ von $V = K[X]$ ist die Produktfamilie $(X^{i+j})_{(i,j) \in \mathbb{N}^2}$ keine Basis des Raums $W = K[X]$.

⚠ Die beiden Familien $(X^{i+j})_{(i,j) \in \mathbb{N}^2}$ und $(X^k)_{k \in \mathbb{N}}$ beschreiben zwar dieselben Teilmengen in $K[X]$, sind aber dennoch sehr verschieden. Die zweite ist linear unabhängig, sogar eine Basis, die erste nicht!

☺ Die Forderungen an das Tensorprodukt bedeuten anschaulich: (1) Es ist tatsächlich ein Produkt, also bilinear. (2) Dieses Produkt ist so frei wie möglich, das heißt, es erfüllt die universelle Basiseigenschaft.

Polynome als vertrautes Beispiel

Die Frage, wann ein Tensor einfach ist, führt uns in diesem konkreten Beispiel zurück auf die allgegenwärtige Zerlegung von Polynomen:

Übung: Wir betrachten $\cdot : K[X]_{\le p} \times K[X]_{\le q} \to K[X]_{\le p+q}$ für $p = q = 1$.
(1) Über $K = \mathbb{Q}$: Das Polynom $R = X^2 - 2 \in \mathbb{Q}[X]_{\le 2}$ liegt nicht im Bild.
(2) Über $K = \mathbb{R}$: Das Polynom $S = X^2 + 1 \in \mathbb{R}[X]_{\le 2}$ liegt nicht im Bild.

Polynome als vertrautes Beispiel

Die universelle Basiseigenschaft S1B(2) bezieht sich auf eine Basis, und diese dürfen wir willkürlich wählen. War ist mit anderen Basen?

Übung: Gilt die universelle Basiseigenschaft für ein Paar von Basen, so gilt sie für jedes Paar von Basen. Wie zeigen Sie das geschickt? (Die universelle Abbildungseigenschaft S2E klärt dies elegant!)

Exkurs: von klassischer Physik zur Quantenphysik

Die **klassische Physik** beschreibt den Zustand eines Systems durch deterministische Größen, etwa Massen $m_k \in \mathbb{R}$ mit Positionen $u_k \in \mathbb{R}^3$ und Geschwindigkeiten $v_k \in \mathbb{R}^3$ wie in **Newtons Himmelsmechanik:**

$$\dot{u}_k = v_k, \quad \dot{v}_k = f_k(u) := \sum\nolimits_{j \neq k} \gamma\, m_j \frac{u_j - u_k}{\|u_j - u_k\|^3}.$$

Warum diese Beschreibung? Weil sie erfolgreich erklärt und vorhersagt!

Die **Quantenmechanik** beschreibt jeden Zustand durch einen Vektor

$$s \in V$$

in einem $\mathbb{C}$–Vektorraum und Operationen durch lineare Abbildungen. Als kleinste Informationseinheit kann ein **Bit** nur den Zustand 0 oder 1 annehmen. Die Quantenmechanik erlaubt zudem **Überlagerungen**:

$$s = \alpha\,|0\rangle + \beta\,|1\rangle$$

Warum diese Beschreibung? Weil sie erfolgreich erklärt und vorhersagt!

Exkurs: von klassischer Physik zur Quantenphysik

Die Koeffizienten $\alpha, \beta \in \mathbb{C}$ mit $(\alpha, \beta) \neq (0,0)$ heißen auch **Amplituden**. Alle möglichen Zustandsvektoren $s = \alpha\,|0\rangle + \beta\,|1\rangle$ und der Nullvektor 0 bilden somit einen zweidimensionalen $\mathbb{C}$–Vektorraum $V \cong \mathbb{C}^2$.

Dieser trägt zudem ein Skalarprodukt mit Orthonormalbasis $(|0\rangle, |1\rangle)$. Die Absolutquadrate $|\alpha|^2$ und $|\beta|^2$ entsprechen Wahrscheinlichkeiten, dazu normieren wir den Zustand s, sodass $1 = \|s\| = |\alpha|^2 + |\beta|^2$ gilt.

Damit modelliert die Quantenmechanik Zufall, Superposition, Wellenphänomene und vieles mehr. Die hierzu verwendeten Rechenregeln formulieren wir in S1F als **Postulate der Quantenmechanik**.

Wie die Physiker/innen historisch, experimentell und theoretisch zu dieser raffinierten und sonderbaren Beschreibung gekommen sind, ist eine faszinierende Geschichte, leider auch lang und gewunden.

Ich will hier die Vogelperspektive einnehmen und mit ein paar groben Pinselstrichen skizzieren, wie unsere mathematischen Werkzeuge der Linearen Algebra hierbei wunderbar zum Einsatz kommen.

Anwendung: Tensoren im Quantencomputing

Wir betrachten zwei Teilchen, jedes mit zwei **Basiszuständen** $|0\rangle, |1\rangle$. Das zusammengesetzte Zwei-Teilchen-System hat vier Basiszustände:

$$|00\rangle,\ |01\rangle,\ |10\rangle,\ |11\rangle$$

Jeder Zustand ist eine **Überlagerungen** dieser Basiszustände, also

$$s = \alpha\,|00\rangle + \beta\,|01\rangle + \gamma\,|10\rangle + \delta\,|11\rangle$$

mit Koeffizienten $(\alpha, \beta, \gamma, \delta) \in \mathbb{C}^4 \setminus \{0\}$. Die möglichen Zustände und der Nullvektor bilden somit einen vierdimensionalen $\mathbb{C}$–Vektorraum.

☺ Allgemein: Aus n einzelnen Teilchen mit Zustandsräumen $V_1, \dots, V_n$ entsteht das n–Teilchen-System mit Zustandsraum $V = V_1 \otimes \dots \otimes V_n$.

⚠ Hierbei werden die Dimensionen nicht addiert, sondern multipliziert. Die Zustände der Teilchen bestehen nicht unabhängig nebeneinander, wie in $V_1 \oplus \dots \oplus V_n$, sondern sind i.A. verschränkt (engl. *entangled*).

⚠ Superposition ist ein fundamentales Prinzip der Quantenmechanik ohne klassische Entsprechung und daher ohne anschauliche Intuition.

Anwendung: Tensoren im Quantencomputing

Klassisch genügt zur Zustandsbestimmung von n Bits die Kenntnis der n Zustände $z_1, \dots, z_n \in \{0, 1\}$. Dies ist ein Vektor $z \in \{0, 1\}^n$. Die Menge der möglichen Zustände hat somit $N = 2^n$ Elemente, Damit arbeiten klassische, binäre Computer sehr erfolgreich.

In der Quantenmechanik ist der Zustandsraum eines Systems ein $\mathbb{C}$–Vektorraum V. Ist $(v_1, \dots, v_N)$ eine Basis, so können nicht nur diese Basiszustände auftreten, sondern auch Überlagerungen (Superpositionen), also alle $\mathbb{C}$–Linearkombinationen $s = \sum_{k=1}^{N} \alpha_k v_k$.

Die Zustandsbestimmung von n Quantenbits (Qubits) bedeutet somit die Kenntnis aller $N = 2^n$ komplexen Koeffizienten $\alpha_1, \dots, \alpha_N \in \mathbb{C}$. Man kann sich fragen, wo die Natur die zusätzliche Speicher- und Rechenkapazität dafür hernimmt, aber so sind wohl die Spielregeln.

Dieses exponentielle Wachstum macht die Berechnung oder Simulation von Quantensystemen mit klassischen Computern nahezu unmöglich. Umgekehrt schlug Richard Feynman 1982 vor, dies zur Entwicklung von Quantencomputern zu nutzen. Sie werden es wohl noch erleben.

Sind Anwendungsbeispiele ein Fluch oder ein Segen?

So manche/r stöhnt vermutlich: „Jetzt übertreibt der Prof: Wir müssen nicht nur Lineare Algebra lernen, jetzt auch noch Quantenmechanik!“

Ja, bitte nehmen Sie mögliche Anwendungen interessiert zur Kenntnis! Der Blick über den Tellerrand lohnt sich, gerade an der Uni Stuttgart.

Sind solche Exkurse übertrieben? ablenkend? oder gar überfordernd? Ich nenne drei Gründe für akademische Neugier und Abenteuerlust:

1. Wir sollten uns freuen, dass mathematische Begriffe und Techniken so überraschend vielseitig anwendbar sind: Was Sie in einem Gebiet lernen, können Sie in vielen anderen Gebieten wiederverwenden.

2. Speziell im Fall des Quantencomputing ist realistisch absehbar, dass dies Ihr Leben nachhaltig beeinflussen wird. Ich will daher nicht kommentarlos an diesem wichtigen Anwendungsbeispiel vorübergehen.

3. Das Studium mathematischer Grundlagen ist fundamental wichtig. Wir wollen abstrahieren, um zahlreiche Einzelfälle zusammenzufassen, zugleich auch motivieren, konkretisieren, illustrieren. Nur so geht es!

Sind Anwendungsbeispiele ein Fluch oder ein Segen?

Der Spagat zwischen diesen beiden Polen ist eine fragile Kunst:
Einerseits die abstrakten Grundlagen: einfach, elegant, allgemein.
Andererseits die konkreten Einzelfälle: verwirrend, sonderbar, speziell.

Manche vertauschen hier die Adjektive, das halte ich für einen Fehler.
Abstraktion strukturiert und vereinfacht: Eine allgemeine Tatsache ist oft leichter zu verstehen und zu erklären als ihre zahlreichen Spezialfälle.

Studierende fordern zu Recht motivierende Beispiele, Querverbindung zwischen den Themen und auch Bezug zu möglichen Anwendungen.
Wir sollten uns also über jede Gelegenheit freuen, wo dies gelingt!

Be careful what you wish for, it might just come true!

Wird der Wunsch nach Beispielen und Anwendungen ernsthaft erfüllt, so fangen erfahrungsgemäß leider auch immer einige an zu jammern:
Der Ruf nach Anwendungsbeispielen entsprang nicht ihrer Neugier, sondern der naiven Sehnsucht nach Einfachheit. Die ist oft unerfüllbar:
Ehrliche Anwendungen sind meist nicht leichter zu verstehen, sondern komplizierter als die übergeordnete Theorie. Sie benötigen beides!

Sind Anwendungsbeispiele ein Fluch oder ein Segen?

Ich bin überzeugt, naturwissenschaftlich-technische Neugier und Allgemeinbildung sind wichtige Bildungsziele unserer Universität. Dazu möchte ich Gelegenheiten bieten, als wohlmeinendes Angebot.

Man kann das Pferd zum Wasser führen, aber nicht zum Trinken zwingen.

Ich weiß auch, dass unser Bildungssystem zumeist nicht Neugier und Lernfreude fördert, sondern Auswendiglernen und Nachbeten erzwingt.

Intrinsische Motivation gelingt nur aufwändig mit viel Mühe und Geduld, extrinsische Zwänge hingegen sind sofort und kostengünstig zu haben.

Auch hier macht die Mischung den Erfolg. Einseitige Lerndressur führt junge Menschen leider zu fatalem Desinteresse an ihrer Bildung:

Ein gutes Pferd springt nur so hoch, wie es muss.

Kurzum, bevor Sie fragen: Dieser Abschnitt ist nicht klausurrelevant. Bitte nutzen Sie Ihre Freiheiten weise und entscheiden Sie selbst!

Habe Mut, dich deines eigenen Verstandes zu bedienen!

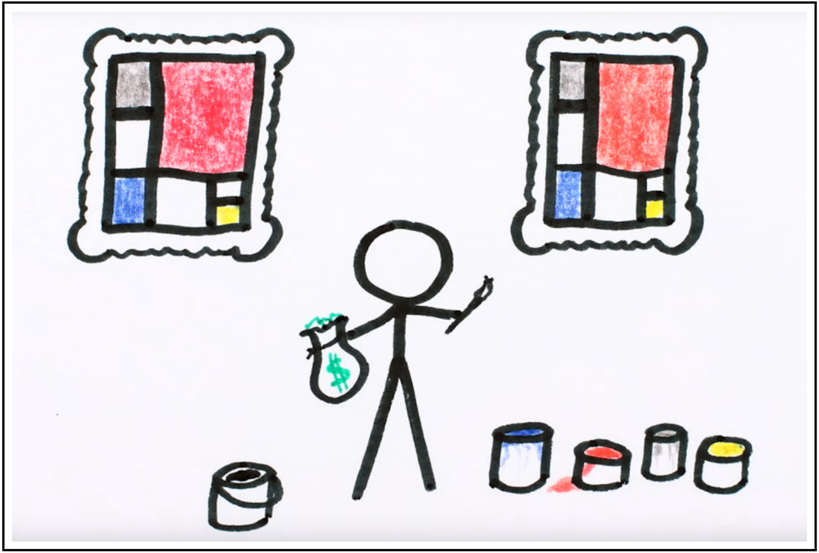
Zum Spannungsfeld von Theorie und Anwendung, zur Wechselwirkung von Grundlagenforschung und technischer Umsetzung zitiere ich Albert Einstein aus seiner Rede anlässlich der Eröffnung der 7. Deutschen Funkausstellung in Berlin 1930 (siehe `youtu.be/VqvGRVM2jk8`).

Verehrte An- und Abwesende!
Wenn Ihr den Rundfunk höret, so denkt auch daran, wie die Menschen in den Besitz dieses wunderbaren Werkzeuges der Mitteilung gekommen sind. Der Urquell aller technischen Errungenschaften ist die göttliche Neugier und der Spieltrieb des bastelnden und grübelnden Forschers und nicht minder die konstruktive Phantasie des technischen Erfinders.

[. . .]

Sollen sich auch alle schämen, die gedankenlos sich der Wunder der Wissenschaft und Technik bedienen und nicht mehr davon geistig erfasst haben als die Kuh von der Botanik der Pflanzen, die sie mit Wohlbehagen frisst.

Anwendung: das No-Cloning-Theorem



Als wunderschönes Video von MinutePhysics, `youtu.be/owPC60Ue0BE`.

Anwendung: das No-Cloning-Theorem

> **A single quantum cannot be cloned**
>
> **If a photon of definite polarization encounters an excited atom, there is typically some nonvanishing probability that the atom will emit a second photon by stimulated emission. Such a photon is guaranteed to have the same polarization as the original photon. But is it possible by this or any other process to amplify a quantum state, that is, to produce several copies of a quantum system (the polarized photon in the present case) each having the same state as the original? If it were, the amplifying process could be used to ascertain the exact state of a quantum system: in the case of a photon, one could determine its polarization by first producing a beam of identically polarized copies and then measuring the Stokes parameters[1]. We show here that the linearity of quantum mechanics forbids such replication and that this conclusion holds for all quantum systems.**

Abstract des vielzitierten Artikels von W.K. Wootters, W.H. Zurek:
A single quantum cannot be cloned. Nature 299 (1982) 802–803

In *Nature* ist selbst eine kurze Notiz eine beachtliche Publikation;
zur Geschichte siehe `en.wikipedia.org/wiki/No-cloning_theorem`.

Anwendung: das No-Cloning-Theorem

Das berühmte **No-Cloning-Theorem** der Quantenmechanik besagt:

Es ist unmöglich, ein quantenmechanisches System perfekt auf ein zweites zu kopieren, ohne dabei das erste zu verändern.

Dies beruht auf einer einfachen Rechnung der Linearen Algebra:

Satz S1E: das No-Cloning-Theorem

Sei V ein K–Vektorraum der Dimension $\dim V \ge 2$ und $e \in V \smallsetminus \{0\}$.
Dann existiert keine K–lineare Abbildung $T : V \otimes V \to V \otimes V$ mit

$$T\big[\,|v\rangle \otimes |e\rangle\,\big] = |v\rangle \otimes |v\rangle$$

für alle möglichen Zustände $v \in V$.

🙂 Es genügen die grundlegenden Rechenregeln aus Definition S1B.

Die *Rechnung* ist erschütternd einfach, die physikalische *Interpretation* ebenso erschütternd tiefsinnig, siehe hierzu das oben zitierte Video.

Übung: Versuchen Sie es bitte zunächst selbst, bevor Sie weiterlesen.

Anwendung: das No-Cloning-Theorem

Beweis: Angenommen, es gäbe solch eine lineare Abbildung T.
Wir wählen eine Basis $(a, b, \dots)$ von V und betrachten den Zustand

$$s = (|a\rangle + |b\rangle) \otimes |e\rangle \qquad \in V \otimes V.$$

Rechenweg 1: Wir multiplizieren erst aus und wenden dann T an.

$$\begin{aligned} T(s) &\overset{\text{(a)}}{=} T\big[\, (|a\rangle + |b\rangle) \otimes |e\rangle \,\big] \\ &\overset{\text{Bil}}{=} T\big[\, (|a\rangle \otimes |e\rangle) + (|b\rangle \otimes |e\rangle) \,\big] \\ &\overset{\text{Lin}}{=} T\big[\, |a\rangle \otimes |e\rangle \,\big] + T\big[\, |b\rangle \otimes |e\rangle \,\big] \\ &\overset{\text{Def}}{=} (|a\rangle \otimes |a\rangle) + (|b\rangle \otimes |b\rangle) \end{aligned}$$

Rechenweg 2: Wir wenden erst T an und multiplizieren dann aus.

$$\begin{aligned} T(s) &\overset{\text{(a)}}{=} T\big[\, (|a\rangle + |b\rangle) \otimes |e\rangle \,\big] \\ &\overset{\text{Def}}{=} (|a\rangle + |b\rangle) \otimes (|a\rangle + |b\rangle) \\ &\overset{\text{Bil}}{=} (|a\rangle \otimes |a\rangle) + (|a\rangle \otimes |b\rangle) + (|b\rangle \otimes |a\rangle) + (|b\rangle \otimes |b\rangle) \end{aligned}$$

Der Koeffizientenvergleich (dank S1B) führt zum Widerspruch. $\boxed{\text{QED}}$

EPR: das Einstein–Podolsky–Rosen–Paradox

Wir betrachten unser Zwei-Teilchen-System $V = V_1 \otimes V_2$ im Zustand

$$s = \alpha\,|00\rangle + \beta\,|01\rangle + \gamma\,|10\rangle + \delta\,|11\rangle.$$

Die Basiszustände bilden die Orthonormalbasis $(|00\rangle, |01\rangle, |10\rangle, |11\rangle)$.

$A = \begin{bmatrix} 0&0&0&0\\ 0&0&0&0\\ 0&0&1&0\\ 0&0&0&1 \end{bmatrix}$

Alice misst $a \in \{0,1\}$.

Wkt	0	1
0	$\lvert\alpha\rvert^2$	$\lvert\beta\rvert^2$
1	$\lvert\gamma\rvert^2$	$\lvert\delta\rvert^2$

$B = \begin{bmatrix} 0&0&0&0\\ 0&1&0&0\\ 0&0&0&0\\ 0&0&0&1 \end{bmatrix}$

Bob misst $b \in \{0,1\}$.

Mit A misst Alice das erste Teilchen und erhält den Messwert $a \in \{0,1\}$.

Das zerlegt $V = \mathrm{E}_0 \oplus \mathrm{E}_1$ in die Eigenräume $\mathrm{E}_a = \mathbb{C}|a\rangle \otimes V_2$ und somit $s = s_0 + s_1$ in $s_0 = \alpha\,|00\rangle + \beta\,|01\rangle \in \mathrm{E}_0$ und $s_1 = \gamma\,|10\rangle + \delta\,|11\rangle \in \mathrm{E}_1$.

Postulate der Messung (S1F): Alice misst den Wert $a \in \{0,1\}$ mit Wkt

$$p_a = \frac{\|s_a\|^2}{\|s\|^2} = \frac{|\alpha|^2 + |\beta|^2 \text{ bzw. } |\gamma|^2 + |\delta|^2}{|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2}.$$

Durch die Messung des Wertes a kollabiert der Zustand von s zu s_a.

EPR: das Einstein–Podolsky–Rosen–Paradox

Sie kennen solche **Vierfeldertafeln** aus der Schule. In der Stochastik stellen Sie so die gemeinsamen Wkten von zwei Ereignissen A und B übersichtlich dar, hier entsprechend für unsere beiden Zufallsvariablen. Damit beantworten Sie Fragen zu Korrelation und Unabhängigkeit.

In unserem Beispiel werden jeweils nur zwei Werte angenommen, nämlich 0 oder 1; diese Miniatur ist der kleinste interessante Fall. Ganz genauso beschreiben wir jedes endlich-dimensionale System; den unendlich-dimensionalen Fall führt die Funktionalanalysis fort.

Die komplexen Koeffizienten $\alpha, \beta, \gamma, \delta \in \mathbb{C}$ werden auch **Amplituden** genannt und entsprechen den Wahrscheinlichkeiten $|\alpha|^2, |\beta|^2, |\gamma|^2, |\delta|^2$. Damit diese sich zu 1 summieren, dividieren wir durch das Normquadrat

$$\|s\|^2 = |\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2.$$

Das kann ganz am Ende geschehen. Oft nimmt man dies schon anfangs vorweg, indem man den Zustandsvektor s auf Länge $\|s\| = 1$ normiert. Genau so habe ich für unsere obige Tabelle alles vorsorglich normiert.

EPR: das Einstein–Podolsky–Rosen–Paradox

Wir betrachten zwei unabhängige Teilchen, also einen **reinen Tensor**:

$$\begin{aligned} s = u \otimes v &= (\alpha_0|0\rangle + \alpha_1|1\rangle) \otimes (\beta_0|0\rangle + \beta_1|1\rangle) \\ &= \alpha_0\beta_0|00\rangle + \alpha_0\beta_1|01\rangle + \alpha_1\beta_0|10\rangle + \alpha_1\beta_1|11\rangle \end{aligned}$$

Zur Vereinfachung normieren wir zu $|\alpha_0^2| + |\alpha_1^2| = |\beta_0^2| + |\beta_1^2| = 1$.
Das erste Teilchen transportieren wir zu Alice, das zweite zu Bob.

$A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Alice misst $a \in \{0, 1\}$.

Wkt	0	1
0	$\lvert\alpha_0^2\rvert \cdot \lvert\beta_0^2\rvert$	$\lvert\alpha_0^2\rvert \cdot \lvert\beta_1^2\rvert$
1	$\lvert\alpha_1^2\rvert \cdot \lvert\beta_0^2\rvert$	$\lvert\alpha_1^2\rvert \cdot \lvert\beta_1^2\rvert$

$B = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Bob misst $b \in \{0, 1\}$.

Alice misst den Zustand und findet den Wert $a \in \{0, 1\}$ mit Wkt $|\alpha_a|^2$.
Dadurch kollabiert der Zustand von s zu $s_a = \alpha_a(\beta_0|a0\rangle + \beta_1|a1\rangle)$.
Daraufhin misst Bob und findet den Wert $b \in \{0, 1\}$ mit Wkt $|\beta_b|^2$.
Fazit: Die Messergebnisse von Alice und Bob sind unabhängig.

☺ Das entspricht unserer Intuition: klassisch, unabhängig.

EPR: das Einstein–Podolsky–Rosen–Paradox

Wir bringen unsere beiden Teilchen in den **verschränkten Zustand**

$$s = \frac{1}{\sqrt{2}} \cdot |00\rangle + 0 \cdot |01\rangle + 0 \cdot |10\rangle + \frac{1}{\sqrt{2}} \cdot |11\rangle.$$

Dieser Zustand lässt sich nicht als reiner Tensor $u \otimes v$ darstellen.
Das erste Teilchen transportieren wir zu Alice, das zweite zu Bob.

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Alice misst $a \in \{0, 1\}$.

Wkt	0	1
0	50%	0%
1	0%	50%

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Bob misst $b \in \{0, 1\}$.

Alice misst den Zustand und findet $a \in \{0, 1\}$, jeweils mit gleicher Wkt.
Bob misst den Zustand und findet $b \in \{0, 1\}$, jeweils mit gleicher Wkt.
Hier sind die Messergebnisse von Alice und Bob immer identisch!
Wie kann das sein, instantan über beliebig große Distanzen?

⚠ Der Wesenskern der Quantenmechanik ist die Superposition.
Dafür gibt es keine klassische Entsprechung oder Anschauung.

Warum gilt diese einfache Rechnung als paradox?
Das liegt ganz allein an der physikalischen Interpretation:

Alice und Bob können beliebig weit entfernt sein! Alice' Messung entspricht einem Münzwurf, mit Ergebnis 0 für Kopf und 1 für Zahl. Auch Bobs Messung entspricht einem Münzwurf. Diese sind jedoch nicht unabhängig, sondern beide Ergebnisse stimmen immer überein.

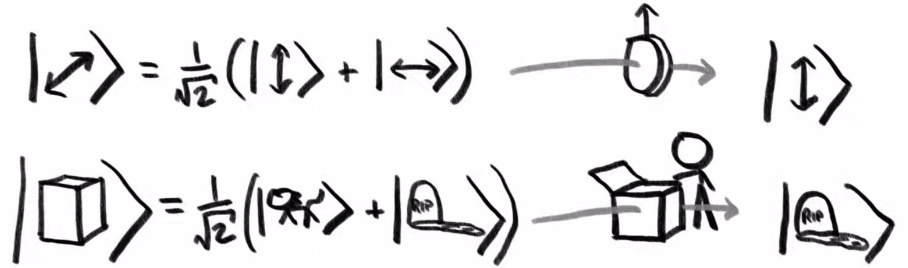
Wie kann das sein? Wird hier Information von Alice zu Bob übertragen, und zwar instantan und somit schneller als Lichtgeschwindigkeit? Einstein wollte sich mit dieser „spukhaften Fernwirkung“ nicht abfinden. Aus dieser Sicht wird die obige Rechnung als „paradox“ bezeichnet.

Das Phänomen ist sicherlich unerwartet und unintuitiv,
aber Experimente bestätigen genau diese Vorhersage.
Information wird hier nicht übertragen, denn Alice kann
ihren Wert zwar messen, doch nicht willkürlich vorgeben.

Lineare Algebra und Quantenmechanik

Auch hier ist die *Rechnung* erschütternd einfach, vor allem dank des zugrundeliegenden, sehr eleganten Tensorkalküls. Die physikalische *Interpretation* jedoch ist ebenso erschütternd tiefsinnig und schwierig.

> *If people do not believe that mathematics is simple,*
> *it is only because they do not realize how complicated life is.*
> John von Neumann (1903–1957)



Video von 3Blue1Brown und MinutePhysics, `youtu.be/MzRCDLre1b4`, und ebenso sehenswert von Veritasium, `youtu.be/ZuvK-od647c`.

Lineare Algebra und Quantenmechanik

MAY 15, 1935 PHYSICAL REVIEW VOLUME 47

Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?

A. EINSTEIN, B. PODOLSKY AND N. ROSEN, *Institute for Advanced Study, Princeton, New Jersey*

(Received March 25, 1935)

In a complete theory there is an element corresponding to each element of reality. A sufficient condition for the reality of a physical quantity is the possibility of predicting it with certainty, without disturbing the system. In quantum mechanics in the case of two physical quantities described by non-commuting operators, the knowledge of one precludes the knowledge of the other. Then either (1) the description of reality given by the wave function in quantum mechanics is not complete or (2) these two quantities cannot have simultaneous reality. Consideration of the problem of making predictions concerning a system on the basis of measurements made on another system that had previously interacted with it leads to the result that if (1) is false then (2) is also false. One is thus led to conclude that the description of reality as given by a wave function is not complete.

Warum erzähle ich Ihnen das hier, in der Linearen Algebra?
Ist Quantenmechanik nicht viel zu schwer und unanschaulich?

Ja, Quantenmechanik ist unanschaulich, doch von fundamentaler Bedeutung: physikalisch, mathematisch, technisch, philosophisch, …

Vor allem aber können wir an solchen Anwendungen viel lernen, vielleicht sogar etwas Physik, ganz sicher aber viel Mathematik!

Ich präsentiere Ihnen diesen faszinierenden Exkurs, um zu illustrieren, dass Sie die mathematischen Grundlagen Ihrer Linearen Algebra sehr vielseitig anwenden können, in der Physik und auch überall sonst.

Ich betone dabei nochmals: Die Mathematik ist zunächst nicht einfach, doch in Ihrer Reichweite und mit Übung gut erlernbar und verständlich. Es lohnt sich daher in solide mathematische Grundlagen zu investieren.

Abstraktion strukturiert und vereinfacht: Eine allgemeine Tatsache ist oft leichter zu verstehen und zu erklären als ihre zahlreichen Spezialfälle. In den Anwendungen hingegen tobt das Leben, prall und verwirrend.

Die obigen Miniaturen zum No-Cloning-Theorem und dem EPR-Paradox zeigen eindrücklich: Die mathematischen Grundlagen der Rechnungen sind „einfache Lineare Algebra“, wie Physiker/innen gerne sagen.

Und das zu Recht. Die (abstrakten) mathematischen Werkzeuge muss man nur einmal erlernen, sie kosten anfangs etwas Geduld und Mühe, doch das ist gut investiert: sie lassen sich überall (konkret) anwenden.

Definition S1F: Postulate der Quantenmechanik

Die Quantenmechanik wird durch folgende Postulate zusammengefasst:

1. Der Zustand eines Systems wird beschrieben durch einen Vektor $\psi \in V$ in einem Hilbert–Raum V über $\mathbb{C}$, genannt Zustandsraum.
2. Jede physikalisch messbare Größe, kurz Observable genannt, wird beschrieben durch einen hermiteschen Operator $A : V \to V$.
3. Das Ergebnis jeder Messung von A ist ein Eigenwert $a \in \mathbb{R}$ von A. (Wenn wir $\dim V < \infty$ annehmen, so wissen wir $V = \bigoplus_{a \in \mathbb{R}} \mathrm{E}_a$.)
4. Die Wkt des Messwerts a im Zustand ψ ist $\mathbf{P}(a|\psi) = \|\psi_a\|^2 / \|\psi\|^2$ gemäß der orthogonalen Zerlegung $V = \mathrm{E}_a \oplus \mathrm{E}_a^\perp$ und $\psi = \psi_a + \psi_a'$.
5. Nach der Messung von a befindet sich das System im Zustand ψ_a. Das ist die orthogonale Projektion von V auf den Eigenraum E_a.
6. Die Schrödinger–Gleichung $\mathrm{i}\hbar\, \partial_t \psi(t) = H(t)\psi(t)$ beschreibt die zeitliche Entwicklung; die Observable $H(t)$ ist die Gesamtenergie.

Lineare Algebra und Quantenmechanik

☺ Die Postulate (1–5) präzisieren die grundlegenden Rechenregeln der Quantenmechanik in der vertrauten Sprache der Linearen Algebra!

So formulierte Werner Heisenberg 1925 seine **Matrizenmechanik** und Erwin Schrödinger im selben Jahre seine **Wellenmechanik**; dafür erhielten sie 1932 bzw. 1933 den Nobelpreis für Physik.

⚠ Zur Vereinfachung möchte ich im Folgenden $\dim V < \infty$ annehmen. Die zur Quantenmechanik nötige Mathematik liegt dann vollständig im Rahmen der **Linearen Algebra**. Der allgemeine Fall ist noch viel faszinierender, benötigt aber stärkere Werkzeuge der **Analysis**.

Diese Postulate der Quantenmechanik kann man wohl kaum intuitiv verstehen, wohl aber gut nutzen: Sie beschreiben die Experimente!

I think I can safely say that nobody understands quantum mechanics.

Richard P. Feynman, *The Character of Physical Law* (1965)

Ich betone nochmals: Die wesentliche Schwierigkeit liegt nicht in der mathematischen Formulierung, dazu haben Sie nun alle Werkzeuge. Das Verrückte ist einzig und allein die physikalische Interpretation.

Aufgabe: Lösen Sie die Schrödinger–Gleichung im Falle $\dim V < \infty$. Der Schrödinger–Hamilton–Operator $H : V \to V$ sei zeitlich konstant:

$$\boxed{\mathrm{i}\hbar\,\partial_t \psi(t) = H\psi(t), \quad \psi(t_0) = \psi^0}$$

Lösung: Dank Spektralsatz existiert eine Orthonormalbasis $(v_1, \dots, v_n)$ von V aus Eigenvektoren, also $Hv_k = E_k v_k$ mit $E_k \in \mathbb{R}$ für $k = 1, \dots, n$.

Wir untersuchen nun $\psi : \mathbb{R} \to V : t \mapsto \psi(t)$ als zeitabhängigen Zustand. Zu jedem Zeitpunkt $t \in \mathbb{R}$ zerlegen wir den Zustand $\psi(t) = \sum_k \psi_k(t) v_k$ als Summe orthogonaler Eigenvektoren, also $\psi_k(t) = \langle\, v_k \mid \psi(t)\,\rangle \in \mathbb{C}$.

Die allgemeine Schrödinger–Gleichung $\mathrm{i}\hbar\,\partial_t \psi(t) = H(t)\psi(t)$ wird nun entkoppelt zu $\mathrm{i}\hbar\,\partial_t \psi_k(t) = E_k \psi_k(t)$. Aus $\psi(t_0) = \sum_k \psi_k^0 v_k$ folgt somit:

$$\boxed{\psi(t) = \sum_k \mathrm{e}^{-\mathrm{i}(t-t_0)E_k/\hbar}\, \psi_k^0 v_k}$$

Alternative: Mit der Matrix-Exponentialfunktion (N3B) können wir die eindeutige Lösung zusammenfassen durch $\psi(t) = \mathrm{e}^{\mathrm{i}(t-t_0)H/\hbar}\, \psi(t_0)$.

Lineare Algebra und Quantenmechanik

🙂 Die explizite Lösung gelingt überraschend leicht: Wir benötigen dazu nur eine Orthonormalbasis $(v_1, \dots, v_n)$ von V aus Eigenvektoren von H.

Wie der Schrödinger–Operator H genau aussieht, hängt vom konkreten Modell ab. Die Lösungsmethode hingegen ist universell anwendbar!

🙂 Wir erkennen hier eine erstaunlich simple zeitliche Entwicklung: Die Amplitude $\psi_k(t) = \mathrm{e}^{-\mathrm{i}(t-t_0)E_k/\hbar}\,\psi_k^0$ bleibt im Betrag immer gleich, nur die Phase ändert sich um den Faktor $\mathrm{e}^{-\mathrm{i}(t-t_0)E_k/\hbar} \in \mathbb{S}^1 = \mathrm{U}_1\,\mathbb{C}$.

Die Geschwindigkeit $E_k/\hbar$ dieser Phasenänderung ist die Energie E_k des Eigenzustands v_k geteilt durch eine universelle Naturkonstante, das Plancksche Wirkungsquantum $\hbar = h/2\pi \approx 1.055 \cdot 10^{-34}\mathrm{Js}$.

🙂 Die Wkten $|\psi_k(t)|^2 \in [0,1]$ ändern sich nicht, lediglich die Phase! Die Phase ist nicht direkt messbar, spielt aber dennoch eine wichtige Rolle: Bei der Addition (Überlagerung, Superposition) führt sie zu Verstärkung oder Auslöschung; man spricht hier von Interferenz. Die Wkten $|\psi_k(t)|^2 \in [0,1]$ alleine genügen dafür nicht!

QM ♡ LA: Erwartung und Varianz

Aufgabe: Sei V ein $\mathbb{C}$–Vektorraum mit Skalarprodukt $\langle - \mid - \rangle$ und $\dim V = n < \infty$ und hierauf $A : V \to V$ ein hermitescher Operator.

Wir interpretieren V als Zustandsraum und A als Observable. Berechnen Sie gemäß den Postulaten der Quantenmechanik (S1F) Erwartung und Varianz des Operators A im gegebenen Zustand $\psi \in V$.

Lösung: Dank Spektralsatz existiert eine Orthonormalbasis $(v_1, \dots, v_n)$ von V aus Eigenvektoren, also $Av_i = a_i v_i$ mit $a_i \in \mathbb{R}$ für $i = 1, \dots, n$.

Wir zerlegen den Zustand $\psi = \sum_{i=1}^n \psi_i v_i$ in eine Summe orthogonaler Eigenvektoren. Die Koeffizienten hierfür sind $\psi_i = \langle v_i \mid \psi \rangle \in \mathbb{C}$ (P1O).

Die Betragsquadrate $|\psi_i^2|$ entsprechen Wkten. Wir normieren daher den Zustand ψ und erhalten so die Gesamtwkt $1 = \|\psi\|^2 = \sum_{i=1}^n |\psi_i^2|$.

Bei vorgegebenem Zustand $\psi \in V$ mit $\|\psi\| = 1$ interpretieren wir die Messung durch den hermiteschen Operator A als Zufallsvariable.

(Etwas nachlässig doch bequem bezeichne ich im Folgenden beides, den Operator A und die Zufallsvariable A, mit demselben Symbol.)

QM ♡ LA: Erwartung und Varianz

Die Erwartung μ der Messung durch A ist demnach:

$$\begin{aligned}
\mu = \mathbf{E}(A) &= \textstyle\sum_i a_i |\psi_i^2| = \sum_i a_i \langle\, \psi_i \mid \psi_i \,\rangle = \sum_i a_i \langle\, \psi \mid \psi_i \,\rangle \\
&= \textstyle\sum_i \langle\, \psi \mid A\psi_i \,\rangle = \big\langle\, \psi \,\big|\, \sum_i A\psi_i \,\big\rangle = \big\langle\, \psi \,\big|\, A \sum_i \psi_i \,\big\rangle \\
&= \langle\, \psi \mid A\psi \,\rangle = \langle\, \psi \mid A \mid \psi \,\rangle =: \langle\, A \,\rangle
\end{aligned}$$

Hierbei ist $\langle\, A \,\rangle$ die bequeme Kurzschreibweise zum gegebenen Zustand ψ. Die Erwartung von A^2 ist entsprechend:

$$\mathbf{E}(A^2) = \langle\, A^2 \,\rangle = \langle\, \psi \mid A^2 \mid \psi \,\rangle = \langle\, A\psi \mid A\psi \,\rangle$$

Daraus erhalten wir wie üblich die Varianz der Zufallsvariable A als Erwartung der quadratischen Abweichung vom Mittelwert:

$$\begin{aligned}
\sigma^2 = \mathbf{V}(A) &= \mathbf{E}((A-\mu)^2) = \mathbf{E}(A^2 - 2\mu A + \mu^2) \\
&= \mathbf{E}(A^2) - 2\mu\mathbf{E}(A) + \mu^2 = \mathbf{E}(A^2) - \mathbf{E}(A)^2 = \langle\, A^2 \,\rangle - \langle\, A \,\rangle^2
\end{aligned}$$

Hier verbinden sich Lineare Algebra, Skalarprodukte und Wahrscheinlichkeitsrechnung wunderbar einfach und elegant.

QM ♡ LA: Erwartung und Varianz

🙂 Zusammenfassend erhalten wir das folgende schöne Ergebnis:

Satz S1G: Erwartung und Varianz

Sei V ein Hilbert–Raum über $\mathbb{C}$; zur Vereinfachung gelte $\dim V < \infty$.

Gegeben sei ein hermitescher Endomorphismus $A : V \to V$ und ein Vektor $\psi \in V$. Nach den Postulaten der Quantenmechanik (S1F) interpretieren wir die Messung durch A als eine Zufallsvariable.

(1) Erwartung und Varianz sind dann gegeben durch:

$$\boxed{\begin{aligned} \mu = \mathbf{E}(A) &= \langle\, A \,\rangle := \langle\, \psi \mid A \mid \psi \,\rangle \\ \mathbf{E}(A^2) &= \langle\, A^2 \,\rangle = \langle\, \psi \mid A^2 \mid \psi \,\rangle = \langle\, A\psi \mid A\psi \,\rangle \\ \sigma^2 = \mathbf{V}(A) &= \mathbf{E}(A^2) - \mathbf{E}(A)^2 = \langle\, A^2 \,\rangle - \langle\, A \,\rangle^2 \end{aligned}}$$

Dieser Satz gilt sinngemäß genauso in unendlicher Dimension; die dazu nötige technische Ausführung diskutieren wir hier nicht.

🙂 Mögliche Messwerte sind die Eigenwerte von A, und ihre Wkten sind verteilt gemäß ψ. Der Erwartungswert $\mu \in \mathbb{R}$ ist der Schwerpunkt dieser Verteilung, und die Streuung $\sigma \in \mathbb{R}_{\geq 0}$ misst die typische Breite.

🙂 Bitte beachten Sie den geschickten und fließenden Übergang von „koordinatengebundener“ zu „koordinatenfreier“ Formulierung: Unsere Rechnungen nutzen eine Eigenbasis von A, das Ergebnis jedoch nicht!

⚠ Auf dem Zustandsraum V arbeiten wir im Allgemeinen mit mehreren verschiedenen hermiteschen Operatoren $A, B, \dots$. Es hat daher meist keinen rechten Sinn, sich auf eine Basis festzulegen: Jeder hermitesche Operator erlaubt zwar eine Eigenbasis, doch bei jedem Wechsel des Operators müssen wir im Allgemeinen auch die Basis wechseln.

🙂 In konkreten Rechnungen ist meist die Wahl einer Basis hilfreich; genau so haben wir oben die Erwartung definiert und berechnet.

🙂 Für die Ergebnisse sind basisfreie Formulierungen am schönsten; genau so haben wir den obigen Satz S1G formuliert.

QM ♡ LA: Heisenbergs Unschärferelation

Wann können wir mit einem „scharfen“ Messergebnis rechnen?
Darauf antwortet die folgende Ergänzung des obigen Satzes:

Satz S1G: Erwartung und Varianz

(2) Genau dann gilt $\mathbf{V}(A) = 0$, wenn ψ ein Eigenzustand von A ist. In diesem Fall ist $\mathbf{E}(A) = \mu$ der zugehörige Eigenwert, also $A\psi = \mu\psi$.

Aufgabe: Beweisen Sie diese Äquivalenz.

Lösung: „$\Leftarrow$“: Angenommen, $\psi \in V \smallsetminus \{0\}$ erfüllt $A\psi = \mu\psi$ mit $\mu \in \mathbb{R}$. Zudem gelte $\|\psi\| = 1$. Dann gilt $\mathbf{E}(A) = \langle\, \psi \mid A \mid \psi \,\rangle = \mu$ und ebenso $\mathbf{E}(A^2) = \langle\, \psi \mid A^2 \mid \psi \,\rangle = \mu^2$, also $\mathbf{V}(A) = \mathbf{E}(A^2) - \mathbf{E}(A)^2 = 0$.

„$\Rightarrow$“: Als Kontraposition zeigen wir: Ist ψ kein Eigenzustand, so gilt $\mathbf{V}(A) > 0$. Zunächst normieren wir $\|\psi\| = 1$. Da A hermitesch ist und $\dim V < \infty$, zerfällt V in die orthogonale Summe von Eigenräumen. Wir zerlegen $\psi = \sum_{k=1}^{m} \psi_k$ in Eigenzustände $\psi_k \ne 0$ mit $A\psi_k = \lambda_k\psi_k$ und $\lambda_k \ne \lambda_\ell$ für $\mathrm{k} \ne \ell$. Gilt dabei $m \ge 2$, so folgt für die Varianz $\mathbf{V}(A) = \mathbf{E}((A-\mu)^2) = \sum_{k=1}^{m} (\lambda_k - \mu)^2 \|\psi_k^2\| > 0$.

QM ♡ LA: Heisenbergs Unschärferelation

Satz S1H: Unschärferelation, Heisenberg 1927, Robertson 1929

Sei V ein Hilbert–Raum über $\mathbb{C}$; zur Vereinfachung gelte $\dim V < \infty$. Gegeben seien hermitesche Operatoren $A, B : V \to V$. Ihr Kommutator $[A, B] = AB - BA$ misst somit die Abweichung von der Kommutation.

Für jeden Zustand $\psi \in V$ gilt dann die Ungleichung

$$\boxed{\mathbf{V}(A)\,\mathbf{V}(B) \;\ge\; \frac{1}{4}\big|\langle\,[A,B]\,\rangle\big|^2.}$$

Im Falle $\mathbf{E}(A) = \mathbf{E}(B) = 0$ bedeutet das ausgeschrieben:

$$\langle\,\psi \mid A^2 \mid \psi\,\rangle\langle\,\psi \mid B^2 \mid \psi\,\rangle \;\ge\; \frac{1}{4}\big|\langle\,\psi \mid [A,B] \mid \psi\,\rangle\big|^2$$

Beweis: Zur Vereinfachung der Rechnung zentrieren wir A und B zu $A' = A - \mathbf{E}(A)$ und $B' = B - \mathbf{E}(B)$. Damit verschieben wir die beiden Erwartungswerte zu 0, doch Varianz und Kommutator ändern sich nicht. Wir können und werden im Folgenden $\mathbf{E}(A) = \mathbf{E}(B) = 0$ annehmen.

QM ♡ LA: Heisenbergs Unschärferelation

Wir nutzen Hermitizität und die Cauchy–Schwarz–Ungleichung:

$$\begin{aligned}\langle A^2 \rangle \langle B^2 \rangle &= \langle \psi \mid A^2\psi \rangle \langle \psi \mid B^2\psi \rangle = \langle A\psi \mid A\psi \rangle \langle B\psi \mid B\psi \rangle \\ &\ge \big|\langle A\psi \mid B\psi \rangle\big|^2 = \big|\langle \psi \mid A\,B\,\psi \rangle\big|^2 = \big|\langle AB \rangle\big|^2.\end{aligned}$$

Wir zerlegen $AB = \frac{1}{2}\{A,B\} + \frac{1}{2}[A,B]$ in die Summanden

$$\{A,B\} = AB + BA \quad \text{und} \quad [A,B] = AB - BA.$$

Der Antikommutator $\{A,B\}$ ist hermitesch, also $\langle \{A,B\} \rangle$ reell, der Kommutator $[A,B]$ ist antihermitesch, also $\langle [A,B] \rangle$ imaginär.

$$\big|\langle AB \rangle\big| = \Big|\frac{1}{2}\langle \{A,B\} \rangle + \frac{1}{2}\langle [A,B] \rangle\Big| \ge \frac{1}{2}\big|\langle [A,B] \rangle\big|$$

Zusammengefasst erhalten wir die allgemeine Unschärferelation:

$$\boxed{\langle A^2 \rangle \langle B^2 \rangle \ge \frac{1}{4}\big|\langle [A,B] \rangle\big|^2}$$

Das beweist die behauptete Ungleichung. QED

QM ♡ LA: Heisenbergs Unschärferelation

Auch hier ist die mathematische Behandlung vergleichsweise leicht, die physikalische Interpretation jedoch erwies sich als revolutionär:

Falls die beiden hermiteschen Operatoren A und B kommutieren, also $[A, B] = 0$ erfüllen, so lassen sie sich simultan diagonalisieren (M4B).

Andernfalls gilt $[A, B] \neq 0$, und die Unschärferelation S1H besagt, dass beide Messwerte nicht gleichzeitig exakt bestimmt sind.

Diese Unschärfe beruht nicht auf einem technisch behebbaren Mangel eines ungenügenden Messinstruments, sondern ist prinzipieller Natur!

Heisenbergs ursprüngliche Formulierung einer Unschärferelation betraf Ort x und Impuls p eines Teilchens: In geeigneter Darstellung gilt dabei $p = (\hbar/\mathrm{i})\partial_x$, daraus folgt $[x, p] = \mathrm{i}\hbar$, also $\sigma(x) \cdot \sigma(p) \geq \hbar/2$. Das heißt:

Es ist prinzipiell unmöglich, den Ort und den Impuls eines Teilchens gleichzeitig beliebig genau zu messen. Anders gesagt bedeutet das:

Es ist unmöglich, einen quantenmechanischen Zustand herzustellen, bei dem der Ort und der Impuls beliebig genau definiert sind.

Motivierendes Beispiel: die Multiplikation von Matrizen

Unser Vorbild ist die **Multiplikation** von Matrizen passender Größe:

$$\mu : K^{a\times b} \times K^{b\times c} \to K^{a\times c} : (u,v) \mapsto w = u \bullet v, \quad w_{i,k} = \sum_{j=1}^{b} u_{i,j} \cdot v_{j,k}$$

Hier ist $U = K^{a\times b}$ ein linearer Raum über dem Ring $R = K^{a\times a}$ von links und über dem Ring $S = K^{b\times b}$ von rechts, und beide sind kompatibel:

$$(r \bullet u) \bullet s = r \bullet (u \bullet s) \quad \text{für alle } r \in R,\ u \in U,\ s \in S$$

Wir sagen U ist ein (R,S)**–linearer Raum** oder ein (R,S)**–Bimodul**.
Ebenso $V = K^{b\times c}$ über (S,T) und $W = K^{a\times c}$ über (R,T) mit $T = K^{c\times c}$.
Die Produktabbildung μ ist biadditiv, sogar (R,S,T)**–bilinear**:

$$\begin{aligned}
\mu(u+u',v) &= (u+u')\bullet v = (u\bullet v)+(u'\bullet v) &&= \mu(u,v)+\mu(u',v),\\
\mu(u,v+v') &= u\bullet(v+v') = (u\bullet v)+(u\bullet v') &&= \mu(u,v)+\mu(u,v'),\\
\mu(r\bullet u,v) &= (r\bullet u)\bullet v \quad = r\bullet(u\bullet v) &&= r\bullet\mu(u,v),\\
\mu(u,v\bullet t) &= u\bullet(v\bullet t) \quad = (u\bullet v)\bullet t &&= \mu(u,v)\bullet t\\
\mu(u\bullet s,v) &= (u\bullet s)\bullet v \quad = u\bullet(s\bullet v) &&= \mu(u,s\bullet v),
\end{aligned}$$

für alle Vektoren $u,u' \in U$, $v,v' \in V$ und Skalare $r \in R$, $s \in S$, $t \in T$.

Motivierendes Beispiel: die Multiplikation von Matrizen

Wir betrachten Matrizen als zentrales Beispiel und extrahieren daraus die folgenden allgemeinen Begriffe. Die so entstehende Darlegung zu Tensorprodukten wird einerseits recht „einfach und konkret“ verlaufen: Was wir am zentralen Beispiel erkennen, formulieren wir nun allgemein.

Sie können, wenn Sie möchten, die folgenden Definitionen und Sätze auch ohne diese Einbettung in den konkreten Beispielkontext lesen. Das macht logisch keinen Unterschied, psychologisch hingegen schon: Die Darstellung wirkt dann vermutlich eher „abstrakt und schwierig“.

Ich plädiere dafür, wie bei nahezu allen mathematischen Themen, zunächst von einem gut gewählten, konkreten Beispiel auszugehen. Daraus schälen wir dann den allgemeingültigen, wahren Kern heraus. Von diesem höheren Standpunkt schreiten wir dann zu Anwendungen.

Allgemein gilt der Erfahrungsgrundsatz: Ein gut verstandenes Beispiel nützt Ihnen mehr als drei schlecht verstandene Sätze. Umgekehrt gilt: Ein gut verstandener Satz bündelt 1000 Beispiele. Nutzen Sie beides!

Beidseitig lineare Räume aka Bimoduln

Definition S2A: beidseitig linearer Raum aka Bimodul

(1) Seien R, S Ringe. Ein (R,S)**–linearer Raum** $(U, +, \bullet, \bar{\bullet})$ ist linkslinear über R und rechtslinear über S und beide Operationen sind kompatibel:

$$(r \bullet u) \,\bar{\bullet}\, s = r \bullet (u \,\bar{\bullet}\, s) \quad \text{für alle } r \in R,\ u \in U,\ s \in S$$

Wir nennen U auch **linear über** (R,S) oder einen (R,S)**–Bimodul**. Zur Betonung schreiben wir ${}_R U_S$ für den (R,S)–linearen Raum U.

(2) Seien U und U' lineare Räume über (R,S). Eine (R,S)**–lineare Abbildung** $f : U \to U'$ ist linkslinear über R und rechtslinear über S:

$$f(u+v) = f(u) + f(v), \quad f(r \bullet u) = r \bullet f(u), \quad f(u \,\bar{\bullet}\, s) = f(u) \,\bar{\bullet}\, s$$

für alle $u, v \in U$, $r \in R$, $s \in S$. Zusammenfassend schreiben wir:

$$\boxed{\mathrm{Hom}_{(R,S)}(U, U') = \{\, f : U \to U' \text{ linear über } (R,S) \,\}}$$

Es gelten die üblichen Rechenregeln für Homomorphismen (I1G).

Beidseitig lineare Räume aka Bimoduln

Ausführlich haben wir hier zwei Skalarmultiplikationen (I1B):

$$\begin{aligned} \bullet &: R \times U \to U &&: (r,u) \mapsto r \bullet u \\ \bar{\bullet} &: U \times S \to U &&: (u,s) \mapsto u \mathbin{\bar{\bullet}} s \end{aligned}$$

Wir fordern einerseits, dass $(U,+,\bullet)$ linkslinear über $(R,+,\cdot)$ ist:

$$\begin{aligned} r \bullet (u + u') &= (r \bullet u) + (r \bullet u'), & 1_R \bullet u &= u, \\ (r + r') \bullet u &= (r \bullet u) + (r' \bullet u), & (r \cdot r') \bullet u &= r \bullet (r' \bullet u). \end{aligned}$$

Wir fordern andererseits, dass $(U,+,\bar{\bullet})$ rechtslinear über $(S,+,\cdot)$ ist:

$$\begin{aligned} (u + u') \mathbin{\bar{\bullet}} s &= (u \mathbin{\bar{\bullet}} s) + (u' \mathbin{\bar{\bullet}} s), & u \mathbin{\bar{\bullet}} 1_S &= u, \\ u \mathbin{\bar{\bullet}} (s + s') &= (u \mathbin{\bar{\bullet}} s) + (u \mathbin{\bar{\bullet}} s'), & u \mathbin{\bar{\bullet}} (s' \cdot s) &= (u \mathbin{\bar{\bullet}} s') \mathbin{\bar{\bullet}} s. \end{aligned}$$

Dies gelte für alle Skalare $r, r' \in R$, $s, s' \in S$ und alle Vektoren $u, u' \in U$.

Zudem fordern wir die Verträglichkeit $(r \bullet u) \mathbin{\bar{\bullet}} s = r \bullet (u \mathbin{\bar{\bullet}} s)$.

Die Linearität über (R,S) ist demnach äquivalent zur Linkslinearität über dem Ring $R \times S^{\mathrm{op}}$ und ebenso zur Rechtslinearität über $R^{\mathrm{op}} \times S$.

Beispiel S2B: Links-Rechts-Symmetrie

Sei $(R, +, \cdot)$ ein kommutativer Ring und $(U, +, \bullet)$ linkslinear über R mit

$$\bullet : R \times U \to U \;:\; (r, u) \mapsto r \bullet u.$$

Dann ist $(U, +, \bar{\bullet})$ ein rechtslinearer Raum über R mit

$$\bar{\bullet} : U \times R \to U \;:\; (u, s) \mapsto u \,\bar{\bullet}\, s := s \bullet u.$$

Dies ist eine Rechtsoperation, denn für alle $u \in U$ und $s, s' \in R$ gilt

$$(u \,\bar{\bullet}\, s') \,\bar{\bullet}\, s \overset{\text{Def}}{=} s \bullet (s' \bullet u) \overset{\text{lOp}}{=} (s \cdot s') \bullet u \overset{\text{Com}}{=} (s' \cdot s) \bullet u \overset{\text{Def}}{=} u \,\bar{\bullet}\, (s' \cdot s).$$

Damit ist $(U, +, \bullet, \bar{\bullet})$ sogar ein (R, R)–linearer Raum (Bimodul), denn

$$(r \bullet u) \,\bar{\bullet}\, s \overset{\text{Def}}{=} s \bullet (r \bullet u) \overset{\text{lOp}}{=} (s \cdot r) \bullet u \overset{\text{Com}}{=} (r \cdot s) \bullet u \overset{\text{lOp}}{=} r \bullet (s \bullet u) \overset{\text{Def}}{=} r \bullet (u \,\bar{\bullet}\, s).$$

Dasselbe gilt, wenn wir mit einem rechtslinearen Raum beginnen.

Wir nennen die Links-Rechts-Operation mit $r \bullet u = u \,\bar{\bullet}\, r$ **kommutativ**.
Das ist die Standardkonvention, sofern nichts anderes vereinbar wird.

Beidseitig lineare Räume aka Bimoduln

In Beispiel S2B haben wir eine Linksoperation vorausgesetzt mit

$$
\begin{aligned}
r \bullet (u + u') &= (r \bullet u) + (r \bullet u'), & 1_R \bullet u &= u, \\
(r + r') \bullet u &= (r \bullet u) + (r' \bullet u), & (r \cdot r') \bullet u &= r \bullet (r' \bullet u).
\end{aligned}
$$

Wir definieren die Rechtsoperation $u \mathbin{\bar{\bullet}} s := s \bullet u$ und folgern daraus

$$
\begin{aligned}
(u + u') \mathbin{\bar{\bullet}} s &= (u \mathbin{\bar{\bullet}} s) + (u' \mathbin{\bar{\bullet}} s), & u \mathbin{\bar{\bullet}} 1_S &= u, \\
u \mathbin{\bar{\bullet}} (s + s') &= (u \mathbin{\bar{\bullet}} s) + (u \mathbin{\bar{\bullet}} s'), & u \mathbin{\bar{\bullet}} (s' \cdot s) &= (u \mathbin{\bar{\bullet}} s') \mathbin{\bar{\bullet}} s.
\end{aligned}
$$

Die ersten drei sind leicht nachzurechnen: Versuchen Sie es als Übung! Oben habe ich nur die letzte und einzig kritische Gleichung gezeigt, denn hier benötigen wir entscheidend die Kommutativität von R.

Über jedem kommutativen Ring, speziell über jedem Körper, ist es bequem, die Skalare auf beiden Seiten schreiben zu können. Die obige Überprüfung rechtfertigt, dass wir dies problemlos dürfen.

So reiht sich die vertraute kommutative Notation harmonisch in den allgemeinen, nicht-notwendig-kommutativen Kontext ein.

Beidseitig lineare Räume aka Bimoduln

🙂 Es gibt gute Gründe, auch nicht-kommutative Ringe zu betrachten:

Beispiel: Sei K ein Ring. Dann ist $U = K^{p\times q}$ ein (R,S)–linearer Raum über dem Ring $R = K^{p\times p}$ von links und dem Ring $S = K^{q\times q}$ von rechts.

Beispiel: $V = K^{n\times 1}$ ist rechtslinear über K und linkslinear über $K^{n\times n}$. Ebenso ist $K^{1\times n}$ ist linkslinear über K und rechtslinear über $K^{n\times n}$.

🙂 Diese erste Beispielgruppe nutzen wir schon lange erfolgreich in der Matrizenrechnung: Hier unterscheiden wir natürlich links und rechts!

Als Skalare können wir immer den Ring $\mathbb{Z}$ der ganzen Zahlen einsetzen:

Beispiel: Ist U linkslinear über R, so auch $(R,\mathbb{Z})$–linear. (I1K)
Ist R kommutativ, so ist U linear über (R,R) dank $u \bullet r := r \bullet u$. (S2B)

Beispiel: Ist U rechtslinear über R, so auch $(\mathbb{Z},R)$–linear. (I1K)
Ist R kommutativ, so ist U linear über (R,R) dank $r \bullet u := u \bullet r$. (S2B)

🙂 Letzteres nutzen wir ebenso intuitiv: Über einem kommutativen Ring, etwa einem Körper, müssen wir links und rechts nicht unterscheiden.

Beispiel S2C: Links-Rechts-Vertauschung

(1) Seien R, S kommutative Ringe und $(U, +, \bullet, \bullet)$ linear über (R, S). Dann erhalten wir den (S, R)–linearen Raum $\overline{{}_R U_S} = {}_S \overline{U}_R$ mit

$$\begin{aligned} \bar{\bullet} &: S \times U \to U &&: (s, u) \mapsto s \,\bar{\bullet}\, u := u \bullet s, \\ \bar{\bullet} &: U \times R \to U &&: (u, r) \mapsto u \,\bar{\bullet}\, r := r \bullet u. \end{aligned}$$

(2) Allgemein seien Anti-Involutionen (R2H) auf R und S gegeben:

$$^{-} : R \to R : r \mapsto \overline{r} \quad \text{und} \quad ^{-} : S \to S : s \mapsto \overline{s}.$$

Dann erhalten wir den (S, R)–linearen Raum $\overline{{}_R U_S} = {}_S \overline{U}_R$ mit

$$\begin{aligned} \bar{\bullet} &: S \times U \to U &&: (s, u) \mapsto s \,\bar{\bullet}\, u := u \bullet \overline{s}, \\ \bar{\bullet} &: U \times R \to U &&: (u, r) \mapsto u \,\bar{\bullet}\, r := \overline{r} \bullet u. \end{aligned}$$

Für jeden kommutativen Ring ist die Identität eine Anti-Involution. Wir erhalten so die Vertauschung (1) als Spezialfall von (2).

Bilineare Abbildungen und interner Ausgleich

Definition S2D: bilineare Abbildung und interner Ausgleich

Gegeben seien Ringe R, S, T und lineare Räume ${}_R U_S$, ${}_S V_T$, ${}_R W_T$.
Eine Abbildung $B : U \times V \to W$ heißt (R, T)**–bilinear**, falls gilt:

1. Für jedes $v \in V$ ist $U \to W : u \mapsto B(u, v)$ linkslinear über R.
2. Für jedes $u \in U$ ist $V \to W : v \mapsto B(u, v)$ rechtslinear über T.

Wir nennen B **ausgeglichen** über S (engl. *balanced*), falls gilt:

3. $B(u \cdot s, v) = B(u, s \cdot v)$ für alle $u \in U$, $v \in V$, $s \in S$.

Gelten alle drei Bedingungen, so nennen wir B kurz (R, S, T)**–bilinear**:

$$\boxed{\mathrm{Bil}_{(R,S,T)}(U, V; W) = \{\, B : U \times V \to W \text{ bilinear über } (R, S, T) \,\}}$$

Beispiel: Sei R kommutativ und U, V, W lineare Räume über R (S2B).
Ist $B : U \times V \to W$ bilinear über (R, R), dann auch über (R, R, R):

$$\begin{aligned} B(u \cdot r, v) &\overset{\text{Sym}}{=} B(r \cdot u, v) \overset{\text{lLin}}{=} r \cdot B(u, v) \\ &\overset{\text{Sym}}{=} B(u, v) \cdot r \overset{\text{rLin}}{=} B(u, v \cdot r) \overset{\text{Sym}}{=} B(u, r \cdot v) \end{aligned}$$

Bilineare Abbildungen und interner Ausgleich

Beispiel: Die Matrixmultiplikation ist bilinear, wie auf Seite S201 erklärt; das war unser motivierendes Beispiel, das wir als Leitbild voranstellen:

$$\mu \;:\; K^{a\times b} \times K^{b\times c} \to K^{a\times c} \;:\; (u,v) \mapsto w = u \bullet v, \quad w_{i,k} = \textstyle\sum_{j=1}^{b} u_{i,j} \cdot v_{j,k}$$

Hier operieren die Ringe $R = K^{a\times a}$ und $S = K^{b\times b}$ und $T = K^{c\times c}$ auf den Räumen ${}_R U_S = K^{a\times b}$ und ${}_S V_T = K^{b\times c}$ und ${}_R W_T = K^{a\times c}$. Die Produktabbildung $\mu : U \times V \to W$ ist (R,S,T)–bilinear, denn es gilt

$$\begin{aligned} \mu(u+u',v) &= \mu(u,v) + \mu(u',v), & \mu(r \bullet u, v) &= r \bullet \mu(u,v), \\ \mu(u,v+v') &= \mu(u,v) + \mu(u,v'), & \mu(u, v \bullet t) &= \mu(u,v) \bullet t, \\ & & \mu(u \bullet s, v) &= \mu(u, s \bullet v) \end{aligned}$$

für alle Vektoren $u, u' \in U$, $v, v' \in V$ und Skalare $r \in R$, $s \in S$, $t \in T$.

🙂 Wir wollen nun jede bilineare Abbildung $\mu : U \times V \to Z$ linearisieren: Uns genügt *eine* einzige, universelle bilineare Abbildung $\tau : U \times V \to W$, um *jede* bilineare Abbildung $\mu : U \times V \to Z$ auf eine lineare Abbildung $f : W \to Z$ zu reduzieren, sodass $\mu = f \circ \tau$ gilt.

Das Tensorprodukt und seine universelle Eigenschaft

Definition S2E: Tensorprodukt

Gegeben seien Ringe R, S, T und lineare Räume ${}_R U_S$ und ${}_S V_T$.

(1) Ein **Produkt** des Paars (U, V) über (R, S, T) mit Werten in ${}_R Z_T$ ist eine (R, S, T)–bilineare Abbildung $\mu : U \times V \to Z$.

(2) Ein **Tensorprodukt** des Paars (U, V) über (R, S, T) ist ein Produkt $\tau : U \times V \to W$ mit der folgenden universellen Abbildungseigenschaft:

$$\begin{tikzcd}
U \times V \arrow[r, "\tau", "\text{bilinear}"'] \arrow[dr, "\text{bilinear}", "\mu"'] & W \arrow[d, dashed, "\exists! f"', "\text{linear}"] \\
 & Z
\end{tikzcd}$$

Jedes Produkt $\mu : U \times V \to Z$ faktorisiert eindeutig über τ, das heißt es existiert genau eine (R, T)–lineare Abbildung $f : W \to Z$ mit $\mu = f \circ \tau$.

Diese Forderung ist äquivalent zur Bijektivität der natürlichen Abbildung

$$\boxed{\Phi_\tau^Z \;:\; \mathrm{Hom}_{(R,T)}(W, Z) \to \mathrm{Bil}_{(R,S,T)}(U, V; Z) \;:\; f \mapsto \mu = f \circ \tau.}$$

Das Tensorprodukt und seine universelle Eigenschaft

🙂 In anderen Worten bedeutet das: (1) τ ist ein Produkt und (2) jedes Konkurrenzprodukt μ faktorisiert eindeutig über τ.

🙂 Die Bijektivität von Φ_τ^Z bedeutet, dass wir jede bilineare Abbildung $\mu : U \times V \to Z$ in eine lineare Abbildung $f : W \to Z$ umrechnen können.

🙂 In Satz S2H zeigen wir die Eindeutigkeit bis auf Isomorphie. Für *das* Tensorprodukt von U, V über (R, S, T) schreiben wir dann

$$\boxed{\otimes \;:\; U \times V \to U \otimes_S V \;:\; (u, v) \mapsto u \otimes v.}$$

In dieser Schreibweise gilt $\mu(u, v) = f(u \otimes v)$ für alle $u \in U$ und $v \in V$.

⚠ Das Symbol „$\otimes$“ hat hier zwei Bedeutungen: Einerseits ist $u \otimes v$ das Tensorprodukt der beiden Vektoren $u \in U$ und $v \in V$ im Raum $U \otimes_S V$. Andererseits ist $U \otimes_S V$ das Tensorprodukt der beiden Räume U und V.

Wenn der Ring S aus dem Kontext klar ist, müssen wir ihn nicht explizit wiederholen. Erfahrungsgemäß ist dies jedoch anfangs meist hilfreich.

Definition S2E: einfache Tensoren

(3) Für Tensorprodukte vereinbaren wir folgende Sprechweisen.
Jedes Element $w \in U \otimes V$ des Tensorraums nennen wir einen **Tensor**.

Hat $w \in U \otimes V$ die besonders einfache Form $w = u \otimes v$ mit $u \in U$ und $v \in V$, so nennen wir den Tensor w **einfach** oder **elementar** oder **rein**.

Anders gesagt, die Menge der **einfachen Tensoren** ist das Bild der Produktabbildung $\otimes : U \times V \to U \otimes V$. Diese ist i.A. nicht surjektiv!

Der Tensorraum entsteht aus allen **Summen** einfacher Tensoren; solche Summen sind im Allgemeinen keine einfachen Tensoren.

Wir haben dies zuvor bereits im Anschluss an Definition S1B diskutiert und illustriert. Wir werden diese Frage später wiederholt aufgreifen.

Das Tensorprodukt und seine universelle Eigenschaft

☺ Über einem Körper K können wir jedes Tensorprodukt sehr konkret erklären durch die **universelle Basiseigenschaft**, siehe Definition S1B.

☹ Für lineare Räume über beliebigen Ringen verfügen wir meist nicht über Basen, daher ist diese Sichtweise hier nicht mehr hilfreich.

☺ Wir vollziehen daher den nötigen Perspektivwechsel und nutzen zur allgemeinen Definition S2E die **universelle Abbildungseigenschaft**.

⚠ Statt einer expliziten Definition durch Konstruktion (wie etwa in S1A) folgen wir der inneren Logik und den äußeren Zwängen und formulieren eine implizite Definition durch eine charakterisierende Eigenschaft.

Das ist mathematisch gesehen der beste Weg: effizient und elegant. Beim ersten Kontakt ist es ungewohnt und bedarf der Gewöhnung; wie immer gelingt dies am besten durch aktive Einübung.

Die elegante Definition S2E mittels UAE zieht sofort Arbeit nach sich: Existenz und Eindeutigkeit müssen anschließend noch gezeigt werden. Auch mit diesen Konstruktionen üben Sie Verständnis und Anwendung.

Das Tensorprodukt und seine universelle Eigenschaft

Beispiel S2F: Polynommodell

Sei K ein kommutativer Ring. Wir betrachten $U = V = K[X]$ und vergleichen zwei Produkte nach $W = K[X,Y]$ und $Z = K[X]$:

$$\begin{aligned} \tau &: K[X] \times K[X] \to K[X,Y] : (u,v) \mapsto u(X)v(Y). \\ \mu &: K[X] \times K[X] \to K[X] \quad : (u,v) \mapsto u(X)v(X). \end{aligned}$$

Das erste Produkt τ hat die universelle Abbildungseigenschaft (S2G). Wir wenden dies speziell auf das Konkurrenzprodukt μ an:

$$\begin{tikzcd} U \times V \arrow[r, "\tau", "\text{bilinear}"'] \arrow[dr, "\text{bilinear}", "\mu"'] & W \arrow[d, dashed, "\exists! f"', "\text{linear}"] & X^i Y^j \arrow[d, mapsto] \\ & Z & X^{i+j} \end{tikzcd}$$

Das zweite Produkt μ faktorisiert eindeutig über τ gemäß $\mu = f \circ \tau$.

Für jedes solche f gilt $f(X^i Y^j) = f(\tau(X^i, X^j)) = \mu(X^i, X^j) = X^{i+j}$.
Dank PLF (K1B) existiert $f : W \to Z : X^i Y^j \mapsto X^{i+j}$ linear über K.

🙂 Wir rechnen gleich in Satz S2G allgemein nach, dass unser Beispiel $\tau : U \times V \to W$ tatsächlich die universelle Abbildungseigenschaft hat.

Zur Illustration vergleichen wir τ mit dem zweiten Produkt $\mu : U \times V \to Z$ und konstruieren explizit die lineare Abbildung $f : W \to Z$ mit $\mu = f \circ \tau$.

Alle Daten liegen explizit vor, daher gelingt uns diese Rechnung leicht: Sowohl die Eindeutigkeit als auch die Existenz von f sind offensichtlich.

🙂 Hinter diesem schön einfachen Beispiel steht ein nützliches Prinzip über jedem kommutativen Ring, insbesondere über jedem Körper:

Für Räume mit Basis ist die universelle Abbildungseigenschaft S2E des Tensorprodukts äquivalent zur universellen Basiseigenschaft S1B.

Der folgende Satz formuliert dies detailliert aus.

Das Prinzip der bilinearen Fortsetzung

Satz S2G: Prinzip der bilinearen Fortsetzung

Gegeben seien Ringe R, T sowie lineare Räume ${}_R U$ und V_T und ${}_R W_T$.

(1) Gegeben seien Basen $(u_i)_{i \in I}$ von U und $(v_j)_{j \in J}$ von V sowie eine beliebige Familie $(w_{i,j})_{(i,j) \in I \times J}$ in W, die Gram–Matrix. Dann existiert genau eine bilineare Abbildung $\tau : U \times V \to W : (u_i, v_j) \mapsto w_{i,j}$, nämlich

$$\boxed{\tau\Big(\sum_{i \in I} r_i u_i, \sum_{j \in J} v_j t_j \Big) = \sum_{(i,j) \in I \times J} r_i \, w_{i,j} \, t_j}$$

für alle Linearkombinationen mit Koeffizienten $r \in R^{(I)}$ und $t \in T^{(J)}$.

(2) Sei zudem $R = T$ kommutativ und U, V, W linear über R (S2B). Genau dann ist $\tau : U \times V \to W$ ein Tensorprodukt, erfüllt also die universelle Eigenschaft S2E, wenn $(w_{i,j})_{(i,j) \in I \times J}$ eine Basis von W ist.

⚠ Warnung vor Missverständnis: Ist $\tau : U \times V \to W$ bilinear, so bildet τ die Basisvektoren $(u_i, 0)$ und $(0, v_j)$ von $U \times V$ alle auf 0 in W ab.

Das Prinzip der bilinearen Fortsetzung

Beweis: (1a) Eindeutigkeit: Sind $\tau, \tau' : U \times V \to W : (u_i, v_j) \mapsto w_{i,j}$ bilinear, so folgt $\tau = \tau'$, denn für alle $u = \sum_i r_i u_i$ und $v = \sum_j v_j t_j$ gilt:

$$\begin{aligned}
\tau\Big(\sum_i r_i u_i, \sum_j v_j t_j\Big) &\overset{\text{Bil}}{=} \sum_i \sum_j r_i \tau(u_i, v_j) t_j \overset{\text{Vor}}{=} \sum_{i,j} r_i \, w_{i,j} \, t_j \\
&\overset{\text{Vor}}{=} \sum_i \sum_j r_i \tau'(u_i, v_j) t_j \overset{\text{Bil}}{=} \tau'\Big(\sum_i r_i u_i, \sum_j v_j t_j\Big)
\end{aligned}$$

(1b) Existenz: Diese Formel definiert eine Abbildung $\tau : U \times V \to W$, sie erfüllt $(u_i, v_j) \mapsto w_{i,j}$ für alle $(i, j) \in I \times J$ und ist (R, T)–bilinear.

(2a) „$\Leftarrow$“: Sei $(w_{i,j})_{(i,j)\in I\times J}$ eine Basis von W. Sei $\mu : U \times V \to Z$ bilinear. Dank PLF (K1B) existiert genau eine R–lineare Abbildung $f : W \to Z$ mit $f(w_{i,j}) = \mu(u_i, v_j)$. Dank (1a) gilt dann $\mu = f \circ \tau$.

(2b) „$\Rightarrow$“: Angenommen $\tau : U \times V \to W$ erfüllt die UAE (S2E). Der R–lineare Raum $W' = R^{(I\times J)}$ hat die R–Basis $(e_{i,j})_{(i,j)\in I\times J}$. Dank (1b) existiert $\tau' : U \times V \to W' : (u_i, v_j) \mapsto e_{i,j}$ bilinear über R. Dank (2a) ist τ' ein Tensorprodukt, erfüllt also ebenfalls die UAE.

Der folgende Satz konstruiert den R–Isomorphismus $f' : W' \xrightarrow{\sim} W$ mit $\tau = f' \circ \tau'$. Somit ist $(w_{i,j} = f'(e_{i,j}))_{(i,j)\in I\times J}$ eine R–Basis von W. $\boxed{\text{QED}}$

Das Tensorprodukt ist eindeutig bis auf Isomorphie.

$$
\begin{tikzcd}
 & W \arrow[dd, shift right=2, "\exists! f"'] \arrow[dd, phantom, "\cong"] \\
U \times V \arrow[ur, "\tau"] \arrow[dr, "\tau'"'] & \\
 & W' \arrow[uu, shift right=2, "\exists! f'"']
\end{tikzcd}
$$

Eindeutigkeit trotz Wahlfreiheit: Alice und Bob konstruieren ihre Tensorprodukte wie sie mögen. Je zwei Tensorprodukte lassen sich eindeutig übersetzen.

Satz S2H: Eindeutigkeit des Tensorprodukts

Gegeben seien Ringe R, S, T sowie lineare Räume ${}_R U_S$ und ${}_S V_T$.

Dann sind je zwei Tensorprodukte $\tau : U \times V \to W$ und $\tau' : U \times V \to W'$ eindeutig isomorph: Es existiert genau ein (R,T)–Isomorphismus $(f, f') : W \cong W'$, für den $\tau' = f \circ \tau$ und $\tau = f' \circ \tau'$ gilt.

Für *das* Tensorprodukt von U, V über (R, S, T) schreiben wir fortan

$$
\boxed{\otimes \;:\; U \times V \to U \otimes_S V \;:\; (u, v) \mapsto u \otimes v.}
$$

Wir schreiben kurz $U \otimes V$, falls der Ring S aus dem Kontext hervorgeht.

Das Tensorprodukt ist eindeutig bis auf Isomorphie.

Beweis: Wir konstruieren (f, f') durch vierfache Anwendung der UAE:

$$
\begin{tikzcd}
 & W \arrow[dd, shift right, "\exists! f"'] \\
U \times V \arrow[ur, "\tau"] \arrow[dr, "\tau'"'] & \\
 & W' \arrow[uu, shift right, "\exists! f'"']
\end{tikzcd}
\qquad
\begin{tikzcd}
 & W \arrow[dd, shift right, "\mathrm{id}_W"'] \arrow[dd, shift left, "g"] \\
U \times V \arrow[ur, "\tau"] \arrow[dr, "\tau"'] & \\
 & W
\end{tikzcd}
\qquad
\begin{tikzcd}
 & W' \\
U \times V \arrow[ur, "\tau'"] \arrow[dr, "\tau'"'] & \\
 & W' \arrow[uu, shift left, "\mathrm{id}_{W'}"] \arrow[uu, shift right, "g'"']
\end{tikzcd}
$$

(a) Dank der UAE des Tensorprodukts τ angewendet auf τ' existiert genau eine (R, T)–lineare Abbildung $f : W \to W'$ mit $\tau' = f \circ \tau$.

(b) Dank der UAE des Tensorprodukts τ' angewendet auf τ existiert genau eine (R, T)–lineare Abbildung $f' : W' \to W$ mit $\tau = f' \circ \tau'$.

(c) Für $g = f' \circ f : W \to W$ gilt $g \circ \tau = \tau = \mathrm{id}_W \circ \tau$, also $g = \mathrm{id}_W$ dank Eindeutigkeit / UAE des Tensorprodukts τ angewendet auf τ.

(d) Für $g' = f \circ f' : W' \to W'$ gilt $g' \circ \tau' = \tau' = \mathrm{id}_{W'} \circ \tau'$, also $g' = \mathrm{id}_{W'}$ dank Eindeutigkeit / UAE des Tensorprodukts τ' angewendet auf τ'.

Somit gilt $f' \circ f = \mathrm{id}_W$ und $f \circ f' = \mathrm{id}_{W'}$, also $(f, f') : W \cong W'$. $\boxed{\text{QED}}$

Erinnerung: die Rechenregeln für Tensoren

Gegeben seien Ringe R, S, T und lineare Räume ${}_R U_S$ und ${}_S V_T$.

Wir wissen, dass das Tensorprodukt $U \otimes V$ eindeutig ist, genauer: Je zwei Tensorprodukte sind eindeutig isomorph. Wir wollen nun die Existenz eines solchen Tensorprodukts zeigen. Hierzu müssen wir eine Konstruktion durchführen. Zuvor erinnern wir uns nochmal an unser Ziel:

(1) Wir nennen $\otimes : U \times V \to W$ ein **Produkt**, falls gilt:

$$
\begin{aligned}
(u + u') \otimes v &= u \otimes v + u' \otimes v, & \qquad (r \bullet u) \otimes v &= r \bullet (u \otimes v), \\
u \otimes (v + v') &= u \otimes v + u \otimes v', & \qquad u \otimes (v \bullet t) &= (u \otimes v) \bullet t, \\
(u \bullet s) \otimes v &= u \otimes (s \bullet v)
\end{aligned}
$$

für alle Vektoren $u, u' \in U$, $v, v' \in V$ und Skalare $r \in R$, $s \in S$, $t \in T$.

(2) Für ein **Tensorprodukt** fordern wir genau diese Eigenschaften, doch sonst keine weiteren, also überflüssigen Relationen (UAE).

🙂 Das ist die Forderung der universellen Abbildungseigenschaft! Genau so wollen wir daher in der folgenden Konstruktion vorgehen.

Erinnerung: die Rechenregeln für Tensoren

Satz S2I: Existenz des Tensorprodukts

Zu (U, V) über (R, S, T) existiert ein Tensorprodukt $\tau : U \times V \to W$.

Dieser abstrakte Existenzsatz ist zwar beruhigend, doch für sich alleine noch längst nicht so hilfreich, wie wir es gerne hätten und benötigen.

Hermann Weyl schrieb hierzu in seinem Artikel *Über die neue Grundlagenkrise der Mathematik* 1921 die weisen Worte:

Ein Existenzsatz verkündet
„das Vorhandensein eines Schatzes,
ohne jedoch zu verraten, an welchem Ort. [...]
Nicht das Existenztheorem ist das Wertvolle,
sondern die im Beweise geführte Konstruktion.“

☺ Die folgende Konstruktion erlaubt uns, den Schatz zu heben! Sie liefert eine explizite Präsentation durch Erzeuger und Relationen, und erlaubt uns damit eine genauere Untersuchung der Feinstruktur.

Existenz des Tensorprodukts durch explizite Konstruktion

Satz S2I: Existenz des Tensorprodukts

Zu (U,V) über (R,S,T) existiert ein Tensorprodukt $\tau : U \times V \to W$.

Wir betrachten die Menge $U \times V$ aller Paare (u,v) mit $u \in U$ und $v \in V$. Sei $F = \mathbb{Z}^{(U\times V)}$ der $\mathbb{Z}$–lineare Raum mit Basis $U \times V$: Jedes Element $\sigma \in F$ ist eine $\mathbb{Z}$–Linearkombination $\sigma = \sigma_1 \cdot (u_1, v_1) + \dots + \sigma_n \cdot (u_n, v_n)$ von Paaren $(u_i, v_i) \in U \times V$ mit Koeffizienten $\sigma_i \in \mathbb{Z}$. Hierin setzen wir

$$G := \left\langle \begin{array}{l|l} (u+u',v)-(u,v)-(u',v) & u,u' \in U \\ (u,v+v')-(u,v)-(u,v') & v,v' \in V \\ (u \bullet s, v)-(u, s \bullet v) & s \in S \end{array} \right\rangle_{\mathbb{Z}} \le F.$$

(0) Wir erhalten die Quotientenabbildung $q : F \twoheadrightarrow W := F/G : \sigma \mapsto [\sigma]$.
(1) Die abelsche Gruppe W wird zu einem (R,T)–linearen Raum durch

$$\begin{aligned} \bullet &: R \times W \to W \;:\; r \bullet [(u,v)] = [(r \bullet u, v)], \\ \bullet &: W \times T \to W \;:\; [(u,v)] \bullet t = [(u, v \bullet t)]. \end{aligned}$$

(2) Damit ist $\tau : U \times V \to W : (u,v) \mapsto u \otimes v := [(u,v)]$ ein Produkt
(3) und erfüllt zudem die universelle Abbildungseigenschaft S2E.

Existenz des Tensorprodukts durch explizite Konstruktion

🙂 Wir präsentieren hier $W = F/G$ durch **Erzeuger** und **Relationen**: Dazu beginnen wir mit einem freien Objekt F, ohne jegliche Relationen, und erzwingen die gewünschten Relationen G durch Quotientenbildung.

⚠ Wir müssen nun noch sorgsam und geduldig nachrechnen, dass diese Konstruktion das gewünschte Ziel tatsächlich erreicht. Genau dies ist Gegenstand des folgenden Beweises.

🙂 Das ist eigentlich Routinearbeit – nachdem man es ein paar Mal durchexerziert hat. In diesem Zusammenhang schrieb Hermann Weyl:

Nicht im Beweis einer gegebenen Konstruktion,
sondern in der Erfindung der Konstruktion liegt
in den meisten Fällen die eigentliche Schwierigkeit.

Da dieser routinierte Nachweis für Sie eine der ersten Gelegenheiten ist, und Sie genau diese Sorgfalt und Routine erst noch erlernen sollen, möchte ich hier ein gutes Vorbild geben und dies für Sie vorführen. Zudem können wir so Verständnisschwierigkeiten klären.

Lösung: (1) Auf $F = \mathbb{Z}^{U \times V}$ haben wir die Linksoperation

$$\bullet : R \times F \to F \;:\; r \bullet (u, v) = (r \bullet u, v).$$

Für jeden Skalar $r \in R$ gilt $r \bullet G \subseteq G$, denn dies gilt auf den Erzeugern:

$$\begin{aligned} r \bullet [(u + u', v) - (u, v) - (u', v)] &= (r \bullet u + r \bullet u', v) - (r \bullet u, v) - (r \bullet u', v) \\ r \bullet [(u, v + v') - (u, v) - (u, v')] &= (r \bullet u, v + v') - (r \bullet u, v) - (r \bullet u, v') \\ r \bullet [(u \bullet s, v) - (u, s \bullet v)] &= ((r \bullet u) \bullet s, v) - ((r \bullet u), s \bullet v) \end{aligned}$$

Somit erhalten wir auf $W = F/G$ die wohldefinierte Linksoperation

$$\bullet : R \times W \to W \;:\; r \bullet [(u, v)] = [(r \bullet u, v)].$$

Ebenso operiert der Ring T von rechts auf $W = F/G$ vermöge

$$\bullet : W \times T \to W \;:\; [(u, v)] \bullet t = [(u, v \bullet t)].$$

Damit ist W ein (R, T)–Modul: Rechnen Sie es nach!

Existenz des Tensorprodukts durch explizite Konstruktion

(1a) Die Linksoperation $\bullet : R \times F \to F : r \bullet (u,v) = (r \bullet u, v)$ erfüllt

$$1_R \bullet (u,v) = (u,v) \quad \text{und} \quad (r \cdot r') \bullet (u,v) = r \bullet (r' \bullet (u,v)).$$

Demnach folgt für $\bullet : R \times W \to W$ ebenso

$$1_R \bullet [(u,v)] = [(u,v)] \quad \text{und} \quad (r \cdot r') \bullet [(u,v)] = r \bullet (r' \bullet [(u,v)]).$$

Die Operation ist zudem links-distributiv auf F, also auch auf W:

$$r \bullet [\sigma + \sigma'] = [r \bullet \sigma + r \bullet \sigma'] = [r \bullet \sigma] + [r \bullet \sigma'] = r \bullet [\sigma] + r \bullet [\sigma']$$

Die Operation ist *nicht* rechts-distributiv auf F, sondern erst auf W:

$$\begin{aligned}(r + r') \bullet [(u,v)] &= [(r \bullet u + r' \bullet u, v)] \\ &= [(r \bullet u, v)] + [(r' \bullet u, v)] = r \bullet [(u,v)] + r' \bullet [(u,v)]\end{aligned}$$

(1b) Für die Rechtsoperationen $\bullet : F \times T \to F : (u,v) \bullet t = (u, v \bullet t)$ und $\bullet : W \times T \to W : [(u,v)] \bullet t = [(u, v \bullet t)]$ gilt entsprechend dasselbe.

(1c) Beide sind kompatibel, offensichtlich auf F, somit auch auf W.

Existenz des Tensorprodukts durch explizite Konstruktion

(2) Unser $\tau : U \times V \to W : (u,v) \mapsto u \otimes v := [(u,v)]$ ist ein Produkt:
Für alle $u, u' \in U$, $v, v' \in V$, $r \in R$, $s \in S$, $t \in T$ gilt nach Konstruktion

$$\begin{aligned} (u+u') \otimes v &= u \otimes v + u' \otimes v, & (r \bullet u) \otimes v &= r \bullet (u \otimes v), \\ u \otimes (v+v') &= u \otimes v + u \otimes v', & u \otimes (v \bullet t) &= (u \otimes v) \bullet t, \\ (u \bullet s) \otimes v &= u \otimes (s \bullet v). \end{aligned}$$

(3) Unser Produkt τ hat die universelle Abbildungseigenschaft S2E:

$$\begin{tikzcd} U \times V \arrow[rr, "\tau", "\text{bilinear}"'] \arrow[drr, "\text{bilinear}", "\mu"'] & & W \arrow[d, dashed, "\exists! f"', "\text{linear}"] \\ & & Z \end{tikzcd}$$

Sei $\mu : U \times V \to Z$ ein weiteres Produkt. Wir zeigen, dass genau eine (R,T)–lineare Abbildung $f : W \to Z$ mit $\mu = f \circ \tau$ existiert.

Existenz des Tensorprodukts durch explizite Konstruktion

(3a) Eindeutigkeit: Vorgelegt seien $f, g : W \to Z$ mit $\mu = f \circ \tau = g \circ \tau$, also $f(u \otimes v) = f(\tau(u,v)) = g(\tau(u,v)) = g(u \otimes v)$ für $(u,v) \in U \times V$. Daraus folgt $f = g$, denn W wird erzeugt von $(u \otimes v)_{(u,v) \in U \times V}$ (K1A).

$$\begin{tikzcd}
 & F \arrow[dr, two heads, "q"] \arrow[dd, "\exists! p"] & \\
U \times V \arrow[ur, "\iota : (u,v) \mapsto (u,v)"] \arrow[rr, "\tau"] \arrow[dr, "\mu"'] & & W \arrow[dl, "\exists! f"] \\
 & Z &
\end{tikzcd}$$

(3b) Wir nutzen $F = \mathbb{Z}^{(U \times V)}$ mit Basis $U \times V$ über $\mathbb{Z}$. Dank PLF (K1B) existiert genau eine $\mathbb{Z}$–lineare Abbildung $p : F \to Z : (u,v) \mapsto \mu(u,v)$. Da μ ein Produkt ist, gilt $\ker(p) \supseteq G$. Demnach induziert p die $\mathbb{Z}$–lineare Abbildung $f : W \to Z$ mit $p = f \circ q$ (I2E). Somit gilt $f(u \otimes v) = \mu(u,v)$ für alle $u \in U$ und $v \in V$. Insbesondere ist f damit linear über (R,T), denn $f(r \cdot (u \otimes v) \cdot t) = \mu(r \cdot u, v \cdot t) = r \cdot \mu(u,v) \cdot t = r \cdot f(u \otimes v) \cdot t$. Homogenität gilt für alle einfachen Tensoren, also für alle Tensoren.

Die einfachen Tensoren erzeugen den Tensorraum.

Korollar S2J: Einfache Tensoren erzeugen alle Tensoren.

Vorgelegt seien Ringe R, S, T sowie lineare Räume ${}_R U_S$ und ${}_S V_T$.

(1) Die einfachen Tensoren $u \otimes v$ erzeugen den Tensorraum $U \otimes_S V$.

Jeder Tensor $w \in U \otimes_S V$ ist eine endliche Summe $w = \sum_{k=1}^{\ell} u_k \otimes v_k$ einer Länge $\ell \in \mathbb{N}$ mit Faktoren $u_1, \dots, u_\ell \in U$ und $v_1, \dots, v_\ell \in V$.

Beweis: (1) Dies folgt aus der Konstruktion von Satz S2I. $\boxed{\text{QED}}$

⚠ Diese Darstellung von w ist im Allgemeinen nicht eindeutig!

⚠ Nicht einmal die Länge ℓ dieser Summe ist eindeutig!

$$\begin{aligned} w = u \otimes v &= (u_1 + u_2) \otimes (v_1 + v_2) \\ &= (u_1 \otimes v_1) + (u_1 \otimes v_2) + (u_2 \otimes v_1) + (u_2 \otimes v_2). \end{aligned}$$

☺ Minimale Länge $\ell = 0$ entspricht dem Nullelement $w = 0$ in $U \otimes_S V$.

☺ Minimale Länge $\ell = 1$ entspricht einfachen Tensoren $w = u \otimes v \neq 0$.

Zu $w \in U \otimes V$ ist der **Tensorrang** $\operatorname{rang}(w)$ die Minimale Länge $\ell \in \mathbb{N}$ aller Summendarstellungen $w = \sum_{k=1}^{\ell} u_k \otimes v_k$ durch einfache Tensoren.

Die einfachen Tensoren erzeugen den Tensorraum.

Korollar S2J: Einfache Tensoren erzeugen alle Tensoren.

(2) Angenommen, $(u_i)_{i\in I}$ erzeugt U über R und $(v_j)_{j\in J}$ erzeugt V über T, dann wird der Tensorraum $U\otimes_S V$ über (R,T) erzeugt von

$$(u_i\otimes v_i)_{(i,j)\in I\times J}.$$

Jeder Tensor $w\in U\otimes_S V$ ist eine Summe $w=\sum_{k=1}^{\ell} r_k u_{i_k}\otimes v_{j_k} t_k$ mit endlicher Länge $\ell\in\mathbb{N}$ sowie $r_1,\dots,r_\ell\in R$ und $t_1,\dots,t_\ell\in T$.

Beweis: (2) Gemäß (1) ist jeder Tensor $w\in U\otimes_S V$ eine Summe einfacher Tensoren $u\otimes v$. Dabei gilt $u=\sum_{m=1}^{M} r_m u_{i_m}$ mit $r_m\in R$ und $v=\sum_{n=1}^{N} v_{j_n} t_n$ mit $t_n\in T$, also $u\otimes v=\sum_{m,n} r_m u_{i_m}\otimes v_{j_n} t_n$. QED

☺ So konstruieren wir geeignete Erzeugendensysteme von $U\otimes_S V$, am besten möglichst kurz und somit effizient für unsere Rechnungen.

☺ In günstigen Fällen erhalten wir sogar eine Basis (S2G), insbesondere über einem kommutativen Ring $R=S=T$.

Universelle Eigenschaft, Eindeutigkeit und Existenz

Das Tensorprodukt illustriert ein allgemein sehr nützliches Vorgehen:

1. Die universelle Eigenschaft S2E präzisiert, was wir eigentlich wollen.
2. Die Eindeutigkeit S2H zeigt, dass unser Ziel damit charakterisiert ist.
3. Die Konstruktion S2I zeigt, dass / wie unser Wunsch erfüllbar ist.

Damit verfügen wir über zwei sich ergänzende Werkzeuge:

- ☺ Die Präsentation F/G ist konkret, leider meist schwer zu nutzen.
- ☺ Die universelle Eigenschaft ist abstrakt, dafür meist leicht zu nutzen.

Mathematische Objekte (wie lineare Räume, . . .) lassen sich meist am besten verstehen und nutzen, wenn wir beide Sichtweisen anwenden: intern, konkret, Rechnen mit Elementen (Basen, Matrizen, . . .) und extern, abstrakt, mit Abbildungen (Homomorphismen, PLF, . . .).

Je nach Anwendung nutzen wir mehr die eine oder die andere, meist jedoch benötigen Sie beide Sichtweisen für die erfolgreiche Lösung.

Universelle Eigenschaft, Eindeutigkeit und Existenz

Mathematische Objekte betrachten wir oft auf diese zwei Arten, da sie zueinander komplementär sind und sich wunderbar ergänzen. Das gilt nicht nur, aber ganz besonders für algebraische Objekte, wie Monoide und Gruppen, Ringe und Körper, oder hier lineare Räume:

Menge X mit Struktur	**Objekt X mit Morphismen**
interne Beschaffenheit von X	externe Eigenschaften von X
Beziehung zwischen Elementen	Beziehung zu anderen Objekten
Datenstruktur, Implementierung	Verhalten, Semantik, Axiome
Was ist X? Wie ist X aufgebaut?	Was tut X? Was leistet X für uns?
konkret und theoretisch	abstrakt und angewandt
explizite Bauanleitung	bequeme Bedienungsanleitung
Erzeuger und Relationen	universelle Eigenschaft
Konstruktion $\Rightarrow$ Existenz	Definition $\Rightarrow$ Eindeutigkeit

Assoziativität des Tensorprodukts

Satz S2K: Assoziativität des Tensorprodukts

Gegeben seien drei lineare Räume ${}_R U_S$ und ${}_S V_T$ und ${}_T W_Q$.

(0) Genau wie für zwei Faktoren (U, V) bzw. (V, W) erklären wir auch für drei Faktoren (U, V, W) das Tensorprodukt:

$$\tau \,:\, U \times V \times W \to U \otimes_S V \otimes_T W$$

(1) Hierzu existieren eindeutig die natürlichen (R, Q)–Isomorphismen

$$\begin{array}{ccccc} (U \otimes_S V) \otimes_T W & \cong & U \otimes_S V \otimes_T W & \cong & U \otimes_S (V \otimes_T W), \\ (u \otimes v) \otimes w & \rightleftarrows & u \otimes v \otimes w & \rightleftarrows & u \otimes (v \otimes w). \end{array}$$

Die *Eindeutigkeit* ist klar (K1A), denn die einfachen Tensoren erzeugen den Tensorraum (S2J). Die *Existenz* der ersehnten Abbildungen muss jedoch sorgsam nachgewiesen werden, und zwar durch Konstruktion!

⚠ Wir können nicht einfach irgend etwas auf den einfachen Tensoren vorschreiben, sondern müssen zudem auch alle Relationen erfüllen. Ein drastisches Gegenbeispiel ist das No-Cloning-Theorem S1E!

Assoziativität des Tensorprodukts

Aufgabe: Beweisen Sie die Existenz durch Konstruktion dank UAE!

$$
\begin{tikzcd}
(U \times V) \times W \arrow[rr, shift left, "\tilde\alpha_\ell"] \arrow[d, "\tau_{U,V} \times \mathrm{id}_W"'] \arrow[ddrr, "q"' near end] & & U \times V \times W \arrow[ll, shift left, "\tilde\alpha'_\ell"] \arrow[rr, shift right, "\tilde\alpha'_r"'] \arrow[ddll, "p" near start] \arrow[dd, "\tau_{U,V,W}"] \arrow[ddrr, "p" near start] & & U \times (V \times W) \arrow[ll, shift right, "\tilde\alpha_r"'] \arrow[d, "\mathrm{id}_U \times \tau_{V,W}"] \arrow[ddll, "q" near end] \\
(U \otimes_S V) \times W \arrow[d, "\tau_{U \otimes V,W}"'] \arrow[drr, "r"] & & & & U \times (V \otimes_T W) \arrow[d, "\tau_{U,V \otimes W}"] \arrow[dll, "r"'] \\
(U \otimes_S V) \otimes_T W \arrow[rr, shift left, "\alpha_\ell"] & & U \otimes_S V \otimes_T W \arrow[ll, shift left, "\alpha'_\ell"] \arrow[rr, shift right, "\alpha'_r"'] & & U \otimes_S (V \otimes_T W) \arrow[ll, shift right, "\alpha_r"']
\end{tikzcd}
$$

Wir haben $(\tilde\alpha_\ell, \tilde\alpha'_\ell) : (U \times V) \times W \cong U \times V \times W : ((u,v),w) \leftrightarrows (u,v,w)$.
Die Komposition $p = \tau_{U\otimes V,W} \circ (\tau_{U,V} \times \mathrm{id}_W) \circ \tilde\alpha'$ ist (R,S,T,Q)–trilinear, induziert dank UAE also $\alpha'_\ell : u \otimes v \otimes w \mapsto (u \otimes v) \otimes w$, wie ersehnt.

Die Komposition $q = \tau_{U,V,W} \circ \tilde\alpha$ ist (R,S,T)–bilinear in $U \times V$, induziert dank UAE also r. Letzteres ist (R,T,Q)–linear, induziert dank UAE also $\alpha_\ell : (u \otimes v) \otimes w \mapsto u \otimes v \otimes w$, wie ersehnt. Demnach gilt $\alpha'_\ell \circ \alpha_\ell = \mathrm{id}$ und $\alpha_\ell \circ \alpha'_\ell = \mathrm{id}$ auf den einfachen Tensoren, also auf allen Tensoren (S2J).

Assoziativität des Tensorprodukts

Ausführung

Assoziativität des Tensorprodukts

Wir diskutieren vertraute Rechenregeln, hier für das Tensorprodukt: Assoziativität, Kommutativität, Neutrales und Distributivität. Unter diesen ist die Assoziativität die aufwändigste; wir führen sie detailliert aus, die weiteren Rechnungen sind dann leichtere Übungen.

Kommutativität des Tensorprodukts

Satz S2L: Kommutativität des Tensorprodukts

Gegeben seien Ringe R, S, T und lineare Räume ${}_R U_S$ und ${}_S V_T$.

(1) Ist $R = S = T$ kommutativ, so haben wir den Isomorphismus

$$U \otimes_S V \cong V \otimes_S U \;:\; u \otimes v \leftrightarrows v \otimes u.$$

(2) Zu gegebenen Anti-Involutionen auf R, S, T haben wir ebenso

$$\overline{U \otimes_S V} \cong \overline{V} \otimes_S \overline{U} \;:\; u \otimes v \leftrightarrows v \otimes u.$$

Typische Anwendungsbeispiele dieses allgemeineren Falls sind die Konjugation auf $\mathbb{C}$ oder $\mathbb{H}$ oder die Transposition auf Matrizen.

Aufgabe: Konstruieren Sie diese Isomorphismen dank UAE.

Kommutativität des Tensorprodukts

Lösung: Es genügt statt (1) gleich den allgemeineren Fall (2) zu zeigen. Die Konstruktion verläuft von oben nach unten in folgendem Diagramm:

$$
\begin{tikzcd}
\overline{U \times V} \arrow[r, shift left, "\tilde{\sigma}"] \arrow[d, "\tau_{U,V}"] \arrow[dr, "q" near start] & \overline{V} \times \overline{U} \arrow[l, shift left, "\tilde{\sigma}'"] \arrow[d, "\tau_{\overline{V},\overline{U}}"] \arrow[dl, "p"' near start] \\
\overline{U \otimes_S V} \arrow[r, shift left, "\sigma"] & \overline{V} \otimes_S \overline{U} \arrow[l, shift left, "\sigma'"]
\end{tikzcd}
$$

Neutrales zum Tensorprodukt

Satz S2M: Neutrales zum Tensorprodukt

Gegeben seien Ringe R, S, T und lineare Räume ${}_RU_S$ und ${}_SV_T$.
Dann wirkt der Raum ${}_SS_S$ neutral bezüglich des Tensorprodukts:

$$U \otimes_S S \cong U \;:\; u \otimes s \leftrightarrows u \cdot s$$

$$S \otimes_S V \cong V \;:\; s \otimes v \leftrightarrows s \cdot v$$

Aufgabe: Konstruieren Sie diese Isomorphismen dank UAE.

Neutrales zum Tensorprodukt

Lösung: Die Konstruktion verläuft von oben nach unten:

$$
\begin{tikzcd}
U \times S \arrow[d, "\tau_{U,S}"'] \arrow[dr, "\tilde\mu : (u,s) \mapsto u \bullet s"] & \\
U \otimes_S S \arrow[r, shift left, "\mu"] & U \arrow[l, shift left, "\mu'"]
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
& S \times V \arrow[d, "\tau_{S,V}"] \arrow[dl, "s \bullet v \mapsfrom (s,v) : \tilde\mu"'] \\
V \arrow[r, shift right, "\mu'"'] & S \otimes_S V \arrow[l, shift right, "\mu"']
\end{tikzcd}
$$

Zum (R,S)–linearen Raum U ist die Skalarmultiplikation von rechts $\tilde\mu : U \times S \to U : (u,s) \mapsto u \bullet s$ bilinear über (R,S,S). Dank UAE S2E induziert sie die (R,S)–lineare Abbildung $\mu : U \otimes_S S : u \otimes s \mapsto u \bullet s$.

Umgekehrt ist auch $\mu' : U \to U \otimes_S S : u \mapsto u \otimes 1_S$ linear über (R,S), denn es gilt $\mu'(r \bullet u) = (r \bullet u) \otimes 1_S = r \bullet (u \otimes 1_S) = r \bullet \mu'(u)$ und $\mu'(u \bullet s) = (u \bullet s) \otimes 1_S = u \otimes s = \mu'(u) \bullet s$ für $u \in U$, $r \in R$, $s \in S$.

Nach Konstruktion gilt $\mu' \circ \mu = \mathrm{id}$ und $\mu \circ \mu' = \mathrm{id}$ auf den einfachen Tensoren, also auf allen Tensoren (S2J).

Distributivität des Tensorprodukts

Satz S2N: Distributivität des Tensorprodukts

Gegeben seien Ringe R, S, T und lineare Räume ${}_R(U_i)_S$ und ${}_S(V_j)_T$.

(1) Dann ist das Tensorprodukt distributiv über die direkte Summe:

$$(U_1 \oplus U_2) \otimes V \cong (U_1 \otimes V) \oplus (U_2 \otimes V)$$
$$U \otimes (V_1 \oplus V_2) \cong (U \otimes V_1) \oplus (U \otimes V_2)$$

(2) Dies gilt ganz allgemein sogar für beliebige direkte Summen:

$$\Big(\bigoplus\nolimits_{i\in I} U_i\Big) \otimes \Big(\bigoplus\nolimits_{j\in J} V_j\Big) \quad\cong\quad \bigoplus\nolimits_{(i,j)\in I\times J}(U_i \otimes V_j)$$
$$\sum\nolimits_{(i,j)\in I\times J}(u_i \otimes v_j) \quad\leftrightarrows\quad \sum\nolimits_{(i,j)\in I\times J}(u_i \otimes v_j)$$

Alle Tensorprodukte werden hier über dem mittleren Ring S gebildet. Ich schreibe statt $\otimes_S$ lieber $\otimes$, um die Symmetrie zu $\oplus$ zu betonen.

Distributivität des Tensorprodukts

Eigenschaften des Tensorprodukts

Eigenschaften der direkten Summe

Übung: Auch die direkte Summe $(U, V) \mapsto U \oplus V$ ist bis auf Isomorphie assoziativ und kommutativ, und der Nullraum 0 ist beidseitig neutral.

Eigenschaften der direkten Summe

Eigenschaften der direkten Summe

Funktorialität des Tensorprodukts

Satz S2O: Funktorialität

(0) Sei $f : U \to U'$ linear über (R,S) und $g : V \to V'$ linear über (S,T). Dann induzieren f und g eine eindeutige (R,T)–lineare Abbildung

$$f \otimes g \;:\; U \otimes_S V \to U' \otimes_S V' \;:\; u \otimes v \mapsto f(u) \otimes g(v).$$

Beweis: Die Konstruktion nutzt die universelle Eigenschaft S2E:

$$\begin{tikzcd}
U \times V \arrow[r, "h : (u,v) \mapsto (f(u), g(v))"] \arrow[d, "\tau : (u,v) \mapsto u \otimes v"] \arrow[dr, "\tau' \circ h : (u,v) \mapsto f(u) \otimes g(v)"] & U' \times V' \arrow[d, "\tau' : (u',v') \mapsto u' \otimes v'"] \\
U \otimes V \arrow[r, "f \otimes g : u \otimes v \mapsto f(u) \otimes g(v)"] & U' \otimes V'
\end{tikzcd}$$

Die Komposition $\tau' \circ h$ ist bilinear über (R,S,T) und induziert dank der UAE von τ die ersehnte (R,T)–lineare Abbildung $f \otimes g$. QED

Funktorialität des Tensorprodukts

$$
\begin{tikzcd}
U \otimes_S V \arrow[r, shift left, "\mathrm{id}_U \otimes \mathrm{id}_V"] \arrow[r, shift right, "\mathrm{id}_{U\otimes V}"'] & U \otimes_S V \arrow[r, "f\otimes g"] \arrow[rr, bend right, "(f'\otimes g')\circ(f\otimes g)"'] & U' \otimes_S V' \arrow[r, "f'\otimes g'"] & U'' \otimes_S V''
\end{tikzcd}
$$

Satz S2O: Funktorialität

(1) Für die Identitäten gilt $\mathrm{id}_U \otimes \mathrm{id}_V = \mathrm{id}_{U\otimes V}$.

(2) Seien $f : U \to U'$ und $f' : U' \to U''$ linear über (R, S) sowie $g : V \to V'$ und $g' : V' \to V''$ linear über (S, T). Dann gilt:

$$(f' \otimes g') \circ (f \otimes g) = (f' \circ f) \otimes (g' \circ g)$$

Aufgabe: Was ist hier zu zeigen? Zeigen Sie es!

Lösung: Die beiden (R, T)–linearen Abbildungen links und rechts sind bereits konstruiert, wir müssen sie nur noch vergleichen. Die Gleichheit gilt auf den einfachen Tensoren, also auf allen Tensoren (S2J).

Das Kronecker–Produkt von Matrizen

Definition S2P: Kronecker–Produkt

Für $A \in R^{m' \times m}$ und $B \in R^{n' \times n}$ definieren wir das **Kronecker–Produkt**

$$A \otimes B := \begin{bmatrix} a_{1,1}B & \dots & a_{1,m}B \\ \vdots & & \vdots \\ a_{m',1}B & \dots & a_{m',m}B \end{bmatrix} \in R^{(m'n') \times (mn)}.$$

Ausgeschrieben bedeutet das:

$$A \otimes B = \begin{bmatrix} a_{1,1}b_{1,1} & \dots & a_{1,1}b_{1,n} & \dots & a_{1,m}b_{1,1} & \dots & a_{1,m}b_{1,n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{1,1}b_{n',1} & \dots & a_{1,1}b_{n',n} & \dots & a_{1,m}b_{n',1} & \dots & a_{1,m}b_{n',n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m',1}b_{1,1} & \dots & a_{m',1}b_{1,n} & \dots & a_{m',m}b_{1,1} & \dots & a_{m',m}b_{1,n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m',1}b_{n',1} & \dots & a_{m',1}b_{n',n} & \dots & a_{m',m}b_{n',1} & \dots & a_{m',m}b_{n',n} \end{bmatrix}$$

Das Kronecker–Produkt von Matrizen

Übung: Es gibt Spezialfälle, in denen das Kronecker–Produkt gleich dem üblichen Matrixprodukt ist. Welche sind das? (Matrixmodell S1C)

Satz S2P: Kronecker–Produkt

Sei $R = S = T$ kommutativ, $\mathcal{A} = (u_1, \dots, u_m)$ eine Basis von U und $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V. Dank S2G erhalten wir daraus die Produktbasis $\mathcal{A} \otimes \mathcal{B} = (u_1 \otimes v_1, \dots, u_1 \otimes v_n, \dots, u_m \otimes v_1, \dots, u_m \otimes v_n)$.

Für Basen $\mathcal{A}' = (u'_1, \dots, u'_{m'})$ von U' und $\mathcal{B}' = (v'_1, \dots, v'_{n'})$ von V' nutzen wir ebenso die lexikographische Ordnung für die Produktbasis $\mathcal{A}' \otimes \mathcal{B}' = (u'_1 \otimes v'_1, \dots, u'_1 \otimes v'_{n'}, \dots, u'_{m'} \otimes v'_1, \dots, u'_{m'} \otimes v'_{n'})$.

Für je zwei R–lineare Abbildungen $f : U \to U'$ und $g : V \to V'$ gilt dann:

$$\mathrm{M}^{\mathcal{A}' \otimes \mathcal{B}'}_{\mathcal{A} \otimes \mathcal{B}}(f \otimes g) = \mathrm{M}^{\mathcal{A}'}_{\mathcal{A}}(f) \otimes \mathrm{M}^{\mathcal{B}'}_{\mathcal{B}}(g)$$

🙂 Damit haben wir für $f \otimes g$ eine konkrete Darstellung als Matrix. Für Matrizen haben wir unsere bewährten, effizienten Werkzeuge.

Das Kronecker–Produkt von Matrizen

Übung: Es gilt $A \otimes (B \otimes C) = (A \otimes B) \otimes C$ aber i.A. $A \otimes B \neq B \otimes A$; Gleichheit gilt nach geeigneter Permutation der Basis (Zeilen&Spalten).

Bei Transposition gilt $(A \otimes B)^\intercal = A^\intercal \otimes B^\intercal$, bei gleicher Reihenfolge!

Beim Kronecker–Produkt erfüllt die Spur $\operatorname{tr}(A \otimes B) = \operatorname{tr}(A) \cdot \operatorname{tr}(B)$. Für $A \in K^{m\times m}$ und $B \in K^{n\times n}$ gilt $\det(A \otimes B) = \det(A)^n \det(B)^m$.

Für die Spektren gilt: Aus $\sigma(A) = \{\lambda_1, \dots, \lambda_m\}$ und $\sigma(B) = \{\mu_1, \dots, \mu_n\}$ folgt $\sigma(A \otimes B) = \{\, \lambda_i \mu_j \mid i = 1, \dots, m,\ j = 1, \dots, n \,\}$.

Für die zugehörigen Eigenräume gilt $\mathrm{E}(A; \lambda) \otimes \mathrm{E}(B; \mu) \subseteq \mathrm{E}(A \otimes B; \lambda\mu)$, somit gilt $\mathrm{E}(A \otimes B; \nu) \supseteq \bigoplus_{\lambda\mu=\nu} \mathrm{E}(A; \lambda) \otimes \mathrm{E}(B; \mu)$, aber i.A. nicht „$\subseteq$“.

Gegenbeispiel: Wir betrachten $A = B = \left[\begin{smallmatrix} \lambda & 1 \\ 0 & \lambda \end{smallmatrix}\right]$. Dann ist neben $e_1 \otimes e_1$ auch $e_1 \otimes e_2 - e_2 \otimes e_1$ ein Eigenvektor von $A \otimes B$ zum Eigenwert λ^2.

Das mehrfache Tensorprodukt

Genau wie für zwei Faktoren erklären wir das n–fache Tensorprodukt

$$\boxed{T = {}_{R_0}(U_1)_{R_1} \otimes {}_{R_1}(U_2)_{R_2} \otimes \cdots \otimes {}_{R_{n-1}}(U_n)_{R_n}.}$$

Gegeben seien hierzu Ringe $R_0, R_1, R_2, \dots, R_n$ und lineare Räume $U_1, U_2, \dots, U_n$, wobei U_i linear über (R_{i-1}, R_i) ist, wie oben vermerkt.

Übung zur Wiederholung: Formulieren Sie die Definitionen S2D / S2E und die anschließenden Sätze für das mehrfache Tensorprodukt.

🙂 Alle Definitionen und Sätze verlaufen parallel zum vorigen Fall $n = 2$.

Warum, so fragen Sie, habe ich das nicht gleich allgemein ausgeführt? Nun ja, beim ersten Durchgang kostet schon der einfachste Fall $n = 2$ die wissbegierig Lernenden viel Mühe, und die zusätzliche Schreibarbeit für den allgemeinen Fall $n \in \mathbb{N}$ ist dabei nur Ballast ohne Mehrwert.

🙂 Beim zweiten Durchgang ist der allgemeine Fall dann leicht, daher formuliere ich die Ausführung als lehrreiche Übung zur Wiederholung.

Das mehrfache Tensorprodukt

Definition S2Q: multilineare Abbildung und interner Ausgleich

Sei $R = (R_0, R_1, \dots, R_n)$ eine Familie von Ringen und $U = (U_1, \dots, U_n)$ eine Familie linearer Räume über R, wobei U_i linear über (R_{i-1}, R_i) ist. Sei $\mu : \prod_{i=1}^n U_i \to W$ eine Abbildung in einen (R_0, R_n)–linearen Raum.

(1) Wir nennen μ **multiadditiv**, falls μ additiv in jedem U_i ist:

$$\mu(\dots, u_i + u_i', \dots) = \mu(\dots, u_i, \dots) + \mu(\dots, u_i', \dots)$$

(2) Wir nennen μ **multihomogen** über $R = (R_0, \dots, R_n)$ falls gilt:

$$\begin{aligned}
\mu(r_0 \bullet u_1, u_2, \dots, u_n) &= r_0 \bullet \mu(u_1, u_2, \dots, u_n) \\
\mu(\dots, u_i \bullet r_i, u_{i+1}, \dots) &= \mu(\dots, u_i, r_i \bullet u_{i+1}, \dots) \\
\mu(u_1, \dots, u_{n-1}, u_n \bullet r_n) &= \mu(u_1, \dots, u_{n-1}, u_n) \bullet r_n
\end{aligned}$$

in jeder Stelle i, für alle Skalare $r_i \in R_i$ und Vektoren $u_i \in U_i$. Gilt (1) und (2), so nennen wir μ kurz **R–multilinear**:

$$\boxed{\mathrm{Mul}_R(U; W) = \left\{\, \mu : \prod_{i=1}^n U_i \to W \text{ multilinear über } R \,\right\}}$$

Das mehrfache Tensorprodukt

Beispiel: Gegeben sei ein kommutativer Ring R und lineare Räume $U_1, \dots, U_n, W$ über R (S2B). Wir betrachten eine n–stellige Abbildung

$$\mu \,:\, U_1 \times U_2 \times \dots \times U_n \to W.$$

Genau dann ist μ multihomogen über $(R, R, \dots, R)$, wenn gilt:

$$\mu(\dots, r \bullet u_i, \dots) = r \bullet \mu(\dots, u_i, \dots)$$

in jeder Stelle i, für alle Skalare $r \in R$ und Vektoren $u_i \in U_i$.

Übung: Beweisen Sie diese Äquivalenz (zunächst für $n = 2, 3, \dots$).

☺ Über jedem kommutativen Ring R erhalten so den vertrauten Begriff der Multilinearität, wie wir ihn etwa von der Determinante (L2A) kennen.

⚠ Für nicht-kommutative Ringe müssen wir behutsamer vorgehen. Hier ist die oben formulierte Definition S2D bzw. S2Q die richtige.

☺ Im kommutativen Fall stimmen beide Definitionen überein.

Definition S2R: das mehrfache Tensorprodukt

Sei $R = (R_0, \dots, R_n)$ eine Familie von Ringen und $U = (U_1, \dots, U_n)$ eine Familie linearer Räume, wobei U_i linear über (R_{i-1}, R_i) ist.

(1) Ein **Produkt** $\mu : \prod_{i=1}^n U_i \to Z$ ist eine R–multilineare Abbildung.

(2) Ein **Tensorprodukt** von U über R ist ein Produkt $\tau : \prod_{i=1}^n U_i \to W$ mit der folgenden universellen Abbildungseigenschaft:

$$\begin{tikzcd}
U_1 \times \cdots \times U_n \arrow[r, "\tau", "\text{multilinear}"'] \arrow[dr, "\text{multilinear}", "\mu"'] & W \arrow[d, dashed, "\exists! f"', "\text{linear}"] \\
& Z
\end{tikzcd}$$

Jedes Produkt $\mu : \prod_{i=1}^n U_i \to Z$ faktorisiert eindeutig über τ:
Es gibt genau eine (R_0, R_n)–lineare Abbildung $f : W \to Z$ mit $\mu = f \circ \tau$.

Diese Forderung ist äquivalent zur Bijektivität der natürlichen Abbildung

$$\boxed{\Phi_\tau^Z \;:\; \mathrm{Hom}_{(R_0,R_n)}(W, Z) \to \mathrm{Mul}_R(U; Z) \;:\; f \mapsto \mu = f \circ \tau.}$$

Satz S2S: Prinzip der multilinearen Fortsetzung

Seien $U_1, \dots, U_n, W$ lineare Räume über dem kommutativen Ring R. Zu $i = 1, \dots, n$ sei $(u_{i,j})_{j \in J_i}$ eine Basis von U_i sowie $(w_j)_{j \in J}$ eine beliebige Familie in W, indiziert durch $J = J_1 \times \dots \times J_n$.

(1) Dann existiert genau eine R–multilineare Abbildung

$$\boxed{\begin{aligned} &\tau : U_1 \times \dots \times U_n \to W : (u_{1,j_1}, \dots, u_{n,j_n}) \mapsto w_{j_1,\dots,j_n}, \quad \text{nämlich} \\ &\tau\Big(\sum_{j_1 \in J_1} \lambda_{1,j_1} u_{1,j_1}, \dots, \sum_{j_n \in J_n} \lambda_{n,j_n} u_{n,j_n} \Big) = \sum_{j \in J} \lambda_{1,j_1} \cdots \lambda_{n,j_n} w_j \end{aligned}}$$

für alle Linearkombinationen mit Koeffizienten $\lambda_i \in R^{(J_i)}$ für $i = 1, \dots, n$.

(2) Genau dann ist τ ein Tensorprodukt, erfüllt also die universelle Abbildungseigenschaft S2E, wenn $(w_j)_{j \in J}$ eine Basis von W ist.

Übung: Beweisen Sie dies nach dem Vorbild von Satz S2G. Schließen Sie daraus das folgende schöne Anwendungsbeispiel.

Beispiel S2T: Polynommodell des mehrfachen Tensorprodukts

Sei K ein kommutativer Ring. Wir betrachten $U_1 = \dots = U_n = K[X]$ und hierzu die vertraute Polynommultiplikation

$$\begin{aligned} \tau \;:\; K[X] \times \dots \times K[X] &\to K[X_1, \dots, X_n] \\ (u_1, \dots, u_n) &\mapsto u_1(X_1) \cdots u_n(X_n). \end{aligned}$$

Dieses Produkt τ hat die universelle Abbildungseigenschaft (S2S).

Satz S2U: Eindeutigkeit des mehrfachen Tensorprodukts

Sei $R = (R_0, \dots, R_n)$ eine Familie von Ringen und $U = (U_1, \dots, U_n)$ eine Familie linearer Räume, wobei U_i linear über (R_{i-1}, R_i) ist.

Dann sind zu diesen Räumen je zwei Tensorprodukte $\tau : \prod_{i=1}^n U_i \to W$ und $\tau' : \prod_{i=1}^n U_i \to W'$ eindeutig isomorph: Es existiert genau ein (R_0, R_n)–Isomorphismus $(f, f') : W \cong W'$ mit $\tau' = f \circ \tau$ und $\tau = f' \circ \tau'$.

Übung: Beweisen Sie dies nach dem Vorbild von Satz S2H.

Das mehrfache Tensorprodukt

Satz S2V: Existenz des mehrfachen Tensorprodukts

Sei $R = (R_0, \dots, R_n)$ eine Familie von Ringen und $U = (U_1, \dots, U_n)$ eine Familie linearer Räume, wobei U_i linear über (R_{i-1}, R_i) ist.

Zu U über R existiert ein Tensorprodukt $\tau : \prod_{i=1}^n U_i \times V \to W$.

Übung: Beweisen Sie dies nach dem Vorbild von Satz S2I.
Aus dieser länglichen, aber expliziten Konstruktion folgt sofort:

Korollar S2W: Einfache Tensoren erzeugen alle Tensoren.

Die einfachen Tensoren $u_1 \otimes \dots \otimes u_n$ mit $u_1 \in U_1, \dots, u_n \in U_n$ erzeugen den Tensorraum $U_1 \otimes \dots \otimes U_n$. Ausführlich bedeutet das:

Jeder beliebige Tensor $w \in U_1 \otimes \dots \otimes U_n$ ist eine endliche Summe $w = \sum_{k=1}^{\ell} u_{1,k} \otimes \dots \otimes u_{n,k}$ mit einer gewissen Länge $\ell \in \mathbb{N}$ und geeigneten Faktoren $u_{i,k} \in U_i$ für $i = 1, \dots, n$ und $k = 1, \dots, \ell$.

Das mehrfache Tensorprodukt

Beispiel S3A: Reellifizierung und Komplexifizierung

(1a) Jeder $\mathbb{C}$–Vektorraum $V_{\mathbb{C}}$ wird zu einem $\mathbb{R}$–Vektorraum $V_{\mathbb{R}}$ durch Einschränkung der Skalare von $\mathbb{C}$ zu $\mathbb{R}$, also die Einschränkung von

$$\bullet : V \times \mathbb{C} \to V \qquad \text{zu} \qquad \bullet : V \times \mathbb{R} \to V.$$

(1b) Ist $(v_j)_{j \in J}$ eine $\mathbb{C}$–Basis von V, so ist $(v_j \bullet 1, v_j \bullet \mathrm{i})_{j \in J}$ eine $\mathbb{R}$–Basis von V. Für die Dimensionen folgt $\dim_{\mathbb{R}}(V) = 2 \dim_{\mathbb{C}}(V)$.

(1c) Jede $\mathbb{C}$–lineare Abbildung $f : V \to V'$ ist auch $\mathbb{R}$–linear.

(2a) Zu jedem $\mathbb{R}$–Vektorraum $U_{\mathbb{R}}$ erhalten wir den $\mathbb{C}$–Vektorraum

$$\boxed{V_{\mathbb{C}} := (U_{\mathbb{R}}) \otimes_{\mathbb{R}} ({}_{\mathbb{R}}\mathbb{C}_{\mathbb{C}}).}$$

(2b) Ist $(u_j)_{j \in J}$ eine $\mathbb{R}$–Basis von $U_{\mathbb{R}}$, so ist $(v_j = u_j \otimes 1)_{j \in J}$ eine $\mathbb{C}$–Basis von $V_{\mathbb{C}}$. Für die Dimensionen folgt $\dim_{\mathbb{R}}(U) = \dim_{\mathbb{C}}(V)$.

(2c) Jede $\mathbb{R}$–lineare Abbildung $f_{\mathbb{R}} : U_{\mathbb{R}} \to U'_{\mathbb{R}}$ induziert eine zugehörige $\mathbb{C}$–lineare Abbildung $g_{\mathbb{C}} = f_{\mathbb{R}} \otimes \mathrm{id}_{\mathbb{C}} : V_{\mathbb{C}} = U_{\mathbb{R}} \otimes_{\mathbb{R}} \mathbb{C}_{\mathbb{C}} \to U'_{\mathbb{R}} \otimes_{\mathbb{R}} \mathbb{C}_{\mathbb{C}} = V'_{\mathbb{C}}$.

Reellifizierung und Komplexifizierung

Die Übergang von $\mathbb{C}$ nach $\mathbb{R}$ ist leicht, hier genügt die Einschränkung. Umgekehrt ist die Erweiterung von $\mathbb{R}$ nach $\mathbb{C}$ keineswegs offensichtlich!

Warum wollen wir das? Der Körper $\mathbb{C}$ hat bessere Eigenschaften, insbesondere zerfällt jedes Polynom über $\mathbb{C}$ in Linearfaktoren.

Diesen Trick haben wir daher schon oft angewendet, mangels allgemeiner Werkzeuge jedoch zunächst noch recht umständlich:

Wir stellen $f_{\mathbb{R}} : U_{\mathbb{R}} \to U'_{\mathbb{R}}$ als eine reelle Matrix $A \in \mathbb{R}^{m\times n}$ dar, und betrachten diese dann als eine komplexe Matrix $A \in \mathbb{C}^{m\times n}$.

Hier sehen Sie nun, was wirklich dahinter steckt: Mit dem Tensorprodukt können wir die Komplexifizierung allgemein und basisfrei erklären!

So erhalten wir nicht nur die komplexifizierte Matrix, sondern auch die $\mathbb{C}$–lineare Abbildung $g_{\mathbb{C}} = f_{\mathbb{R}} \otimes \mathrm{id}_{\mathbb{C}} : V_{\mathbb{C}} = U_{\mathbb{R}} \otimes_{\mathbb{R}} \mathbb{C}_{\mathbb{C}} \to U'_{\mathbb{R}} \otimes_{\mathbb{R}} \mathbb{C}_{\mathbb{C}} = V'_{\mathbb{C}}$

Dieser Trick funktioniert allgemein für jede Ringerweiterung $S \le R$. Dies führt uns zum folgenden allgemeinen Satz.

Erweiterung des Grundrings

Satz S3B: Erweiterung des Grundrings

Seien $S \le R$ Ringe und U_S ein (rechts)linearer Raum über S.

(1) Die **Erweiterung der Skalare** ergibt einen linearen Raum über R:

$$\boxed{V_R := (U_S) \otimes_S ({}_S R_R)}$$

(2) Wir haben die natürliche S–lineare Abbildung (i.A. keine Einbettung)

$$\iota \,:\, U_S \xrightarrow{\sim} U_S \otimes_S S \xrightarrow{\mathrm{id}\otimes\mathrm{inc}} U_S \otimes_S R = V_R \,:\, u \mapsto v = u \otimes 1_R.$$

(3) Freie Räume bleiben frei: Ist $(u_i)_{i\in I}$ eine Basis von U_S über S, dann ist $(v_i = u_i \otimes 1_R)_{i\in I}$ eine Basis von V_R über R, und ι ist injektiv.

(4) Jede S–lineare Abbildung $f_S : U_S \to U'_S$ induziert eine zugehörige R–lineare Abbildung $g_R = f_S \otimes \mathrm{id}_R : V_R = U_S \otimes_S R_R \to U'_S \otimes_S R_R = V'_R$.

(5) Die darstellenden Matrizen sind dieselben in $S^{m\times n} \le R^{m\times n}$.

Dasselbe gilt entsprechend für linkslineare Räume.

Die Erweiterung des Grundrings ist funktoriell.

$$\begin{tikzcd}
U \arrow[r, "\mathrm{id}_U"] & U \arrow[r, "f"] \arrow[rr, bend right, "f'\circ f"'] & U' \arrow[r, "f'"] & U''
\end{tikzcd}$$

$$\begin{tikzcd}
U \otimes_S R \arrow[r, shift left, "\mathrm{id}_U \otimes \mathrm{id}_R"] \arrow[r, shift right, "\mathrm{id}_{U\otimes R}"'] & U \otimes_S R \arrow[r, "f\otimes\mathrm{id}_R"] \arrow[rr, bend right, "(f'\otimes\mathrm{id}_R)\circ(f\otimes\mathrm{id}_R)=(f'\circ f)\otimes\mathrm{id}_R"'] & U' \otimes_S R \arrow[r, "f'\otimes\mathrm{id}_R"] & U'' \otimes_S R
\end{tikzcd}$$

Satz S3B: Die Erweiterung des Grundrings ist funktoriell.

(6) Für die Identität $\mathrm{id}_U : U \to U$ gilt

$$\mathrm{id}_U \otimes \mathrm{id}_R = \mathrm{id}_{U\otimes R} = \mathrm{id}_V : V \to V.$$

(7) Für $f : U \to U'$ und $f' : U' \to U''$ und $f' \circ f : U \to U''$ gilt

$$(f' \otimes \mathrm{id}_R) \circ (f \otimes \mathrm{id}_R) = (f' \circ f) \otimes \mathrm{id}_R\,.$$

Übung: Alle Daten liegen explizit vor. Rechnen Sie es nach!

Tensorieren mit $\mathbb{Q}$

Aufgabe: Welche $\mathbb{Q}$–Dimension hat $\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Q}$? allgemein $(\mathbb{Z}/n) \otimes_{\mathbb{Z}} \mathbb{Q}$?

Satz S3C: Tensorieren mit $\mathbb{Q}$

(1) Dank S2M haben wir den Isomorphismus

$$\boxed{\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Q} \cong \mathbb{Q} \;:\; (a \otimes b) \mapsto ab.}$$

(2) Für alle $n \in \mathbb{N}_{\geq 1}$ hingegen gilt

$$\boxed{(\mathbb{Z}/n) \otimes_{\mathbb{Z}} \mathbb{Q} = 0.}$$

Beweis: (2) Die einfachen Tensoren erzeugen den Tensorraum (S2J) und $[a] \otimes b = [a] \otimes (n \bullet b/n) = ([a] \bullet n) \otimes (b/n) = 0 \otimes (b/n) = 0$. QED

Anwendungsbeispiel: Für jede endlich erzeugte abelsche Gruppe

$$V \cong \mathbb{Z}^r \oplus \mathbb{Z}/n_1 \oplus \cdots \oplus \mathbb{Z}/n_t$$

mit $n_1, \dots, n_t \geq 2$ extrahiert Tensorieren mit $\mathbb{Q}$ den freien Anteil:

$$V \otimes_{\mathbb{Z}} \mathbb{Q} \cong \mathbb{Q}^r \quad \text{und} \quad r = \dim_{\mathbb{Q}}(V \otimes_{\mathbb{Z}} \mathbb{Q}).$$

Tensorieren mit $\mathbb{Q}$

🙂 Unter diesem Isomorphismus wird die kanonische Basis $(e_1, \dots, e_r)$ von $\mathbb{Z}^r$ abgebildet auf die kanonische Basis $(e_1, \dots, e_r)$ von $\mathbb{Q}^r$.

Der Torsionsteil $\mathbb{Z}/n_1 \oplus \dots \oplus \mathbb{Z}/n_t$ hingegen wird auf 0 abgebildet. Hierzu nutzen wir (2) und die Distributivität des Tensorprodukts (S2N).

⚠ Die natürliche Abbildung ι ist im Allgemeinen keineswegs injektiv. Wir illustrieren Ringerweiterungen hier mit dem Übergang von $\mathbb{Z}$ zu $\mathbb{Q}$:

Über $\mathbb{Z}$ kann das Tensorieren mit $\mathbb{Q}$ einige Information vernichten. Manchmal möchten wir genau das, so wie hier illustriert.

🙂 Über einem Körper wie $\mathbb{Q}$ hat jeder lineare Raum eine Basis. Über dem Ring wie $\mathbb{Z}$ hingegen gilt diese Eigenschaft leider nicht.

Zur Vereinfachung können wir daher über $\mathbb{Z}$ mit $\mathbb{Q}$ tensorieren; wir verlieren dabei Information, aber gewinnen stärkere Sätze.

Diese grundlegende Beobachtung führt uns zu folgender Definition.

Der Rang einer abelschen Gruppe

Wie können wir den „Rang“ einer abelschen Gruppen erklären?

Definition S3D: Rang einer abelschen Gruppe

Sei $(V, +)$ eine abelsche Gruppe, also ein $\mathbb{Z}$–linearer Raum.
Der **Rang** von V, genauer der **$\mathbb{Q}$–Rang**, ist dann definiert durch

$$\boxed{\operatorname{rang}_{\mathbb{Q}}(V) := \dim_{\mathbb{Q}}(V \otimes_{\mathbb{Z}} \mathbb{Q}).}$$

Dasselbe gilt für jeden Körper $\mathbb{K}$, etwa $\mathbb{K} = \mathbb{F}_p = \mathbb{Z}/p$ mit p prim:

$$\boxed{\operatorname{rang}_{\mathbb{K}}(V) := \dim_{\mathbb{K}}(V \otimes_{\mathbb{Z}} \mathbb{K}).}$$

Anwendungsbeispiel: Für jede endlich erzeugte abelsche Gruppe

$$V \cong \mathbb{Z}^r \oplus \mathbb{Z}/n_1 \oplus \dots \oplus \mathbb{Z}/n_t$$

mit $n_1, \dots, n_t \ge 2$ gilt $\operatorname{rang}_{\mathbb{Q}}(V) = r$ wie oben und zudem

$$\operatorname{rang}_{\mathbb{F}_p}(V) = r + \sharp\big\{\, i \in \{1, \dots, t\} \;\big|\; p \mid n_i \,\big\}.$$

Der Rang einer abelschen Gruppe

Wir gehen hier von $\mathbb{Z}$ zu einem Körper $\mathbb{K}$ über, die natürlichen Kandidaten hierzu sind neben dem Bruchkörper $\mathbb{Q} = \operatorname{Frac}(\mathbb{Z})$ vor allem die Quotientenkörper $\mathbb{Z} \twoheadrightarrow \mathbb{Z}/p = \mathbb{F}_p$ für p prim.

Der Vorteil eines Körpers $\mathbb{K}$ ist, wie oben motiviert, dass wir allgemein über Basen verfügen und von der Dimension über $\mathbb{K}$ sprechen könnnen. Unsere gründliche Arbeit zu Vektorräumen zahlt sich hier aus!

Beispiel: Für $V = \mathbb{Z}^7 \oplus \mathbb{Z}_2^4 \oplus \mathbb{Z}_8 \oplus \mathbb{Z}_3^2 \oplus \mathbb{Z}_9$ gilt $\operatorname{rang}_{\mathbb{Q}}(V) = 7$ sowie $\operatorname{rang}_{\mathbb{F}_2}(V) = 12$ und $\operatorname{rang}_{\mathbb{F}_3}(V) = 10$ und $\operatorname{rang}_{\mathbb{F}_p}(V) = 7$ für alle $p \geq 5$.

Für die Rechnung benötigen wir das Tensorprodukt $(\mathbb{Z}/p) \otimes_{\mathbb{Z}} (\mathbb{Z}/q)$. Dies führt uns zu dem nächsten schönen Satz.

Das Tensorprodukt $(\mathbb{Z}/p) \otimes_{\mathbb{Z}} (\mathbb{Z}/q)$

Aufgabe: Wie viele Elemente hat $(\mathbb{Z}/5) \otimes_{\mathbb{Z}} (\mathbb{Z}/9)$?

Lösung: Wir finden $5 \cdot 2 \equiv 1 \bmod 9$. In $(\mathbb{Z}/5) \otimes (\mathbb{Z}/9)$ folgt daraus:

$$[a] \otimes [b] = [a] \otimes (5 \cdot 2 \cdot [b]) = ([a] \cdot 5) \otimes (2 \cdot [b]) = 0 \otimes (2 \cdot [b]) = 0$$

Daher gilt $(\mathbb{Z}/5) \otimes_{\mathbb{Z}} (\mathbb{Z}/9) = 0$. Was steckt allgemein dahinter?

Satz S3E: das Tensorprodukt $(\mathbb{Z}/p) \otimes_{\mathbb{Z}} (\mathbb{Z}/q)$

Für $p, q \in \mathbb{N}$ und $r = \operatorname{ggT}(p, q)$ haben wir den Isomorphismus

$$\boxed{(\mathbb{Z}/p) \otimes_{\mathbb{Z}} (\mathbb{Z}/q) \;\cong\; (\mathbb{Z}/r) \;:\; [a]_p \otimes [b]_q \mapsto [ab]_r, \quad [c]_p \otimes [1]_q \leftarrow\!\shortmid [c]_r}$$

Wichtiger Spezialfall: Gilt $\operatorname{ggT}(p, q) = 1$, so folgt $(\mathbb{Z}/p) \otimes_{\mathbb{Z}} (\mathbb{Z}/q) = 0$.

Aufgabe: Warum ist die Umkehrfunktion $[c]_p \otimes [1]_q \leftarrow\!\shortmid [c]_r$ wohldefiniert?

Lösung: Dank Bézout (A2I) existieren $u, v \in \mathbb{Z}$ mit $r = pu + qv$, also
$[r]_r \mapsto [r]_p \otimes [1]_q = ([v]_p \bullet q) \otimes [1]_q = [v]_p \otimes (q \bullet [1]_q) = [v]_p \otimes [0]_q = 0$.

Das Tensorprodukt $\mathbb{Z}/p \otimes_{\mathbb{Z}} \mathbb{Z}/q$

☺ Beim Tensorprodukt multiplizieren sich die Dimensionen (S1B).

Naiv würde man daher für $(\mathbb{Z}/5) \otimes_{\mathbb{Z}} (\mathbb{Z}/9)$ vielleicht $5 \cdot 9 = 45$ Elemente vermuten; das gilt tatsächlich für das kartesische Produkt $(\mathbb{Z}/5) \times (\mathbb{Z}/9)$.

⚠ Das Tensorprodukt hingegen verhält sich deutlich anders! Hierzu rechnen wir sorgfältig mit Erzeugern und Relationen.

⚠ Dieselbe Notation $[\dots]$ für Äquivalenzklassen bedeutet hier dreimal Verschiedenes! Zur Betonung habe ich die Indizes p, q, r hinzugefügt.

Aufgabe: Führen Sie die Konstruktion des ersehnten Isomorphismus $(\mu, \mu') : (\mathbb{Z}/p) \otimes_{\mathbb{Z}} (\mathbb{Z}/q) \cong (\mathbb{Z}/r) : [a] \otimes [b] \leftrightarrows [ab]$ detailliert aus.

Das Tensorprodukt $\mathbb{Z}/p \otimes_{\mathbb{Z}} \mathbb{Z}/q$

Beweis: Wir nutzen auch hier die universelle Abbildungseigenschaft:

$$
\begin{tikzcd}
\mathbb{Z} \times \mathbb{Z} \arrow[rr, shift left, "\tilde\mu : (a,b) \mapsto ab"] \arrow[d, "\mathrm{quo} \times \mathrm{quo}"'] \arrow[ddrr, "g"] & & \mathbb{Z} \arrow[ll, shift left, "(c,1) \mapsfrom c : \tilde\mu'"] \arrow[dd, "\mathrm{quo}"] \arrow[ddll, "f"'] \\
(\mathbb{Z}/p) \times (\mathbb{Z}/q) \arrow[d, "\tau"'] \arrow[drr, "h"] & & \\
(\mathbb{Z}/p) \otimes_{\mathbb{Z}} (\mathbb{Z}/q) \arrow[rr, shift left, "\mu"] & & (\mathbb{Z}/r) \arrow[ll, shift left, "\mu'"]
\end{tikzcd}
$$

Die Abbildung $\tilde{\mu} : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z} : (a, b) \mapsto ab$ ist bilinear. Die Komposition $g = \mathrm{quo} \circ \tilde{\mu}$ erfüllt zudem $g(p, 0) = g(0, q) = 0$. Somit induziert g die Abbildung $h : (\mathbb{Z}/p) \times (\mathbb{Z}/q) \to (\mathbb{Z}/r) : ([a], [b]) \mapsto [ab]$, ebenfalls bilinear. Dank UAE erhalten wir $\mu : (\mathbb{Z}/p) \otimes_{\mathbb{Z}} (\mathbb{Z}/q) \to (\mathbb{Z}/r) : [a] \otimes [b] \mapsto [ab]$.

Die Abbildung $f : \mathbb{Z} \to (\mathbb{Z}/p) \otimes_{\mathbb{Z}} (\mathbb{Z}/q) : c \mapsto [c] \otimes [1]$ erfüllt $\ker(f) \supseteq r\mathbb{Z}$, wie wir oben dank Bézout–Koeffizienten bereits nachgerechnet haben. Sie induziert somit $\mu' : (\mathbb{Z}/r) \to (\mathbb{Z}/p) \otimes_{\mathbb{Z}} (\mathbb{Z}/q) : [c] \mapsto [c] \otimes [1]$. Damit gilt $\mu' \circ \mu = \mathrm{id}$ und $\mu \circ \mu' = \mathrm{id}$. (Nachrechnen!) QED

Darstellung von Homomorphismen

Satz S3F: Darstellung von Homomorphismen

Seien U, V rechtslineare Räume über dem Ring R. Wir betrachten dazu

$$\Phi \,:\, V \otimes_R U^* \to \operatorname{Hom}_R(U,V) \,:\, \Phi(v \otimes \varphi)(u) = v \bullet \varphi(u).$$

(1) Hat U eine endliche Basis $(u_1, \dots, u_n)$, so ist Φ bijektiv mit Inverser

$$\Psi \,:\, \operatorname{Hom}_R(U,V) \to V \otimes_R U^* \,:\, f \mapsto \textstyle\sum_{j=1}^n f(u_j) \otimes u_j^*.$$

(2) Hat V eine endliche Basis $(v_1, \dots, v_m)$, so ist Φ bijektiv mit Inverser

$$\Psi \,:\, \operatorname{Hom}_R(U,V) \to V \otimes_R U^* \,:\, f \mapsto \textstyle\sum_{i=1}^m v_i \otimes f^*(v_i^*).$$

(3) Gilt (1) und (2) mit R kommutativ, so ist $(e_{i,j} = \Phi(v_i \otimes u_j^*))_{i=1,\dots,m}^{j=1,\dots,n}$ eine Basis von $\operatorname{Hom}_R(U,V) \cong R^{m \times n}$, wie von Matrizen vertraut.

(4) Ist R ein Körper, so ist Φ injektiv, und das Bild in $\operatorname{Hom}_R(U,V)$ sind genau die Homomorphismen mit endlichem Rang:

$$\textstyle\sum_{i=1}^r v_i \otimes \varphi_i \;\mapsto\; \sum_{i=1}^r \Phi(v_i \otimes \varphi_i)$$

Darstellung von Homomorphismen

🙂 Diese Schreibweise ist übersichtlich und effizient, insbesondere wenn der Rang r klein ist gegenüber den Dimensionen m und n.

Beispiel S3G: Singulärwertzerlegung

Die obige Darstellung kennen wir von der Singulärwertzerlegung:

$$\sum_{i=1}^{r} \sigma_i (v_i \otimes u_i^*) \;\mapsto\; \sum_{i=1}^{r} \sigma_i \, \Phi(v_i \otimes u_i^*)$$

Hierzu betrachten wir eine lineare Abbildung $f : U \to V$ über $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Die beiden Vektorräume U und V tragen jeweils ein Skalarprodukt.

Zu f sind $\sigma_1 \ge \dots \ge \sigma_r \ge 0$ die Singulärwerte, die Singulärvektoren $v_1, \dots, v_r \in V$ und $u_1, \dots, u_r \in U$ bilden zwei orthonormale Familien, und $u_i^* = \langle\, u_i \mid - \,\rangle \in U^*$ ist die zu u_i gehörige Linearform.

Damit schreibt sich die lineare Abbildung f gemäß

$$f \;:\; U \to V \;:\; u \mapsto \sum_{i=1}^{r} \sigma_i v_i \cdot \langle\, u_i \mid u \,\rangle.$$

Speziell für die euklidischen Räume $U = \mathbb{K}^m$ und $V = \mathbb{K}^n$ erhalten wir die Matrixdarstellung $\mathrm{M}(f) = \sum_{i=1}^{r} \sigma_i v_i u_i^\dagger$ bezüglich der Standardbasen.

Darstellung von Homomorphismen

Aufgabe: Alle Daten liegen explizit vor. Rechnen Sie Satz S3F nach!
(0) Warum ist Φ wohldefiniert? (1,2) Gilt $\Phi \circ \Psi = \mathrm{id}$ und $\Psi \circ \Phi = \mathrm{id}$?
(3) Warum sind Φ und Ψ nun R–linear? (4) Was sind Kern und Bild?