

# Kapitel 5: Lineare Gleichungssysteme

## 5.1 Grundlagen

Das vorliegende Kapitel enthält eine erste größere Anwendung der linearen Algebra. Die linearen Gleichungssysteme bilden eine Art Paradebeispiel einer solchen Anwendung. Über die lineare Algebra hinausgehende Strukturen werden nicht benötigt, anders als in den Anwendungen, die in den beiden nachfolgenden Kapiteln behandelt werden. Andererseits liegt durchaus eine anspruchsvolle Anwendung vor, so dass das Kapitel unabhängig von der Bedeutung des Stoffes als Übung im Umgang mit einem komplexen Sachverhalt zu sehen ist. Das gilt vorzugsweise über den zweiten Teil, der den Dualraum einführt.

### 5.1.0 Vorbemerkung

(1.0.1) Wir beschränken in diesem Abschnitt die Diskussion auf Vektorräume, wobei wir zusätzlich bei den linearen Abbildungen einen endlichdimensionalen Urbildraum voraussetzen wollen.  $\dim_K V = n < \infty$ . Dann ist  $\text{Bild } f \subset W$  automatisch auch endlichdimensional mit Dimension  $\leq n$ .

(1.0.2) Zu einer linearen Gleichung gehört stets eine lineare Abbildung. Genauer: Die Umstände, die zur Aufstellung eines linearen Gleichungssystems führen, erlauben zugleich die Konstruktion eines Vektorraumhomomorphismus  $\lambda : V \rightarrow W$  und eines Vektors  $\vec{b} \in W$ , derart, dass das Ausgangssystem zur Bestimmungsgleichung  $\lambda(\vec{x}) = \vec{b}$  gleichwertig ist.

Man kann nun den gesamten Apparat der linearen Vektoralgebra einsetzen, um die Probleme der linearen Gleichungen zu behandeln. Und das erweist sich als vorteilhaft. Zunächst haben wir bereits eine größere Zahl verfügbarer Resultate, die darauf warten, auf den Problemkreis der linearen Gleichungssysteme angewandt zu werden. Wir brauchen auch nicht aufwendig einen neuen Begriffsapparat (für Gleichungssysteme) einzuführen, sondern verwenden und ergänzen das, was wir bereits kennen, nämlich den im ersten Kapitel entwickelten **Begriffsapparat für Abbildungen**. Und für weitergehende Probleme weist uns die lineare Algebra meist erfolgversprechende Richtungen zur Problemlösung.

(1.0.3) Die linearen Gleichungssysteme haben beträchtliche Bedeutung für die unterschiedlichsten Disziplinen, so dass eine etwas tiefergehende Einarbeitung sicher wichtig ist. Wir bringen am Ende des Teilkapitels ein Beispiel aus der Physik, die Statik.

(1.0.4) Im anschließenden zweiten Teil des Kapitels verwenden wir die Analyse der Gleichungssysteme, um die Einführung eines schwierigeren Begriffs der linearen Algebra - den Dualraum (eines Vektorraumes) - zu begründen.

### 5.1.1 Zusammenstellung bereits verfügbarer Resultate.

(1.1.1) Wir stellen nochmals die bereits in Kapitel 4.1.4 besprochenen allgemeinen Resultate zu den linearen Gleichungen zusammen:

Es sei	$f(\vec{x}) = \vec{0}$ eine lineare homogene Gleichung mit zugehöriger Abbildung $f : V \rightarrow W$ , wobei $V$ und $W$ endlichdim. Vektorräume über dem Körper $K$ mit $\dim_K V = n$ , $\dim_K W = m$ .
Dann	ist $\mathbb{L} = \text{Kern}(f)$ . Die Lösungsmenge ist immer Teilraum von $V$ .
◆	Ist $\dim \text{Kern}(f) = k < \infty$ und $a$ eine Basis von $\text{Kern}(f)$ , so ist $L_a = (K^k, (\alpha_i) \mapsto \sum \vec{a}_i \alpha_i, V)$ eine Parametrisierung der Lösungsmenge $\mathbb{L}$ . D.h. zur vollen Angabe von $\mathbb{L}$ benötigt man nur die $n$ Vektoren einer Basis und die Abbildung $f$ .

Und jetzt der inhomogene Fall:

Es sei  $f(\vec{x}) = \vec{b}$  eine lineare inhomogene Gleichung mit  $\vec{b} \in W$ .  
 Dann ist die Gleichung unlösbar, wenn  $\vec{b} \notin \text{Bild}(f)$  ist.  
 Ist dagegen  $\vec{b} \in \text{Bild}(f)$ , dann ist die Gleichung lösbar. Ist  $\vec{x} \in V$  irgendeine derartige Lösung, so gilt für die Lösungsmenge

$\mathbb{L} = \vec{x}_S + \mathbb{L}_h = \vec{x}_S + \text{Kern}(f)$

Ist  $a$  Basis von  $\text{Kern}(f)$ , so ist  $(K^k, (\alpha_i) \mapsto \sum \vec{a}_i \alpha_i, V)$  Paramtrisierung von  $\mathbb{L}_h$ .

◆ Inhomogene Gleichungen sind daher entweder unlösbar, oder ihre Lösungsmenge ist eine Parallelverschiebung von  $\text{Kern}(f)$ .

◆ Ist  $\text{Kern}(f)$  nicht endlichdimensional, dann ist  $K^k$  überall durch  $\mathcal{E}(I, K)$  zu ersetzen. Damit gelten dieselben Resultate. Vgl. Kap4.2.4.

(1.1.2) Zur homogenen Gleichung  $f(\vec{x}) = \vec{0}$  ist nichts zu beweisen. Für den Rest wiederholen wir den Beweis aus 4.1.5. Die Aussagen zur inhomogenen Gleichung ergeben sich wie folgt:

Ist  $\vec{a} \in (\vec{x}_S + \text{Kern}(f))$  d.h.  $\vec{a} = \vec{x}_S + \vec{k}$  mit  $f(\vec{k}) = \vec{0}$ , so folgt

$$f(\vec{x}_S + \vec{k}) = f(\vec{x}_S) + f(\vec{k}) = \vec{b} + \vec{0} = \vec{b}.$$

D.h.  $\vec{a}$  ist Lösung. Ist umgekehrt  $\vec{A}$  irgendeine Lösung von  $f(\vec{x}) = \vec{b}$ , so können wir im Rahmen der Vektorrechnung schreiben:  $\vec{A} = \vec{x}_S + (\vec{A} - \vec{x}_S)$ . Nun ist  $f(\vec{A} - \vec{x}_S) = f(\vec{A}) - f(\vec{x}_S) = \vec{b} - \vec{b} = \vec{0}$ . D.h.  $(\vec{A} - \vec{x}_S) \in \text{Kern}(f)$  und  $\vec{A} \in (\vec{x}_S + \text{Kern}(f))$  wie behauptet.

(1.1.3) Die zusammengestellten Aussagen zu den Lösungsmengen linearer Gleichungen sind für die praktische Arbeit von großem Nutzen. Darüber sollte die Einfachheit des soeben gegebenen Beweises nicht hinwegtäuschen. Insbesondere sollte man immer an den folgenden Sachverhalt denken:

Um den Teilraum  $\text{Kern}(f)$  zu bekommen, genügt es, wenn man sich eine Basis dieses Raumes verschafft. Kennt man die Dimension des Kernes, benötigt man nur ebensoviele unabhängige Elemente, also Lösungen der homogenen Gleichung.

(1.1.4) Im Falle der linearen Differentialgleichungen ist zwar  $\dim V = \infty$ , aber man kann die endliche (!) Dimension des Kernes relativ leicht aus der Struktur der Gleichung ablesen. Dies werden wir in Kap. 6 besprechen. Dann gelten die Resultate entsprechend.

(1.1.5) Nachfolgend wollen wir uns mit dem Fall eines endlichdimensionalen  $V$  befassen, also dem traditionellen Problem der linearen Gleichungssysteme. Die zugehörige Gleichung läßt sich - spätestens nach Basiswahl - als Matrixgleichung  $M\vec{x} = \vec{b}$  schreiben. Hier ist die genauere Bestimmung der Kerndimension  $k$  nicht immer so leicht. Entsprechend werden wir später häufiger auf spezielle ergänzende Resultate zu diesem Problem - Bestimmung der Dimension des Kernes - stoßen. Vielfach geht es darum, herauszufinden, ob  $k = 0$  gilt, d. h. ob die Abbildung injektiv ist, was bedeutet, dass die zugehörige Gleichung **höchstens eine Lösung** besitzt.

(1.1.6) Ein Hauptziel der Arbeit in diesem Kapitel wird sein, neben  $\text{Kern}(f)$  einer zweiten  $\ell$ -dimensionalen Raum zu konstruieren mit  $k + \ell = n$ . Dieser Raum soll die *Bedingungen erfassen, die das Ausgangssystem ausmachen*, aber auch alle Bedingungen, die daraus folgen. Wir werden sehen, dass ein Raum der angegebenen Dimension  $n - k$  als Teilraum des Dualraums  $V^* = \text{Hom}_K(V, K)$  von  $V$  eindeutig bestimmt ist. Überdies werden wir sehen, wie man die Zahl  $\ell$  direkt aus der Matrix  $M$  oder der Abbildung  $f$  des Systems gewinnen kann. Und mit  $\ell$  hat man über  $k = n - \ell$  auch  $k$ , die Zahl, die eigentlich interessiert.

(1.1.7) Kurz und griffig fassen wir zusammen:

Ein Gleichungssystem umfaßt Bedingungen und Lösungen. Die Bedingungen werden durch einen  $\ell$ -dimensionalen Raum in  $V^*$  und die Lösungen - im homogenen Fall - durch einen  $k$ -dimensionalen Raum in  $V$  bestimmt. Es gilt  $k + \ell = n$ .

(1.1.8) Zusätzlich möchte man auch noch genauer verstehen, **wieso** ein lineares Gleichungssystem diese bemerkenswerte Struktur erzeugt und wie daraus die Rechenverfahren entstehen, mit deren Hilfe man üblich solche Systeme löst.

(1.1.9) Die Frage nach dem Warum werden wir in mehreren Anläufen behandeln. Zunächst diskutieren wir ein Beispiel, das uns den Weg weisen wird und zugleich einige Definitionen motiviert. Dann versuchen wir eine allgemeine Behandlung ohne Dualraum und schließlich eine mit diesem.

(1.1.10) Ist  $k = \dim \text{Kern} f = 0$ , dann ist  $f$  injektiv und auf  $\text{Bild} f$  invertierbar. Ist also  $\vec{b} \in \text{Bild} f$ , dann erhält man die (dann eindeutige) Lösung von  $f(\vec{x}) = \vec{b}$  mit Hilfe der inversen Abbildung  $f^{-1}$  zu  $\vec{x} = f^{-1}(\vec{b})$ . Beachten Sie den Rollenwechsel: In der ersten Gleichung ist  $\vec{x}$  Unbestimmte, in der zweiten *Bezeichnung* für den eindeutigen Lösungsvektor. Oder auch: Verfügt man über die inverse Abbildung, dann erhält man damit (im injektiven Fall) den Lösungsvektor.

(1.1.11) Liegt speziell eine Matrixgleichung  $M\vec{x} = \vec{b}$  vor, dann kann man die Gleichung algebraisch auflösen, indem man von links mit der inversen Matrix multipliziert.  $\vec{x} = M^{-1}\vec{b}$ . Man wird also in dem Fall  $k = 0$  generell danach streben, die zu  $M$  inverse Matrix zu finden.

*Wir beginnen die genaueren Ausführungen mit zwei Abschnitten vorbereitenden Charakters.*

## 5.1.2 Unterschiedliche Darstellungsformen linearer Gleichungssysteme.

(1.2.1) Die Vektorrechnung stellt uns eine Reihe von Darstellungsformen linearer Gleichungen zur Verfügung.  $\lambda : V \rightarrow W$  sei die zugehörige lineare Abbildung. Vier dieser Darstellungsformen sind besonders wichtig:

Die **absolute (koordinaten freie) Form**  $\lambda(\vec{x}) = \vec{b}$ .  
 Diese bevorzugen wir fast immer, wenn es um "Verständnis" geht.  
 Sie ist verfügbar, sobald  $\lambda$  als Abbildung erklärt ist, insbesondere natürlich wenn  $\lambda$  als Matrix eingeführt ist.

Die üblichste Form ist

Die **Matrixform**  $A\vec{x}^K = \vec{b}^K$   
 Sie entsteht durch Vorgabe von je einer Basis in  $V$  und in  $W$ .  $A = \lambda^K$  ist die beschreibende Matrix. Diese Form erlaubt die Verwendung des gesamten Matrixkalküls.

Mehr aus der üblichen Vektorrechnung kommt die folgende Form:

Die **Vektorform oder genauer Spaltenvektorform:**

$$\vec{S}_1 x_1 + \vec{S}_2 x_2 + \dots + \vec{S}_n x_n = \vec{b}$$

Das ist eine Vektorgleichung im Wertraum  $W$ . Sie folgt aus der absoluten Form, wenn man nur im Urbildraum eine Basis  $a$  einführt und  $\vec{x} = \sum \vec{a}_i x_i$  setzt.

Natürlich ist  $\vec{S}_i = \lambda(\vec{a}_i)$ . Führt man zusätzlich in  $W$  eine Basis ein, so ist  $\lambda^K(\vec{a}_i) = \vec{S}_i^K$  die  $i$ -te Spalte der beschreibenden Matrix  $A = \lambda^K$ .

Besonders in der elementaren Vektorrechnung und bei der Anwendung physikalischer Gesetze tauchen lineare Gleichungen gern in dieser Form auf. Man denke an Schnittmengenbestimmungen und Kräftegleichgewichte. Die Methode der reziproken Basis bezieht sich auf diese Form.

An die übliche elementare Gestalt linearer Gleichungssysteme als Zusammenstellung von  $m$  Einzelgleichungen knüpft die vierte Form an:

Die **System- oder Zeilenvektorform:**  
 $\lambda_1(\vec{x}) = b_1, \lambda_2(\vec{x}) = b_2, \dots, \lambda_m(\vec{x}) = b_m$   
 Man erhält das System aus der absoluten Form, indem man im Wertraum eine Basis einführt und dann für  $\lambda(x) = b$  Koeffizientenvergleich vornimmt.  
 $\lambda_r(\vec{x})$  ist also die Komponente von  $\lambda(\vec{x}) \in W$  in Richtung des  $r$ -ten Basisvektors.

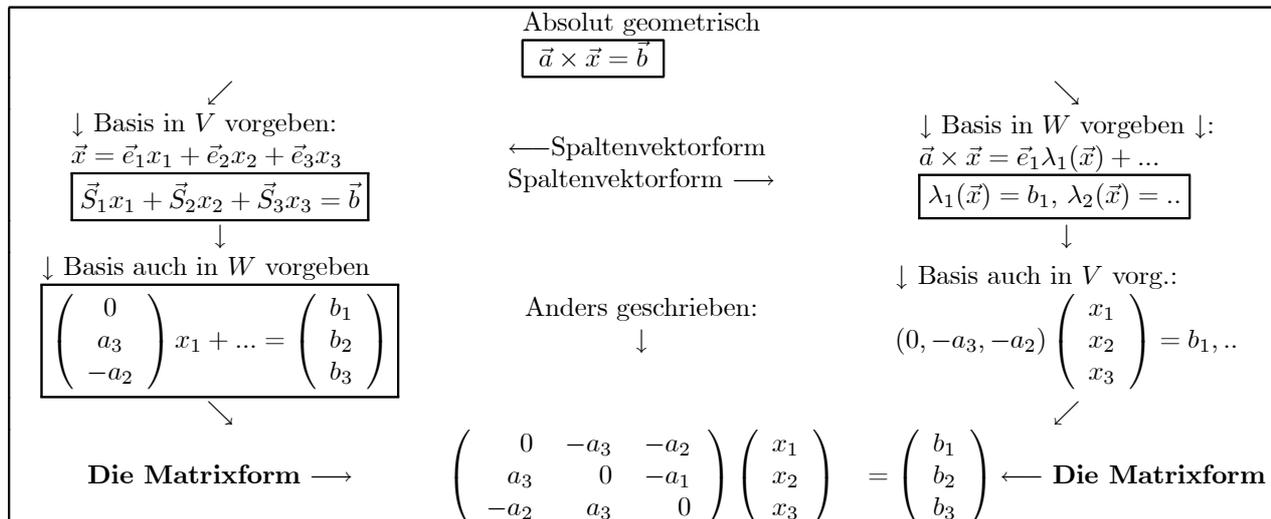
Die Zuordnung  $\vec{x} \mapsto \lambda_r(\vec{x})$  definiert eine Linearform  $\lambda_r = (V, \vec{x} \mapsto \lambda_r(\vec{x}), K) \in V^*$ . Wir haben also ein Gleichungssystem vorliegen, das aus  $m$  gegebenen Linearformen entsteht. Führt man auch in  $V$  eine Basis ein, so ist  $\lambda_r^K$  gerade die  $r$ -te Zeile der beschreibenden Matrix  $A$ . (Daher Zeilenvektorform!).

Die übliche Form eines linearen Gleichungssystems entpuppt sich hier als ein System für Linearformen. (Zu den Bezeichnungen vgl. 4.(4.4.3).

(□ F. ) Welcher Unterschied besteht zwischen  $\vec{x}$ ,  $\vec{x}^K$  und  $x_i$ . Was ist  $(x_i)$ ?

(1.2.2) Ein Beispiel für die unterschiedlichen Darstellungsformen:

In der Statik treten Gleichungen des Typs  $\vec{a} \times \vec{x} = \vec{b}$  (Vektorprodukt) auf mit gesuchtem  $\vec{x}$ . Die zugehörige Abbildung  $\lambda = (V_0^3, \vec{x} \mapsto \vec{a} \times \vec{x}, V_0^3)$  ist infolge der Bilinearität des Vektorproduktes linear. Gesucht ist  $\mathbb{L} = \lambda^{-1}(\{\vec{b}\})$  wobei  $\vec{a}$  äußerer Parameter ist. Wie sehen die zugehörigen verschiedenen Darstellungsformen dieser Gleichung aus?



(1.2.3) Man sollte lineare Gleichungen in all ihren Verkleidungen erkennen und auch in der Lage sein, sie je nach Bedarf ineinander umzuwandeln. Vollziehen Sie daher die einzelnen Schritte des Beispiels sorgfältig nach.

(□ F. ) Wie sieht eine lineare Matrixgleichung aus, wenn man sie im Indexkalkül formuliert?

### 5.1.3 Problemtypen, die im Zusammenhang mit einer Gleichung auftreten.

(1.3.1) Einige Probleme sind im Zusammenhang mit einer Bestimmungsgleichung immer von besonderem Interesse. Sie steuern das weitere Vorgehen und legen in gewisser Weise die Bedeutung von Resultaten fest. Wir nennen fünf derartige Fragerichtungen:

(1.3.2) **Die Struktur der Lösungsmenge  $\mathbb{L}$ .** Man möchte möglichst viele und tiefgehende Kenntnisse über  $\mathbb{L}$  erlangen, ohne  $\mathbb{L}$  selbst besitzen zu müssen, berechnet zu haben. Eine typische Frage: Wie kann sich die Struktur der Lösungsmenge ändern, etwa deren Dimension, wenn äußere Parameter auftreten?

(1.3.3) **Auflösungskalküle.** Man möchte  $\mathbb{L}$  bei Bedarf kalkülartig mit bestimmten Zusatzbedingungen (wenig Aufwand, kontrollierbare Ungenauigkeit, gute Verallgemeinerbarkeit ...) bestimmen.

(1.3.4) **Zulässige Umformungen (des Gleichungssystems).** In Auflösungskalkülen nimmt man Umformungen des Gleichungssystems vor, von denen man meint, dass sie die Lösungsmenge nicht oder in kontrollierbarer Weise beeinflussen. Wie rechtfertigt man derartige Manipulationen? Was für zulässige Umformungen gibt es überhaupt? Gelten sie auch noch, wenn man äußere Parameter zulässt?

(1.3.5) **Ursachenverständnis.** Warum hat eine Lösungsmenge eine bestimmte Struktur? Warum sind gewisse Umformungen zulässig, andere nicht? Zugehöriges Verständnis ist wichtig, sobald man mit neuartigen Situationen konfrontiert wird und sich fragt, ob die bisherigen Resultate übertragbar sind oder nicht. (Etwa lineare Gleichungen in Moduln,  $\dim V = \infty$ , Ungleichungen, nichtlineare Gleichungen....)

**(1.3.6) Einbettung in Problemsituationen.** Die Gleichungsstruktur soll in den unterschiedlichen Problemsituationen der Anwendungsbereiche wiedererkannt werden. Dabei kommt es einerseits darauf an, die Verwendbarkeit allgemeiner Resultate zu erkennen, andererseits aber auch, zu erkennen, wann und weshalb die die allgemeine Struktur noch nicht reicht, das anstehende Problem zu lösen.

**(1.3.7)** Die Behandlung dieser Fragen sollte einem bei längeren Überlegungen immer als eine Art geistiger Kompaß dienen. Es sind also keine Fragen, für die wir eine bestimmte eindeutige Antwort erhalten werden, aber sie liefern einem Orientierungen der Art: "Jetzt sind wir dabei, einen Teilaspekt der Lösungsmengenstruktur zu beweisen bzw. zu verstehen." Oder: "Dies ist ein Resultat, das für einen Lösungskalkül wichtig werden kann." Usw. Die nachfolgenden Überlegungen versuchen immer, Resultate und Formalismen aus Kap. 4 mit der Behandlung dieser Fragen zu verbinden.

- Können Sie die folgenden Fragen beantworten, die eine konkrete Illustration zu (1.34-5) darstellen: "Warum haben die folgenden beiden Gleichungssysteme dieselben Lösungen?" Und "Welche weiteren linearen Gleichungssysteme haben noch diese Lösungsmenge?"

$\begin{aligned} 3x+z &= 0 \\ 3y-7z &= 0 \\ 7x+y &= 0 \end{aligned}$	$\begin{aligned} 2x-y+3z &= 0 \\ x+y-2z &= 0 \end{aligned}$
--	---

Beachten Sie: Die Warum-Frage wird nicht dadurch beantwortet, dass Sie die Gleichheit der beiden Lösungen verifizieren. Am Ende diese Kapitels können Sie die Antwort geben und begründen.

### 5.1.4 Zeilenrang und Spaltenrang

**(1.4.1)** Es sei  $f : V \rightarrow W$  lineare Abbildung endlichdimensionaler Vektorräume und  $f(x)=b$  eine zugehörige lineare Gleichung. Wir setzen  $k = \dim \text{Kern}(f)$  und  $\ell = \dim \text{Bild}(f)$ . Nach dem Dimensionssatz für Homomorphismen gilt  $k + \ell = n = \dim V$ . Wir wissen, dass beide Zahlen gerade auch für Gleichungen wichtige Kenngrößen bilden.

**(1.4.2)** Führen wir jetzt in  $V$  und  $W$  Basen ein, so wird  $f$  durch die zugehörige Darstellungsmatrix  $M = M_f^K$  beschrieben, und unsere Gleichung geht in die gleichwertige Matrixgleichung  $M\vec{x}^K = \vec{b}^K$  über.  $K$  bezeichnet unser Koordinatensystem. Wir stellen uns die Frage: **Kann man  $k$  oder  $\ell$  direkt aus der Matrix  $M$  bestimmen?** Bei positiver Antwort hätten wir ein wichtiges Resultat zu unserem ersten Problemtyp (1.3.2): Wir hätten eine relevante Aussage über die Lösung, ohne diese bereits zu kennen. Es zeigt sich, dass  $\ell$  die über die Matrix leichter zugängliche Größe ist.

**(1.4.3)**  $\text{Bild}(f)$  ist isomorph zu  $\text{Bild}M$ . Wir können daher  $\ell$  bestimmen, indem wir die Dimension von  $\text{Bild}M$  bestimmen.  $M : K^n \rightarrow K^m$ . Nun nehmen wir die kanonische Basis  $e^K$  von  $K^n$  und bilden sie mit  $M$  ab. Wir wissen einerseits, dass  $M(\vec{e}^K)$  die  $i$ -te Spalte von  $M$  ist, und andererseits, dass die Familie  $M(\vec{e}^K)$  ein **Erzeugendensystem** von  $\text{Bild}M$  ist.

**(1.4.4)** Wir können Spalten streichen, die sich als Linearkombination der verbleibenden darstellen lassen und erhalten - Basisauswahlsatz 4(3.2.10) - schließlich eine **Basis** von  $\text{Bild}M$ . **Die Anzahl der verbleibenden Spalten ergibt die gesuchte Dimension  $\ell$ .**

**(1.4.5)** Als Rezept:

Sei  $M$  eine  $m \times n$ -Matrix über  $K$ . Zerlege  $M$  in Spalten.

- ◆ Streiche Spalten, die sich durch die verbleibenden ausdrücken lassen.
- ◆ Es verbleibt eine linear unabhängige Familie von Spaltenvektoren.
- ◆ Deren Anzahl nennen wir den **Spaltenrang** von  $M$ . Symbolisch  $rg_S(M)$ .

Dann gilt:  $\text{Spaltenrang}(M) = \dim \text{Bild}(M) = \ell$   
Und  $k = n - \ell$ .

**(1.4.6)** Da jeder Spaltenvektor in  $K^m$  liegt und es  $n$  Spaltenvektoren gibt, gilt immer  $rg_S(M) \leq \min(n, m)$ .

**(1.4.7)** Der Spaltenrang ist in vielen Fällen per Inspektion erhältlich. Dann hat man über  $\ell$  auch das zugehörige  $k$ .

(1.4.8) Wir haben die Matrix in Spalten zerlegt. Es liegt nahe, zu fragen, ob man  $M$  nicht auch in Zeilen zerlegen sollte, um damit eine entsprechende Konstruktion vorzunehmen. Dies haben wir bereits immer getan, wenn wir die Gleichung als System von Gleichungen interpretieren. Jede Einzelgleichung gehört zu einer Zeile der Matrix. Die größtmögliche Zahl unabhängiger (homogener) Gleichungen ist also die größtmögliche Zahl unabhängiger Zeilen.

(1.4.9) Wieder als Rezept:

Sei	$M$ eine $m \times n$ -Matrix über $K$ . Zerlege $M$ in Zeilen.
◆	Streiche Zeilen, die sich durch die verbleibenden ausdrücken lassen.
◆	Es verbleibt eine linear unabhängige Familie von Zeilenvektoren.
◆	Deren Anzahl nennen wir den <b>Zeilenrang von <math>M</math></b> . Symbolisch $rgz(M)$ .
Dann	repräsentiert der Zeilenrang die Zahl der unabhängigen, nicht fortlaßbaren Bedingungen im homogenen Gleichungssystem.

(1.4.10) Erneut gilt natürlich  $rgz(M) \leq \min(n,m)$ . Aber ansonsten scheinen Zeilenrang und Spaltenrang einer Matrix wenig miteinander zu tun zu haben. Dass sie völlig unterschiedliche Eigenschaften der Gleichung repräsentieren, verstärkt diesen Eindruck.

(1.4.11) Bestimmt man jedoch für einzelne Beispielmatrix beide Ränge explizit, so findet man, dass beide Ränge immer gleich sind. Dies gilt insbesondere auch, wenn man in die Matrix äußere Parameter einbaut und sich bemüht, damit unterschiedliche Werte der beiden Ränge zu produzieren. Nach einigen Beispielen gewinnt man den Verdacht, dass die beiden Ränge tatsächlich immer übereinstimmen könnten. Man fragt nach einem Beweis. Wir werden im Laufe des Kapitels zwei Beweise dieses Sachverhaltes geben.

(1.4.12) Zunächst akzeptieren wir das Resultat **Zeilenrang = Spaltenrang**. Vereinfachend ist es üblich, diese Zahl den **Rang** (der Matrix, des Homomorphismus,...) zu nennen. Nochmals: Dieser Rang wird vielfach einfach durch Inspektion der Matrix bestimmt. (Vgl. etwa (1.4.5) oder (1.5.49).) Oder er wird geeignet ausgerechnet (etwa über  $\dim \text{Bild}$ ) oder man verschafft ihn sich anderweitig. Welchen Nutzen hat das? Wir formulieren die Antwort kurz rezeptartig.

(1.4.13) **Welchen Nutzen bringt die Kenntnis des Ranges bei der Behandlung eines konkreten Gleichungssystems?**

- Gegeben ein lineares Gleichungssystem  $M\vec{x} = \vec{b}$ .
- Man verschafft sich den zugehörigen Rang, also  $\ell = \dim \text{Bild} M$ .
- Es folgt  $k = n - \ell = \dim \text{Kern} M$ .
- Dann enthält die Lösung der homogenen Gleichung  $k$  freie Parameter ( ist ein  $k$ -dimensionaler Teilraum von  $V$ ).  $k = 0$  bedeutet (höchstens) eindeutige Lösbarkeit.
- Ist eine zugehörige inhomogene Gleichung lösbar, so enthält auch hier die allgemeine Lösung  $k$  freie Parameter.

(1.4.14) In gewissen Fällen - etwa bei elektrischen Schaltungen - wird man erwarten und möchte es daher beweisen, dass ein Gleichungssystem eindeutig lösbar ist (genau eine Lösung hat). Dann wird man so vorgehen, dass man zeigt, dass der Rang gleich  $n = \dim V$  ist und dass  $\vec{b}$  im Bild liegt.

(1.4.13) Wie entsteht die Lösungsstruktur eines linearen Gleichungssystems? Was tragen  $k$  und  $\ell$  - insbesondere also auch Zeilenrang und Spaltenrang - zum Verständnis der Lösungsmengenstruktur bei? Diese Frage behandeln wir jetzt in einem ersten längeren Anlauf mit Hilfe der Methode des Basiswechsels. D.h. wir beschreiben die gegebene Gleichung **in einem neuen Koordinatensystem, das wir jeweils geeignet wählen**. Genauer suchen wir analog zum Beispiel (1.2.2) einmal eine Basis für den Werteraum, einmal für den Urbildraum und schließlich für beide Räume gemeinsam. Die Resultate aus Kap. 4.5 sind anzuwenden.

## 5.1.5 Basiswechsel im Werteraum

(1.5.1) Wir wählen im Werterraum - und nur dort - eine neue Basis. Die alte Basis sei  $c$  und die neue  $C$ . Die zugehörige Transformationsmatrix  $S$  wird durch  $c = CS$  bestimmt. Unser Ausgangssystem ist gegeben durch  $f(\vec{x}) = \sum \vec{c}_k M_{kr} x_r = \sum \vec{c}_k b_k$ . Auf den Index  $A$  für "alt" verzichten wir hier. Koeffizientenvergleich gibt unser Gleichungssystem.

(1.5.2) Jetzt führen wir in der üblichen Weise die neue Basis ein.

$f(\vec{x}) = \sum \vec{c}_k M_{kr} x_r$	$= \sum \vec{C}_j S_{jk} M_{kr} x_r$	$\vec{b} = \sum \vec{c}_k b_k = \sum \vec{C}_j S_{jk} b_k$
	$= \sum \vec{C}_j (\sum S_{jk} M_{kr} x_r)$	$= \sum \vec{C}_j (\sum S_{jk} b_k)$
	$= \sum \vec{C}_j (\sum (SM)_{jr} x_r)$	$= \sum \vec{C}_j (Sb)_j$

(1.5.3) Koeffizientenvergleich liefert das neue, aber gleichwertige Gleichungssystem

$\sum (SM)_{jr} x_r = (Sb)_j$	$(SM)\vec{x} = \vec{Sb}$
Indexform	Matrixform

(1.5.4) Basiswechsel im Werterraum bedeutet daher einfach, dass man das Ausgangssystem  $M\vec{x} = \vec{b}$  mit einer invertierbaren  $m \times m$ -Matrix von links multipliziert. D.h. insbesondere: Die üblichen elementaren Zeilenoperationen sind als derartige Basiswechsel interpretierbar! Oder auch: Jede Einzelgleichung wird durch eine Linearkombination der Gleichungen ersetzt, wobei die Invertierbarkeit von  $S$  die Äquivalenz mit den Ausgangsbedingungen sicherstellt.

(□ F. ) Zu welchem Fragentyp aus (1.3.2-6) liefert uns (1.5.4) einen Einstieg?

(1.5.5) Was kann man hinsichtlich der Form der neuen Matrix  $SM$  erreichen? Nun, wir können zuerst eine Basis von  $Bild(f)$  wählen und diese sodann zu einer Basis von ganz  $W$  ergänzen:  $C_1, \dots, C_\ell$  und  $C_{\ell+1}, \dots, C_m$ . Dann ist  $M(\vec{e}_j)$  nur Linearkombination der ersten  $\ell$  Basisvektoren, die ja  $Bild(M)$  aufspannen. D.h. die letzten  $m - \ell$  Zeilen von  $SM$  sind lauter Nullen! Oder: Die linken Seiten dieser Gleichungen sind Null. Wie steht es mit den rechten Seiten? Nun, wenn  $\vec{b} \in Bild(M)$  gilt, dann sind auch die entsprechenden Komponenten von  $\vec{b}$  alle Null, und die letzten Gleichungen sind alle trivial:  $0 = 0$ . **Wir können sie streichen.** Damit haben wir übrigens auch Zeilenrang  $\leq \ell$ .

(1.5.6) Ist dagegen  $b \notin Bild(M)$ , dann hat  $\vec{b}$  mindestens ein  $b_i \neq 0$  für ein  $i > \ell$ . Und dann ist das System widersprüchlich, unlösbar. Das entspricht voll unseren Erwartungen.

Am Ende dieses Teilkapitels in 5.1.11 entwickeln wir mit den Resultaten dieses Abschnittes ein Lösungskalkül für die Bestimmung der inversen Matrix bzw. zur Lösung linearer Gleichungssysteme.

### 0.0.1 5.1.6 Basiswechsel im Urbildraum

(1.6.1) Jetzt wählen wir im Urbildraum  $V$  - und nur hier - eine neue Basis. Die Ausgangsbasis sei  $e$ , die neue Basis bezeichnen wir mit  $E$ . Die zugehörige Transformationsmatrix ist gegeben durch  $e = ET$ .

(1.6.2) Für den allgemeinen Vektor  $\vec{x}$  des Urbildraumes haben wir zwei Darstellungen:

$$\vec{x} = \sum \vec{e}_i x_i = \sum \vec{E}_j^K x_j^N.$$

D.h. wir führen einen Satz von neuen Unbestimmten  $x_j^N$  ein, mit deren Hilfe wir die alten Bedingungen beschreiben wollen. Oder auch

$x_i^N = \sum T_{ik} x_k,$	$x_k = \sum T_{ki}^{-1} x_i^N$	$x_i^N$ in der Rolle von Hilfsgrößen
----------------------------	--------------------------------	--------------------------------------

(1.6.3) Das Problem lautet: **Wie sehen die Bedingungsgleichungen aus, wenn man sie mit Hilfe der neuen Koordinaten schreibt?** Was kann man durch geeignete Basiswahl erreichen? (Hat man etwa die Gleichung  $2(x + y) = 3$ , so kann man die Hilfsvariable  $x^N = x + y$  einführen und erhält für sie die Gleichung  $2x^N = 3$ ).

(1.6.4) Die übliche Rechnung gibt die Antwort: ( $c$  sei feste Basis in  $W$ ) :

$$\begin{aligned} \vec{y} &= f(\vec{x}) = \sum \vec{c}_r M_{rs}^N x_s^N \quad \text{einerseits per Definition} \\ &= \sum \vec{c}_r M_{rs} x_s = \sum \vec{c}_r M_{rs} T_{si}^{-1} x_i^N \quad \text{andererseits über (1.6.2).} \end{aligned}$$

(1.6.5) Koeffizientenvergleich liefert die gesuchten Matrixbeziehungen

$$M^N = MT^{-1} \quad M^N \vec{x}^N = \vec{b} \quad \vec{x} = T^{-1} \vec{x}^N.$$

(□ **F.** ) Die Formel für  $M^N$  folgt auch sofort aus der allgemeinen Transformationsformel  $M^N = SM^A T^{-1}$  aus Kap 4. Begründen sie das.

(1.6.6) Mit (1.6.5) haben wir das gleichwertige Gleichungssystem für die als Hilfsvariablen eingeführten neuen Koordinaten: Die erste Gleichung bestimmt  $M^N$ , die Lösung der zweiten gibt  $\vec{x}^N$  und die dritte liefert schließlich die ursprünglichen Unbestimmten.

(1.6.7) Was kann man auf diese Weise erreichen? D.h. wie sollte man gleichungsspezifisch die neue Basis  $E$  von  $V$  wählen? Wir wählen zuerst eine Basis  $\vec{E}_1, \dots, \vec{E}_k$  von  $Kern(f)$ , die wir sodann zu einer Basis von ganz  $V$  ergänzen. Dies bedeutet, dass  $f(\vec{E}_i) = \vec{0}$  ist für  $i = 1, 2, \dots, k$ . Und das heißt, dass die ersten  $k$  Spalten der beschreibenden Matrix  $M^N = MT^{-1}$  alle Null sind. Im zugehörigen Gleichungssystem treten daher die Koordinaten  $x_1^N, \dots, x_k^N$  überhaupt nicht auf. Man kann sie als freie Parameter wählen! Die übrigen neuen Koordinaten werden durch die Bedingungsgleichung bestimmt - sofern das System nicht widersprüchlich ist. Dann kann man alles über  $\vec{x}^A = \vec{x} = T^{-1}\vec{x}^N$  wieder durch die ursprünglichen Größen ausdrücken.

### 5.1.7 Allgemeiner Basiswechsel und die Lösungsstruktur des Gleichungssystems

(1.7.1) Mit Hilfe der bisher gefundenen Resultate können wir die Struktur der Lösungsmengen linearer Gleichungen insgesamt besser verstehen. Wir tun das, indem wir die Abbildung  $f$  als Zusammensetzung von drei Abbildungen darstellen, wobei jeder Faktor ganz bestimmte Phänomene der Lösungsmengenstruktur liefert und deren Auftreten verständlich macht.

(1.7.2) Wir bestimmen  $Kern(f)$  und wählen eine Basis  $(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_k)$  dieses Teilraumes. Nun ergänzen wir diese Basis  $a$  zu einer Basis  $ax$  von ganz  $V$ . Also  $ax = (\vec{a}_1, \dots, \vec{a}_k, \vec{a}_{k+1}, \dots, \vec{a}_n)$ . Das gibt Anlaß zu einer direkten Zerlegung von  $V$ . Nämlich  $V = Kern(f) \oplus U$ . Dabei ist  $U$  ein  $\ell$ -dimensionaler Raum mit Basis  $(\vec{a}_{k+1}, \dots, \vec{a}_n)$ . Wir wissen - Dimensionssatz für Homomorphismen -, dass  $U$  isomorph zu  $Bild(f)$  ist und dass die Restriktion von  $f$  auf  $U$  die Isomorphie realisiert. Wir nennen die Restriktion  $\varphi$ , also  $\varphi = f|_U$ . Dann ist  $(\varphi(\vec{a}_{k+1}), \dots, \varphi(\vec{a}_n))$  eine Basis von  $Bild(f)$ . Oder auch:  $\varphi$  ist **bijektive Parametrisierung von  $Bild(f)$** . Für jedes  $\vec{y} \in Bild(f)$  haben wir ein eindeutiges  $\vec{u} = \vec{a}_{k+1}u_1 + \dots + \vec{a}_n u_\ell$  aus  $U$  mit  $\varphi(\vec{u}) = \vec{y}$ . Oder

$$\varphi(\vec{u}) = \varphi(\vec{a}_{k+1}u_1 + \dots + \vec{a}_n u_\ell) = \varphi(\vec{a}_{k+1})u_1 + \dots + \varphi(\vec{a}_n)u_\ell.$$

Weiter hat jedes  $\vec{x} \in V$  eine eindeutige Darstellung  $\vec{x} = \vec{x}_0 + \vec{u}$  mit  $\vec{x}_0 \in Kern(f)$  und  $\vec{u} \in U$ . **Wir haben**  $f(\vec{x}) = f(\vec{x}_0 + \vec{u}) = f(\vec{x}_0) + f(\vec{u}) = \vec{0} + \varphi(\vec{u})$ .

(1.7.3) Das -  $f(\vec{x}) = \varphi(\vec{u})$  - ist eine wichtige Beziehung. Denn damit können wir unsere Zuordnung  $\vec{x} \mapsto f(\vec{x})$  wie folgt zerlegen:

$$\vec{x} = \vec{x}_0 + \vec{u} \xrightarrow{p} \vec{u} \xrightarrow{\varphi} \varphi(\vec{u}) = f(\vec{x}) \xrightarrow{j} f(\vec{x})$$

Dabei ist  $p = (V, \vec{x} \mapsto \vec{u}, U)$  eine lineare Projektionsabbildung mit  $Kern(p) = Kern(f)$ . Weiter ist  $\varphi = (U, \vec{u} \mapsto \varphi(\vec{u}), Bildf)$  ein Isomorphismus, also umkehrbar. Schließlich ist  $j = (Bild(f), \vec{y} \mapsto \vec{y}, W)$  eine Injektion. (Injektiv: höchstens eine Lösung.)

(1.7.4) Zusammen haben wir

$$\boxed{f = j \circ \varphi \circ p}$$

Das ist eine Darstellung, die allgemein für jede lineare Abbildung im endlichdimensionalen Fall gilt und zeigt, wie derartige lineare Abbildungen aufgebaut sind.

(1.7.5) Hier interessiert uns die Lösung der Bestimmungsgleichung  $f(\vec{x}) = \vec{b}$ . Diese Lösung können wir in drei aufeinander folgende Schritte aufspalten, die genau das wiedergeben, was man üblicherweise an Erfahrungen aus Einzelbeispielen gewinnt.  $f(\vec{x}) = \vec{b}$  läßt sich nämlich über die Einführung von zwei Hilfsvariablen  $\vec{u}$  und  $\vec{y}$  schreiben:

$$j(\vec{y}) = \vec{b} \quad \varphi(\vec{u}) = \vec{y} \quad p(\vec{x}) = \vec{u} .$$

Setzt man die Gleichungen ineinander ein, erhält man wieder die Ausgangsgleichung in der Form  $j(\varphi(p(\vec{x}))) = \vec{b}$ . Die Lösungsprozedur läuft entgegen der Richtung von  $f$ .

(1.7.6) Aber das ist nicht, was wir tun wollen. Vielmehr wollen wir die drei Gleichungen aus (1.7.5), die wir so erhalten, nacheinander - von links nach rechts - lösen. Die resultierenden  $\vec{x}$ , sind die Lösungen unserer ursprünglichen Gleichung.

(1.7.7) Beginnen wir mit  $j(\vec{y}) = \vec{b}$ . Da  $j(\vec{y}) = \vec{y}$  ist, hat diese Gleichung genau dann eine eindeutige Lösung, wenn  $\vec{y} \in \text{Bild}(f)$  ist. Also:

$$\boxed{j(\vec{y}) = \vec{b} \quad \text{hat die Lösung } \vec{y} = \vec{b} \text{ falls } \vec{b} \in \text{Bild}f \\ \text{und ist unlösbar, wenn } \vec{b} \notin \text{Bild}f}$$

Die Injektionsabbildung  $j$  entscheidet daher, ob die Gleichung lösbar oder unlösbar ist.

(1.7.8) Sei jetzt  $\vec{b} \in \text{Bild}f$ , die Gleichung also lösbar. Dann ist  $\varphi(\vec{u}) = \vec{b}$  zu lösen. Nun ist  $\varphi$  ein Isomorphismus, also:

$$\boxed{\varphi(\vec{u}) = \vec{b} \quad \text{mit } \vec{b} \in \text{Bild}f \text{ hat genau einen Lösung } \vec{u} = \varphi^{-1}(\vec{b}) \in U.}$$

$\varphi$  verschafft uns mithin eine spezielle Lösung des Ausgangssystems. Denn oben haben wir ja gesehen, dass  $p(\vec{u}) = f(\vec{x})$  ist. Das so gefundene  $\vec{u}$  ist der Vektor, den wir sonst mit  $\vec{x}_S$  (=spezielle Lösung der inhomogenen Gleichung) bezeichnet haben, nur dass wir jetzt eine Konstruktionsmethode für ihn haben.

(1.7.9) Das Lösen von  $p(\vec{x}) = \vec{u}$  verbleibt als letzter Schritt. Da  $p$  eine Projektionsabbildung mit  $\text{Kern}(p) = \text{Kern}(f)$  ist und wir ja eine Basis von  $\text{Kern}(f)$  gewählt haben, können wir die Lösung sofort angeben:

$$\boxed{p(\vec{x}) = \vec{u} \text{ hat folgende allgemeinen Lösung:} \\ \vec{x}_{\mathbb{L}}(\alpha_1, \dots, \alpha_k) = \vec{u} + \vec{a}_1\alpha_1 + \dots + \vec{a}_k\alpha_k \\ \text{oder} \quad \mathbb{L} = \vec{u} + \text{Kern}f}$$

Die Abbildung  $p$  verschafft uns somit die durch  $\text{Kern}(f)$  festgelegte Vieldeutigkeit der allgemeinen Lösung gleich in Form einer Parametrisierung.

(1.7.10) Wir sehen, wie die vertraute Struktur der Lösungsmengen linearer Gleichungssysteme zustande kommt: Der Freiheit bei der Wahl des Aufpunktes  $\vec{x}_S$  entspricht die Nichteindeutigkeit des zu  $\text{Kern}(f)$  supplementären Raumes  $U$ .

**5.1.7a Die korrespondierende Zerlegung der Matrix** (1.7.11) Wir haben unsere Basiswahl benutzt, um  $f$  zu zerlegen. Aber über die Basiswahl wird natürlich auch die beschreibende Matrix  $M^N$  festgelegt. Diese Matrix hat korrespondierend zur Einfachheit der drei Gleichungen eine besonders einfache Form. Sie repräsentiert die ideale Endform, die die Lösungskalküle für die Systemmatrix anstreben. Überlegen wir uns, wie sie aussieht. Natürlich ist das Kochrezept aus 4.(4.3.20) einfach stur anzuwenden.

(1.7.12) Ausgangspunkt unserer Konstruktion war die den Kern ergänzende Basis  $a$  von  $V$ . In  $W$  haben wir dann die Basis von  $\text{Bild}(f)$ , nämlich  $\vec{f}_1 = \varphi(\vec{a}_{k+1}), \dots, \vec{f}_\ell = \varphi(\vec{a}_n)$ . Diese ergänzen wir zu einer Basis  $\vec{f}_1, \dots, \vec{f}_m$  von ganz  $W$ . Das von den beiden Basen bestimmte Koordinatensystem indizieren wir mit  $N$ . **Wie sieht dann die  $M$  bezüglich  $N$  beschreibende neue Matrix  $M^N$  aus?**

(1.7.13) Nach dem Rezept müssen wir  $f(\text{Basis in } V)$  bilden und durch die Basis in  $W$  darstellen. Zunächst haben wir  $f(\vec{a}_i) = \vec{0}$  für  $i = 1, \dots, k$ . Diese Vektoren spannen ja  $\text{Kern}(f)$  auf. Das ergibt  $k$  Spalten mit Nullen. Dann haben wir  $f(\vec{a}_{k+j}) = \varphi(\vec{a}_{k+j}) = \vec{f}_j$  für  $j = 1, \dots, \ell$ . Aber  $f_j$  ist selbst ein Vektor unserer Basis für  $W$ . Die Basisdarstellung steht daher bereits da. Das Rezept gibt eine Spalte mit Nullen bis auf die  $j$ -te Stelle mit einer 1.

(1.7.14) Zusammen finden wir eine beschreibende Matrix der folgenden Form:

$$M^N = \begin{pmatrix} 0 & \dots & 0 & 1 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & & & & \vdots \\ \vdots & \vdots & \vdots & & & & & & & & \vdots \\ 0 & \dots & 0 & 0 & 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & \dots & & & & 0 & & 0 & & \dots & 0 \\ \vdots & \vdots & & & & & & & & \vdots & \vdots \\ \vdots & \vdots & & & & & & & & \vdots & \vdots \\ 0 & \dots & 0 & & & & & 0 & 0 & \dots & 0 \end{pmatrix}$$

Und damit:  $\text{Zeilenrang}(M^N) = \text{Spaltenrang}(M^N) = \ell$ .

Sollte Ihnen immer noch nicht klar sein, dass dies die ideale Endform eines Gleichungssystems ist, müssen Sie nur einmal darangehen, das erhaltene System  $M^N \vec{x}^N = \vec{b}^N$  zu lösen.

(1.7.15) Natürlich kann man die Zerlegung von  $f$ , also  $f = j \circ \varphi \circ p$  auch als Matrixprodukt darstellen. Man überlegt sich sofort:

$$M^N = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

## 5.1.8 Ein Beispiel

(1.8.1) Die Diskussion von Beispielen sollte nicht nach einem festgelegten Schema geschehen. Man sollte fallspezifische Besonderheiten berücksichtigen. Jedoch wird man immer bestrebt sein, entweder geeignete neue Basen zu konstruieren oder aber gleichwertig die beiden Transformationsmatrizen  $T$  und  $S$ . Hierzu gibt es verschiedene Möglichkeiten. Im jetzt folgenden Beispiel gehen wir so vor. Die Basis im Werteraum wird dabei etwas anders gewählt als soeben beschrieben. Derartige Konkretisierungen allgemeiner Verfahren sind in der Regel zur praktischen Lösungsbestimmung nicht besonders geeignet. Sie sollen eher die allgemeinen Resultate verdeutlichen.

(1.8.2)

$$M\vec{x} = \vec{b} \quad \text{mit} \quad M = \begin{pmatrix} 2 & 3 & -1 & 1 \\ 1 & -1 & -1/2 & -1/3 \\ 3 & 2 & -3/2 & 2/3 \end{pmatrix} \quad \text{und} \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

(1.8.3) Wir beobachten für die Zeilen  $\vec{z}_3 = \vec{z}_1 + \vec{z}_2$  und für die Spalten  $\vec{s}_1 = -2\vec{s}_3$  und  $\vec{s}_2 = 3\vec{s}_4$ . Der Rang von  $M$  ist offensichtlich 2. Es folgt  $k = \dim \text{Kern}M = 2$ .

(1.8.4) In Spaltenform schreibt sich die Gleichung  $\vec{s}_1 x_1 + \vec{s}_2 x_2 + \vec{s}_3 x_3 + \vec{s}_4 x_4 = \vec{b}$ . Oder unter Ausnutzung der Spaltenabhängigkeiten:

$$\vec{s}_3(-2x_1 + x_3) + \vec{s}_4(3x_2 + x_4) = \vec{b}$$

(1.8.5) Das legt die Einführung neuer Variabler nahe:

$$x_3^N = -2x_1 + x_3 \quad x_4^N = 3x_2 + x_4 \quad \text{und} \quad x_1^N, x_2^N \quad \text{noch nicht festgelegt.}$$

(1.8.6) Da  $\vec{x}^N = T\vec{x}^A$  gelten soll, wissen wir

$$T = \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ -2 & 0 & 1 & 0 \\ 0 & 3 & 0 & 1 \end{pmatrix}.$$

(1.8.7) Nun ist es leicht, mit Hilfe der neuen Spaltenform den Kern zu bestimmen. Wir erhalten nichttriviale Linearkombinationen der Ausgangsspalten, wenn wir  $x_3^N = 0$  und  $x_4^N = 0$  fordern, was ja die unmittelbar erkennbaren Abhängigkeiten der Spalten der Ausgangsmatrix  $M$  beschreibt. Also  $x_3 = 2x_1$  und  $x_4 = -3x_2$ .

(1.8.8) Das gibt für  $\text{Kern}M$ :

$$\vec{x}_K(x_1, x_2) = \begin{pmatrix} x_1 \\ x_2 \\ 2x_1 \\ -3x_2 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ -3 \end{pmatrix} = \vec{a}_1 x_1 + \vec{a}_2 x_2.$$

(1.8.9)  $(\vec{a}_1, \vec{a}_2)$  bildet eine Basis von  $\text{Kern}M$ . Wir ergänzen Sie zu einer Basis von ganz  $V$ . Beispielsweise durch  $\vec{a}_3 = \vec{e}_3$  und  $\vec{a}_4 = \vec{e}_4$ . Für die zugehörige Transformationsmatrix  $T$  (des Urbildraumes) lesen

wir unmittelbar ab (neue Basis durch alte, also  $T^{-1}$ ):

$$T^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 0 & -3 & 0 & 1 \end{pmatrix} \quad \text{Inversion gibt} \quad T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & 0 & 1 & 0 \\ 0 & 3 & 0 & 1 \end{pmatrix}$$

(1.8.10) Damit ist für die Urbildseite unseres Diagramms alles bestimmt.

(1.8.11) Auf der Werteseite wissen wir, dass die dritte Zeile gleich der Summe der ersten beiden ist. Wir führen als neue Variable anstelle von  $y_3$  daher  $y_3^N = y_3 - y_1 - y_2$  ein. Die beiden anderen Variablen verändern wir nicht:  $y_2^N = y_2$  und  $y_1^N = y_1$ . Aus  $\vec{y}^N = S\vec{y}^A$  lesen wir die zugehörige Transformationsmatrix  $S$  ab:

$$S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix} \quad \text{Und damit:} \quad S^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

(1.8.12) Jetzt können wir die neue Abbildungsmatrix  $M^N = SMT^{-1}$  bestimmen. Wir finden:

$$M^N = SMT^{-1} = \begin{pmatrix} 0 & 0 & -1 & 1 \\ 0 & 0 & -1/2 & -1/3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

(1.8.14) Die ersten beiden Spalten sind Null. Das repräsentiert den zweidimensionalen Kern. Die letzte Zeile ist Null, das ist der Teil von  $W = \mathbb{R}^3$ , der nicht im Bild liegt. Die  $2 \times 2$ -Matrix, die  $\varphi$  repräsentiert, ist nicht die Einheitsmatrix, weil wir in  $W$  darauf verzichtet haben,  $\varphi(\vec{a}_3)$  und  $\varphi(\vec{a}_4)$  als Basis von  $BildM$  zu nehmen. Stattdessen haben wir die kanonische Basis so modifiziert, dass die beobachtete Zeilenabhängigkeit ausgenutzt wurde und die Nullen der letzten Zeile produzierte. Über  $S^{-1}$  sehen wir, wie diese andere Basis von  $W$  aussieht:  $\vec{f}_1 = \vec{e}_1 + \vec{e}_3$ ,  $\vec{f}_2 = \vec{e}_2 + \vec{e}_3$  und  $\vec{f}_3 = \vec{e}_3$ . Der letzte Vektor  $\vec{f}_3$  liegt nicht in  $BildM$ , erzeugt eine dazu supplementäre Gerade.

(1.8.15) **Wie erhält man das neue System unmittelbar?** Die Ausgangsgleichung  $M\vec{x} = \vec{b}$  wird mit  $S$  multipliziert und wie folgt geschrieben:  $(SMT^{-1})T\vec{x} = S\vec{b}$ . Das entstehende neue System  $M^N\vec{x}^N = \vec{b}^N$  ist leicht zu lösen. Es ist lösbar, wenn  $b_3^N = 0$  ist. D.h. in den alten Koordinaten  $b_3 = b_1 + b_2$ , wie man unmittelbar im Ausgangssystem verifiziert. Bei Lösbarkeit sind weiter  $x_1^N$  und  $x_2^N$  frei wählbar, und die beiden Bedingungen für  $x_3^N$  und  $x_4^N$  geben sofort

$$x_3^N = -\frac{2}{5}(b_1^N + 3b_2^N) \quad \text{und} \quad x_4^N = \frac{3}{5}(b_1^N - 2b_2^N)$$

Mit Hilfe von  $\vec{x} = \vec{x}^A = T^{-1}\vec{x}^N$  und  $\vec{b} = \vec{b}^A = S^{-1}\vec{b}^N$  kann man dann zu den ursprünglichen Variablen zurückkehren und die Lösung hinschreiben. .

## 5.1.9 Zeilenrang gleich Spaltenrang. Erster Beweis.

(1.9.1) Abschließend können wir als allgemeine Anwendung der dargestellten Methode des Basiswechsels einen ersten Beweis der Gleichheit der beiden Ränge geben. Wir argumentieren in zwei Schritten.

(1.9.2) Zuerst zeigen wir, dass ein Basiswechsel im Werteraum den Zeilenrang nicht beeinflusst. (Den Spaltenrang natürlich auch nicht, denn dieser ist basisunabhängig die Dimension des Bildes.) Im zweiten Schritt wählen wir dann eine spezielle Werteraumbasis, für deren beschreibende Matrix wir Zeilenrang = Spaltenrang nachweisen können. Zusammen folgt unsere Behauptung.

(1.9.3) Den ersten Schritt beweisen wir unter Verwendung des Indexkalküls. Was ist zu zeigen? Angenommen es existiert eine nichttriviale Beziehung  $\sum \alpha_i \vec{z}_i^A = \vec{0}$  zwischen den Zeilen der Ausgangsmatrix. Dann müssen wir daraus eine nichttriviale Beziehung  $\sum \beta_i \vec{z}_i^N = \vec{0}$  der Zeilen der transformierten Matrix konstruieren und umgekehrt. (Als Folge darf man für alt und neu stets die **gleiche Zahl von Zeilen** streichen). Sei  $S$  die Transformationsmatrix des Basiswechsels, also  $M^N = SM^A$ . Weiter ist der  $i$ -te Zeilenvektor die  $i$ -te Zeile der Matrix, also  $(\vec{z}_i)_r = M_{ir}$ . ( $i$  ist hier konstant,  $r$  frei.). Dann folgt ( $\sum \alpha_i \vec{z}_i^A = \vec{0}$  ist Voraussetzung,  $\sum \beta_i \vec{z}_i^N = \vec{0}$  das Resultat):

$$0 = \vec{0}_r = \sum \alpha_i (\vec{z}_i^A)_r = \sum \alpha_i M_{ir}^A = \sum \alpha_i S_{is}^{-1} M_{sr}^N = \sum (\sum_i \alpha_i S_{is}^{-1}) M_{sr}^N = \sum \beta_s M_{sr}^N = \sum \beta_s (\vec{z}_s^N)_r.$$

Mit  $\beta_s = \sum_i \alpha_i S_{is}^{-1}$ . Und die  $\beta_s$  können nicht alle Null sein, weil andernfalls  $S^{-1}$  nicht invertierbar wäre. Vertauschung von  $A$  und  $N$  gibt die andere Richtung.

(1.9.4) Damit ist der erste Teil gezeigt: **Basiswechsel im Werteraum ändert den Zeilenraum nicht**. Das entspricht unserer Vorstellung, dass der Zeilenrang die Anzahl der unabhängigen Bedingungen des homogenen Systems erfaßt.

(1.9.5) Jetzt wählen wir eine Basis von  $BildM$  und ergänzen sie zu einer Basis von  $W$ . Die neue Matrix hat daher denselben Zeilenrang wie die alte. Da die letzten  $m - \ell$  Zeilen alle Null sind, ist  $rg_z(M) \leq \ell$ . Aber der Zeilenrang kann nicht  $< \ell$  sein. Die  $n \times \ell$ -Matrix, die durch Streichen der letzten  $m - \ell$  Zeilen entsteht, gehört zur Gleichung  $\varphi \circ p(\vec{x}) = \vec{b}$ , die für alle  $\vec{b} \in BildM$  lösbar ist. Wären die Zeilen noch abhängig, wäre die Gleichung nur lösbar, wenn zwischen den Komponenten von  $\vec{b}$  dieselbe Beziehung bestünde, und das muß nicht der Fall sein. **Damit ist Zeilenrang gleich Spaltenrang bewiesen.**

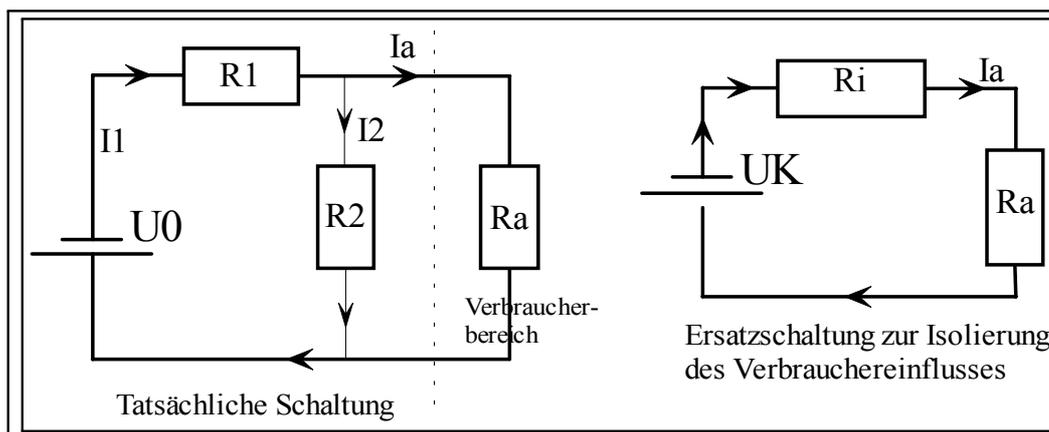
(1.9.6) Ganz befriedigend ist der Beweis nicht. Der Spaltenrang hat eine geometrische Bedeutung. Für den Zeilenrang haben wir eine solche Interpretation bisher nicht, und der Beweis liefert eine solche nicht ohne weiteres. Wir haben bereits angedeutet, dass wir hierzu den Dualraum benötigen und werden diese Interpretation im nächsten Kapitel entwickeln.

(1.9.7) Zugleich werden wir noch eine weitere Lücke in unserem Verständnis linearer Gleichungen schließen. Bisher haben wir alle Äquivalenzumformungen unseres Systems als Basiswechsel interpretiert. Aber was ist, wenn wir die Zahl der Gleichungen verändern? Es gibt auch solche Äquivalenzumformungen. Wie sehen diese aus, welche sind zulässig? Sind es nur die, die wir üblicherweise verwenden oder gibt es noch andere trickreichere? Die geometrische Interpretation des Zeilenranges wird auch hierauf eine Antwort geben.

(1.9.8) Gibt es noch überhaupt noch andere eventuell nützliche Äquivalenzumformungen? Die Bedeutung dieser Frage sollte man nicht unterschätzen und damit auch nicht die Leistung des nachfolgenden Teilkapitels, die alle überhaupt möglichen Umformungen liefert. Um die Bedeutung dieser Leistung zu verdeutlichen, bringen wir noch ein Beispiel von Umformungen mit bemerkenswerten Konsequenzen.

## 5.1.10 Fortgeschrittene Umformungen eines Gleichungssystems

(1.10.1) Wir betrachten ein Beispiel aus der Elektrotechnik. Es geht um das Problem, komplexe elektronische Schaltungen für einen bestimmten Typ von Anwendungen durch eine einfache Ersatzschaltung zu ersetzen. Wir wählen dazu ein relativ einfaches Beispiel. Die Methode läßt sich problemlos verallgemeinern, nur dass die Formulierung der Verallgemeinerung relativ viel Formalismus verlangt, was wir hier vermeiden wollen.



(1.10.2) Wir wissen: Legt man im Netzwerk der linken Seite eine bestimmte Spannung  $U_0$  an, dann bewirkt das in jedem der Zweige einen Strom eindeutig bestimmter Stärke. Formal erhält man dies Resultat, indem man mit Hilfe der Kirchhoffschen Regeln das System durch ein lineares Gleichungssystem (für die Stromstärken) ersetzt, das eindeutig lösbar ist, also zu  $k = 0$  gehört.

(1.10.3) Die Situationen, die uns interessieren, sind solche, in denen  $R_a$  als Verbraucher interpretiert wird. Der Rest der Schaltung sei stärker fixiert. D.h. es interessiert, welche Ströme durch  $R_a$  fließen und wie groß der zugehörige Spannungsabfall ist (in Abhängigkeit von  $R_a$ ). Die oben gegebene Ersatzschaltung

leistet das in einfacher und naheliegender Weise. Das Verfahren ist so, dass man auch für weitgehend beliebige andere Netzwerke eine analoge Ersatzschaltung erhält. Zwei Probleme entstehen:

- 1) Ist die Behauptung wirklich korrekt: Hat die Ersatzschaltung dasselbe Verhalten wie das Ausgangssystem?
- 2) Wie erhält man die beiden Hilfsgrößen  $R_i$  und  $U_K$ , die man ja als Parameter der Ersatzschaltung benötigt?

(1.10.4) Inspiziert man die vorgegebene Schaltung, so erscheint die erste Behauptung doch überraschend: Kann man den parallelen Zweig mit  $R_2$  einfach dadurch beseitigen, dass man die Parameter ändert? Und das soll sogar immer möglich sein, gleichgültig wie kompliziert die Gesamtschaltung aufgebaut ist?

(1.10.5) Wir wollen das über die Kirchhoffschen Regeln entstehende Gleichungssystem jetzt nicht einfach lösen, schon gar nicht schematisch, sondern es so umformen, dass dadurch beide Fragen beantwortet werden. Kurz, wir nehmen Umformungen des Gleichungssystems vor, die nicht das Ziel haben, die Lösung zu bestimmen, sondern die statt dessen die obigen Ersatzsysteme liefern.

(1.10.6) Jetzt die Umformungen:

$$\begin{array}{r|l} I_1 & -I_2 = I_a & +0 \\ I_1 R_1 & +I_2 R_2 = 0 & +U_0 \\ \hline & I_2 R_2 = I_a R_a & \end{array}$$

In unserem Beispiel sieht das Gleichungssystem wie angegeben aus. Dabei sind alle Widerstände äußere Parameter, ebenso wie die angelegte Spannung  $U_0$ . Alle Ströme sind unbestimmt. Für nichtnegative Widerstände sollte es eine eindeutige Lösung geben.

Mit diesem System führen wir eine Reihe von Manipulationen aus, die durch die Darstellung bereits angedeutet sind:

- Die letzte Gleichung wird abgetrennt. Das verbleibende System ist dann unterbestimmt. Dies kompensieren wir dadurch, dass wir  $I_a$  zu einem äußeren Parameter machen. Dann sollte das Restsystem erneut eindeutig lösbar sein.
- Aus dem Restsystem machen wir zwei Systeme, indem wir den Inhomogenitätenvektor in der angedeuteten Weise aufteilen. Die gesuchte Lösung folgt per Superposition aus den beiden Einzellösungen.
- Nimmt man die abgetrennte Gleichung hinzu, haben wir aus einem System drei Systeme gemacht, die zusammen gleichwertig zum Ausgangssystem sind.

Wir wollen zeigen, dass jedes der beiden neuen Gleichungssysteme eine der gesuchten Größen  $U_K$  und  $R_i$  festlegt. **Und die abgetrennte Gleichung wird zum Ohmschen Gesetz für die Ersatzschaltung.**

**Das Leerlaufsystem:** (Die zugehörigen Ströme bezeichnen wir mit  $I'$ ) Dies ist das alte System mit geöffnetem Verbraucherzweig ( $I_a = 0$ )

$$\begin{array}{r|l} I'_1 & -I'_2 = 0 & +0 \\ R_1 I'_1 & +R_2 I'_2 = 0 & +U_0 \\ & \dots & \end{array}$$

Dies System kann man lösen und damit den Leerlaufspannungsabfall zwischen den Klemmen des Verbrauchers bestimmen, hier  $U_K = I'_2 R_2$ .

Die Rechnung ergibt:  $U_K = \frac{U_0 R_2}{R_1 + R_2}$ . Die Ersatzspannung hängt also nur von den Größen des reduzierten Systems ab, nicht von  $R_a$ , wie man natürlich bereits der Systemmatrix entnehmen kann.

**Das invertierte deaktivierte System:** (die zugehörigen Ströme bezeichnen wir mit  $I''$ .)

$$\begin{array}{r|l} I''_1 & -I''_2 = I_a & +0 \\ R_1 I''_1 & +R_2 I''_2 = 0 & 0 \\ & \dots & \end{array}$$

Im alten System wird die Spannungsquelle kurzgeschlossen ( $U_0 = 0$ ). Statt des Verbrauchers wird an den Ausgangsklemmen eine Spannung  $U_a$  angelegt, die gerade einen zugehörigen Strom  $I_a$  bewirkt. (Rollenverteilung: Zunächst  $I_a$  äußerer Parameter,  $U_a$  unbestimmt.)

Für dieses System ergibt sich der innere Widerstand  $R_i$  des Systems über  $U_a = R_i I_a$ . Den inneren Widerstand kann man mit den üblichen Methoden bestimmen. Der Spannungsabfall zwischen den anstelle des Verbrauchers nach außen gehenden Klemmen ist  $I_2'' R_2$ . Also haben wir  $I_2'' R_2 = (-1) R_i I_a$ .

Für unser Beispiel ergibt die Rechnung  $R_i = \frac{R_1 R_2}{R_1 + R_2}$ .

**Das Ersatzsystem.** Abschließend kehren wir durch Hinzunahme der fehlenden Gleichung  $R_2 I_2 = R_a I_a$  zum Gesamtgleichungssystem zurück.  $I_a$  erhält in dieser letzten Gleichung erneut die Rolle einer Unbestimmten. Und  $I_2$  hängt von dieser Unbestimmten ab:  $I_2 = I_2' + I_2''$ . erhalten wir per Superposition. Ergebnis:

$$R_2 I_2 = R_2 I_2' + R_2 I_2'' = U_K - R_i I_a$$

Anwenden der letzten Gleichung gibt  $U_K - I_a R_i = I_a R_a$ . Oder als Endergebnis  $U_K = (R_a + R_i) I_a$ . **Das ist aber gerade das Ohmsche Gesetz für die Ersatzschaltung wie oben behauptet.** Diese Gleichung legt wie verlangt den Verbraucherstrom  $I_a$  fest.

(1.10.7) Das Ganze können wir auch so interpretieren: Durch die Gleichungsmanipulation wurde die Abhängigkeit des Stromes  $I_2$  von  $I_a$  in die Form  $R_2 I_2 = U_K - R_i I_a$  gebracht, und die Hilfsgrößen  $U_K$  und  $R_i$  wurden über zwei Hilfsschaltungen bestimmt.

(1.10.8) Die vorgenommenen Umformungen des Gleichungssystems - Aufteilung in drei Gleichungssysteme usw. - sind komplizierter als die üblichen. Sie liefern am Ende auch das übliche Resultat, aber in einer für das Problem viel geeigneteren Form.

(□ F. ) Überlegen Sie sich, wie man die einzelnen vorgeführten Schritte auf andere Netzwerke übertragen kann.

### 5.1.11 Eine Methode zur Berechnung der inversen Matrix

Wir wollen jetzt noch kurz andeuten, wie man mit Hilfe der Resultate inverse Matrizen berechnet. Eine einfache Modifikation dieser Überlegungen liefert übliche Lösungskalküle für lineare Gleichungssysteme. Im Gegensatz zum Beweis von "Zeilenrang=Spaltenrang" verwenden wir jetzt den Matrixkalkül.

Ausgangspunkt sind einige einfache Beobachtungen zur Matrixmultiplikation, die wir am Beispiel  $n = 3$  illustrieren. Sei  $M$  eine beliebige  $3 \times 3$ -Matrix, deren Komponenten wir zum Zwecke besserer Übersicht wie folgt ansetzen:

$$M = \begin{pmatrix} a & b & c \\ A & B & C \\ x & y & z \end{pmatrix}.$$

Jetzt betrachten wir zwei spezielle Matrizen  $E_1$  und  $E_2$  und bilden die (sofort nachrechenbaren) Matrixprodukte  $E_1 M$  und  $E_2 M$ :

$E_1 = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$	$E_1 M = \begin{pmatrix} \alpha a & \alpha b & \alpha c \\ x & y & z \\ A & B & C \end{pmatrix}$
$E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \alpha & \beta & 1 \end{pmatrix}$	$E_2 M = \begin{pmatrix} a & b & c \\ A & B & C \\ x + \alpha a + \beta A & y + \alpha b + \beta B & z + \alpha c + \beta C \end{pmatrix}$

$E_2 M$  etwa entsteht wie folgt: Die neue dritte Zeile erhält man, indem zur alten das  $\alpha$ -fache der ersten und das  $\beta$ -fache der zweiten hinzugezählt wird. Die beiden anderen Zeilen bleiben unverändert.

Inspektion dieser Beispiele zeigt: Es gibt spezielle und invertierbare Matrizen  $E$ , für die man bei gegebenem  $M$  das Produkt  $EM$  durch eine der folgenden "elementaren Zeilenoperationen" gewinnt:

- Eine Zeile von  $M$  wird mit einem konstanten Faktor  $\alpha \neq 0$  multipliziert
- Zwei Zeilen werden vertauscht

- Zu einer Zeile wird eine Linearkombination der übrigen Zeilen hinzugezählt.

Nochmals. Die Aussage lautet: **Zu jeder dieser Zeilenoperationen läßt sich eine invertierbare Matrix E finden, so dass EM die beschriebene Gestalt hat, egal, wie M sonst aussieht.** Oder auch: **Jede Matrixumformung der beschriebenen Art läßt sich interpretieren als Multiplikation von links mit einer geeigneten invertierbaren Matrix E.**

- (□ **F.** ) Wie sehen diese Matrizen  $E$  für die beschriebenen 3 Operationen für allgemeines  $n$  aus?
- (□ **F.** ) Kann man entsprechende Resultate für die Produktreihenfolge  $ME$  herleiten?
- (□ **F.** ) Zeigen sie, dass alle diese Matrizen wirklich invertierbar sind.

**Jetzt die eigentliche Idee** (zur Bestimmung der inversen Matrix).

- Sei  $X$  eine invertierbare quadratische  $n \times n$ -Matrix. Dann gilt  $XX^{-1} = \mathbb{I}_n$  wobei  $\mathbb{I}_n$  die zugehörige Einheitsmatrix ist. Diese Gleichung multiplizieren wir nacheinander von links mit geeigneten Matrizen des beschriebenen Typs. Also zuerst:

$$E(XX^{-1}) = (EX)X^{-1} = E\mathbb{I}_n = E. \quad \text{Oder} \quad (EX)X^{-1} = E.$$

Dabei haben wir das Assoziativgesetz der Matrixmultiplikation benutzt.

- Nun wissen wir aber, wie man  $EX$  aus dem gegebenen  $X$  erhält: Durch eine elementare Zeilenoperation. Wir multiplizieren so lange weiter, bis wir  $E_k E_{k-1} \dots E_1 X = \mathbb{I}_n$  erhalten. Dann wird aus der Ausgangsgleichung insgesamt:

$$(E_k E_{k-1} \dots E_1 X) X^{-1} = X^{-1} = E_k E_{k-1} \dots E_1 \quad \text{Oder} \quad \boxed{X^{-1} = E_k E_{k-1} \dots E_1}$$

**D.h. das Produkt der benutzten E-s ist die gesuchte inverse Matrix!**

Das Gesagte läßt sich leicht zu einem Kalkül ausbauen:

- Wähle drei Spalten mit ausreichend vielen Zeilen. Links oben wird  $X$  eingefügt, rechts oben  $\mathbb{I}_n$ . In der Mitte wird eine auszuführende Zeilenoperation notiert.
- Die notierte Zeilenoperation wird auf beide Matrizen angewandt, was den zugehörigen Matrixeintrag der nächsten Zeile liefert.
- In die (noch freie) Mitte wird eine weitere Zeilenoperation notiert. Usw.

**Wie sind die Zeilenoperationen zu wählen?** So dass die Matrix der linken Spalte sich der Einheitsmatrix nähert. Ist diese erreicht, steht rechts die gewünschte inverse Matrix.

**Mit welcher Strategie sollte man die Zeilenoperationen wählen?** Von einer Ecke, die nicht zur Diagonalen gehört, muß man sich einen ausbreitenden Bereich von Nullen unter oder über der Diagonalen verschaffen. Dann dasselbe für die andere Seite der Diagonalen. Erst am Ende die Diagonalglieder zu 1 machen. Mit etwas Geschick lassen sich mehrere Operationen auf einmal durchführen. Nach spätestestens  $n^2$  Schritten ist man fertig.

Brüche möglichst lange meiden, möglichst über Faktormultiplikation beseitigen.

Falls möglich, ganz zu Anfang durch Vertauschen der Reihenfolge der Zeilen in einer Ecke bereits einen Nullbereich schaffen.

Ein erstes Beispiel mit vertrautem Resultat mit der Hilfsgrößen  $D = ad - bc$ . Die Notation der mittleren Spalte sollte sich von selbst erklären. Sei  $D \neq 0$ .

$X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$	1 $a(2) - c(1)$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
$\begin{pmatrix} a & b \\ 0(!) & D \end{pmatrix}$	$D(1) - b(2)$ (2)	$\begin{pmatrix} 1 & 0 \\ -c & a \end{pmatrix}$
$\begin{pmatrix} Da & 0(!) \\ 0 & D \end{pmatrix}$	(1)/ $Da$ (2)/ $D$	$\begin{pmatrix} D - bc & -ba \\ -c & a \end{pmatrix}$
$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	Fertig!	$X^{-1} = \frac{1}{D} \begin{pmatrix} c & -b \\ -c & a \end{pmatrix}$

Wichtig ist, in der Mitte eine Notation zu verwenden, die genau angibt, was zu tun ist, so dass man bei Rechenfehlern die Rechnung später kontrollieren kann. Beachten Sie: Vereinzelte Nullen taugen nichts, sie müssen sich von einer Ecke her ausbreiten!

Nehmen wir als weiteres Beispiel eine einfache  $3 \times 3$ -Matrix

$X = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 3 & 0 & 0 \end{pmatrix}$	$\begin{matrix} (3) \\ 2(1) - (2) \\ (2) \end{matrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
$\begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$	$\begin{matrix} (1)/3 \\ (2)/2 \\ (3)/2 \end{matrix}$	$\begin{pmatrix} 0 & 0 & 1 \\ 2 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$
$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	Fertig	$X^{-1} = \begin{pmatrix} 0 & 0 & 1/3 \\ 1 & -1/2 & 0 \\ 0 & 1/2 & 0 \end{pmatrix}$

(□ **F.** ) Wie ist die Methode abzuwandeln, wenn man ein Gleichungssystem  $M\vec{x} = \vec{b}$  mit quadratischem und invertierbarem  $M$  lösen möchte? (Nur die rechte Spalte ist zu ändern!)

(□ **F.** ) Wie ist die Methode abzuändern, wenn  $M$  nicht invertierbar ist? Dann kann man die linke Spalte nicht zu einer Einheitsmatrix machen. Aber welche Matrixendform genügt für die linke Spalte, um das Gleichungssystem zu lösen? ?

## 5.2 Der Dualraum ( eines endlichdimensionalen Vektorraums )

### 5.2.0 Vorbemerkung

Um die in 5.1.9 aufgeworfenen Probleme sinnvoll beantworten zu können, benötigen wir den Formalismus des Dualraumes. Dabei handelt es sich um eine wichtige Ergänzung zu Kapitel 4, die allerdings einige geistige Arbeit erfordert. Insbesondere deshalb, weil hier eine Struktur vorliegt, die nicht leicht und unmittelbar an die Anschauung angekoppelt werden kann. Andererseits erweist sie sich als sehr wichtig und nützlich. Wir werden ihr in den Kapiteln 9-11 wieder begegnen. Im Bereich der Differentialgeometrie kommt ihr schließlich (unter dem Stichwort Differentialformen ) allergrößte Bedeutung zu.

### 5.2.1 Zur Notwendigkeit der Dualraumstruktur

(2.1.1) In 4.(4.4.2) Haben wir allgemein die Vektorräume  $\text{Hom}_K(V, W)$  der Homomorphismen  $V \rightarrow W$  nebst ihren Matrixdarstellungen eingeführt. Wir wollen uns mit einem Spezialfall dieser Räume befassen, nämlich  $\text{Hom}_K(V, K)$ , eigentlich  $\text{Hom}_K(V, K^1)$ . Dies sind die linearen Abbildungen  $V \rightarrow K^1$ . Dass wir das im Zusammenhang mit den linearen Gleichungen tun, ist plausibel. Denn eine Einzelgleichung ist ein  $1 \times n$ -System und damit ein Element dieses Raumes. Ein  $m \times n$ -System besteht dann aus  $m$  derartigen Elementen, wie mit dem Stichwort Zeilenvektorform in (1.2.1) bereits dargelegt worden ist.

(2.1.2) Zuerst stellen wir zugehörige Definitionen und Bezeichnungen zusammen.  $K$  sei kommutativer Körper,  $V$  Rechtsvektorraum über  $K$ , endlichdimensional mit  $\dim V = n < \infty$ . Dann hat der zugehörige Dualraum  $V^* = \text{Hom}_K(V, K)$  dieselbe Dimension  $n$ . Wir schreiben auch  $V^*$  als Rechtsvektorraum, was nicht für alle Zwecke optimal ist.

Die Elemente von  $V^*$  werden "Linearformen von  $V$ " oder "lineare Funktionale auf  $V$ " genannt.  $\lambda \in V^*$  kann in zwei Rollen auftreten: Einerseits als Abbildung  $V \rightarrow K$ , andererseits aber auch als Vektor im Sinne eines Elementes aus einem Vektorraum.

(□ **F.** ) Der Vektorraum  $\text{Hom}_K(K^1, V)$  läßt sich kanonisch mit  $V$  identifizieren. Was nimmt man als Identifikationsabbildung? Welche geometrische Interpretation hat diese Identifikation?

(2.1.3) Zwischen  $V$  und  $V^*$  gibt es generell keinen kanonischen Isomorphismus, der eine Identifizierung der Elemente erlauben würde. **Stattdessen gibt es eine bijektive Beziehung zwischen den Teilräumen von  $V^*$  und  $V$ .** Diese Beziehung ist so angelegt, dass jedem  $k$ -dimensionalen Teilraum des einen Raumes

ein  $(n - k)$ -dimensionaler Teilraum des anderen entspricht. Wir haben es mit einer bijektiven Abbildung  $\mathcal{T}(V) \rightarrow \mathcal{T}(V^*)$  zu tun.

(2.1.4) Diese Beziehung spiegelt sich in einer Vielzahl von Erscheinungen und Anwendungen wieder. Ein vertrautes Beispiel zur anschaulichen Verdeutlichung wird durch die Beziehung "Gerade - dazu senkrechte Ebene" im euklidischen  $V_0^3$  geliefert. Das Beispiel kann man meist dazu verwenden, eine sinnvolle Vermutung über zu erwartende Resultate entwickeln.

(2.1.5) Leider schießt das Beispiel in anderer Hinsicht über das Ziel hinaus, insofern es die Vorstellung fördert, Raum und Dualraum zu identifizieren, was im Beispiel angemessen und möglich ist. Komplexere Anwendungen fordern jedoch eine Trennung der beiden Räume. Dieser Unterschied zur anschaulichen Vorstellungswelt dürfte einer von mehreren Gründen sein, weshalb der Dualraumbegriff Anfängern Schwierigkeiten bereitet.

(2.1.6) Neben der Beziehung zwischen den Teilräumen besteht eine Beziehung zwischen den Basen. Zu jeder Basis in  $V$  gehört eine eindeutig bestimmte Basis von  $V^*$  und umgekehrt. Diese *zueinander dualen Basen* erweisen sich als ein für viele Zwecke nützliches Hilfsmittel. Dualen Basen entsprechen den reziproken Basen im  $V_0^3$ .

(2.1.7) Wir wollen jetzt einige inhaltliche Argumente geben, weshalb man Raum und Dualraum auseinanderhalten sollte.

(2.1.8) ♦ Im unendlichdimensionalen Fall ist eine Identifizierung nicht mehr möglich, weil beide Räume nicht isomorph sind.

(2.1.9) ♦ In der Physik findet man Vektoren unterschiedlichen Typs - wie Geschwindigkeitsvektoren und Verschiebungsvektoren - für die eine Addition sicherlich unsinnig ist. Sie sollten also **nicht** in demselben Vektorraum untergebracht werden. Andererseits besteht eine Beziehung vom Typ eines Skalarproduktes zwischen ihnen (Arbeit), und das Paar Raum-Dualraum erlaubt die Formalisierung dieser Beziehung.

♦ Ein ähnliches Beispiel verdeutlicht Bedeutung und Häufigkeit der durch die Arbeit eingeführten Struktur: Eine Menge von Waren sei mit  $1, 2, \dots, N$  indiziert. Ein erster Vektor  $(p_1, \dots, p_N)$  gebe einen Einzelpreis für jede Ware. Ein zweiter Vektor  $(n_1, \dots, n_N)$  gebe für jede Ware eine Anzahl. Addition beider Vektoren ist offensichtlich sachlich unsinnig, obwohl die Dimensionen übereinstimmen und sogar in beiden Fällen eine kanonische Basis vorliegt. Aber die Verknüpfungsgröße  $\sum n_i p_i$ , mit der Bedeutung eines Gesamtpreises ist möglich und sachlich wichtig. Wir werden sehen, dass auch hier die Räume dual zueinander sind.

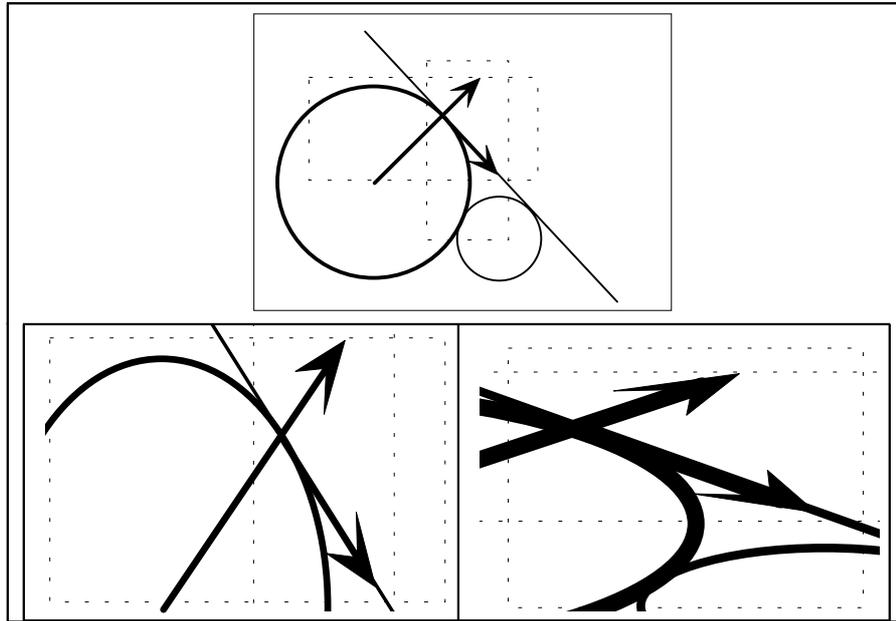
(2.1.10) ♦ Stellen wir uns ein Computerhirn vor, das Ausschnitte einer physikalisch-geometrischen Ebene mitsamt beschreibenden Vektorpfeilen auf einem Bildschirm darstellt. Nehmen wir an, dass u.a. eine Kurve mit einem Geschwindigkeitsvektor und einem Normalenvektor an die Kurve dargestellt werden soll. Und zwar soll der Computer diese Konfiguration über zwei verschiedene rechteckige **nicht ähnliche** Ausschnitte auf dem Bildschirm darstellen. Falls er beim Wechsel des Ausschnittes beide Vektorpfeile gleich behandelt, kommt offensichtlicher Unfug heraus. Entweder erscheint der eine Pfeil nicht mehr senkrecht auf dem zweiten, oder der Geschwindigkeitspfeil erscheint nicht mehr tangential zur Kurve. Die Figuren illustrieren das. Hält man Raum und Dualraum auseinander, so findet man, dass der Geschwindigkeitsvektor zu  $V_0^2$ , der Normalenvektor dagegen zum Dualraum  $(V_0^2)^*$  gehört. Sobald der Computer dies berücksichtigt, ergeben sich zu jeder Ausschnittwahl korrekte Bilder. Wie das genau geschieht, werden wir im Verlaufe dieses Kapitels sehen. Offensichtlich liegt wieder ein Basiswechsel vor.

Etwa genauer: Der ursprüngliche Bildschirmausschnitt werde durch die beiden Seitenvektoren  $\vec{b}$  und  $\vec{h}$  begrenzt, der neue durch  $\vec{B} = \vec{b}\beta$  und durch  $\vec{H} = \vec{h}\gamma$ . Dann lautet die Transformationsmatrix (von alt nach neu) gemäß Kochrezept:

$$T = \begin{pmatrix} \frac{1}{\beta} & 0 \\ 0 & \frac{1}{\gamma} \end{pmatrix}. \quad \text{Also } \vec{x}^N = T\vec{x}^A.$$

Die beiden Vektorpfeile  $\vec{x}^A = \begin{pmatrix} 7 \\ 2 \end{pmatrix}$  und  $\vec{y}^A = \begin{pmatrix} -2 \\ 7 \end{pmatrix}$  stehen im alten Bildschirmausschnitt aufeinander senkrecht. Ihre Bilder  $\vec{x}^N = \begin{pmatrix} \frac{7}{\beta} \\ \frac{2}{\gamma} \end{pmatrix}$  und  $\vec{y}^N = \begin{pmatrix} -\frac{2}{\beta} \\ \frac{7}{\gamma} \end{pmatrix}$  im neuen Bildschirmausschnitt stehen für  $\beta \neq \gamma$  nicht aufeinander senkrecht. Bildet man dagegen die Matrix

$T^{-1} = \begin{pmatrix} \beta & 0 \\ 0 & \gamma \end{pmatrix}$  und transformiert  $\vec{y}$  damit, so folgt  $\vec{y}^N = T^{-1}\vec{y}^A = \begin{pmatrix} -2\beta & 0 \\ 0 & 7\beta \end{pmatrix}$ , so folgt  $(\vec{x}^N \cdot \vec{y}^N) = 0$ . D.h. die so transformierten Pfeile stehen auch im neuen verzerrten Bild aufeinander senkrecht.



### 5.2.1a Übersicht über die weitere Behandlung der Dualraumstruktur

**(2.1.11)** In einem ersten Teil dieses Kapitels wollen wir uns mit Schreibweisen und Formalismen befassen, um die Resultate besser merken und oder gar intuitiv erahnen zu können und andererseits, um den Rechenkalkül zu verbessern und den Umgang damit zu üben. Insbesondere müssen wir den Dualraumbegriff an den Matrix- und Indexformalismus anschließen.

Danach führen wir den Begriff der *zueinander dualen Basen* ein und schließlich behandeln wir die angekündigte Beziehung zwischen den Teilräumen. Beim Beweis ist eine nichttriviale Idee erforderlich.

### 5.2.2 Das kanonische Skalarprodukt

**(2.2.1)** Neben der inhaltlichen Begründung, wieso man zwischen  $V$  und  $V^*$  unterscheiden sollte, gibt es noch die rein formale Unterscheidung und den damit verbundenen Kalkül.  $\lambda \in V = \text{Hom}_K(V, K)$  bedeutet ja, dass  $\lambda = (V, \vec{x} \mapsto \lambda(\vec{x}), K)$  eine Abbildung ist, also ein Automat, **der aus einem eingegebenen Vektor eine Zahl macht**, und keineswegs ein gerichteter Pfeil.

**(2.2.2)** Besonders deutlich und einfach wird der formale Unterschied im Falle  $V = \mathbb{R}^3$ . Dann kann man  $V^* = (\mathbb{R}^3)^*$  mit dem Raum der Zeilenvektoren identifizieren, die ja zumindest **formal** etwas anderes als die Spaltenvektoren sind. Typisch gilt im Rahmen der Matrixschreibweise:

$$\lambda = (1, 3, -2) \in (\mathbb{R}^3)^* \quad \lambda = \left( \mathbb{R}^3, \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto 1x + 3y - 2z, \mathbb{R} \right)$$

Das Beispiel des  $\mathbb{R}^3$  samt Unterscheidung zwischen Zeilen- und Spaltenvektoren kann und sollte der Leser immer wählen, wenn er Aussagen über den Dualraum formal und einfach konkretisieren möchte. Wir werden das Beispiel später genauer ausführen.

(2.2.3) Interpretiert man  $\lambda \in V^*$  als Abbildung, so sieht es aus, als wäre  $\lambda$  stärker festgelegt, weniger variabel als das Argument  $\vec{x}$ . Das ist aber weder formal erforderlich, noch gibt es die Situation vieler Anwendungen angemessen wieder. Dort können beide Bestandteile gleichermaßen veränderlich sein.

(2.2.4) Um diesen Sachverhalt zu verdeutlichen, führt man die folgende Abbildung ein:

$\langle \cdot   \cdot \rangle = (V^* \times V, (\lambda, \vec{x}) \mapsto \langle \lambda   \vec{x} \rangle = \lambda(\vec{x}), K)$	<b>Kanonisches Skalarprodukt</b> zwischen $V$ und $V^*$
--	--

(2.2.5) Diese Abbildung ähnelt stark dem euklidischen Skalarprodukt in  $V$ , nur dass die Faktoren dort aus gleichen, hier aus unterschiedlichen Räumen kommen: Eine Linearform und ein Vektor werden gepaart und ergeben eine Zahl, den Wert  $\lambda(\vec{x})$ . Und da die Faktoren aus unterschiedlichen Räumen kommen, ist bereits die **Frage** nach dem Kommutativgesetz sinnlos. Das Skalarprodukt selbst ist kommutativ. Die Formel für die Arbeit aus der elementaren Physik, also  $(\vec{F}, \Delta\vec{x}) \mapsto A = \vec{F} \cdot \Delta\vec{x}$  gehört in den Bereich dieses kanonischen Produktes, da die Faktoren verschiedener Art sind. Vgl. auch (2.1.9).

(2.2.6) Aber die andere wichtige Rechenregel für das Skalarprodukt - die **Bilinearität** - ist erfüllt:

<p><b>Das kanonische Skalarprodukt <math>\langle \cdot   \cdot \rangle</math> ist bilinear.</b> D.h. es gilt:</p> $\langle \lambda   \vec{x}\alpha + \vec{y}\beta \rangle = \langle \lambda   \vec{x} \rangle \alpha + \langle \lambda   \vec{y} \rangle \beta \quad \text{und} \quad \langle \lambda\alpha + \mu\beta   \vec{x} \rangle = \langle \lambda   \vec{x} \rangle \alpha + \langle \mu   \vec{x} \rangle \beta$ <p>für alle <math>\alpha, \beta \in K</math> und <math>\lambda, \mu \in V^*</math> und <math>\vec{x}, \vec{y} \in V</math>.</p>
--

(2.2.7) Der Beweis der beiden Gleichungen ist (anders als in vielen sonstigen Fällen) unterschiedlich! Die erste drückt aus, dass  $\lambda$  linear ist. Die zweite dagegen beschreibt die werteweise Verknüpfung der Elemente von  $V^*$ , die allgemein in  $\text{Hom}_K(V, W)$  eingeführt war.  $(\lambda + \mu)(\vec{x}) = \lambda(\vec{x}) + \mu(\vec{x})$  schreibt sich in der neuen Form:  $\langle \lambda + \mu | \vec{x} \rangle = \langle \lambda | \vec{x} \rangle + \langle \mu | \vec{x} \rangle$ .

(2.2.8) Die Bilinearität erweist sich als besonders nützlich beim Rechnen mit dem Indexkalkül. Die folgende Rechnung ist als Konsequenz der Bilinearität zulässig. Sie beinhaltet die **Fundamentalidentität für Skalarprodukte**.

<p><b>Seien</b> <math>\lambda = \sum_m \sigma_m \lambda_m</math> und <math>\vec{x} = \sum_p \vec{a}_p x_p</math> mit <math>\lambda, \sigma_m \in V^*</math> und <math>\vec{x}, \vec{a}_p \in V</math> und <math>\lambda_m, x_p \in K</math>.</p> <p><b>Dann</b> gilt <math>\langle \lambda   \vec{x} \rangle = \langle \sum_m \sigma_m \lambda_m   \sum_p \vec{a}_p x_p \rangle = \sum_{mp} \langle \sigma_m   \vec{a}_p \rangle \lambda_m x_p</math></p>
---

Entgegen unseren Regeln haben wir hier zwei Faktoren  $\lambda_m$  und  $\vec{a}_p$  vertauscht. Alle Zahlfaktoren werden nach rechts herausgezogen. Formen und Vektoren müssen an ihrem Ort innerhalb der spitzen Klammern verbleiben.

Ein Konkretisierungsbeispiel für eine derartige Rechnung:

<p><math>\lambda = \alpha 3 + \beta 2 + \gamma</math> mit <math>\alpha, \beta, \gamma \in V^*</math> und <math>\vec{x} = \vec{a}5 + \vec{b}7 + \vec{c}(-3)</math>. Das gibt:</p> $\langle \lambda   \vec{x} \rangle = \langle \alpha 3 + \beta 2 + \gamma   \vec{a}5 + \vec{b}7 + \vec{c}(-3) \rangle = \langle \alpha   \vec{a} \rangle 15 + \dots + \langle \gamma   \vec{c} \rangle (-3).$ <p>Das Endergebnis ist ein Körperelement.</p>
---

(2.2.9) Was die Quantifizierung von Linearformen anbelangt, so haben wir zwei Möglichkeiten: In ihrer Rolle als Homomorphismen werden sie quantifiziert durch die jeweilige beschreibende Matrix. Die ist vom Typ  $1 \times n$ , d.h. eine einzelne Zeile. Um diese Quantifizierung zu erhalten, benötigen wir eine Basis  $b$  von  $V^*$ . (Für den Werteraum  $K$  nehmen wir stets die kanonische Basis  $e = (1)$ ).

(2.2.10) Zu ihrer Rolle als Vektor dagegen müssen wir noch eine Basis  $b^*$  von  $V^*$  vorgeben, die Form entwickeln und dann zum zugehörigen Komponentenvektor übergehen. Das ist nach unseren Vereinbarungen ein **Spaltenvektor**. Man möchte natürlich eine Beziehung zwischen den beiden Quantifizierungen haben. Die nachfolgende Einführung der dualen Basis leistet das, wie wir sehen werden, aber eben basisabhängig.

## 5.2.3 Die duale Basis (einer gegebenen Basis von $V$ )

(2.3.1) Sei  $a = (\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n)$  eine Basis von  $V$ . Wir wollen mit Hilfe von  $a$  eine eindeutig bestimmte Basis  $a^* = (a_1^*, a_2^*, \dots, a_n^*)$  von  $V^*$  konstruieren. Für jedes  $i$  gilt  $a_i^* \in V^*$ . Gibt es lineare Funktionale, die durch die Vorgabe der Basis  $a$  von  $V$  besonders ausgezeichnet sind? Nun, die Hauptbedeutung jeder

Basis  $a$  besteht in der Möglichkeit der zugehörigen eindeutigen Komponentendarstellung  $\vec{x} = \sum \vec{a}_i x_i$  für  $\vec{x} \in V$ . Es liegt nahe, die Zuordnung  $\vec{x} \mapsto x_i$  für festes  $i$  zu betrachten: **Dem Vektor wird seine  $i$ -te Komponente bezüglich der Basis zugeordnet.**

(2.3.2) Als volle Abbildung gibt das:

$$\vec{a}_i^* = (V, \vec{x} = \sum \vec{a}_r x_r \mapsto x_i, K) \quad \text{für } i = 1, 2, \dots, n \quad (\text{laufender Index}).$$

Das ist offensichtlich für jedes  $i$  ein Element aus  $V^*$ .

(2.3.3) Zur Festlegung eines Homomorphismus - und  $a_i^*$  ist einer! - genügt (nach der zugehörigen Fundamentalidentität) die Angabe der Werte auf einer Basis des Urbildraumes. In unserem Fall also der Werte für  $a$ . Hierfür findet man sofort:

$$a_i^*(\vec{a}_k) = \langle a_i^* | \vec{a}_k \rangle = \delta_{ik} = \begin{cases} 1 & \text{für } i = k \\ 0 & \text{für } i \neq k. \end{cases}$$

(2.3.4) Hier steht  $\delta_{ik}$  für die Komponenten der Einheitmatrix  $\mathbb{I}_n$ . Es handelt sich um ein besonders auch in der Physik verbreitetes Symbol, das für die Einheitsmatrix die Verbindung zum Indexkalkül herstellt. Man hat ja  $\mathbb{I}_n = (\delta_{ik})$ . Man sagt auch gern,  $\delta_{ik}$  sei "das Kroneckersymbol". Meist wird in den Lehrbüchern  $a_i^*$  mittels der zweiten Beziehung (2.3.3) eingeführt und die erste (2.3.1) daraus abgeleitet.

(2.3.5) Unsere Einführung der dualen Basis ergibt die folgende Argumentationskette, die unbedingt zu verstehen und zu beachten ist:

◆	$V^*$ ist als Vektorraum konkret konstruiert, sobald man eine Basis besitzt.
◆	$a^*$ ist eine Basis (siehe unten).
◆	Jeder Basisvektor $a_i^*$ ist als Abbildungstriplett $V \rightarrow K$ bekannt und eindeutig festgelegt.

**Die immer wieder gestellte Frage "Wie rechne ich denn die  $a_i^*$  aus?" ist daher ohne Zutatenpräzisierung unsinnig.**

Sinnvoll ist nur: Wie kann ich diesen festgelegten Vektor  $a_i^*$  noch in einer anderen (als der gegebenen) Basis von  $V^*$  darstellen? Insbesondere hat man im Falle  $V = K^n$  die kanonische Basis  $e$  mit ihrer Dualbasis  $e^*$ . **Dann** ist es sinnvoll zu fragen: Wie rechne ich  $a^*$  als Linearkombination der Basis  $e^*$  aus?

(2.3.6) Wir haben zunächst nur eine **Familie**  $a^* = (a_1^*, \dots, a_n^*)$  von  $n$  Linearformen bestimmt. Bildet diese Familie aber auch tatsächlich eine Basis? Sie tut es. Unter der Quantifizierung geht  $a^*$  in die kanonische Basis von  $K^{n*}$  über, und das ist nur möglich, wenn  $a^*$  selbst Basis ist.

(2.3.8) Aber man kann auch leicht unmittelbar zeigen, dass die Familie  $a^*$  linear unabhängig ist. Aus Dimensionsgründen ist sie dann automatisch Basis.

(2.3.7) Klar ist: Dies Argument gilt so nur im endlichdimensionalen Fall. Im unendlichdimensionalen Fall ist die Familie linear unabhängig, aber nicht mehr Erzeugendensystem, also Basis. Das Dimensionsargument ist nicht länger verfügbar. Auch (2.3.6) ist nicht verallgemeinerbar.

(2.3.8) Zur Übung im Umgang mit  $a^*$  führen wir den Nachweis der linearen Unabhängigkeit aus.

Sei also  $\sum a_i^* \alpha_i = 0$ . Rechts steht die Nullabbildung. D.h. es muß  $\langle \sum a_i^* \alpha_i | \vec{x} \rangle = 0$  für alle  $\vec{x} \in V$  gelten. Oder wegen der Bilinearität:  $\sum \langle a_i^* | \vec{x} \rangle \alpha_i = 0$ . Jetzt wählen wir  $\vec{x} = \vec{a}_k$ . Dann folgt  $\sum_i \delta_{ik} \alpha_i = 0$  wobei über  $i$  summiert wird. Nun sind aber alle Summanden mit  $i \neq k$  gleich Null, weil der Faktor  $\delta_{ik}$  verschwindet. Nur der Summand mit  $i = k$  bleibt. Obige Bedingung reduziert sich auf  $\alpha_k = 0$ . Da wir für  $k$  alle Werte der Indexfamilie der Basis  $a$  nehmen können, **ist die lineare Unabhängigkeit gezeigt**. Auch für den unendlichdimensionalen Fall.

Ist  $V$  endlichdimensional, so folgt die Basiseigenschaft, da wir eine Familie aus  $n$  Vektoren haben.

(2.3.11) Zu den Bezeichnungen:

Bezeichnung:	Die so entstandene Basis $a^*$ von $V$ nennt man
	<b>die zu <math>a</math> duale Basis</b> oder <b>Dualbasis von <math>a</math>.</b>
Nie vergessen:	$a^*$ hängt von der Basis $a$ von $V$ ab.
Wir vereinbaren:	<i>Wahl eines Koordinatensystems</i> $K$ soll bedeuten, dass man bei der
	Wahl einer Basis $a$ in $V$ im Dualraum die Dualbasis $a^*$ zu $a$ wählt!

## 5.2.4 Darstellungen und Quantifizierungen von Linearformen und Dualbasen

Wir kehren zur *Quantifizierungsfrage für Linearformen* zurück.

(2.4.1) Sei also  $a$  Basis von  $V$  und  $a^*$  zugehörige duale Basis von  $V^*$ . Dann können wir jedes  $\lambda \in V^*$  auf **drei Weisen** darstellen:

1. Als Linearkombination bezüglich der Basis  $a^*$
2. Als beschreibende Matrix  $M_\lambda$  bezüglich  $a$  für  $V$
3. Als von  $a^*$  erzeugten Koordinatenvektor der Koeffizienten aus 1)

(2.4.2) Fangen wir mit der Basisdarstellung an: Da  $a^*$  Basis von  $V^*$  ist, können wir stets eine eindeutige Darstellung  $\lambda = \sum a_i^* \lambda_i$  angeben, erhalten also 1). Hiermit gewinnen wir 2): Denn die zugehörige beschreibende Matrix  $M_\lambda$ , entsteht nach dem Kochrezept, indem wir  $\lambda(\vec{a}_k) = \sum a_i^*(\vec{a}_k) \lambda_i = \lambda_k$  berechnen. Also  $M_\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$ . **Der aus den Komponenten gebildete Zeilenvektor ist die beschreibende Matrix von  $\lambda$ .** Der zugehörige unter 3) genannte Koordinatenvektor dagegen ist wie immer der aus den Komponenten gebildete **Spaltenvektor**. Der Übergang zwischen den beiden Rollen Abbildung und Vektor ist im Rahmen der Quantifizierung daher einfach der Übergang Zeile in Spalte, die Transposition (mit gleichzeitigem Wechsel vom Rechtsvektorraum zum Linksvektorraum). Also:

$\vec{\lambda}^K = {}^t M_\lambda^K$	Übergang Abbildungsrolle/Vektorrolle
--------------------------------------	---

Zur Transposition vgl. (2.5.11).

(2.4.3) Zusammenfassende Übersicht zur Quantifizierung von Vektoren und Linearformen:

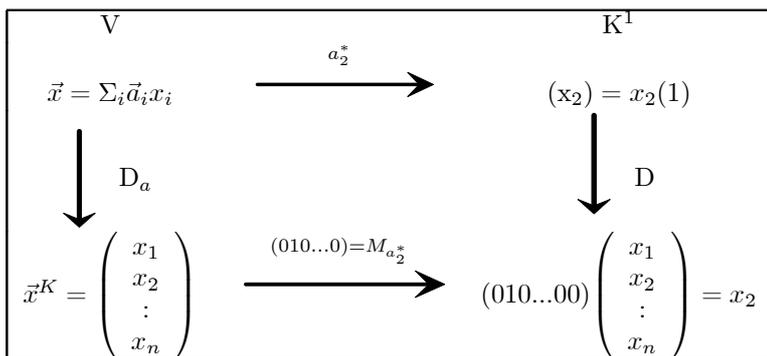
Absolute Darstellung <i>Vektor</i> : $\vec{x} = \vec{a}_1 x_1 + \dots + \vec{a}_n x_n$	Koordinatenvektor $\vec{x}^K = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$
Absolute Darstellung <i>Linearform</i> : $\lambda = a_1^* \lambda_1 + \dots + a_n^* \lambda_n$	Koordinatenvektor $\lambda^K = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$
$\vec{\lambda}^K = {}^t M_\lambda^K$	Beschr. Matrix $M_\lambda^K = (\lambda_1, \lambda_2, \dots, \lambda_n)$

(2.4.4) **Weshalb lassen wir die Beschreibung als Koordinatenvektor nicht einfach fort?** Das geht nicht, weil wichtige physikalische Problemsituationen diese Beschreibung erfordern! Dabei faßt man gewisse Meßdaten zu einem Zahlentupel zusammen, weil man vermutet, dass sie zusammengehören, dass sie gemeinsam eine Komponentendarstellung einer physikalischen Größe bilden. Man weiß aber noch nicht, welcher Art diese Größe genauer ist, das verlangt weitere Überlegungen und Experimente. **Die dann in jedem Fall verwendete Einstiegsform ist der Koordinatenvektor.** Insbesondere kann es sein, dass man zunächst nicht weiß, ob ein solches  $n$ -Tupel zu  $V$  oder zu  $V^*$  gehört. Daher werden die in (2.4.3) gegebenen Darstellungen alle benötigt.

### 5.2.4a Ein Beispiel für den Darstellungsformalismus

(2.4.5) Wir führen ein Beispiel einer beschreibenden Matrix etwas aus. In  $V$  sei die Basis  $a$  gegeben. In  $K^1$  ist (1) die kanonische Basis. Wir erinnern an unsere Vereinbarung, dass in  $K^1$  Vektoren  $(x) = x(1)$  geschrieben werden. Nun können wir sofort die  $a_i^*$  beschreibende Matrix bestimmen. Zur Verdeutlichung

wählen wir  $i = 2$ . Es ergibt sich folgendes Bild:



Die beschreibende Matrix für  $a_2^*$  ist gerade der Zeilenvektor  $e_2^* = (010\dots0)$

Das Diagramm verdeutlicht die Verhältnisse für allgemeines  $V$  und allgemeine Basiswahl. Der entstehende Darstellungsisomorphismus zwischen Linearformen und Zeilenvektoren ist basisabhängig. Ist dagegen  $V = K^n$ , dann hat man mehr Struktur. Insbesondere kann man  $\lambda \in V^*$  immer als Linearkombination der Dualbasis der kanonischen Basis von  $K^n$  darstellen. Wählt man oben und unten im Diagramm die kanonischen Basen, dann kann man obere und untere Zeile identifizieren. Im Ergebnis ist die Basisdarstellung von  $\lambda$  bezüglich  $e^*$  gleich der beschreibenden Matrix. Wegen der Bedeutung führen wir dies Resultat genauer aus:

### 5.2.4b Die kanonische Darstellung von Linearformen des $K^n$ . (Linearformen als Zeilenvektoren).

(2.4.6) Sei  $V = K^n$ =Spaltenraum mit kanonischer Basis. Im Indexkalkül ein Rechtsraum, **nicht** der Tupelraum des Matrixkalküls! Dann kann man den Dualraum  $V^* = K^{n*}$  mit dem Zeilenraum der  $n$ -Tupel identifizieren. Jede Linearform ist gleich ihrer beschreibenden Matrix bezüglich der kanonischen Basis  $e^*$  und gleich der Linearkombination bezüglich der kanonischen Basis des Zeilenraumes. 1) und 2) aus (2.4.1) sind identifiziert. **Die kanonische Zeilenbasis  $e^*$  ist die duale Basis zur Spaltenbasis  $e$ .** Kurz:

**In der kanonischen Darstellung sind die Linearformen des  $K^n$  gerade die Zeilenvektoren.**

(2.4.7) Der Übergang von  $\lambda \in K^{n*}$  zum beschreibenden Koordinatenvektor  $\vec{\lambda}^K$  geschieht durch Umwandlung in den zugehörigen Spaltenvektor. Erst an dieser Stelle erfolgt der Übergang zu einem Linksvektorraum.

(2.4.8) Wählt man jetzt in  $K^n$  eine andere als die kanonische Basis- nennen wir sie  $a$  -, so gehört dazu auch eine andere duale Basis  $a$ , deren Vektoren wir natürlich als Zeilenvektoren, d.h. als Linearkombination der kanonischen Basis  $e^*$  ausdrücken können! (Das ist gerade die in (2.3.5)) als sinnvoll herausgearbeitete Frage.)

(2.4.9) Als Bedingungen zur Bestimmung (dieser Linearkombinationen) muß man die  $n^2$  Gleichungen  $a_i^*(\vec{a}_k) = \delta_{ik}$  heranziehen, welche die beiden Basen verbinden. Manchmal kann man die Lösungen direkt raten. Ein solches Beispiel zum  $\mathbb{R}^3$  :

$$\vec{a}_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \quad \vec{a}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \vec{a}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{Raten: } \begin{aligned} a_1^* &= \frac{1}{2}(-1, 1, 1) \\ a_2^* &= \frac{1}{2}(1, -1, 1) \\ a_3^* &= \frac{1}{2}(1, 1, -1) \end{aligned}$$

Man verifiziert die 9 Gleichungen  $\langle a_i^* | \vec{a}_k \rangle = \delta_{ik}$ .

(2.4.10) Allgemein erhält man die duale Basis rezeptartig für die beschriebene spezielle Situation wie folgt:

### 5-2.4c Darstellung einer allgemeinen dualen Basis von $K^{n*}$ durch die kanonische Basis:

<b>Es sei</b>	$V = K^n$ und $a$ eine Basis von $V$ . Die Vektoren $\vec{a}_i$ seien als Spaltenvektoren gegeben, d. h. durch ihre Darstellung bezüglich der kanonischen Basis $e$ . Wir bilden aus diesen Spaltenvektoren die quadratische Matrix $T$ . Das ist die Transformationsmatrix von $a$ nach $e =$ kanonisch, also eine invertierbare Matrix.
<b>Sei</b>	$T^{-1}$ die zugehörige inverse Matrix.
<b>Dann</b>	bilden die Zeilen von $T^{-1}$ die zu $a$ duale Basis. Genauer die Darstellung der dualen Basis bezüglich der kanonischen Basis $e^*$ von $(K^n)^*$ . Wir haben daher: $k$ -te kanonische Komponente des Dualbasisvektors $a_i^* = (a_i^*)_k = T_{ik}^{-1}$ . Oder auch $a_i^* = \sum_k e_k^* T_{ik}^{-1}$ Der Übergang zum Koordinatenvektor erfolgt durch Transposition.

Man kann auch umgekehrt von einer Basis im Dualraum ausgehen. Dann erhält man das folgende Rezept:

<b>Sei</b>	$a^*$ eine aus Zeilenvektoren gebildete Basis von $K^{n*}$ und $M$ die aus diesen Zeilen gebildete Matrix. Dies ist die Transformationsmatrix von $a^*$ nach $e^*$ . Daher ist $M$ invertierbar. <b>Sei</b> $M^{-1}$ die zugehörige inverse Matrix.
<b>Dann</b>	bilden die Spalten von $M^{-1}$ eine Basis von $V$ , deren Dualbasis gleich $a^*$ ist.

**(2.4.11)** Beweis: Es gilt  $T^{-1}T = \mathbb{I}_n =$  Einheitsmatrix. Dann folgt nach dem allgemeinen Schema für die Matrixmultiplikation:

$$(i\text{-te Zeile von } T^{-1}) \times (j\text{-te Spalte von } T) = (ij)\text{-te Komponente von } \mathbb{I}_n, \text{ also } \delta_{ij}$$

Das ist aber gerade die Definitionsgleichung für den  $i$ -ten Vektor der Dualbasis in der betrachteten Darstellung als Zeilenvektor.

Zum zweiten Teil: Die Spalten bilden eine Basis, da  $M^{-1}$  invertierbar ist und folglich der Spaltenrang maximal sein muß. Der Rest folgt wieder aus der Matrixgleichung  $MM^{-1} = \mathbb{I}_n$ .

Die in (2.4.10) beschriebene Konstruktion ist recht einfach auszuführen. Sie können sie leicht am Beispiel (2.4.9) bestätigen.

(□ **F.** ) Wählen Sie eine einfache invertierbare  $4 \times 4$ -Matrix, lassen Sie sich von einem Computeralgebraprogramm die inverse Matrix ausgeben und lesen Sie dann die Dualbasis ab. Testen Sie die definierenden Beziehungen  $\langle a_i^* | \vec{a}_k \rangle$ .

## 5.2.4d Lineare Gleichungen und Superpositionprinzip

**(2.4.12)** Jetzt ist eine interessante und wichtige Anwendung der Dualbasis im Bereich der linearen Gleichungen möglich: Sei  $M\vec{x}^e = b^e$  eine in Matrixform gegebene lineare Gleichung. D.h. als Gleichung zwischen den Koordinaten. Die Matrix  $M$  sei invertierbar, so dass die Gleichung eindeutig lösbar ist durch  $\vec{x}^e = M^{-1}\vec{b}^e$ . Nach unseren Konventionen besagt der Index  $e$ , dass Koordinatenvektoren bezüglich der kanonischen Basis vorliegen. Durch Linearkombinationsbildung bezüglich der kanonischen Basis  $e$  wird daraus eine absolute Vektorgleichung, wie in (1.2.1) dargestellt:  $M(\vec{x}) = \sum \vec{e}_i M_{ik} x_k = \sum \vec{e}_i b_i = \vec{b}$ . Die Darstellung verwendet den Indexkalkül. Anwenden der inversen Abbildung  $M^{-1}$  mit beschreibender Matrix  $(M_{ik}^{-1})$  ergibt den absoluten Lösungsvektor:

$$\vec{x} = M^{-1}\vec{b} = \sum M^{-1}(\vec{e}_j) b_j = \sum \vec{e}_k M_{kj}^{-1} b_j = \sum_k \vec{e}_k (\sum_j M_{kj}^{-1} b_j) = \sum_k \vec{e}_k (\sum_j M_{kj}^{-1} \delta_{jr}) b_r.$$

**(2.4.13)** Wir führen die Vektoren  $\vec{X}_r = \sum_k \vec{e}_k (\sum_j M_{kj}^{-1} \delta_{jr})$  ein. Damit folgt dreierlei:

1) Die Lösung  $\vec{x}$  unserer Ausgangsgleichung ist eine Linearkombination dieser Vektoren:

$$\vec{x} = \sum_r \vec{X}_r b_r$$

Insbesondere bildet die Familie der  $\vec{X}_r$  eine Basis. ( $M$  ist Isomorphismus!)

2) Die duale Basis zu der Spaltenvektorbasis der  $\vec{X}_r$  wird nach (2.4.10) durch die Zeilenvektoren der Matrix  $M$  gebildet. D.h. die speziellen in 1) benötigten Lösungen  $\vec{X}_r$  lassen sich per Dualität als Spalten von  $M^{-1}$  bestimmen.

3)  $\vec{X}_r$  ist der eindeutige Lösungsvektor des inhomogenen Gleichungssystems mit  $b_i = \delta_{ir}$ . D.h. alle  $b_i$  sind Null, nur  $b_r$  ist gleich 1.

(2.4.14) Die Resultate 1) und 3) ergeben zusammen ein Superpositionsprinzip. (Vgl. Kap.4 (1.6.11).)

Die gefundene Struktur erweist für viele Anwendungen als nützlich und enorm verallgemeinerungsfähig.

Die typische zugehörige Argumentation zugehörige sieht wie folgt aus:

(2.4.15) Wir interpretieren die Zuordnung  $\vec{b} \mapsto M^{-1}\vec{b}$  als "Ursache-Wirkungskette". Der Inhomogenitätsvektor  $\vec{b}$  ist die Ursache, die Quelle, die etwas anderes, nämlich  $\vec{x}$ , eindeutig erzeugt. Und  $M\vec{x} = \vec{b}$  ist die festlegende (lineare) Bedingung. Oder auch:  $\vec{b}$  ist so etwas wie eine felderzeugende Ladung. Der Konfigurationsraum sind die  $n$  Tupelplätze. Und  $\vec{x}$  ist das von der Ladung erzeugte Feld.

Wird nun eine einzelne Ladung der Stärke 1 am Ort der  $r$ -ten Komponente angebracht, dann entsteht nach 3) die Wirkung  $\vec{X}_r$ , als Feld im gesamten Konfigurationsraum. Bringt man an mehreren Stellen zugleich Ladungen an, so ergibt sich die resultierende Wirkung durch Superposition, also über Vektoraddition der Einzelfelder.

Und nach 2) erhält man die speziellen Lösungen  $X$  rechnerisch wie folgt:

Die Ursache-Wirkungsbeziehung wird durch die Gleichung  $M\vec{x} = \vec{b}$  beschrieben. Bilde zu  $M$  die inverse Matrix und zerlege diese in Spalten. Dies ergibt die  $\vec{X}_r$ . . Diese speziellen Lösungen  $\vec{X}_r$  nennt man auch gerne suggestiv *Elementarlösungen*. Wichtig ist, dass die Ausgangsgleichung eindeutig lösbar und linear ist.

Ein naheliegender Anschlußgedanke ist, dass man eine physikalische inhaltliche Annahme über die Gleichartigkeit der einzelnen Quellorte besitzt, typisch in Form einer Gruppenoperation. Dann wird man daraus Rückschlüsse über die Form von  $M$  ziehen können, also das zugehörige Naturgesetz.

(□ **F.** ) Bestimmen Sie auch die Vektoren  $\vec{X}_r$ , für ein Beispielsystem. Veranschaulichen Sie sich das Resultat vom Feldstandpunkt.

## 5.2.5 Das Rechnen mit kanonischem Skalarprodukt und Dualbasen: Komponentenformel und Fundamentalidentität.

(2.5.1) Wir üben Umgang und Rechnen mit dem eingeführten Begriffssystem an zwei wichtigen Beispielen.

(2.5.2) Zuerst die folgende Rechnung, die an die Definition (2.3.3) anknüpft. Allerdings nehmen wir jetzt an, die duale Basis sei verfügbar.

Sei  $\vec{x} = \sum \vec{a}_p x_p$ . Dann folgt  $\langle a_k^* | \vec{x} \rangle = \langle a_k^* | \sum \vec{a}_p x_p \rangle = \sum \langle a_k^* | \vec{a}_p \rangle x_p = \sum \delta_{kp} x_p = x_k, k = 1, \dots, n$ .

Damit haben wir unseren Ausgangspunkt bei der Definition der dualen Basis wiedergewonnen:  $a_k^*$  projiziert aus  $\vec{x}$  die  $k$ -te Komponente heraus.

(2.5.3) Stellen wir die wichtigsten Formeln für diese Komponentenprojektion zusammen:

$$\begin{aligned} x_p &= \langle a_p^* | \vec{x} \rangle & \vec{x} &= \sum \vec{a}_p \langle a_p^* | \vec{x} \rangle \\ \text{oder} & & a_p^* &= (V, \vec{x} = \sum \vec{a}_i x_i \mapsto x_p, \mathbb{R}) \end{aligned}$$

**Mit diesen nützlichen Formeln verfügen wir zum ersten Mal ganz allgemein über eine Berechnungsformel für die Komponenten!** Bisher hatten wir zunächst nur eine Bezeichnung und waren gezwungen, die  $x_p$  bei Bedarf über ein Gleichungssystem auszurechnen. Von jetzt ab können wir diese Berechnungsformel einsetzen und damit weiterrechnen.

(2.5.4) Als zweites Beispiel berechnen wir die Fundamentalidentität (2.2.8) für das kanonische Skalarprodukt erneut, und zwar für den Fall, dass im Dualraum die duale Basis genommen wird. (Bei unserer ersten Rechnung wurde eine beliebige Basis von  $V^*$  zugelassen).

Vorgegeben sind  $\vec{x} \in V$  und  $\lambda \in V^*$ . Also  $\vec{x} = \sum \vec{a}_i x_i$  und  $\lambda = \sum a_i^* \lambda_i$ . Wir berechnen damit den Wert  $\langle \lambda | \vec{x} \rangle$  des kanonischen Skalarproduktes. Denken Sie daran: **Das Ergebnis ist eine von der Basiswahl unabhängige Zahl.** Was wir jetzt herleiten, ist eine Berechnungsformel, um diese Zahl aus den **basisabhängigen** Komponenten zu gewinnen.

(2.5.5) Nach dem üblichen Schema finden wir die **Fundamentalidentität** bei zueinander dualen Basen:

$$\langle \lambda | \vec{x} \rangle = \langle \sum a_i^* \lambda_i | \sum \vec{a}_k x_k \rangle = \sum \langle a_i^* | \vec{a}_k \rangle \lambda_i x_k = \sum_{ik} \delta_{ik} \lambda_i x_k = \sum_i \lambda_i x_i .$$

(2.5.6) Ein formal außerordentlich ansprechendes Resultat! Interpretiert man  $\lambda$  als lineare Abbildung, so ist  $(\lambda_1, \dots, \lambda_n)$  die zugehörige beschreibende Matrix, und wir finden einfach unsere übliche Regel Zeile  $\times$  Spalte wieder. Interpretieren wir dagegen  $\lambda$  als Vektor und  $\vec{\lambda}^K$  als zugehörigen Koordinatenvektor im  $K^n$ , dann ergibt sich der Rechenausdruck für die Komponentenform des euklidischen Skalarproduktes zwischen  $\vec{\lambda}^K$  und  $\vec{x}^K$ . Der Übergang in dieser Formel zwischen den beiden Rollen ist damit vollständig.

Beachten sie auch Folgendes: Wählt man eine zweite neue Basis  $b$  von  $V$  und ist  $b^*$  die zugehörige duale Basis, so dass  $\vec{x} = \sum \vec{b}_k x_k^N$  und  $\lambda = \sum b_k^* \lambda_k^N$  ist, dann folgt notwendig auch für dieses System  $\langle \lambda | \vec{x} \rangle = \sum_k \lambda_k^N x_k^N$ . Das resultierende Körperelement bleibt dasselbe, auch wenn sich alle einzelnen Komponenten unterscheiden. Das ist ein für die Anwendungen sehr wichtiger Sachverhalt, da er zeigt, wie man aus basisabhängigen Komponenten eine basisunabhängige Größe - einen Skalar (im Sinne der Physiker!) - bei Forminvarianz der definierenden Gleichung erhält. Die Bemerkungen aus (2.4.4) sollten an dieser Stelle erneut beachtet werden.

### 5.2.5a Die Vorgabe von Linearformen durch eine Komponentenformel.

(2.5.7) Liest man die Gleichung (2.5.5) von rechts nach links, so erhält man eine wichtige Methode, Linearformen zu definieren. Denn hat man einen Ausdruck des Typs  $\sum_k \lambda_k x_k$  wobei die  $x_k$  die Komponenten des allgemeinen Vektors  $\vec{x} \in V$  bezüglich der Basis  $a$  sind, dann **ist**  $\lambda = \sum a_i^* \lambda_i$  ein Element aus  $V^*$ , dessen Werte sich über die angegebene Formel berechnen. Erforderlich ist, dass zueinander duale Basen verwendet werden.

(2.5.8) Gegeben sei etwa ein lineares Gleichungssystem  $M\vec{x} = \vec{b}$  - bezogen auf die kanonische Basis  $a$  mit dualer Basis  $a^*$ . Insbesondere darf  $a$  gleich der kanonischen Basis  $e$  sein. In Komponenten lautet das System  $\sum_k M_{ik} x_k = b_i$ . Jetzt benutzen wir (2.5.3), um die Koordinaten  $x_k$  durch  $e_k^*(\vec{x})$  zu ersetzen. (Dies ist ein nützlicher Trick, der sich später besonders in der Tensorrechnung bewährt.) Das ergibt  $\sum_k M_{ik} e_k^*(\vec{x}) = b_i$ . Und das läßt sich wie folgt schreiben:  $\lambda_i(\vec{x}) = b_i$ , mit  $\lambda_i = \sum_k e_k^* M_{ik}$ . Hierdurch ist das Dualraumelement  $\lambda_i$  definiert, das gerade die  $i$ -te Gleichung des Systems repräsentiert bzw. durch diese Gleichung festgelegt wird.

(2.5.9) An dieser Stelle müssen wir die "falsche" vertauschte Indexreihenfolge verwenden, um alle unsere Konventionen einzuhalten. Statt  $M_{ik} e_k^*$  müssen wir  $e_k^* M_{ik}$  schreiben, um die Rechtsraumeigenschaft zu sichern.

(2.5.10) Damit haben wir das Gleichungssystem  $M\vec{x} = \vec{b}$  aus der Matrixform in die Zeilenvektorform umgewandelt und zugleich eine formal korrekte Definition für die Zeilenlinearformen  $\lambda_i$ , erhalten. Nochmals die Formel:  $\lambda_i = \sum_k e_k^* M_{ik}$ .

### 5.2.3f Die transponierte Matrix

(2.5.11) Das Problem der "falschen" Indexreihenfolge tritt häufiger auf. Man löst es mit Hilfe der folgenden bereits mehrfach erwähnten Definition:

Es sei  $M = (M_{ik})$  eine  $m \times n$ -Matrix  
 Dann definiert man dazu eine zweite Matrix  ${}^t M$  durch die Komponentenformel  
 $({}^t M)_{ik} = M_{ki}$  für  $i = 1, \dots, m$  und  $k = 1, \dots, n$ .

D.h. die  $(i, k)$ -te Komponente von  ${}^t M$  ist die  $(k, i)$ -te von  $M$ . Man nennt  ${}^t M$  **die zu  $M$  transponierte Matrix**. Es handelt sich um eine  $n \times m$ -Matrix.

(2.5.12) Beispielsweise ist  $({}^tM)_{12} = M_{21}$ . Und entsprechend folgt allgemein, dass man die transponierte Matrix erhält, **indem man Zeilen und Spalten vertauscht**.

$$\begin{array}{|c|c|}
 \hline
 M = \begin{pmatrix} 1 & 2 \\ a & b \end{pmatrix} & {}^tM = \begin{pmatrix} 1 & a \\ 2 & b \end{pmatrix} \\
 \hline
 \end{array}
 \quad
 \begin{array}{|c|c|}
 \hline
 P = \begin{pmatrix} 1 & 2 & 3 \\ a & b & c \end{pmatrix} & {}^tP = \begin{pmatrix} 1 & a \\ 2 & b \\ 3 & c \end{pmatrix} \\
 \hline
 \end{array}$$

$$\begin{array}{|c|c|}
 \hline
 \vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} & {}^t\vec{x} = (x, y, z) \\
 \hline
 \end{array}$$

Die Transposition macht aus einem Zeilenvektor einen Spaltenvektor und umgekehrt. Dieser Sachverhalt wird gerne benutzt, um im Text das Schreiben von Spaltenvektoren zu umgehen. Meist schreibt man mit Klammerersparnis  ${}^tM_{ik}$ , anstelle von  $({}^tM)_{ik} = (i,k)$ -te Komponente der Matrix  ${}^tM$ .

(2.5.13) Jetzt können wir in der Formel aus (2.5.10) erneut die korrekte Indexreihenfolge herstellen: Kurz: Durch Einführen der transponierten Matrix kann man die Reihenfolge zweier Matrixindizes vertauschen, sofern dies erforderlich ist.

(2.5.14) In Kap.9 werden wir klären, was die Matrixtransposition inhaltlich bedeutet. Bis dahin verwenden wir sie als nützliches Rechenhilfsmittel. Die nachfolgende Frage gibt eine Reihe zugehöriger Rechenregeln, die sich alle gut mit Hilfe des Indexkalküls beweisen lassen.

(□ **F.** ) Zeigen Sie:

a)  ${}^t(A+B) = {}^tA + {}^tB$     b)  ${}^{tt}M = M$     c)  ${}^t(AB) = {}^tB {}^tA$  (Reihenfolge!)    d)  $({}^tM)^{-1} = {}^t(M^{-1})$ .

(Man gehe dabei immer von folgender Denkfigur des Indexkalküls aus: Zwei Matrizen gleichen Typs sind gleich, wenn alle ihre Komponenten gleich sind.)

## 5.2.6 Basiswechsel und Transformationsmatrix

(2.6.1) Die Resultate aus Kap.4.4.6 gelten natürlich auch für den Dualraum. Geht man in  $V$  von  $a$  zu einer neuen Basis  $b$  über, werden sich die Koordinaten der **Linearform  $\lambda$  bezüglich der zugehörigen dualen Basis** ändern. Und zwar so, dass gilt:

$$\lambda = \sum a_i^* \lambda_i^A = \sum b_i^* \lambda_i^N$$

Nach der Regeln aus Kap.4.(4.6.7) benötigen wir die Transformationsmatrix  $X$  für die  $a^* = b^* X$  gilt ("die alte Basis durch die neue ausdrücken"). Mit ihr wird wie gewünscht  $\vec{\lambda}^N = X \vec{\lambda}^A$

(2.6.2) Beim Übergang von  $A$  nach  $N$  transformieren sich die Koordinatenvektoren von Linearformen mit  $X$ , nicht mit  $T$ , wie die Vektorkomponenten. Dies  $X$  ist die andere Matrix, die das Computergehirn aus (2.2.10) kennen muß, um die Normalenvektoren korrekt zu transformieren.

(2.6.3) Die Bestimmung von  $X$  ist leicht, wenn wir  $\lambda$  als lineare Abbildung interpretieren. Es seien  $M^A$  und  $M^N$  die zugehörigen beschreibenden Matrizen. Das sind die Zeilenvektoren, die in Spalten umgewandelt gerade  $\vec{\lambda}^A$  und  $\vec{\lambda}^N$  ergeben. Nun wissen wir aber, dass  $M^N = S M T^{-1}$  gilt. Dabei ist  $T$  die Transformationsmatrix der Vektorkomponenten. Für  $S$  ist die Einheitmatrix zu nehmen, denn  $S$  gehört zu  $K^1$  wo wir stets  $e=(1)$  als Basis wählen. Also  $M^N = M^A T^{-1}$ . Oder in Komponenten:  $\lambda_i^N = \sum \lambda_k^A T_{ki}^{-1}$ . Das müssen wir als Matrix schreiben, die auf den Spaltenvektor  $\vec{\lambda}^N$  angewandt wird, was nur durch Einführen der transponierten Matrix geht:  $\lambda_i^N = \sum_k T_{ki}^{-1} \lambda_k^A = \sum_k ({}^tT^{-1})_{ik} \lambda_k^A$ . Ergebnis:

$$\begin{array}{|c|c|c|c|}
 \hline
 X = {}^tT^{-1} & \vec{\lambda}^N = {}^tT^{-1} \vec{\lambda}^A & \text{und} & a^* = b^* {}^tT^{-1} \\
 & \vec{x}^N = T \vec{x}^A & \text{und} & M_\lambda^N = M_\lambda^A T^{-1} \\
 & & & a = b T \\
 \hline
 \end{array}$$

Also: Alles was in  $V$  mit Hilfe von  $T$  gemacht wird, muß in  $V^*$  mit Hilfe von  ${}^tT^{-1}$  gemacht werden. Und sobald man  $T$  kennt, ist auch  ${}^tT^{-1}$  festgelegt. Ausgehend von  $T$  kann man somit alle Umrechnungen vornehmen. Die Matrix  ${}^tT^{-1}$  wird auch als "die zu  $T$  kontragrediente Matrix" bezeichnet.

(2.6.4) Beispiel

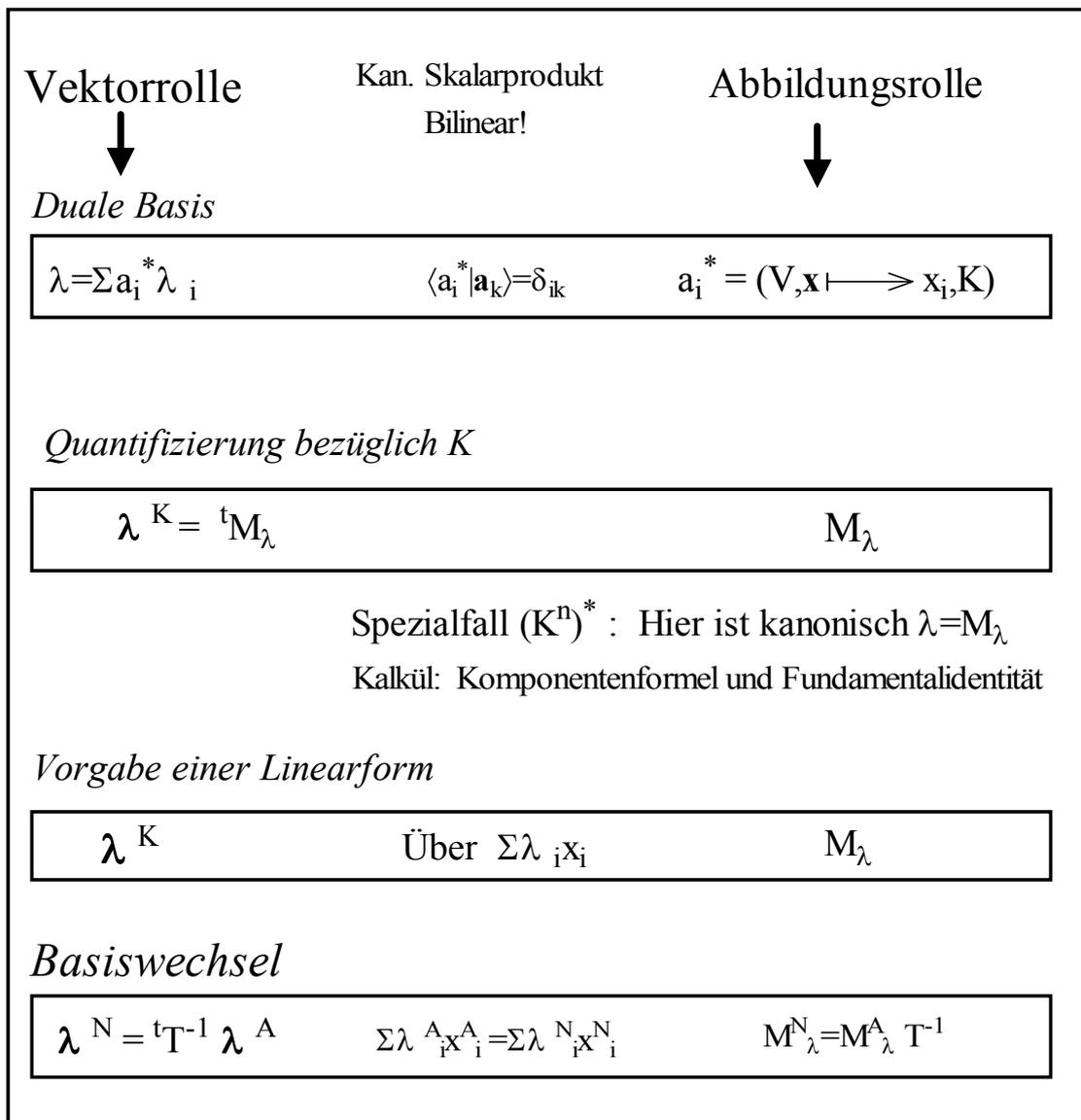
$$T = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 0 \\ 1 & 3 & 2 \end{pmatrix} \quad T^{-1} = \begin{pmatrix} -2 & 1 & 1 \\ 4 & -1 & -2 \\ -5 & 1 & 3 \end{pmatrix} \quad {}^tT^{-1} = \begin{pmatrix} -2 & 4 & -5 \\ 1 & -1 & 1 \\ 1 & -2 & 3 \end{pmatrix}$$

(□ **F.** ) Beweisen Sie die wichtige Gleichung  $X = {}^t T^{-1}$  erneut auf folgendem Weg: Beginnen Sie mit  $a_i^* = \sum_k b_k^* X_{ki}$ . Bilden Sie für beide Seiten der Gleichung das kanonische Skalarprodukt mit  $\vec{b}_r$ , und drücken Sie dann  $\vec{b}_r$  durch die alte Basis  $a$  aus.

### 5.2.7 Zusammenfassung des formalen Teiles

Damit ist die doch unangenehm lange Einführung des Dualraumformalismus beendet. Der Formalismus ist so aufwendig, weil die Linearformen immer in den zwei Rollen als Vektor und als Abbildung auftreten können. Denkt man jedoch an die jeweilige Rolle und hält man sich stur an die angegebenen Regeln, dann ist der Formalismus fast immer unproblematisch und leicht zu verwirklichen.

Fassen wir unsere Resultate und Überlegungen zur Darstellung und Quantifizierung von Linearformen abschließend schematisch zusammen:



### 5.2.8 Hyperebenen. Zur Geometrie der Linearformen

(2.8.1) Sei  $\lambda \in V^*$  eine Linearform  $\neq 0$  in der Rolle als lineare Abbildung. Wie steht es hier mit der üblichen Frage nach Kern und Bild? Bild  $X$  ist offensichtlich eindimensional, D.h.  $\text{Bild } \lambda = K^1$  für  $\lambda \neq 0$ . Dann ist der Kern  $(n-1)$ -dimensional. Was wird geometrisch durch diesen Kern beschrieben? Die Analyse

dieser Frage bereitet auf die eingangs angekündigte Beziehung zwischen Teilräumen von  $V^*$  und  $V$  vor. Im Falle  $n=3$  ist  $\text{Kern}\lambda$  eine Ebene im Raum. Eine auffällige Eigenschaft einer Ebene ist, dass sie den gesamten Raum in zwei Halbräume teilt. Die Ebene selbst ist die Grenze. Man gelangt von dem einem Halbraum in den anderen nur, indem man die Grenze durchquert, man kann sie nicht umgehen. Die beiden Halbräume bekommt man, indem man die Grenze parallel entlang einer geeigneten Halbgeraden verschiebt. Eine Gerade in der Ebene besitzt eine analoge Teilungseigenschaft, eine Gerade im Raum besitzt sie nicht.

(2.8.2) Und die soeben beschriebene geometrische Struktur finden wir in beliebigen Dimensionen wieder!

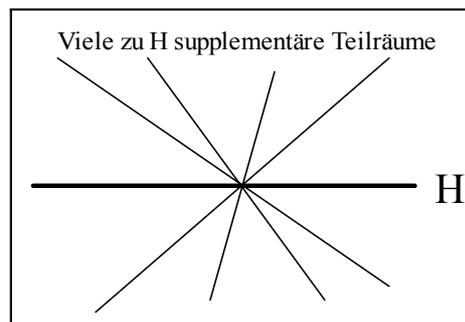
(2.8.3) Ein Teilraum der Dimension  $n-1$  heißt eine **Hyperebene** von  $V$ . Der Kern einer Linearform  $\lambda \neq 0$  ist stets eine Hyperebene.  $\lambda$  selbst spannt einen 1-dimensionalen Teilraum  $G$  in  $V^*$  auf. Das ist unser erstes Beispiel für die Beziehung zwischen den Teilräumen. Es gilt  $\langle \gamma | \vec{x} \rangle = 0$  für alle  $\gamma \in G$  und  $\vec{x} \in H$ .

(2.8.4) Zur Bezeichnung "Hyperebene": Den Wert  $2=3-1$  im dreidimensionalen Konfigurationsraum kann man auf zwei Weisen auf den  $n$ -dimensionalen Fall ausdehnen. Als  $2$  für alle  $n$  (Ebene) oder als  $n-1$  (Hyperebene). D.h. Ebene steht generell für zweidimensional. Und Hyperebene für Raumteiler der Dimension  $n-1$ .

(2.8.5) Was die Verallgemeinerung der dreidimensionalen Anschauung angeht, so überzeugt man sich leicht:  $\text{Kern}\lambda = \{ \vec{x} | \lambda(\vec{x})=0 \}$  übernimmt die Rolle der Grenze.  $H_+ = \{ \vec{x} | \lambda(\vec{x}) > 0 \}$  übernimmt die Rolle des einen "oberen" Halbraumes usw. Die Grenze kann nicht umgangen werden, da der Wert  $\lambda(\vec{x})$  aus Stetigkeitsgründen bei Vorzeichenwechsel den Wert  $0$  annehmen muß.  $\lambda(\vec{x})=c$  liefert eine Parallelverschiebung der Hyperebene.

Hat man insbesondere eine einzige nichttriviale lineare Gleichung für  $n$  Unbestimmte, so liegt genau diese geometrische Konfiguration vor.

(2.8.6) Ist  $H$  eine Hyperebene von  $V$ , so können wir wie üblich eine Basis  $\vec{a}_1, \dots, \vec{a}_{n-1}$  von  $H$  durch  $\vec{a}_n$  zu einer Basis  $a$  von  $V$  ergänzen.



Die übliche Skizze erinnert daran, dass der supplementäre Vektor in keiner Weise eindeutig ist. Es werden eine Vielzahl verschiedener Geraden (in  $V$ ) erzeugt. Unsere zu Anfang formulierte Behauptung war, dass dies in  $V^*$  anders sei. Dort wird durch  $H$  eine einzige Gerade ausgezeichnet. Wir können jetzt bereits eine Vermutung aufstellen: Zu erwarten ist, dass dies die durch  $a_n^*$  erzeugte Gerade  $G$  ist, also das Gegenüber von  $a$  in der zu  $a$  dualen Basis.

(2.8.7) Als weiteres Beispiel betrachten wir ein einfaches  $2 \times 3$  Gleichungssystem.

$$\boxed{\begin{array}{l} -x+y+z=0 \\ +x-y+z=0 \end{array}} \quad \text{Lösung:} \quad \vec{x}_L(\alpha) = \alpha \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$$

Die zugehörigen Zeilenvektoren sind gerade  $2a_1^*$  und  $2a_2^*$  aus der in (2.4.9) konstruierten Basis von  $(\mathbb{R}^3)^*$ . Löst man dieses System, so findet man  $\vec{x}_L(\alpha) = \alpha \vec{a}_3$ . D.h. zwischen Zeilenraum (der die Bedingungen beschreibt, Indizes 1,2) und Lösungsraum (Index 3) besteht gerade eine Beziehung der gewünschten Art. Setzen wir  $\lambda = (-1, 1, 1)$  und  $\mu = (1, -1, 1)$ , so besteht die Lösungsmenge des Systems aus allen  $\vec{x}$ , die  $\lambda(\vec{x})=0$  und  $\mu(\vec{x})=0$  erfüllen.

(□ F. ) Bestimme den Kern, also alle  $\vec{x}$  mit  $\lambda(\vec{x})=0$ . Das ergibt eine Hyperebene.

## 5.2.9 Annihilatoren und Dualität

(2.9.1) Die gefundene Beziehung zwischen Hyperebenen und Gleichungssystemen legt die Definition zu-  
mindest nahe, mit der wir hoffen dürfen, die angekündigte Paarung der Teilräume zu realisieren:

Definition:		Sei $M \subset V$ nicht leere Teilmenge. Dann ist der <i>Annihilator von M</i> die folgende Menge: $M^\perp = \{\varphi \mid \varphi \in V^* \text{ und } \langle \varphi   \vec{x} \rangle = 0 \text{ für alle } \vec{x} \in M\}$ .
$V^*$	$V$	Sei $U^* \subset V^*$ nicht leere Teilmenge. Dann ist analog der <i>Annihilator von <math>U^*</math></i> definiert durch ${}^\perp U^* = \{\vec{x} \mid \vec{x} \in V \text{ und } \langle \varphi   \vec{x} \rangle = 0 \text{ für alle } \varphi \in U^*\}$
$U^*$	$M$	
$\searrow$	$\swarrow$	
$M^\perp$	${}^\perp U^*$	

Jeder nichtleeren Menge von Vektoren aus  $V$  wird hierdurch eine Menge von Linearformen zugeordnet und umgekehrt.

(□ F. ) Was insbesondere dem Nullraum, was dem gesamten Raum?

(2.9.2) Wir hatten bereits gesagt, dass die Schreibweise  $\langle \lambda | \vec{x} \rangle$  an das Skalarprodukt von Vektoren erinnern soll. Jetzt gehen wir mit der Analogie einen Schritt weiter: Beim Skalarprodukt bedeutet  $\vec{x} \cdot \vec{y} = 0$ , dass die beiden Vektoren aufeinander senkrecht stehen. Analog dazu haben wir die Gleichung  $\varphi(\vec{x}) = \langle \varphi | \vec{x} \rangle = 0$ . Gilt diese Gleichung, so sagen wir: "Die Form  $\varphi$  steht senkrecht auf dem Vektor  $\vec{x}$ " und umgekehrt "Der Vektor  $\vec{x}$  steht senkrecht auf der Form  $\varphi$ ". Das ist wieder eine Sprechweise, die Gedächtnis und Intuition unterstützen soll.

Unsere Annihilatordefinitionen besagen:  $M^\perp$  besteht aus allen Formen, die auf ganz  $M$  senkrecht stehen. Und  ${}^\perp U^*$  aus allen Vektoren, die auf ganz  $U^*$  senkrecht stehen.

(2.9.3) Ist  $M\vec{x} = \vec{0}$  ein lineares homogenes Gleichungssystem und ist  $Z$  die Menge der zugehörigen Zeilenvektoren, so ist die Lösungsmenge  $\mathbb{L}$  gerade der Annihilator von  $Z$ , also  $Z^\perp$ . **Beim homogenen Gleichungssystem stehen die Lösungsvektoren senkrecht auf den Bedingungsformen.** Dieser kurze Sachverhalt ist entscheidend für die Behandlung des Gleichungsproblems.

Ist  $\lambda \neq 0$  eine Linearform, so ist der Annihilator  ${}^\perp \{\lambda\}$  gerade die Hyperebene  $\text{Kern } \lambda$ .

(2.9.4) Nach diesen Vorbereitungen stellen wir die für uns wichtigsten Eigenschaften der Annihilatoren zusammen. Beachten Sie, dass es alles Eigenschaften sind, die wir vom üblichen Senkrechtstehen erwarten. Die angekündigte Beziehung zwischen den Teilräumen steckt insbesondere in Punkt d.

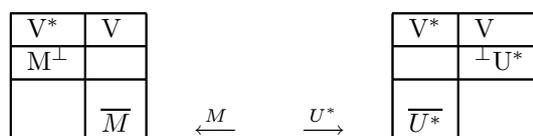
(2.9.5) Zu jeder Aussage gibt es dabei eine spiegelbildliche "duale", die man erhält, indem man  $V$  und  $V^*$  vertauscht. Wir machen das über die Art der Darstellung deutlich.

Satz: Es sei  $V$  endlichdimensionaler Vektorraum und  $M \subset V$  sowie  $U^* \subset V^*$  nicht leere Teilmengen. Und  $\overline{M}$  bzw.  $\overline{{}^\perp U^*}$  seien die daraus erzeugten Teilräume. Dann gilt:

- |  |   |
|--|---|
| a) $M^\perp \subset V^*$ ist Teilraum.                       | a*) ${}^\perp U^* \subset V$ ist Teilraum.  |
| b) $M^\perp = \overline{M^\perp}$                            | b*) $\overline{{}^\perp U^*} = {}^\perp \overline{U^*}$   |
| c) $M_1 \subset M_2 \Rightarrow M_2^\perp \subset M_1^\perp$ | c*) $\overline{U_1} \subset \overline{U_2} \Rightarrow {}^\perp \overline{U_2} \subset {}^\perp \overline{U_1}$ |
| d) $\dim \overline{M} + \dim M^\perp = \dim V$               | d*)   |
| e) ${}^\perp (M^\perp) = \overline{M}$                       | e*)   |

(□ F. ) Ergänzen Sie d\* und e\* selbst.

(2.9.6) Die wichtigeren Teile des Satzes lassen sich in der nachfolgenden Weise diagrammatisch zusammenfassen. Immer liegt die folgende Struktur vor: Eine Menge ist in einem der Räume gegeben, man bildet den daraus erzeugten Teilraum. Und dieser bestimmt auf der anderen "dualen" Seite einen Annihilator in Form eines zweiten Teilraumes, mit gewissen Eigenschaften zur Dimension, die das Diagramm zusammenfaßt.



(2.9.7) Die (nicht eingezeichneten) jeweils supplementären Räume auf der eigenen Seite sind **nicht** eindeutig festgelegt. Man muß immer zur dualen Seite übergehen, um Eindeutigkeit zu erlangen.

(2.9.8) Machen wir uns mit der Bedeutung des Satzes vertraut, indem wir mit seiner Hilfe das Problem der linearen Gleichungssysteme analysieren. ( Die Behandlung dieser Frage war ja der Grund, den Dualraumformalismus in diesem Kapitel zu behandeln. )

(2.9.9) Sei  $A\vec{x} = \vec{b}$  ein solches System und  $U^*$  die Menge der Linearformen der Einzelgleichungen. Laut Definition der Annihilatoren ist dann  ${}^\perp U^* = \text{Kern} A$  (= Raum aller  $\vec{x} \in V$ , die alle Gleichungen von  $A\vec{x} = \vec{0}$  erfüllen). Weiter ist  $Z = \overline{U^*}$  der zugehörige Zeilenraum. Dann gilt  $L = {}^\perp Z$ . D.h. **der Lösungsraum ist der Annihilator des Zeilenraums**. Wegen  $b^*$ ) hängt  $L$  nur von  $Z$ , nicht aber vom gewählten Erzeugendensystem ab. Wählt man anstelle von  $U^*$  ein anderes Erzeugendensystem  $W^*$  von  $Z$  und damit ein anderes Gleichungssystem, so ergibt sich doch dieselbe Lösungsmenge. Dieses Erzeugendensystem kann durchaus auch unendlich viele Elemente haben. Der Zeilenrang ist seiner Konzeption nach die Dimension des Zeilenraumes. Damit erfaßt und rechtfertigt  $d^*$ ) die die erwartete Interpretation der Gleichung  $k = \ell - n$  für lineare Gleichungssysteme:

$(\text{Anzahl der freien Parameter}) + (\text{Zahl der unabhängigen Bedingungen}) \\ = \text{Zahl der Unbestimmten.}$
--

(2.9.10) Und das beantwortet die noch offene Frage nach den allgemeinen Äquivalenzumformungen wie folgt:

**Eine Äquivalenzumformung eines linearen Gleichungssystems im allgemeinsten Sinne bedeutet immer eine andere Wahl einer Erzeugendensystems im Zeilenraum.**

Oder auch: Die Gleichung bestimmt den Zeilenraum. Dieser hat einen eindeutigen Annihilator in  $V$  und das ist der Lösungsraum des homogenen Systems.

(2.9.11) Zum Beweis von (2.7.5): Wir haben hier ein Beispiel, bei dem das Verständnis des Beweises deutlich mehr bringt als die Sicherheit über die Richtigkeit der Aussage Insbesondere zeigt der Beweis auch, wie man den jeweiligen Annihilator finden kann. Etwas, das der Leser bei der Diskussion des Gleichungsproblems vermißt haben könnte. Wir bringen den Beweis von (2.9.5) daher vor der weiteren Diskussion der Aussagen des Satzes.

a) und  $a^*$ ) folgen mit dem Teilraumkriterium aus der Bilinearität von  $\langle \cdot | \cdot \rangle$ . Zu benutzen sind Gleichungen wie  $\langle \lambda + \mu | \vec{x} \rangle = \langle \lambda | \vec{x} \rangle + \langle \mu | \vec{x} \rangle$ . Ist dann eine Seite Null, so ist es auch die andere. Damit folgen auch b) und  $b^*$ )

c) und  $c^*$ ) folgen mit der üblichen Denkfigur. ( Die Bedeutung unbedingt im Diagramm klarmachen. Verkleinerung der einen Seite bedeutet Vergrößerung der anderen)

d) verlangt mehr Überlegung, obwohl man auch hier mit dem Routineansatz durchkommt. Da es um eine Dimension geht, wird man mit einer Basis arbeiten. Wähle eine Basis  $a_1, \dots, a_k$  von  $\overline{M}$ , dem erzeugten Teilraum. Ergänze zu einer Basis  $a_1, \dots, a_n$  von  $V$ . Bilde die zugehörige duale Basis  $a_1^*, \dots, a_n^*$ . Wir vermuten über Diagramminspektion, dass  $a$  Basis von  $M$  ist. Das ist zu beweisen.

Sei also  $N^*$  der von diesen Vektoren aufgespannte Teilraum von  $V^*$ . Zu zeigen ist  $N^* = \overline{M}^\perp$ . Erste Richtung: Sei  $\lambda \in N^*$ . D.h.  $\lambda = a_{k+1}^* \lambda_{k+1} + \dots + a_n^* \lambda_n$ . Laut Definition "duale Basis" folgt  $\langle \lambda | \vec{a}_i \rangle = \lambda(\vec{a}_i) = 0$  für alle  $i$  mit  $1 \leq i \leq k$ . Also  $\lambda(\vec{x}) = 0$  für alle  $\vec{x} \in \overline{M}$ . Und das heißt  $\lambda \in \overline{M}^\perp$

Zweite Richtung: Sei  $\lambda \in \overline{M}^\perp$ . Also  $\lambda(\vec{x}) = 0$  für alle  $\vec{x} \in \overline{M}$ . Andererseits hat  $\lambda$  als Element von  $V^*$  eine Basisdarstellung  $\lambda = a_1^* \lambda_1 + \dots + a_n^* \lambda_n$ . Aus  $\lambda(\vec{a}_i) = 0$  folgt dann natürlich  $\lambda_i = 0$ . Diese Gleichungen gelten für  $i = 1, 2, \dots, k$ . Daher können höchstens die restlichen  $X$ . ungleich Null sein und das heißt  $\lambda \in N$ . Der Beweis ist **fertig bis auf  $d^*$ )**.

Die nachfolgende Skizze zeigt, was der Beweis zusätzlich liefert: Startet man mit einer Basis von  $U$ , die man zu einer von  $V$  ergänzt, dann erhält man in  $V^*$  eine Basis von  $U^\perp$ , die zu einer Basis von  $V$  ergänzt ist. In  $V$  gibt es ohne Skalarprodukt keinen eindeutigen zu  $U$  normalen Raum. Aber in  $V^*$  ist das der Fall und die Basen werden bei Übergang zur Dualbasis mitgeliefert. Die jeweils supplementären Räume sind nicht eindeutig festgelegt. Im Beispiel ist  $\dim V = 10$ ,  $\dim U = 4$  und  $\dim U^\perp = 6$ .

$U^\perp$	}	$a_{10}^*$	$\vec{a}_{10}$
		$a_5^*$	$\vec{a}_5$
		}	$U$
		$a_1^*$	$\vec{a}_1$

(2.9.12) Versucht man den Beweis für d) auf d\*) zu übertragen, so tritt bei sorgfältigem Vorgehen folgendes Problem auf: Jetzt geht man aus von einer Basis  $a^* = (a_1^*, \dots, a_n^*)$  von  $V^*$ . **Und dazu benötigt man eine Basis  $a$  von  $V$ , derart, dass  $a^*$  die duale Basis von  $a$  ist.** Mit diesem Resultat geht der Beweis durch.

(2.7.13) In (2.4.10) haben wir das zum Beweis benötigte Resultat bereits bereitgestellt. (Nicht von den Bezeichnungen irreführen lassen. Die gegebene Basis haben wir so bezeichnet, wie eine duale Basis. Aber damit ist es noch längst keine!) Inspiziert man den zugehörigen Beweis, so sieht man, dass er auf folgender Argumentation beruht: Ist eine Matrix  $M$  invertierbar, so ist auch die inverse Matrix invertierbar. Das sichert die gesuchte Existenz. Die Argumentation verläuft weder koordinatenfrei noch ist sie gut auf den unendlichdimensionalen Fall übertragbar. **Im Anhang gehen wir das Problem - suche zu einer Basis  $b^*$  von  $V^*$  eine Basis  $a$  von  $V$ , so daß  $b^*=a^*$  ist - daher noch auf eine andere Weise an.**

### 0.0.2 5.2.10 Zusammenfassung der Resultate zu den Gleichungssystemen

(2.10.1) Zu jeder homogenen linearen Gleichung gehören zwei Räume: Der  $k$ -dimensionale Lösungsraum  $\mathbb{L}$  und der  $\ell$ -dimensionale Bedingungsraum  $Z=\mathbb{L}^\perp$  mit  $k+\ell=n=\dim V$ . Die Vorgabe der Gleichung bestimmt beide Räume - insbesondere ihre Dimensionen - eindeutig. Die algebraische Verbindung zwischen den beiden Räumen erfolgt durch das kanonische Skalarprodukt, das hier immer den Wert Null ergibt..

(2.10.2) *Vorgeben eines linearen Gleichungssystems* heißt in der Regel: Vorgabe eines Erzeugendensystems des Bedingungsraumes  $Z$  sowie Festlegen der zugehörigen Inhomogenitäten. Lösen eines homogenen Gleichungssystems heißt in der Regel: Bestimmen einer Basis des Lösungsraumes. Hat man eine solche, hat man über die Linearkombinationsabbildung automatisch eine bijektive Parametrisierung der allgemeinen Lösung.

(2.10.3) Der Übergang zu einer äquivalenten Gleichung bedeutet Wahl oder Konstruktion eines anderen Erzeugendensystems des Bedingungsraumes.

(2.10.4) Bei einem **inhomogenen eindeutig lösbaren Gleichungssystem** ist der Bedingungsraum gleich  $V^*$ . Wahl einer Basis von Bedingungen bestimmt eine dazu duale Basis von  $V$ . Das ist eine Basis von Elementarlösungen, mit deren Hilfe man die Lösung für jede Inhomogenität durch Linearkombination erhält.

(2.10.5) Bei einem **inhomogenen vieldeutig lösbaren, unterbestimmten System** ist der Bedingungsraum ein echter  $\ell$ -dimensionaler Teilraum von  $V^*$ . Jede Basis von Bedingungen ist erst noch zu einer Gesamtbasis zu ergänzen. D.h. man hat noch  $k=n-\ell$  zusätzliche unabhängige Bedingungsgleichungen hinzuzunehmen. Da es sich um zusätzliche Bedingungen handelt, muß die Lösung des neuen Systems in der Lösungsmenge des ursprünglichen Systems enthalten sein. Insbesondere dürfen wir die Inhomogenitäten der zusätzlichen Gleichungen frei vorgeben, d. h. als äußeren Parameter ansetzen.

(2.10.6) Zur ergänzten Basis gehört erneut eine duale Basis von Elementarlösungen. Linearkombination der ersten Elementarlösungen ergibt die übliche spezielle Lösung der inhomogenen Ausgangsgleichung. Die ergänzten Gleichungen haben freie Inhomogenitäten und ergeben den Kern, also die allgemeine Lösung der homogenen Gleichung. Jede ergänzte Gleichung entspricht der Wahl eines freien Parameters für eine Koordinate oder eine Kombination von Koordinaten beim Eliminationsprozeß.

(2.10.7) Der Dualitätssatz zeigt, dass und wie jede Lösungsprozedur zu einem Resultat der geläufigen Struktur führt. Überdies sieht man, dass die Ergänzung des Systems durch  $k$  unabhängige Bedingungen mit freier Inhomogenität immer zum Kern und damit zu sämtlichen Lösungen des ursprünglichen Systems führt.

(2.10.8) Als Beispiel behandeln wir die Frage aus (1.3.7). **”Warum haben die nachfolgenden Gleichungssysteme dieselbe Lösungsmenge?”**

$3x+z=0$	$2x-y+3z=0$ $x+y-2z=0$
$3y-7z=0$	
$7x+y=0$	

Das zweite System hat die Zeilenvektoren  $z_1 = (2 - 1, 3)$  und  $z_2 = (1, 1, -2)$ . Sie spannen einen zweidimensionalen Zeilenraum  $Z \subset \mathbb{R}^{3*}$  auf, für den die beiden (einfacheren) Vektoren  $a_1^* = (3, 0, 1) = z_1 + z_2$  und  $a_2^* = (0, 3, -7) = -z_1 + 2z_2$  eine Basis bilden. Wir können durch  $a_3^* = ((0, 1, 0)$  zu einer Basis von  $\mathbb{R}^{3*}$  ergänzen. Nun wird das erste System durch die drei Zeilenvektoren  $u_1 = (3, 0, 1) = a_1^*$ ,  $u_2 = (0, 3 - 7) = a_2^*$  und  $u_3 = (7, 1, 0)$  gebildet. Wegen  $7u_1 + u_2 = 3u_3$  erzeugen  $u_1, u_2, u_3$  erneut den Raum  $Z$ . **Beide Systeme erzeugen denselben Zeilenraum.** und in beiden Fällen besteht der Lösungsraum aus dem orthogonalen Komplement von  $Z$ . Konstruiert man die duale Basis zur gegebenen Basis

$$\vec{a}_1 = \frac{1}{3} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \vec{a}_2^* = \frac{1}{7} \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix} \quad \vec{a}_3 = \frac{1}{7} \begin{pmatrix} -1 \\ 7 \\ 3 \end{pmatrix}$$

so wird der Lösungsraum  $\mathbb{L}$  beider Gleichungen durch  $\vec{a}_3$  erzeugt, was man leicht direkt verifiziert.

### 5.2.11 Anhang: Der Bidualraum von $V$ .

**(2.11.1)** Jeder Vektorraum hat seinen zugehörigen Dualraum. Auch  $V^*$  ist Vektorraum und besitzt daher einen Dualraum  $(V^*)^*$  oder kürzer  $V^{**}$ . Diesen Raum nennen wir den Bidualraum von  $V$ . Seine Elemente sind definitionsgemäß lineare Abbildungen  $V^* \rightarrow K^1$ . Sie machen aus Linearformen Zahlen. Formal:

$$\beta = (V^*, \lambda \mapsto \beta(\lambda), K) \quad \text{mit} \quad \lambda = (V, \vec{x} \mapsto \lambda(\vec{x}), K).$$

**(2.11.2)** Die Bedeutung des Bidualraumes liegt darin, dass es eine kanonische Abbildung zwischen  $V$  und  $V^{**}$  gibt, die wir ja zwischen  $V$  und  $V^*$  gerade nicht hatten. Diese Abbildung sieht wie folgt aus:

$$\iota = (V, \vec{x} \mapsto \iota_{\vec{x}}, V^{**}) \quad \text{mit} \quad \iota_{\vec{x}} = (V^*, \lambda \mapsto \iota_{\vec{x}}(\lambda) = \dots, K)$$

**(2.11.3)** Beachten Sie: **Bis zu dieser Stelle ist alles nur Explikation des Formalismus.** Alles was fehlt, was nicht trivial ist, ist eine durch die Lücke angedeutete Berechnungsformel für den Wert. Und zwar muß der fehlende Term ein Körperelement ergeben. Und da nimmt man naheliegenderweise folgende Formel:

$$\boxed{\iota_{\vec{x}}(\lambda) = \langle \lambda | \vec{x} \rangle}$$

Die rechte Seite ergibt wirklich ein Körperelement, das man aus den beiden Eingabegrößen  $\vec{x}$  und  $\lambda$  bilden kann.

**(2.11.4)** Und jetzt zu den Eigenschaften dieser Abbildung  $\iota$  ( "Jota" ):

**Satz:**  $\iota : V \rightarrow V^{**}$  bildet immer einen **kanonischen Isomorphismus** zwischen  $V$  und  $\text{Bild}(\iota)$ .

**Ist**  $V$  endlichdimensional, so ist  $\iota$  ein kanonischer Isomorphismus zwischen  $V$  und  $V^{**}$ . Andernfalls ist  $\iota$  nur injektiv, aber nicht surjektiv.

Kurz: Dualraum und Bidualraum unterscheiden sich in Hinblick auf die Vektorraumstruktur nicht. Ihre Elemente gehen durch willkürfreie Umbenennung auseinander hervor.

**(2.11.5)** Beweis: a)  $\iota$  ist linear. Zu zeigen ist  $\iota_{\vec{x}+\vec{y}} = \iota_{\vec{x}} + \iota_{\vec{y}}$  und  $\iota_{\alpha\vec{x}} = \alpha\iota_{\vec{x}}$ . Die erste Gleichung verlangt  $\iota_{\vec{x}+\vec{y}}(\lambda) = \iota_{\vec{x}}(\lambda) + \iota_{\vec{y}}(\lambda)$  für alle  $\lambda$  aus  $V^*$ . Das besagt aber laut Definition:  $\langle \lambda | \vec{x} + \vec{y} \rangle = \langle \lambda | \vec{x} \rangle + \langle \lambda | \vec{y} \rangle$ . Diese Beziehung gilt aber als Folger der

Bilinearität des kanonischen Skalarproduktes. Der zweite Teil folgt analog.

b) Zu zeigen:  $\text{Kern}(\iota) = \{\vec{0}\}$ . Für welche  $\vec{x} \in V$  ist  $\iota_{\vec{x}} = 0 = (V^*, \lambda \mapsto 0, K)$ ? D.h.  $0 = \iota_{\vec{x}}(\lambda) = \langle \lambda | \vec{x} \rangle$  für alle  $\lambda \in V^*$ ? Sei  $a$  eine Basis von  $V$  und  $a^*$  die dazu duale Basis (bzw. duale Familie, wenn  $\dim V = \infty$ ). Die Beziehung muß auch für alle  $a_i^*$  aus  $a^*$  gelten. Also  $a_i^*(\vec{x}) = 0$ . D.h.: alle Komponenten von  $x$  bezüglich der Basis  $a$  sind Null. Also  $x = 0$ . Der Kern ist trivial.

c) Ist  $\iota$  surjektiv? Im endlichdimensionalen Fall folgt dies sofort aus Dimensionsgründen:  $\dim \text{Bild}(\iota) = \dim V = n$ . Andererseits wissen wir  $\dim V^{**} = n$  und  $\text{Bild}(\iota) \subset V^{**}$ . Im unendlichdimensionalen Fall geht das Argument nicht durch.

(2.11.6) Was folgt für unser Hauptproblem aus (2.7.12) im Falle  $\dim V = n < \infty$ ? Sei  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$  eine Basis des Dualraumes  $V^*$  und  $\lambda^* = (\lambda_1^*, \dots, \lambda_n^*)$  die zugehörige Dualbasis aus  $V^{**}$ , dem Dualraum von  $V$ . Da  $\iota$  Isomorphismus ist, existiert die inverse Abbildung  $\iota^{-1}$ . Wir bilden die Familie  $a = (\vec{a}_1, \dots, \vec{a}_n)$  mit  $\vec{a}_i = \iota^{-1}(\lambda_i^*)$ . Also  $\iota(\vec{a}_i) = \lambda_i^*$ . Das setzen wir ein in die definierende Gleichung  $\delta_{ij} = \lambda_i^*(\lambda_j)$ . Es folgt  $\delta_{ij} = \lambda_i^*(\lambda_j) = \langle \lambda_j | \vec{a}_i \rangle$ . Die so erhaltene Beziehung  $\delta_{ij} = \langle \lambda_j | \vec{a}_i \rangle$  besagt aber gerade, dass  $\lambda$  die duale Basis von  $a$  ist.

Fassen wir zusammen:

$$\begin{array}{ccccc}
 V^{**} & & V^* & & V \\
 \lambda^* & \xleftarrow{\text{dual}} & \lambda & \boxed{a^* = \lambda!!} & \vec{a} \\
 \searrow & \rightarrow & \xrightarrow{\iota^{-1}} & \nearrow & 
 \end{array}$$

Man gibt  $\lambda$  aus  $V^*$  vor und bildet dazu die biduale Basis  $\lambda^*$  von  $V^{**}$ . mit  $\iota^{-1}$  gelangen wir nach  $V$  und erhalten die Basis  $a$ . Deren Dualbasis ist gerade das gegebene  $\lambda^*$ .

Damit können wir den Beweis des Dualitätssatzes abschließen. Der Beweis von d\*) läuft jetzt analog zu dem von d).

Und was e) und e\*) betrifft, so müssen wir einfach nur die jeweiligen Basen der betrachteten Teilräume verfolgen. Etwa von  $U$  nach  $U^\perp$  und dann weiter nach  ${}^\perp(U^\perp)$ .