



UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

Höhere Mathematik B für Elektrotechniker

Kerstin Hesse

Universität Paderborn, Sommersemester 2023

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \iff \mathbf{A} \vec{x} = \vec{b}.$$

Die Grundlage für dieses Skript bildeten die handschriftlichen Aufzeichnungen zur „Höheren Mathematik B für Elektrotechniker“ (kurz: HM B) von Frau Dr. Cornelia Kaiser aus dem Sommersemester 2014, die von der Autorin im Sommersemester 2015 zu einem Skript (mit gemeinsamer Autorenschaft von Frau Hesse und Frau Kaiser) ausgebaut wurden. Die Autorin dankt Frau Prof. Dr. Andrea Walther für die Nutzung ihrer Vorlesungsmaterialien zur Numerischen Linearen Algebra und dafür, dass sie die Arbeitskraft ihrer Sekretärin Frau Karin Senske zur Unterstützung beim Erstellen des Manuskripts im akademischen Jahr 2014/15 zur Verfügung stellte. Ebenso dankt sie natürlich Frau Senske für ihre Mitarbeit an dem Manuskript.

Das Skript vom Sommersemester 2015 wurde im Winter/Frühjahr 2017 von Frau Hesse überarbeitet und ergänzt. Dabei wurde die gemeinsame Autorenschaft aufgehoben, und Frau Kaiser und Frau Hesse werden basierend auf dem ursprünglichen gemeinsamen Skript für weitere HM B-Durchgänge jeweils ihre eigene Version des Skripts zur Verfügung stellen. Die Version von Frau Hesse ist das vorliegende Skript.

Paderborn, März 2023

Kerstin Hesse

Einleitung

Als Studierende/r der Elektrotechnik, des Computer Engineering, des Wirtschaftsingenieurwesens (Elektrotechnik) oder der Physik müssen Sie die Vorlesungen „Höhere Mathematik A für Elektrotechniker (HM A)“, „Höhere Mathematik B für Elektrotechniker (HM B)“ und „Höhere Mathematik C für Elektrotechniker (HM C)“ hören, um das nötige mathematische Wissen für Ihren Studiengang zu erwerben.

Warum sollten Sie die Zeit investieren, um die mathematischen Inhalte und Techniken dieser Kurse zu lernen und verstehen?

Mathematik ist die Sprache der Naturwissenschaften, denn physikalische und technische Phänomene lassen sich nur mit der Sprache der Mathematik sauber beschreiben und modellieren. So wird der Zerfallsprozess einer radioaktiven Substanz durch eine Differentialgleichungen beschrieben, und die physikalischen Gesetze des Elektromagnetismus werden durch die Maxwellschen Gleichungen, ein System partieller Differentialgleichungen, beschrieben. Wärmeleitung, Wellen und Schwingungen, sowie Strömungsprozesse können nur mit Mathematik formal beschrieben werden.

Diese Beispiele machen deutlich, dass die Mathematik eine unverzichtbare „Sprache“ für jede*n Ingenieur*in und Naturwissenschaftler*in ist. Ohne die nötigen Mathematikkenntnisse werden Sie auch in den Fachvorlesungen Ihres Studiengangs meist wenig verstehen können. Sie sollten daher von Anfang an die nötige Zeit investieren, um die Inhalte der Vorlesungen HM A, HM B und HM C richtig zu verstehen und systematisch zu lernen.

Was wird in der HM B besprochen?

In **Teil IV „Lineare Algebra“** beschäftigen wir uns mit linearen Gleichungssystemen und Matrizen, sowie mit Vektorräumen und ihren Eigenschaften.

Lineare Gleichungssysteme mit zwei (bzw. drei) Gleichungen und zwei (bzw. drei) Unbekannten mit einer eindeutigen Lösung haben wir bereits in der HM A (siehe Kapitel 3 in Skript der HM A) mit der Cramerschen Regel gelöst. Nun betrachten

wir allgemeine Gleichungssysteme mit m Gleichungen und n Unbekannten, und wir lernen diese mit dem Eliminationsverfahren von Gauß zu lösen und erhalten auch eine erste Lösungstheorie für lineare Gleichungssysteme.

Sie kennen aus der HM A (siehe Kapitel 2 und 3 im Skript der HM A) bereits den Vektorraum \mathbb{R}^n sowie quadratische Matrizen. Auch die quadratischen Matrizen bilden einen Vektorraum, und wir werden sehen, dass beispielsweise auch die Menge der stetigen auf \mathbb{R} definierten Funktionen einen Vektorraum bildet. Für Vektoren aus Vektorräumen lernen wir neue Konzepte kennen: lineare Unabhängigkeit und lineare Hülle, sowie die Begriffe der Basis und der Dimension eines Vektorraums. Mit Hilfe der Vektorraumtheorie können wir unsere Lösungstheorie der linearen Gleichungssysteme vervollständigen. Dabei lernen wir den neuen Begriff des Rangs einer Matrix kennen.

Danach interessieren wir uns für quadratische Matrizen (also Matrizen mit gleich vielen Zeilen wie Spalten). Wir lernen die inverse Matrix einer invertierbaren quadratischen Matrix kennen und führen die Determinante einer beliebigen quadratischen Matrix ein. Aus der HM A (siehe Kapitel 3 im Skript der HM A) kennen wir bereits die Determinante von 2×2 - und 3×3 -Matrizen. Schließlich interessieren wir uns noch für Eigenwerte und Eigenvektoren von quadratischen Matrizen. Dieses zunächst vielleicht ungewöhnlich anmutende Konstrukt findet direkt in Teil V der Vorlesung eine praktische Anwendung.

In **Teil V „Lineare Differentialgleichungen“** betrachten wir lineare Differentialgleichungen zweiter und höherer Ordnung und anschließend Systeme linearer Differentialgleichungen. Sie kennen lineare Differentialgleichungen erster Ordnung bereits aus der HM A (siehe Kapitel 9 im Skript der HM A).

In **Teil VI „Differentialrechnung in mehreren Variablen“** lernen wir Funktionen mit mehreren Variablen (z.B. mehrere Ortskoordinaten und die Zeit) kennen und lernen, wie man diese differenziert und mit Hilfe der Ableitungen Extremwertprobleme löst. Dabei tauchen die neuen Begriffe der Jacobi-Matrix (der ersten partiellen Ableitungen) und der Hesse-Matrix (der zweiten partiellen Ableitungen) auf. Hier werden die Ideen der Differentialrechnung in einer Variablen verallgemeinert. Wir lernen noch nicht, wie man Funktionen mit mehreren Variablen integriert; dieses ist ein zentraler Teil des Stoffes der HM C.

Was für mathematisches Wissen wird vorausgesetzt?

Die HM B baut ganz wesentlich auf die Inhalte der HM A auf. Dieses gilt insbesondere für Teil V „Lineare Differentialgleichungen“ und Teil VI „Differentialrechnung in mehreren Variablen“ der Vorlesung. Wenn Sie bei den Inhalten der HM A Defizite haben sollten, so müssen Sie diese unbedingt beheben! Aus der

Schule werden ansonsten eine solide Beherrschung der Rechentechniken der Mittelstufe vorausgesetzt: Klammersetzung, Vorzeichenregeln, binomische Formeln, Assoziativ-, Kommutativ- und Distributivgesetz, Bruchrechnung,

Wie sollte man für die HM B lernen?

- **Kommen Sie immer zu den Vorlesungen und nehmen Sie aktiv an diesen teil:** Bringen Sie die Beamer-Folien in die Vorlesungen mit, schreiben Sie die Beispiele und Herleitungen von der Tafel mit, oder machen sie sich zumindest Notizen, damit Sie die Vorlesungen nacharbeiten können. Wenn Sie das Skript dabei haben, dann können Sie dieses natürlich auch mit Anmerkungen versehen. Denken Sie mit, und versuchen Sie möglichst viel bereits in den Vorlesungen zu verstehen.
- **Lassen Sie sich in den Vorlesungen nicht durch Ihr Smartphone, Tablet oder Handy ablenken!** Nur wenn Sie sich ganz auf die Vorlesungen konzentrieren, haben Sie eine Chance, die mathematischen Inhalte direkt in den Vorlesungen zu verstehen.
- **Gehen Sie immer zu Ihrer Übungsgruppe und bearbeiten Sie die Gruppenübungen** (diese werden in der Übungsgruppe bearbeitet) und die **Hausübungen** (diese sollten Sie nach der Übung zu Hause bearbeiten). Schauen Sie die Übungszettel **vor** dem Besuch der Übungsgruppe an, und versuchen Sie die Gruppenübungen bereits vor dem Besuch der Übungsgruppe zu rechnen, so dass Sie dort, wo Sie Probleme haben, konkret nachfragen können und von der Tutorin bzw. dem Tutor Hilfe bekommen. Wenn Sie unvorbereitet in die Übungsgruppe kommen, ziehen Sie nur einen sehr geringen Nutzen aus Ihrer Übungsgruppe. **Mathematik lernt sich nur durch Übung, d.h. indem man die mathematischen Techniken für Beispiele und Übungsaufgaben anwendet.** Daher ist es unerlässlich, dass Sie die Übungsaufgaben bearbeiten!
- **Wenn Sie die Übungsaufgaben lösen, dann sollten Sie parallel dazu das zugehörige Material aus den Vorlesungen nacharbeiten.** Dieses passiert ganz „natürlich“, denn die Übungsaufgaben sind so konzipiert, dass Sie mit ihnen den Vorlesungsstoff anwenden und üben. Das Nacharbeiten kann mit den Beamer-Folien und Ihren handschriftlichen Notizen der Beispiele und Herleitungen von der Tafel und/oder mit diesem Skript erfolgen. Das Skript ist dabei wesentlich ausführlicher als die Beamer-Folien und der Tafelanschrieb und somit als Ihre handschriftlichen Notizen. Im Skript finden Sie weitere und teilweise andere Beispiele und zusätzliche Erklärungen. Das Skript kann wie ein Lehrbuch verwendet werden.

- **Was machen Sie, wenn Sie etwas nicht verstehen?** Wichtig ist vor allem, zu wissen, dass dieses bei mathematischen Themen völlig normal ist und allen Studierenden hin und wieder passiert! Was können Sie tun, um das Problem zu beheben?
 - Geben Sie nicht auf, sondern befassen Sie sich weiter mit den mathematischen Inhalten. Manche mathematischen Themen muss man mehrfach studieren, bis „der Groschen fällt“.
 - Fragen Sie Ihre Kommilitoninnen und Kommilitonen danach und diskutieren Sie mit ihnen darüber.
 - Fragen Sie die Dozentin in den Vorlesungen und/oder die Tutorin bzw. den Tutor in den Übungen.
 - Schauen Sie die zu dem Material gehörigen Beispiele an: Mathematik lernt sich durch das Verständnis der Beispiele. Wenn Sie das Beispiel verstehen, dann wird die mathematische Technik klarer. Können Sie nun vielleicht ein ähnliches Beispiel selber durchrechnen? Wenn ja, dann sind Sie einen Schritt weiter gekommen.
 - Lesen Sie ein Thema, mit dem Sie Probleme haben, in einem Lehrbuch nach, um eine alternative Darstellung zu bekommen.
- Nutzen Sie die Gelegenheit und trauen Sie sich, in den **Vorlesungen und in den Übungen Fragen zu stellen**. Es gibt keine dummen Fragen, sondern dumm ist nur, wenn man nicht fragt und ignorant bleibt. Die Vorlesungen und die Übungen sind dazu da, Sie beim Lernen zu unterstützen – also machen Sie von der Gelegenheit, Fragen zu stellen, Gebrauch!
- **Gruppenarbeit:** Gruppenarbeit ist nützlich und kann sehr produktiv sein. Übungsaufgaben sind oft leichter zu lösen, wenn verschiedene Personen ihre Ideen beisteuern. Indem Sie sich von anderen etwas erklären lassen, lernen Sie etwas dazu. Wenn Sie anderen etwas erklären, so lernen Sie auch etwas dazu und gewinnen größere Klarheit über das bereits verstandene Material. Wichtig ist aber, dass Sie nach der Gruppenarbeit nun auch in der Lage sind, die gelösten Aufgaben **eigenständig zu rechnen**, denn in der Klausur sind Sie auf sich alleine gestellt und haben keine Gruppe zur Hand.
- **Klausurvorbereitung:** Wenn Sie während des Semesters die Vorlesungen gut nachgearbeitet haben und die Übungsaufgaben erfolgreich gelöst haben, dann sind Sie bereits gut vorbereitet. Wiederholen Sie den Stoff noch einmal, rechnen Sie zu allen Themen passende Übungsaufgaben und lernen Sie das nötige Wissen. (Es gibt in der Klausur keine Formelsammlung, keinen Taschenrechner und keine sonstigen Hilfsmittel!)

Zum Schluss noch eine **Warnung: Mathematische Themen bauen aufeinander auf!** Man kann sich als gutes Modell den Bau einer Mauer vorstellen. In der HM A legen Sie die ersten drei Schichten/Reihen Ziegelsteine der Mauer. Wo

Sie Wissens- und Verständnislücken haben, fehlen Ziegelsteine. Die Mauer kann bereits hier lokal einbrechen. Mit der HM B legen Sie die vierte bis sechste Reihe Ziegelsteine der Mauer und in der HM C die siebte bis neunte Reihe. Wo bereits Lücken in den ersten drei Reihen der Mauer sind, können die vierte bis sechste Reihe Ziegelsteine nicht stabil aufgelegt werden und brechen sogar ein. Erst wenn Sie Ihre Wissens- und Verständnislücken aus der HM A geschlossen haben, können Sie alle Inhalte der HM B richtig verstehen. Analoges gilt für die HM C. Es ist daher ganz wichtig, dass Sie beim Nacharbeiten und Verstehen der Vorlesungsinhalte „am Ball bleiben“, damit Ihre Mauer aus mathematischem Wissen keine Lücken aufweist und Sie in der HM C auf die HM A und HM B aufbauen können.

Ich freue mich auf Ihre Teilnahme an der HM B!

Kerstin Hesse

Paderborn, März 2023

Literaturverzeichnis

Bei der Erstellung der Skripten der Vorlesungen HM A, HM B und HM C wurde die unten aufgelistete Literatur verwendet. Nicht alle dieser Bücher sind als Literatur für die Vorlesungen HM A, HM B und HM C geeignet. Für konkrete Literaturempfehlungen zur der jeweiligen HM-Vorlesung konsultieren Sie bitte das jeweilige Kursdokument, welches in der ersten Vorlesung verteilt wurde und welches Sie jeweils in PANDA herunterladen können.

- [1] Herbert Amann, Joachim Escher: Analysis I. Basel: Birkhäuser Verlag, 2006.
- [2] Herbert Amann, Joachim Escher: Analysis II. Basel: Birkhäuser Verlag, 2008.
- [3] Herbert Amann, Joachim Escher: Analysis III. Basel: Birkhäuser Verlag, 2009.
- [4] Bernd Aulbach: Gewöhnliche Differenzialgleichungen, 2. Auflage. München: Spektrum Akademischer Verlag, 2010.
- [5] Hans-Jochen Bartsch: Kleine Formelsammlung Mathematik, 5. aktualisierte Auflage. München: Fachbuchverlag Leipzig, 2011.
- [6] Ilja N. Bronstein, Konstantin A. Semendjajew, Gerhard Musiol, Heiner Mühlig: Taschenbuch der Mathematik. Haan-Gruiten: Europa-Lehrmittel, 2016.
- [7] Klemens Burg, Herbert Haf, Friedrich Wille: Höhere Mathematik für Ingenieure. Band I: Analysis, 10. Auflage. Wiesbaden: Springer Vieweg, 2013.
- [8] Klemens Burg, Herbert Haf, Friedrich Wille: Höhere Mathematik für Ingenieure. Band II: Lineare Algebra, 7. Auflage. Wiesbaden: Springer+Teubner Verlag, 2012.
- [9] Klemens Burg, Herbert Haf, Friedrich Wille: Höhere Mathematik für Ingenieure. Band III: Gewöhnliche Differentialgleichungen, Distributionen, Integraltransformationen, 6. Auflage. Wiesbaden: Springer Vieweg, 2013.
- [10] Klemens Burg, Herbert Haf, Friedrich Wille, Andreas Meister: Vektoranalysis, 2. Auflage. Wiesbaden: Springer+Teubner Verlag, 2012.

-
- [11] Gerd Fischer: Lineare Algebra, 18. Auflage. Wiesbaden: Springer Spektrum, 2014.
- [12] Otto Forster: Analysis 1. Differential- und Integralrechnung einer Veränderlichen, 12. Auflage. Wiesbaden: Springer Spektrum, 2016.
- [13] Otto Forster: Analysis 2. Differentialrechnung in \mathbb{R}^n , gewöhnliche Differentialgleichungen, 10. Auflage. Wiesbaden: Springer Spektrum, 2013.
- [14] Peter Furlan: Das gelbe Rechenbuch: für Ingenieure, Naturwissenschaftler und Mathematiker; Rechenverfahren der höheren Mathematik in Einzelschritten erklärt; mit vielen ausführlich gerechneten Beispielen; Band 1: Lineare Algebra, Differentialrechnung. Dortmund: Furlan Verlag, 2006.
- [15] Peter Furlan: Das gelbe Rechenbuch: für Ingenieure, Naturwissenschaftler und Mathematiker; Rechenverfahren der höheren Mathematik in Einzelschritten erklärt; mit vielen ausführlich gerechneten Beispielen; Band 2: Integralrechnung, mehrdimensionale Differentialrechnung, mehrdimensionale Integralrechnung. Dortmund: Furlan Verlag, 2007.
- [16] Peter Furlan: Das gelbe Rechenbuch: für Ingenieure, Naturwissenschaftler und Mathematiker; Rechenverfahren der höheren Mathematik in Einzelschritten erklärt; mit vielen ausführlich gerechneten Beispielen; Band 3: Gewöhnliche Differentialgleichungen, Funktionentheorie, Integraltransformationen, partielle Differentialgleichungen. Dortmund: Furlan Verlag, 2012.
- [17] Karl-Heinz Goldhorn, Hans-Peter Heinz: Mathematik für Physiker 1, Grundlagen aus Analysis und Linearer Algebra, 1. Auflage. Berlin, Heidelberg: Springer-Verlag, 2007.
- [18] Karl-Heinz Goldhorn, Hans-Peter Heinz: Mathematik für Physiker 2, Funktionentheorie – Dynamik – Mannigfaltigkeiten – Variationsrechnung, 1. Auflage. Berlin, Heidelberg: Springer-Verlag, 2007.
- [19] Karl-Heinz Goldhorn, Hans-Peter Heinz: Mathematik für Physiker 3, Partielle Differentialgleichungen – Orthogonalreihen – Integraltransformationen, 1. Auflage. Berlin, Heidelberg: Springer-Verlag, 2008.
- [20] Kerstin Hesse: MATH2120 – Mathematical Methods for Differential Equations, Outline Lecture Notes. Sydney: The University of New South Wales, 2006. (Das Copyright dieses Skripts liegt bei der University of New South Wales.)
- [21] Kerstin Hesse: Mathematik für Chemiker (Vorlesungsskript). Universität Paderborn, 2016.
- [22] Harro Heuser: Lehrbuch der Analysis, Teil 1, 17. Auflage. Wiesbaden: Vieweg+Teubner Verlag, 2009.

-
- [23] Harro Heuser: Lehrbuch der Analysis, Teil 2, 14. Auflage. Wiesbaden: Vieweg+Teubner Verlag, 2008.
- [24] R. Kent Nagle, Edward B. Saff, Arthur David Snider: Fundamentals of Differential Equations, 6. Auflage. Pearson Addison Wesley, 2004.
- [25] MATH2011 – Several Variable Calculus (Vorlesungsskript). Sydney: The University of New South Wales, 2003. (Das Copyright dieses Skripts liegt bei der University of New South Wales.)
- [26] Kurt Meyberg, Peter Vachenauer: Höhere Mathematik 1: Differential- und Integralrechnung, Vektor- und Matrizenrechnung, 6. Auflage. Berlin, Heidelberg: Springer-Verlag, 2001.
- [27] Kurt Meyberg, Peter Vachenauer: Höhere Mathematik 2: Differentialgleichungen, Funktionentheorie, Fourier-Analysis, Variationsrechnung, 4. Auflage. Berlin, Heidelberg: Springer-Verlag, 2001.
- [28] Lothar Papula: Mathematik für Ingenieure und Naturwissenschaftler, Band 1: Ein Lehr- und Arbeitsbuch für das Grundstudium, 14. Auflage. Wiesbaden: Springer Vieweg, 2014.
- [29] Lothar Papula: Mathematik für Ingenieure und Naturwissenschaftler, Band 2: Ein Lehr- und Arbeitsbuch für das Grundstudium, 14. Auflage. Wiesbaden: Springer Vieweg, 2015.
- [30] Lothar Papula: Mathematik für Ingenieure und Naturwissenschaftler, Band 3: Vektoranalysis, Wahrscheinlichkeitsrechnung, Mathematische Statistik, Fehler- und Ausgleichsrechnung, 7. Auflage. Wiesbaden: Springer Vieweg, 2016.
- [31] Lothar Papula: Mathematik für Ingenieure und Naturwissenschaftler – Anwendungsbeispiele, 7. Auflage. Wiesbaden: Springer Vieweg, 2015.
- [32] Andrea Walther: Numerische Mathematik 1 (Vorlesungsskript). Universität Paderborn.

Inhaltsverzeichnis

IV	Lineare Algebra	1
11	Lineare Gleichungssysteme	3
11.1	Vektoren und Matrizen	3
11.2	Lineare Gleichungssysteme: Notation	19
11.3	Das Gaußsche Eliminationsverfahren	21
11.4	Lösungstheorie für lineare Gleichungssysteme	31
12	Vektorräume	35
12.1	Vektorräume: Definition und Beispiele	36
12.2	Untervektorräume	44
12.3	Linearkombinationen	48
12.4	Lineare Unabhängigkeit	54
12.5	Basis und Dimension	59
12.6	Koordinaten	67
12.7	Der Rang einer Matrix	72
12.8	Skalarprodukt und Norm	82
12.9	Orthogonalität	86
13	Quadratische Matrizen	95
13.1	Einführung zu quadratischen Matrizen	95
13.2	Invertierbare Matrizen	99
13.3	Die Determinante	107
13.4	Rechenregeln für Determinanten	115
13.5	Einige Anwendungen von Determinanten	123
13.6	Eigenwerte und Eigenvektoren	128

V	Lineare Differentialgleichungen	141
14	Lineare Differentialgleichungen höherer Ordnung	143
14.1	Einführung zu linearen Differentialgleichungen höherer Ordnung . . .	143
14.2	Lösungstheorie für die homogene lineare Differentialgleichung . . .	146
14.3	Lösungstheorie für die inhomogene lineare Differentialgleichung . . .	148
14.4	Homogene lineare DGL mit konstanten Koeffizienten: $n = 2$	150
14.5	Homogene lineare DGL mit konst. Koeffizienten: $n \in \mathbb{N}$ beliebig . . .	159
14.6	Inhomogene lineare DGL mit konstanten Koeffizienten	165
14.7	Anfangswertaufgaben	171
15	Systeme linearer Differentialgleichungen	175
15.1	Wege in \mathbb{R}^n	175
15.2	Allgemeine Systeme linearer Differentialgleichungen	180
15.3	Homogene Systeme mit konstanten Koeffizienten: $n = 2$	190
15.4	Ein Anwendungsbeispiel	201
15.5	Homogene Systeme mit konstanten Koeffizienten: $n \in \mathbb{N}$ beliebig . . .	206
15.6	Zusammenhang mit linearen DGLen höherer Ordnung	214
VI	Differentialrechnung in mehreren Variablen	221
16	Konvergenz, Stetigkeit und Differenzierbarkeit	223
16.1	Offene Teilmengen und Konvergenz in \mathbb{R}^n	223
16.2	Konvergenz und Stetigkeit für Funktionen	237
16.3	Partielle Ableitungen	243
16.4	Totale Differenzierbarkeit	251
16.5	Die Kettenregel	261
16.6	Höhere Ableitungen	267
16.7	Der Satz von Taylor	270
17	Extremwertaufgaben	277
17.1	Lokale Extremalstellen	277
17.2	Symmetrische Matrizen	283
17.3	Hinreichende Bedingung für lokale Extremalstellen	291
17.4	Globale Extrema	298
17.5	Extrema unter Nebenbedingungen	303

Teil IV

Lineare Algebra

KAPITEL 11

Lineare Gleichungssysteme

In der HM A haben wir in Kapiteln 2 und 3 der HM A bereits Vektoren, quadratische Matrizen und lineare Gleichungssysteme mit gleich vielen Gleichungen wie Unbekannten kennengelernt. Wir haben die Determinante von 2×2 -Matrizen und 3×3 -Matrizen eingeführt und haben als Lösungsverfahren für lineare Gleichungssysteme mit zwei (bzw. drei) Gleichungen und zwei (bzw. drei) Unbekannten und mit einer eindeutig bestimmten Lösung die Cramersche Regel kennengelernt. In diesem und den beiden nachfolgenden Kapiteln werden die bereits bekannten Ideen verallgemeinert und weiter ausgebaut.

In Teilkapitel 11.1 führen wir Vektoren und Matrizen mit reellen bzw. komplexen Zahlen als Einträgen ein. In Teilkapitel 11.2 schreiben wir beliebige reelle oder komplexe lineare Gleichungssysteme mit Vektoren und Matrizen, und in Teilkapitel 11.3 lernen wir das Gaußsche Eliminationsverfahren zur Lösung eines linearen Gleichungssystems kennen. In Teilkapitel 11.4 lernen wir erste allgemeine theoretische Aussagen über die Lösbarkeit eines linearen Gleichungssystems kennen.

In diesem und den nachfolgenden Kapitel arbeiten wir entweder reell oder komplex. Der Buchstabe \mathbb{K} bezeichnet im Folgenden also entweder \mathbb{R} oder \mathbb{C} .

11.1 Vektoren und Matrizen

Wir führen zunächst unseren „Vektorraum“ \mathbb{K}^n ein.

Definition 11.1. (Punkte und Zeilenvektoren in \mathbb{R}^n , \mathbb{C}^n und \mathbb{K}^n)

\mathbb{K} ist immer die Menge der reellen Zahlen \mathbb{R} oder die Menge der komplexen Zahlen \mathbb{C} . Wir definieren

$$\mathbb{K}^n := \{(x_1, x_2, \dots, x_n) : x_1, x_2, \dots, x_n \in \mathbb{K}\}$$

und erhalten als Sonderfälle

$$\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) : x_1, x_2, \dots, x_n \in \mathbb{R}\},$$

$$\mathbb{C}^n := \{(z_1, z_2, \dots, z_n) : z_1, z_2, \dots, z_n \in \mathbb{C}\}.$$

(x_1, x_2, \dots, x_n) heißt **Punkt** oder auch **Zeilenvektor** in \mathbb{K}^n .

Zwei Punkte (x_1, x_2, \dots, x_n) und (y_1, y_2, \dots, y_n) heißen **gleich**, wenn $x_1 = y_1$, $x_2 = y_2$, \dots , $x_n = y_n$ gilt.

Als Nächstes definieren wir Matrizen mit Einträgen in \mathbb{K} .

Definition 11.2. (Matrizen)

Seien $m, n \in \mathbb{N}$.

(1) Ein rechteckiges Schema der Form

$$\mathbf{A} := [a_{i,k}] := \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}$$

mit allen $a_{i,k} \in \mathbb{K}$ heißt eine $m \times n$ -**Matrix**. Sie hat m Zeilen und n Spalten. Die $a_{i,k}$ heißen die **Komponenten** oder **Einträge** der Matrix. In $a_{i,k}$ bezieht sich der erste Index (hier i) auf die Zeilennummer und der zweite Index (hier k) auf die Spaltennummer.

(2) Mit $\mathbb{K}^{m \times n}$ bezeichnen wir die Menge aller $m \times n$ -Matrizen mit Einträgen in \mathbb{K} .

(3) Die $m \times n$ -Matrix

$$\mathbf{O} = \mathbf{O}_{m \times n} := \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

heißt die **Nullmatrix** in $\mathbb{K}^{m \times n}$.

- (4) Matrizen in $\mathbb{K}^{m \times 1}$ (d.h. Matrizen mit nur einer Spalte) heißen auch **Spaltenvektoren**.
- (5) Matrizen in $\mathbb{K}^{1 \times n}$ (d.h. Matrizen mit nur einer Zeile) heißen auch **Zeilenvektoren**.

Wie wir es bereits in der HM A (siehe Kapitel 2 im Skript der HM A) kennengelernt haben, kann man **Spaltenvektoren mit Punkten in \mathbb{K}^m identifizieren**:

$$\begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \quad \text{wird mit dem Punkt} \quad (a_1, \dots, a_m) \quad \text{identifiziert.}$$

Deshalb schreiben wir statt $\mathbb{K}^{m \times 1}$ auch \mathbb{K}^m .

Definition 11.3. (gleiche Matrizen)

Zwei Matrizen $\mathbf{A} = [a_{i,k}] \in \mathbb{K}^{m \times n}$ und $\mathbf{B} = [b_{i,k}] \in \mathbb{K}^{p \times q}$ sind **gleich** (also $\mathbf{A} = \mathbf{B}$), wenn gilt:

- (i) $m = p$ und $n = q$ und
- (ii) $a_{i,k} = b_{i,k}$ für alle $i = 1, 2, \dots, m$ und alle $k = 1, 2, \dots, n$.

Betrachten wir einige Beispiele für Matrizen.

Beispiel 11.4. (Matrizen)

(a) Hier sind drei Matrizen

$$\begin{bmatrix} 1 & 2 \\ 3 & -4 \\ -5 & 6 \end{bmatrix} \in \mathbb{R}^{3 \times 2} \quad \text{und} \quad \in \mathbb{C}^{3 \times 2},$$

$$\begin{bmatrix} 1 & j & j-1 \\ 0 & 2 & -1 \end{bmatrix} \in \mathbb{C}^{2 \times 3}, \quad \begin{bmatrix} 1+j & 1-j \\ 2-j & 2+j \end{bmatrix} \in \mathbb{C}^{2 \times 2}$$

und ein Spaltenvektor und ein Zeilenvektor

$$\begin{bmatrix} 7j \\ 16 \\ -3 \\ 1 \end{bmatrix} \in \mathbb{C}^{4 \times 1}, \quad [1 \quad -1 \quad 1 \quad -1] \in \mathbb{R}^{1 \times 4} \quad \text{und} \quad \in \mathbb{C}^{1 \times 4}.$$

(b) Hier sind einige Nullmatrizen:

$$\mathbf{O}_{2 \times 2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{K}^{2 \times 2}, \quad \mathbf{O}_{3 \times 4} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \in \mathbb{K}^{3 \times 4}.$$

Es gilt

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \text{da die Spaltenanzahl verschieden ist.}$$

(c) Die Matrizen

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix} \quad \text{und} \quad \mathbf{B} = [b_{i,k}] \in \mathbb{K}^{3 \times 2} \quad \text{mit} \quad b_{i,k} = (-1)^{i+k}$$

sind gleich, denn

$$\mathbf{B} = \begin{bmatrix} (-1)^{1+1} & (-1)^{1+2} \\ (-1)^{2+1} & (-1)^{2+2} \\ (-1)^{3+1} & (-1)^{3+2} \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix} = \mathbf{A}.$$

Als Nächstes lernen wir die Addition von Matrizen und die Multiplikation einer Matrix mit einem Skalar (also einer Zahl in \mathbb{K}) kennen.

Definition 11.5. (Addition von Matrizen)

Zwei $m \times n$ -Matrizen $\mathbf{A} = [a_{i,k}]$, $\mathbf{B} = [b_{i,k}] \in \mathbb{K}^{m \times n}$ werden **addiert**, indem man die entsprechenden Einträge addiert:

$$\mathbf{A} + \mathbf{B} := [a_{i,k} + b_{i,k}] \in \mathbb{K}^{m \times n}, \quad \text{d.h.}$$

$$\begin{aligned} \mathbf{A} + \mathbf{B} &= \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} + \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,n} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,n} \\ \vdots & \vdots & & \vdots \\ b_{m,1} & b_{m,2} & \cdots & b_{m,n} \end{bmatrix} \\ &= \begin{bmatrix} a_{1,1} + b_{1,1} & a_{1,2} + b_{1,2} & \cdots & a_{1,n} + b_{1,n} \\ a_{2,1} + b_{2,1} & a_{2,2} + b_{2,2} & \cdots & a_{2,n} + b_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} + b_{m,1} & a_{m,2} + b_{m,2} & \cdots & a_{m,n} + b_{m,n} \end{bmatrix}. \end{aligned}$$

Betrachten wir einige Beispiele für die Addition von Matrizen.

Beispiel 11.6. (Addition von Matrizen)

(a) Die Matrizen

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 4 \\ 2 & -1 & 5 \end{bmatrix} \quad \text{und} \quad \mathbf{B} = \begin{bmatrix} 7 & 1 & -6 \\ -1 & 2 & -3 \end{bmatrix}$$

sind beide in $\mathbb{K}^{2 \times 3}$. Ihre Summe ist

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} 1+7 & 0+1 & 4-6 \\ 2-1 & -1+2 & 5-3 \end{bmatrix} = \begin{bmatrix} 8 & 1 & -2 \\ 1 & 1 & 2 \end{bmatrix}.$$

(b) Die Matrizen

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \text{und} \quad \mathbf{B} = \begin{bmatrix} 2 & -1 & 4 \\ 0 & 3 & 2 \end{bmatrix}$$

können nicht addiert werden, da die Matrizen nicht vom selben Typ sind. (\mathbf{A} ist eine 2×2 -Matrix, und \mathbf{B} ist eine 2×3 -Matrix.)

Im nächsten Satz sind die Rechenregeln für die Matrizenaddition festgehalten.

Satz 11.7. (Rechenregeln für die Matrizenaddition)

Seien $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{K}^{m \times n}$. Dann gelten:

- (1) **Kommutativgesetz:** $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
- (2) **Assoziativgesetz:** $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$
- (3) $\mathbf{A} + \mathbf{O}_{m \times n} = \mathbf{O}_{m \times n} + \mathbf{A} = \mathbf{A}$ für die Nullmatrix $\mathbf{O}_{m \times n}$ in $\mathbb{K}^{m \times n}$

Man kann die Rechenregeln für Matrizen durch direktes Nachrechnen mit Hilfe der Rechenregeln für die reellen bzw. komplexen Zahlen nachweisen.

Beweis von Satz 11.7: Es seien

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_{1,1} & \cdots & b_{1,n} \\ \vdots & & \vdots \\ b_{m,1} & \cdots & b_{m,n} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} c_{1,1} & \cdots & c_{1,n} \\ \vdots & & \vdots \\ c_{m,1} & \cdots & c_{m,n} \end{bmatrix}$$

in $\mathbb{K}^{m \times n}$ und die Nullmatrix

$$\mathbf{O}_{m \times n} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix}.$$

(1) Es gilt

$$\begin{aligned}
 \mathbf{A} + \mathbf{B} &= \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix} + \begin{bmatrix} b_{1,1} & \cdots & b_{1,n} \\ \vdots & & \vdots \\ b_{m,1} & \cdots & b_{m,n} \end{bmatrix} \\
 &= \begin{bmatrix} a_{1,1} + b_{1,1} & \cdots & a_{1,n} + b_{1,n} \\ \vdots & & \vdots \\ a_{m,1} + b_{m,1} & \cdots & a_{m,n} + b_{m,n} \end{bmatrix} \\
 &= \begin{bmatrix} b_{1,1} + a_{1,1} & \cdots & b_{1,n} + a_{1,n} \\ \vdots & & \vdots \\ b_{m,1} + a_{m,1} & \cdots & b_{m,n} + a_{m,n} \end{bmatrix} \\
 &= \begin{bmatrix} b_{1,1} & \cdots & b_{1,n} \\ \vdots & & \vdots \\ b_{m,1} & \cdots & b_{m,n} \end{bmatrix} + \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix} = \mathbf{B} + \mathbf{A},
 \end{aligned}$$

wobei wir in der dritten Zeile das Kommutativgesetz der reellen bzw. der komplexen Zahlen genutzt haben.

(2) Es gilt

$$\begin{aligned}
 \mathbf{A} + (\mathbf{B} + \mathbf{C}) &= \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix} + \left(\begin{bmatrix} b_{1,1} & \cdots & b_{1,n} \\ \vdots & & \vdots \\ b_{m,1} & \cdots & b_{m,n} \end{bmatrix} + \begin{bmatrix} c_{1,1} & \cdots & c_{1,n} \\ \vdots & & \vdots \\ c_{m,1} & \cdots & c_{m,n} \end{bmatrix} \right) \\
 &= \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix} + \begin{bmatrix} b_{1,1} + c_{1,1} & \cdots & b_{1,n} + c_{1,n} \\ \vdots & & \vdots \\ b_{m,1} + c_{m,1} & \cdots & b_{m,n} + c_{m,n} \end{bmatrix} \\
 &= \begin{bmatrix} a_{1,1} + b_{1,1} + c_{1,1} & \cdots & a_{1,n} + b_{1,n} + c_{1,n} \\ \vdots & & \vdots \\ a_{m,1} + b_{m,1} + c_{m,1} & \cdots & a_{m,n} + b_{m,n} + c_{m,n} \end{bmatrix} \\
 &= \begin{bmatrix} a_{1,1} + b_{1,1} & \cdots & a_{1,n} + b_{1,n} \\ \vdots & & \vdots \\ a_{m,1} + b_{m,1} & \cdots & a_{m,n} + b_{m,n} \end{bmatrix} + \begin{bmatrix} c_{1,1} & \cdots & c_{1,n} \\ \vdots & & \vdots \\ c_{m,1} & \cdots & c_{m,n} \end{bmatrix} \\
 &= \left(\begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix} + \begin{bmatrix} b_{1,1} & \cdots & b_{1,n} \\ \vdots & & \vdots \\ b_{m,1} & \cdots & b_{m,n} \end{bmatrix} \right) + \begin{bmatrix} c_{1,1} & \cdots & c_{1,n} \\ \vdots & & \vdots \\ c_{m,1} & \cdots & c_{m,n} \end{bmatrix}
 \end{aligned}$$

$$= (\mathbf{A} + \mathbf{B}) + \mathbf{C},$$

wobei wir das Assoziativgesetz für die reellen bzw. die komplexen Zahlen genutzt haben.

(3) Es gilt

$$\begin{aligned} \mathbf{A} + \mathbf{O}_{m \times n} &= \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix} + \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \\ &= \begin{bmatrix} a_{1,1} + 0 & \cdots & a_{1,n} + 0 \\ \vdots & & \vdots \\ a_{m,1} + 0 & \cdots & a_{m,n} + 0 \end{bmatrix} = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix} = \mathbf{A}. \end{aligned}$$

Wegen Satz 11.7 (1) gilt weiter $\mathbf{A} + \mathbf{O}_{m \times n} = \mathbf{O}_{m \times n} + \mathbf{A}$.

Damit haben wir alle drei Rechenregeln für die Matrizenaddition bewiesen. \square

Nun lernen wir die Multiplikation einer Matrix mit einem Skalar kennen.

Definition 11.8. (Multiplikation einer Matrix mit einem Skalar)

Eine Matrix $\mathbf{A} = [a_{i,k}] \in \mathbb{K}^{m \times n}$ wird **mit dem Skalar** $\lambda \in \mathbb{K}$ **multipliziert**, indem man jeden Eintrag von \mathbf{A} mit λ multipliziert:

$$\lambda \mathbf{A} := [\lambda a_{i,k}] \in \mathbb{K}^{m \times n} \quad d.h.$$

$$\lambda \mathbf{A} = \lambda \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} = \begin{bmatrix} \lambda a_{1,1} & \lambda a_{1,2} & \cdots & \lambda a_{1,n} \\ \lambda a_{2,1} & \lambda a_{2,2} & \cdots & \lambda a_{2,n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{m,1} & \lambda a_{m,2} & \cdots & \lambda a_{m,n} \end{bmatrix}.$$

Insbesondere gilt

$$-\mathbf{A} := (-1) \mathbf{A} = [-a_{i,k}] \in \mathbb{K}^{m \times n}.$$

Beispiel 11.9. (Multiplikation einer Matrix mit einem Skalar)

Seien

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} j & 3 & 1+j \\ 2 & -j & 4 \end{bmatrix} \quad \text{und} \quad \lambda = -2, \quad \mu = j.$$

Dann sind

$$\lambda \mathbf{A} = (-2) \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} -2 & -4 \\ -6 & -8 \\ -10 & -12 \\ -14 & -16 \end{bmatrix},$$

$$\mu \mathbf{B} = j \begin{bmatrix} j & 3 & 1+j \\ 2 & -j & 4 \end{bmatrix} = \begin{bmatrix} j^2 & 3j & (1+j)j \\ 2j & -j^2 & 4j \end{bmatrix} = \begin{bmatrix} -1 & 3j & -1+j \\ 2j & 1 & 4j \end{bmatrix}.$$

Satz 11.10. (Rechenregeln für die Multiplikation einer Matrix mit einem Skalar)

Seien $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{m \times n}$ und $\lambda, \mu \in \mathbb{K}$. Dann gelten

(1) **Assoziativgesetz:** $\lambda(\mu \mathbf{A}) = (\lambda \mu) \mathbf{A} = (\mu \lambda) \mathbf{A} = \mu(\lambda \mathbf{A})$

(2) **Distributivgesetze:**

$$(\lambda + \mu) \mathbf{A} = \lambda \mathbf{A} + \mu \mathbf{A} \quad \text{und} \quad \lambda(\mathbf{A} + \mathbf{B}) = \lambda \mathbf{A} + \lambda \mathbf{B}$$

Beweis von Satz 11.10: Es seien

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_{1,1} & \cdots & b_{1,n} \\ \vdots & & \vdots \\ b_{m,1} & \cdots & b_{m,n} \end{bmatrix} \in \mathbb{K}^{m \times n} \quad \text{und} \quad \lambda, \mu \in \mathbb{K}.$$

(1) Es gilt

$$\begin{aligned} \lambda(\mu \mathbf{A}) &= \lambda \begin{bmatrix} \mu a_{1,1} & \cdots & \mu a_{1,n} \\ \vdots & & \vdots \\ \mu a_{m,1} & \cdots & \mu a_{m,n} \end{bmatrix} = \begin{bmatrix} \lambda \mu a_{1,1} & \cdots & \lambda \mu a_{1,n} \\ \vdots & & \vdots \\ \lambda \mu a_{m,1} & \cdots & \lambda \mu a_{m,n} \end{bmatrix} \\ &= (\lambda \mu) \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix} = (\lambda \mu) \mathbf{A} = (\mu \lambda) \mathbf{A}, \end{aligned}$$

wobei wir das Assoziativgesetz und das Kommutativgesetz der Multiplikation der reellen bzw. der komplexen Zahlen ausgenutzt haben. Analog zeigt man $\mu(\lambda \mathbf{A}) = (\mu \lambda) \mathbf{A}$.

(2) Es gilt

$$\begin{aligned}
 (\lambda + \mu) \mathbf{A} &= \begin{bmatrix} (\lambda + \mu) a_{1,1} & \cdots & (\lambda + \mu) a_{1,n} \\ \vdots & & \vdots \\ (\lambda + \mu) a_{m,1} & \cdots & (\lambda + \mu) a_{m,n} \end{bmatrix} \\
 &= \begin{bmatrix} \lambda a_{1,1} + \mu a_{1,1} & \cdots & \lambda a_{1,n} + \mu a_{1,n} \\ \vdots & & \vdots \\ \lambda a_{m,1} + \mu a_{m,1} & \cdots & \lambda a_{m,n} + \mu a_{m,n} \end{bmatrix} \\
 &= \begin{bmatrix} \lambda a_{1,1} & \cdots & \lambda a_{1,n} \\ \vdots & & \vdots \\ \lambda a_{m,1} & \cdots & \lambda a_{m,n} \end{bmatrix} + \begin{bmatrix} \mu a_{1,1} & \cdots & \mu a_{1,n} \\ \vdots & & \vdots \\ \mu a_{m,1} & \cdots & \mu a_{m,n} \end{bmatrix} = \lambda \mathbf{A} + \mu \mathbf{A},
 \end{aligned}$$

wobei wir das Distributivgesetz der reellen bzw. der komplexen Zahlen genutzt haben. Weiter gilt

$$\begin{aligned}
 \lambda (\mathbf{A} + \mathbf{B}) &= \lambda \begin{bmatrix} a_{1,1} + b_{1,1} & \cdots & a_{1,n} + b_{1,n} \\ \vdots & & \vdots \\ a_{m,1} + b_{m,1} & \cdots & a_{m,n} + b_{m,n} \end{bmatrix} \\
 &= \begin{bmatrix} \lambda (a_{1,1} + b_{1,1}) & \cdots & \lambda (a_{1,n} + b_{1,n}) \\ \vdots & & \vdots \\ \lambda (a_{m,1} + b_{m,1}) & \cdots & \lambda (a_{m,n} + b_{m,n}) \end{bmatrix} \\
 &= \begin{bmatrix} \lambda a_{1,1} + \lambda b_{1,1} & \cdots & \lambda a_{1,n} + \lambda b_{1,n} \\ \vdots & & \vdots \\ \lambda a_{m,1} + \lambda b_{m,1} & \cdots & \lambda a_{m,n} + \lambda b_{m,n} \end{bmatrix} \\
 &= \begin{bmatrix} \lambda a_{1,1} & \cdots & \lambda a_{1,n} \\ \vdots & & \vdots \\ \lambda a_{m,1} & \cdots & \lambda a_{m,n} \end{bmatrix} + \begin{bmatrix} \lambda b_{1,1} & \cdots & \lambda b_{1,n} \\ \vdots & & \vdots \\ \lambda b_{m,1} & \cdots & \lambda b_{m,n} \end{bmatrix} = \lambda \mathbf{A} + \lambda \mathbf{B},
 \end{aligned}$$

wobei wir das Distributivgesetz der reellen bzw. der komplexen Zahlen genutzt haben.

Damit haben wir alle Rechenregeln bewiesen. □

In Kapitel 3 der HM A haben wir bereits die **Matrix-Vektor-Multiplikation** kennengelernt, an die wir hier für den Fall einer 3×3 -Matrix noch einmal erinnern: Für

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} \quad \text{und} \quad \vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

ergibt die Matrix-Vektor-Multiplikation („Zeile mal Spalte“)

$$\mathbf{A} \vec{x} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} a_{1,1} x_1 + a_{1,2} x_2 + a_{1,3} x_3 \\ a_{2,1} x_1 + a_{2,2} x_2 + a_{2,3} x_3 \\ a_{3,1} x_1 + a_{3,2} x_2 + a_{3,3} x_3 \end{bmatrix}.$$

Die Matrix-Vektor-Multiplikation ist ein Sonderfall der Matrizenmultiplikation, welche wir nun einführen.

Definition 11.11. (Multiplikation von Matrizen)

Seien $\mathbf{A} = [a_{i,k}] \in \mathbb{K}^{m \times n}$ und $\mathbf{B} = [b_{k,\ell}] \in \mathbb{K}^{n \times p}$. Dann definieren wir

$$\mathbf{C} := \mathbf{A} \cdot \mathbf{B} = [c_{i,\ell}] \in \mathbb{K}^{m \times p}$$

durch

$$c_{i,\ell} := \sum_{k=1}^n a_{i,k} b_{k,\ell} = a_{i,1} b_{1,\ell} + a_{i,2} b_{2,\ell} + \dots + a_{i,n} b_{n,\ell}.$$

Der Eintrag $c_{i,\ell}$ in der i -ten Zeile und der ℓ -ten Spalte von $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$ ist also das **Skalarprodukt** der i -ten Zeilenvektors von A ,

$$[a_{i,1} \quad a_{i,2} \quad \dots \quad a_{i,n}],$$

mit dem ℓ -ten Spaltenvektor von B ,

$$\begin{bmatrix} b_{1,\ell} \\ b_{2,\ell} \\ \vdots \\ b_{n,\ell} \end{bmatrix}.$$

Ausgeschrieben finden wir:

$$\mathbf{C} = \mathbf{A} \cdot \mathbf{B} = \begin{bmatrix} \sum_{k=1}^n a_{1,k} b_{k,1} & \dots & \sum_{k=1}^n a_{1,k} b_{k,p} \\ \vdots & & \vdots \\ \sum_{k=1}^n a_{m,k} b_{k,1} & \dots & \sum_{k=1}^n a_{m,k} b_{k,p} \end{bmatrix} = \begin{bmatrix} c_{1,1} & \dots & c_{1,p} \\ \vdots & & \vdots \\ c_{m,1} & \dots & c_{m,p} \end{bmatrix}.$$

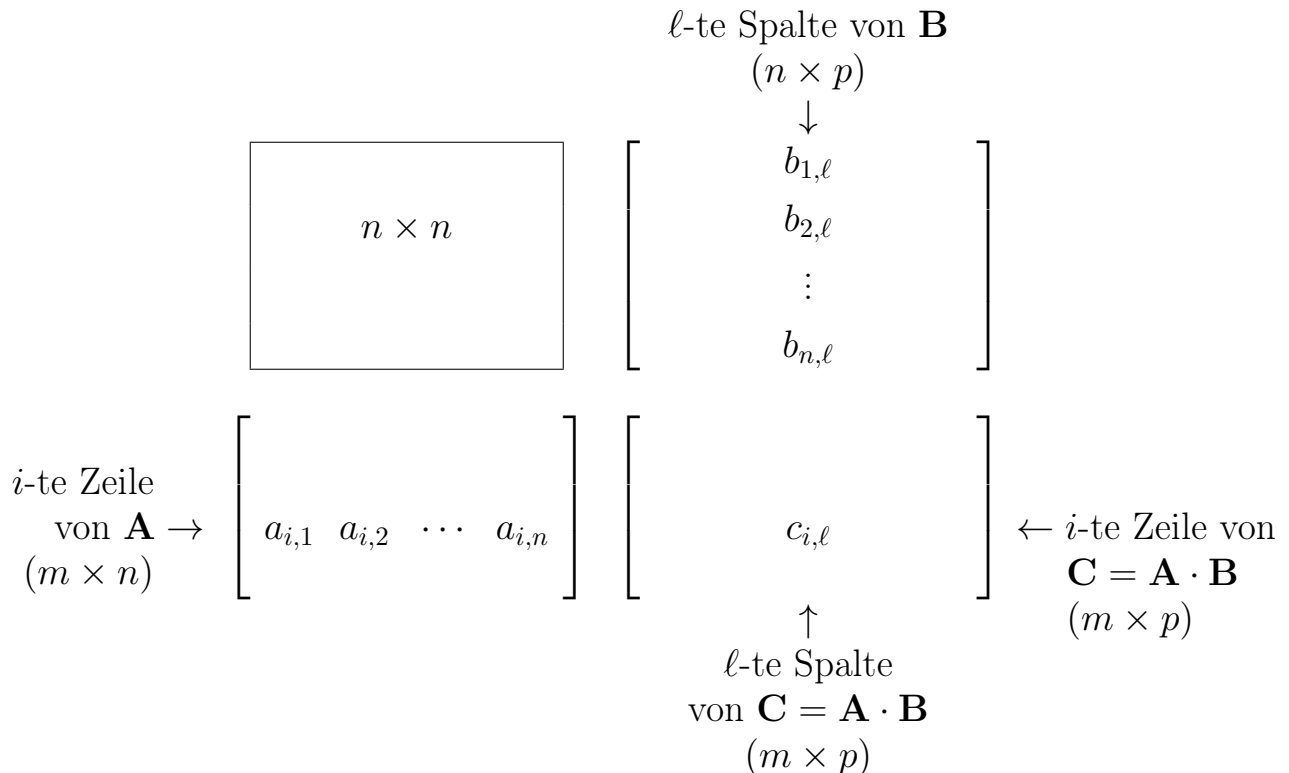
Achtung: Nur wenn gilt

Anzahl der Spalten von \mathbf{A} = Anzahl der Zeilen von \mathbf{B}

kann das Matrizenprodukt $\mathbf{A} \cdot \mathbf{B}$ gebildet werden!

Bemerkung 11.12. (Berechnung der Matrizenprodukts)

Die Berechnung von $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$ merkt man sich als Schema „**Skalarprodukte von Zeilenvektoren von \mathbf{A} und Spaltenvektoren von \mathbf{B}** “, wie in dem folgenden Diagramm illustriert:



So findet man die Formel

$$c_{i,\ell} = a_{i,1} b_{1,\ell} + a_{i,2} b_{2,\ell} + \dots + a_{i,n} b_{n,\ell},$$

die wir aus der Definition 11.11 des Matrizenprodukts kennen.

Beispiel 11.13. (Multiplikation von Matrizen)

(a) Seien

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \text{und} \quad \mathbf{B} = \begin{bmatrix} -1 & 0 & 3 \\ 2 & 1 & -4 \end{bmatrix}.$$

Dann sind $\mathbf{A} \in \mathbb{R}^{2 \times 2}$ und $\mathbf{B} \in \mathbb{R}^{2 \times 3}$. Wir können das Matrizenprodukt $\mathbf{A} \cdot \mathbf{B}$ bilden und $\mathbf{A} \cdot \mathbf{B} \in \mathbb{R}^{2 \times 3}$. $\mathbf{B} \cdot \mathbf{A}$ ist dagegen nicht definiert. Wir finden

$$\mathbf{A} \cdot \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} -1 & 0 & 3 \\ 2 & 1 & -4 \end{bmatrix}$$

$$\begin{aligned}
&= \begin{bmatrix} 1 \cdot (-1) + 2 \cdot 2 & 1 \cdot 0 + 2 \cdot 1 & 1 \cdot 3 + 2 \cdot (-4) \\ 3 \cdot (-1) + 4 \cdot 2 & 3 \cdot 0 + 4 \cdot 1 & 3 \cdot 3 + 4 \cdot (-4) \end{bmatrix} \\
&= \begin{bmatrix} 3 & 2 & -5 \\ 5 & 4 & -7 \end{bmatrix}.
\end{aligned}$$

(b) Betrachten wir die zwei Matrizen

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ -2 & 0 & 4 \end{bmatrix} \quad \text{und} \quad \mathbf{B} = \begin{bmatrix} 0 & -3 \\ 2 & 1 \\ -1 & 4 \end{bmatrix}.$$

Dann ist $\mathbf{A} \in \mathbb{R}^{2 \times 3}$ und $\mathbf{B} \in \mathbb{R}^{3 \times 2}$. Also können wir sowohl $\mathbf{A} \cdot \mathbf{B}$ als auch $\mathbf{B} \cdot \mathbf{A}$ berechnen. Die Matrix $\mathbf{A} \cdot \mathbf{B}$ ist in $\mathbb{R}^{2 \times 2}$, und wir erhalten

$$\begin{aligned}
\mathbf{A} \cdot \mathbf{B} &= \begin{bmatrix} 1 & 2 & 3 \\ -2 & 0 & 4 \end{bmatrix} \begin{bmatrix} 0 & -3 \\ 2 & 1 \\ -1 & 4 \end{bmatrix} \\
&= \begin{bmatrix} 1 \cdot 0 + 2 \cdot 2 + 3 \cdot (-1) & 1 \cdot (-3) + 2 \cdot 1 + 3 \cdot 4 \\ (-2) \cdot 0 + 0 \cdot 2 + 4 \cdot (-1) & (-2) \cdot (-3) + 0 \cdot 1 + 4 \cdot 4 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 11 \\ -4 & 22 \end{bmatrix}.
\end{aligned}$$

Die Matrix $\mathbf{B} \cdot \mathbf{A}$ ist in $\mathbb{R}^{3 \times 3}$, und wir erhalten

$$\begin{aligned}
\mathbf{B} \cdot \mathbf{A} &= \begin{bmatrix} 0 & -3 \\ 2 & 1 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ -2 & 0 & 4 \end{bmatrix} \\
&= \begin{bmatrix} 0 \cdot 1 + (-3) \cdot (-2) & 0 \cdot 2 + (-3) \cdot 0 & 0 \cdot 3 + (-3) \cdot 4 \\ 2 \cdot 1 + 1 \cdot (-2) & 2 \cdot 2 + 1 \cdot 0 & 2 \cdot 3 + 1 \cdot 4 \\ (-1) \cdot 1 + 4 \cdot (-2) & (-1) \cdot 2 + 4 \cdot 0 & (-1) \cdot 3 + 4 \cdot 4 \end{bmatrix} \\
&= \begin{bmatrix} 6 & 0 & -12 \\ 0 & 4 & 10 \\ -9 & -2 & 13 \end{bmatrix}.
\end{aligned}$$

(c) Für die quadratischen 2×2 -Matrizen

$$\mathbf{C} = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \quad \text{und} \quad \mathbf{D} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

können wir sowohl $\mathbf{C} \cdot \mathbf{D}$ als auch $\mathbf{D} \cdot \mathbf{C}$ berechnen. Wir finden

$$\mathbf{C} \cdot \mathbf{D} = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 \cdot 1 + 1 \cdot 0 & 2 \cdot (-1) + 1 \cdot 1 \\ 0 \cdot 1 + 1 \cdot 0 & 0 \cdot (-1) + 1 \cdot 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix},$$

$$\mathbf{D} \cdot \mathbf{C} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 \cdot 2 + (-1) \cdot 0 & 1 \cdot 1 + (-1) \cdot 1 \\ 0 \cdot 2 + 1 \cdot 0 & 0 \cdot 1 + 1 \cdot 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}.$$

Wir sehen also, dass $\mathbf{C} \cdot \mathbf{D} \neq \mathbf{D} \cdot \mathbf{C}$ gilt.

Bemerkung 11.14. (zur Multiplikation von Matrizen)

- (1) $\mathbf{A} \cdot \mathbf{B}$ ist nur dann definiert, wenn die Spaltenanzahl von \mathbf{A} mit der Zeilenanzahl von \mathbf{B} übereinstimmt.
- (2) Selbst wenn $\mathbf{A} \cdot \mathbf{B}$ und $\mathbf{B} \cdot \mathbf{A}$ beide definiert sind, gilt im Allgemeinen $\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A}$, d.h. die Multiplikation von Matrizen ist **nicht kommutativ!**

Im nächsten Satz halten wir die Rechenregeln für die Multiplikation von Matrizen fest.

Satz 11.15. (Rechenregeln für die Multiplikation von Matrizen)

(1) *Assoziativgesetz:*

$$\mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{C}) = (\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{C} \quad \text{für alle } \mathbf{A} \in \mathbb{K}^{m \times n}, \mathbf{B} \in \mathbb{K}^{n \times p}, \mathbf{C} \in \mathbb{K}^{p \times \ell}.$$

$$\lambda(\mathbf{A} \cdot \mathbf{B}) = (\lambda \mathbf{A}) \cdot \mathbf{B} = \mathbf{A} \cdot (\lambda \mathbf{B}) \quad \text{für alle } \lambda \in \mathbb{K} \text{ und} \\ \text{alle } \mathbf{A} \in \mathbb{K}^{m \times n}, \mathbf{B} \in \mathbb{K}^{n \times p}.$$

(2) *Distributivgesetz:*

$$\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C} \quad \text{für alle } \mathbf{A} \in \mathbb{K}^{m \times n}, \mathbf{B}, \mathbf{C} \in \mathbb{K}^{n \times p}.$$

$$(\mathbf{A} + \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot \mathbf{C} + \mathbf{B} \cdot \mathbf{C} \quad \text{für alle } \mathbf{A}, \mathbf{B} \in \mathbb{K}^{m \times n}, \mathbf{C} \in \mathbb{K}^{n \times p}.$$

Wir beweisen Satz 11.15 (1). Den Beweis von Satz 11.15 (2) werden wir in einer Übungsaufgabe durchführen.

Beweis von Satz 11.15 (1): Wir berechnen zunächst die Einträge von $\mathbf{A}\mathbf{B}$ und $\mathbf{B}\mathbf{C}$:

$$\mathbf{A}\cdot\mathbf{B} = [(\mathbf{A}\cdot\mathbf{B})_{i,s}] = \left[\sum_{r=1}^n a_{i,r} b_{r,s} \right] \quad \text{und} \quad \mathbf{B}\cdot\mathbf{C} = [(\mathbf{B}\cdot\mathbf{C})_{r,k}] = \left[\sum_{s=1}^p b_{r,s} c_{s,k} \right].$$

Damit finden wir

$$\begin{aligned} \mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C}) &= [(\mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C}))_{i,k}] = \left[\sum_{r=1}^n a_{i,r} (\mathbf{B}\cdot\mathbf{C})_{r,k} \right] \\ &= \left[\sum_{r=1}^n a_{i,r} \sum_{s=1}^p b_{r,s} c_{s,k} \right] = \left[\sum_{s=1}^p \left(\sum_{r=1}^n a_{i,r} b_{r,s} \right) c_{s,k} \right] \\ &= \left[\sum_{s=1}^p (\mathbf{A}\cdot\mathbf{B})_{i,s} c_{s,k} \right] = [((\mathbf{A}\cdot\mathbf{B})\cdot\mathbf{C})_{i,k}] = (\mathbf{A}\cdot\mathbf{B})\cdot\mathbf{C}, \end{aligned}$$

$$\begin{aligned} \lambda(\mathbf{A}\cdot\mathbf{B}) &= [\lambda(\mathbf{A}\cdot\mathbf{B})_{i,s}] = \left[\lambda \sum_{r=1}^n a_{i,r} b_{r,s} \right] \\ &= \underbrace{\left[\sum_{r=1}^n (\lambda a_{i,r}) b_{r,s} \right]}_{=(\lambda\mathbf{A})\cdot\mathbf{B}} = \underbrace{\left[\sum_{r=1}^n a_{i,r} (\lambda b_{r,s}) \right]}_{=\mathbf{A}\cdot(\lambda\mathbf{B})}. \end{aligned}$$

Damit haben wir Satz 11.15 (1) bewiesen. □

Als letzte neue Begriffe in diesem Teilkapitel lernen wir die transponierte und die adjungierte Matrix kennen.

Definition 11.16. (transponierte Matrix)

Schreibt man die Einträge der Spalten der Matrix $\mathbf{A} = [a_{i,k}] \in \mathbb{K}^{m \times n}$ in die Zeilen einer neuen Matrix \mathbf{B} , so hat diese n Zeilen und m Spalten. Diese Matrix heißt **Transponierte** (oder **transponierte Matrix**) von \mathbf{A} und wird mit \mathbf{A}^T bezeichnet:

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \quad \Longrightarrow \quad \mathbf{A}^T := \begin{bmatrix} a_{1,1} & \cdots & a_{m,1} \\ a_{1,2} & \cdots & a_{m,2} \\ \vdots & & \vdots \\ a_{1,n} & \cdots & a_{m,n} \end{bmatrix}$$

Als Sonderfall erhalten wir aus einem **Spaltenvektor** \vec{x} (also einer Matrix in $\mathbb{K}^{m \times 1}$) einen **Zeilenvektor** \vec{x}^T (also eine Matrix in $\mathbb{K}^{1 \times m}$):

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \implies \vec{x}^T = [x_1 \ x_2 \ \cdots \ x_m]$$

Notation: Wir bezeichnen mit \vec{x} immer nur Spaltenvektoren; wollen wir einen Zeilenvektor darstellen, so schreiben wir \vec{x}^T .

Beispiel 11.17. (transponierte Matrix)

$$(a) \quad \mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \in \mathbb{R}^{2 \times 3} \implies \mathbf{A}^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \in \mathbb{R}^{3 \times 2}$$

$$(b) \quad \mathbf{B} = \begin{bmatrix} j & 0 \\ 1 & -j \end{bmatrix} \in \mathbb{C}^{2 \times 2} \implies \mathbf{B}^T = \begin{bmatrix} j & 1 \\ 0 & -j \end{bmatrix} \in \mathbb{C}^{2 \times 2}$$

$$(c) \quad \vec{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \implies \vec{x}^T = [1 \ 2 \ 3],$$

$$\vec{y}^T = [1 \ -1 \ 0] \implies \vec{y} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

Als Nächstes halten wir die Rechenregeln für die transponierte Matrix fest.

Hilfssatz 11.18. (Rechenregeln für die transponierte Matrix)

Seien $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{m \times n}$, $\mathbf{C} \in \mathbb{K}^{n \times p}$ und $\lambda \in \mathbb{K}$. Dann gelten:

$$(1) \quad (\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$$

$$(2) \quad (\lambda \mathbf{A})^T = \lambda \mathbf{A}^T$$

$$(3) \quad (\mathbf{A} \cdot \mathbf{C})^T = \mathbf{C}^T \cdot \mathbf{A}^T$$

Beweis: Hilfssatz 11.18 wird in einer Übungsaufgabe bewiesen. \square

Definition 11.19. (adjungierte einer komplexen Matrix)

Die zu $\mathbf{A} \in \mathbb{C}^{m \times n}$ *adjungierte Matrix* $\mathbf{A}^* \in \mathbb{C}^{n \times m}$ entsteht, indem man die Transponierte von \mathbf{A} bildet und von jedem Eintrag das konjugiert Komplexe nimmt.

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \in \mathbb{C}^{m \times n} \quad \Longrightarrow \quad \mathbf{A}^* := \begin{bmatrix} \overline{a_{1,1}} & \cdots & \overline{a_{m,1}} \\ \overline{a_{1,2}} & \cdots & \overline{a_{m,2}} \\ \vdots & & \vdots \\ \overline{a_{1,n}} & \cdots & \overline{a_{m,n}} \end{bmatrix}$$

Beispiel 11.20. (adjungierte Matrix)

$$\begin{aligned} \mathbf{A} = \begin{bmatrix} j & 1+j \\ 2-j & 3 \end{bmatrix} & \Longrightarrow \mathbf{A}^T = \begin{bmatrix} j & 2-j \\ 1+j & 3 \end{bmatrix} \\ & \Longrightarrow \mathbf{A}^* = \begin{bmatrix} -j & 2+j \\ 1-j & 3 \end{bmatrix} \end{aligned}$$

Als Letztes halten wir die Rechenregeln für die adjungierte Matrix fest.

Hilfssatz 11.21. (Rechenregeln für die adjungierte Matrix)

Seien $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m \times n}$, $\mathbf{C} \in \mathbb{C}^{n \times p}$ und $\lambda \in \mathbb{C}$. Dann gelten:

- (1) $(\mathbf{A} + \mathbf{B})^* = \mathbf{A}^* + \mathbf{B}^*$
- (2) $(\lambda \mathbf{A})^* = \bar{\lambda} \mathbf{A}^*$
- (3) $(\mathbf{A} \cdot \mathbf{C})^* = \mathbf{C}^* \cdot \mathbf{A}^*$

Beweis: Hilfssatz 11.21 wird in einer Übungsaufgabe bewiesen. \square

11.2 Lineare Gleichungssysteme: Notation

Lineare Gleichungssysteme mit genauso vielen Unbekannten wie Gleichungen haben wir bereits in Kapitel 3 der HM A kennengelernt. Auch dort haben wir diese bereits mit Matrizen und Vektoren mit Hilfe der Matrix-Vektor-Multiplikation geschrieben. Wir wollen nun analog für beliebige lineare Gleichungssysteme vorgehen.

Wir betrachten das **lineare Gleichungssystem (LGS)**

$$\begin{aligned}
 a_{1,1} x_1 + a_{1,2} x_2 + \dots + a_{1,n} x_n &= b_1 \\
 a_{2,1} x_1 + a_{2,2} x_2 + \dots + a_{2,n} x_n &= b_2 \\
 &\vdots \\
 a_{m,1} x_1 + a_{m,2} x_2 + \dots + a_{m,n} x_n &= b_m
 \end{aligned} \tag{11.1}$$

mit m **Gleichungen** und n **Unbekannten** x_1, x_2, \dots, x_n . Vorgegeben sind dabei die Koeffizienten $a_{i,k}$, $i = 1, 2, \dots, m$; $k = 1, 2, \dots, n$, und b_i , $i = 1, 2, \dots, m$, in \mathbb{K} . Gesucht werden $x_1, x_2, \dots, x_n \in \mathbb{K}$, welche die m Gleichungen erfüllen.

In Summenschreibweise lautet das lineare Gleichungssystem (11.1)

$$\begin{aligned}
 \sum_{k=1}^n a_{1,k} x_k &= b_1 \\
 \sum_{k=1}^n a_{2,k} x_k &= b_2 \\
 &\vdots \\
 \sum_{k=1}^n a_{m,k} x_k &= b_m
 \end{aligned}$$

oder kürzer

$$\sum_{k=1}^n a_{i,k} x_k = b_i, \quad i = 1, 2, \dots, m.$$

Wir wollen nun unser lineares Gleichungssystem in **Matrixschreibweise** notieren: Dazu schreiben wir für (11.1) zunächst in Vektorform:

$$\begin{bmatrix} a_{1,1} x_1 + a_{1,2} x_2 + \dots + a_{1,n} x_n \\ a_{2,1} x_1 + a_{2,2} x_2 + \dots + a_{2,n} x_n \\ \vdots \\ a_{m,1} x_1 + a_{m,2} x_2 + \dots + a_{m,n} x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Fassen wir die Koeffizienten $a_{i,k}$ zu der Matrix

$$\mathbf{A} := \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \in \mathbb{K}^{m \times n}$$

und die x_k bzw. b_i jeweils zu einem Spaltenvektor (also einer Matrix mit nur einer Spalte)

$$\vec{\mathbf{x}} := \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{K}^n \quad \text{bzw.} \quad \vec{\mathbf{b}} := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \in \mathbb{K}^m$$

zusammen, so lässt sich das lineare Gleichungssystem schreiben als

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

oder kurz

$$\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}.$$

Die Matrix $\mathbf{A} \in \mathbb{K}^{m \times n}$ heißt die **Koeffizientenmatrix** und $\vec{\mathbf{b}} \in \mathbb{K}^m$ die **rechte Seite** des LGS. Die Matrix

$$[\mathbf{A} | \vec{\mathbf{b}}] := \left[\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} & b_m \end{array} \right] \in \mathbb{K}^{m \times (n+1)}$$

heißt die **erweiterte Koeffizientenmatrix** des LGS $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}$.

Betrachten wir hierzu ein Zahlenbeispiel.

Beispiel 11.22. (LGS in Matrizen Schreibweise)

Wir schreiben das lineare Gleichungssystem zunächst in Matrizen Schreibweise

$$\left[\begin{array}{cccc|c} x_1 + & x_2 - & 3x_3 + & x_4 = & 1 \\ 2x_1 + & x_2 + & x_3 - & x_4 = & 0 \\ & 2x_2 - & 13x_3 + & x_4 = & -1 \end{array} \right]$$

$$\iff \begin{bmatrix} 1 & 1 & -3 & 1 \\ 2 & 1 & 1 & -1 \\ 0 & 2 & -13 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

und dann mit der erweiterten Koeffizientenmatrix

$$\left[\begin{array}{cccc|c} 1 & 1 & -3 & 1 & 1 \\ 2 & 1 & 1 & -1 & 0 \\ 0 & 2 & -13 & 1 & -1 \end{array} \right].$$

11.3 Das Gaußsche Eliminationsverfahren

Seien $\mathbf{A} \in \mathbb{K}^{m \times n}$ und $\mathbf{b} \in \mathbb{K}^m$. Wir bezeichnen die **Lösungsmenge** des linearen Gleichungssystems $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}$ (mit der erweiterten Koeffizientenmatrix $[\mathbf{A} | \vec{\mathbf{b}}]$) mit $\mathbb{L}_{[\mathbf{A} | \vec{\mathbf{b}}]}$, also

$$\mathbb{L}_{[\mathbf{A} | \vec{\mathbf{b}}]} := \{ \vec{\mathbf{x}} \in \mathbb{K}^n : \mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}} \}.$$

Wir werden in den späteren Kapiteln der Vorlesung sehen, dass es für $\mathbb{L}_{[\mathbf{A} | \vec{\mathbf{b}}]}$ **nur die folgenden drei Möglichkeiten** gibt:

- (1) $\mathbb{L}_{[\mathbf{A} | \vec{\mathbf{b}}]} = \emptyset$.
- (2) $\mathbb{L}_{[\mathbf{A} | \vec{\mathbf{b}}]}$ enthält genau einen Vektor; das LGS ist also eindeutig lösbar.
- (3) $\mathbb{L}_{[\mathbf{A} | \vec{\mathbf{b}}]}$ enthält unendlich viele Lösungen.

In diesem Abschnitt geht es darum, wie man $\mathbb{L}_{[\mathbf{A} | \vec{\mathbf{b}}]}$ konkret berechnen kann.

Hilfssatz 11.23. (elementare Zeilenoperationen)

Die Lösungsmenge $\mathbb{L}_{[\mathbf{A} | \vec{\mathbf{b}}]}$ des linearen Gleichungssystems $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}$ ändert sich nicht unter den folgenden **elementaren Zeilenoperationen**:

- (E1) Multiplikation einer Zeile mit $\lambda \in \mathbb{K} \setminus \{0\}$.
- (E2) Ersetzen einer Zeile durch die Summe aus dieser Zeile und dem μ -fachen einer anderen Zeile ($\mu \in \mathbb{K}$).
- (E3) Vertauschen zweier Zeilen.

Das **Gaußsche Eliminationsverfahren** besteht aus der systematischen Anwendung dieser elementaren Zeilenoperationen. Ziel ist es dabei, das lineare Gleichungssystem

chungssystem so zu vereinfachen, dass man die Lösungsmenge leicht ablesen kann. Dies soll jetzt zunächst an zwei Beispielen demonstriert werden.

Notation: Dabei bezeichnen wir die Zeilen der erweiterten Koeffizientenmatrix als Z_1, Z_2, \dots, Z_m (also Z_i ist die i -te Zeile) und **notieren die durchgeführten elementaren Zeilenoperationen** wie folgt:

$Z_i \rightarrow \lambda Z_i$ (wobei $\lambda \neq 0$) bedeutet, dass die i -te Zeile mit λ multipliziert wird.

$Z_i \rightarrow Z_i + \mu Z_j$ bedeutet, dass zu der i -ten Zeile das μ -fache der j -ten Zeile addiert wird.

$Z_i \leftrightarrow Z_j$ (wobei $i \neq j$) bedeutet, dass die i -te und j -te Zeile getauscht werden.

Beispiel 11.24. (Gaußsches Eliminationsverfahren für LGS)

Für welches $\alpha \in \mathbb{R}$ ist das reelle lineare Gleichungssystem

$$\begin{aligned} x_1 + x_2 - 3x_3 + x_4 &= 1 \\ 2x_1 + x_2 + x_3 - x_4 &= 0 \\ 2x_2 - 13x_3 + x_4 &= -1 \\ 2x_1 - x_2 + 14x_3 - 2x_4 &= \alpha \end{aligned}$$

lösbar? Bestimmen Sie gegebenenfalls die Lösungsmenge.

Wir schreiben das lineare Gleichungssystem mit der erweiterten Koeffizientenmatrix

$$[\mathbf{A} | \vec{\mathbf{b}}] = \left[\begin{array}{cccc|c} 1 & 1 & -3 & 1 & 1 \\ 2 & 1 & 1 & -1 & 0 \\ 0 & 2 & -13 & 1 & -1 \\ 2 & -1 & 14 & -2 & \alpha \end{array} \right]$$

und nutzen dann die elementaren Zeilenumformungen (E1), (E2) und (E3) systematisch, um das lineare Gleichungssystem zu lösen.

$$\begin{array}{ccc} \left[\begin{array}{cccc|c} 1 & 1 & -3 & 1 & 1 \\ 2 & 1 & 1 & -1 & 0 \\ 0 & 2 & -13 & 1 & -1 \\ 2 & -1 & 14 & -2 & \alpha \end{array} \right] & \begin{array}{c} Z_2 \rightarrow Z_2 - 2Z_1 \\ Z_4 \rightarrow Z_4 - 2Z_1 \\ \downarrow \\ \Leftrightarrow \end{array} & \left[\begin{array}{cccc|c} 1 & 1 & -3 & 1 & 1 \\ 0 & -1 & 7 & -3 & -2 \\ 0 & 2 & -13 & 1 & -1 \\ 0 & -3 & 20 & -4 & \alpha - 2 \end{array} \right] \\ & & \\ & & \begin{array}{c} Z_2 \rightarrow (-1) \cdot Z_2 \\ \downarrow \\ \Leftrightarrow \end{array} & \left[\begin{array}{cccc|c} 1 & 1 & -3 & 1 & 1 \\ 0 & 1 & -7 & 3 & 2 \\ 0 & 2 & -13 & 1 & -1 \\ 0 & -3 & 20 & -4 & \alpha - 2 \end{array} \right] \end{array}$$

$$\begin{array}{c}
 \begin{array}{l}
 Z_3 \rightarrow Z_3 - 2 Z_2 \\
 Z_4 \rightarrow Z_4 + 3 Z_2 \\
 \downarrow \\
 \Leftrightarrow
 \end{array} \\
 \\
 \begin{array}{c}
 \begin{array}{l}
 Z_4 \rightarrow Z_4 + Z_3 \\
 \downarrow \\
 \Leftrightarrow
 \end{array}
 \end{array}
 \end{array}
 \left[\begin{array}{cccc|c}
 1 & 1 & -3 & 1 & 1 \\
 0 & 1 & -7 & 3 & 2 \\
 0 & 0 & 1 & -5 & -5 \\
 0 & 0 & -1 & 5 & \alpha + 4
 \end{array} \right]$$

$$\left[\begin{array}{cccc|c}
 1 & 1 & -3 & 1 & 1 \\
 0 & 1 & -7 & 3 & 2 \\
 0 & 0 & 1 & -5 & -5 \\
 0 & 0 & 0 & 0 & \alpha - 1
 \end{array} \right]$$

Die erweiterte Koeffizientenmatrix ist nun in **Stufenform**. Damit sieht das lineare Gleichungssystem so aus:

$$\begin{array}{rcl}
 x_1 + x_2 - 3x_3 + x_4 & = & 1 \\
 x_2 - 7x_3 + 3x_4 & = & 2 \\
 x_3 - 5x_4 & = & -5 \\
 0 & = & \alpha - 1
 \end{array}$$

Falls $\alpha \neq 1$, ist $\mathbb{L}_{[\mathbf{A}|\vec{\mathbf{b}}]} = \emptyset$, denn die letzte Gleichung ist dann nicht erfüllbar.

Falls $\alpha = 1$ ist, so ist die letzte Gleichung immer wahr und das LGS reduziert sich auf drei Gleichungen:

$$\begin{array}{rcl}
 x_1 + x_2 - 3x_3 + x_4 & = & 1 \quad \text{(I)} \\
 x_2 - 7x_3 + 3x_4 & = & 2 \quad \text{(II)} \\
 x_3 - 5x_4 & = & -5 \quad \text{(III)}
 \end{array}$$

Wir setzen $x_4 := \lambda \in \mathbb{R}$ und bestimmen die Lösungsmenge $\mathbb{L}_{[\mathbf{A}|\vec{\mathbf{b}}]}$ durch „Rückwärtsrechnen“:

$$\begin{array}{l}
 \text{(III) : } \quad x_3 = -5 + 5x_4 = -5 + 5\lambda \\
 \text{in (II) : } \quad x_2 = 2 + 7x_3 - 3x_4 = 2 + 7(-5 + 5\lambda) - 3\lambda = -33 + 32\lambda \\
 \text{in (I) : } \quad x_1 = 1 - x_2 + 3x_3 - x_4 \\
 \quad \quad \quad = 1 - (-33 + 32\lambda) + 3(-5 + 5\lambda) - \lambda = 19 - 18\lambda
 \end{array}$$

Also finden wir die folgende Lösungsmenge für $\alpha = 1$:

$$\mathbb{L}_{[\mathbf{A}|\vec{\mathbf{b}}]} = \left\{ \left[\begin{array}{c} 19 - 18\lambda \\ -33 + 32\lambda \\ -5 + 5\lambda \\ \lambda \end{array} \right] : \lambda \in \mathbb{R} \right\} = \left\{ \left[\begin{array}{c} 19 \\ -33 \\ -5 \\ 0 \end{array} \right] + \lambda \left[\begin{array}{c} -18 \\ 32 \\ 5 \\ 1 \end{array} \right] : \lambda \in \mathbb{R} \right\}.$$

Diese Lösungsmenge ist eine Gerade in \mathbb{R}^4 .

Wir führen zur Übung noch die Reduktion auf die reduzierte Stufenform durch:

$$\left[\begin{array}{cccc|c} 1 & 1 & -3 & 1 & 1 \\ 0 & 1 & -7 & 3 & 2 \\ 0 & 0 & 1 & -5 & -5 \\ 0 & 0 & 0 & 0 & \alpha - 1 \end{array} \right] \begin{array}{l} Z_2 \rightarrow Z_2 + 7 Z_3 \\ Z_1 \rightarrow Z_1 + 3 Z_3 \\ \downarrow \\ \Leftrightarrow \end{array} \left[\begin{array}{cccc|c} 1 & 1 & 0 & -14 & -14 \\ 0 & 1 & 0 & -32 & -33 \\ 0 & 0 & 1 & -5 & -5 \\ 0 & 0 & 0 & 0 & \alpha - 1 \end{array} \right]$$

$$\begin{array}{l} Z_1 \rightarrow Z_1 - Z_2 \\ \downarrow \\ \Leftrightarrow \end{array} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 18 & 19 \\ 0 & 1 & 0 & -32 & -33 \\ 0 & 0 & 1 & -5 & -5 \\ 0 & 0 & 0 & 0 & \alpha - 1 \end{array} \right] \Rightarrow \left\{ \begin{array}{l} \text{LGS ist nur lösbar, wenn } \alpha = 1. \\ \text{Dann gilt:} \\ x_1 = 19 - 18 x_4 \\ x_2 = -33 + 32 x_4 \\ x_3 = -5 + 5 x_4 \\ x_4 = \lambda \text{ mit } \lambda \in \mathbb{R} \text{ beliebig.} \end{array} \right.$$

Natürlich erhalten wir dieselbe Lösung wie mit „Rückwärtsrechnen“.

Wir halten die Vorgehensweise aus dem Beispiel allgemein fest.

Methoden 11.25. (Gaußsches Eliminationsverfahren)

Wir betrachten ein lineares Gleichungssystem (LGS) $\mathbf{A} \vec{x} = \vec{b}$ mit der Matrix $\mathbf{A} \in \mathbb{K}^{m \times n}$ und der rechten Seite $\vec{b} \in \mathbb{K}^m$.

- (1) Durch elementare Zeilenoperationen lässt sich jede erweiterte Koeffizientenmatrix $[\mathbf{A} | \vec{b}]$ in die sogenannte **Stufenform** bringen

$$\left[\begin{array}{cccccccccccc|cccc} 0 & \dots & 0 & 1 & * & \dots & * & * & * & \dots & * & * & * & \dots & * & * & \dots & * \\ \vdots & & \vdots & 0 & \dots & \dots & 0 & 1 & * & \dots & * & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & & \vdots & \vdots & & & \vdots & 0 & \dots & \dots & 0 & * & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & & \vdots & \vdots & & & \vdots & \vdots & & & \vdots & \dots & 1 & * & \dots & * & * & \vdots \\ \vdots & & \vdots & \vdots & & & \vdots & \vdots & & & \vdots & 0 & \dots & \dots & 0 & d_{r+1} & \vdots & \vdots \\ \vdots & & \vdots & \vdots & & & \vdots & \vdots & & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & \dots & \dots & 0 & 0 & \dots & \dots & 0 & 0 & \dots & \dots & 0 & d_m & \vdots & \vdots \end{array} \right] \left. \begin{array}{l} * \\ \vdots \\ \vdots \\ * \\ d_{r+1} \\ \vdots \\ d_m \end{array} \right\} \begin{array}{l} r \\ m \end{array}$$

$\underbrace{\hspace{15em}}_n$

oder sogar in die sogenannte **reduzierte Stufenform**

$$\left[\begin{array}{cccccccccccc|cccc}
 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * & 0 & * & \cdots & * & * \\
 \vdots & & \vdots & 0 & \cdots & \cdots & 0 & 1 & * & \cdots & * & \vdots & \vdots & & \vdots & \vdots \\
 \vdots & & \vdots & \vdots & & & \vdots & 0 & \cdots & \cdots & 0 & 0 & \vdots & & \vdots & \vdots \\
 \vdots & & \vdots & \vdots & & & \vdots & \vdots & & & \vdots & \cdots & 1 & * & \cdots & * \\
 \vdots & & \vdots & \vdots & & & \vdots & \vdots & & & \vdots & 0 & \cdots & \cdots & 0 & d_{r+1} \\
 \vdots & & \vdots & \vdots & & & \vdots & \vdots & & & \vdots & \vdots & & & \vdots & \vdots \\
 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 & d_m
 \end{array} \right]$$

$\underbrace{\hspace{15em}}_n$
} r
} m

- Bei der Stufenform und der reduzierten Stufenform sind die Zeilen, in denen alle Koeffizienten null sind, in den unteren Zeilen der erweiterten Koeffizientenmatrix angeordnet.
- Bei der **Stufenform** ist jede Zeile der erweiterten Matrix von der Form

$$[0 \ \cdots \ 0 \ 1 \ * \ \cdots \ * \ | \ *],$$

wobei die *-Symbole für beliebige (reelle oder komplexe) Zahlen stehen können. Dabei gilt die folgende Regel für die Anordnung der Zeilen: Wandert man durch die Zeilen der Matrix von oben nach unten, so muss die Eins in einer Zeile immer weiter rechts als in der vorhergehenden Zeile auftreten.

- Bei der **reduzierten Stufenform** handelt es sich um eine Matrix in Stufenform mit zusätzlichen Eigenschaften: Über jeder Eins, die in einer Zeile (von links nach rechts) der erste Eintrag ungleich null ist, sind alle Einträge null. (D.h. in der Spalte dieser Eins sind alle Einträge außer der Eins null.)

(2) Es gilt immer $r \leq \min\{m, n\}$.

(3) **Lösbarkeit:**

Fall 1: $r = m$ oder ($r < m$ und $d_{r+1} = \dots = d_m = 0$)

\implies LGS ist lösbar, d.h. $\mathbb{L}_{[\mathbf{A}|\vec{\mathbf{b}}]} \neq \emptyset$.

Falls $r = n$: LGS hat genau eine Lösung.

Falls $r < n$: LGS hat unendlich viele Lösungen

Anzahl der Parameter: $n - r$

Fall 2: $r < m$ und $d_i \neq 0$ für mindestens ein $i > r$

\implies LGS ist unlösbar, d.h. $\mathbb{L}_{[\mathbf{A}|\vec{\mathbf{b}}]} = \emptyset$.

(4) In Fall 1 erhält man die Lösungsmenge aus der Stufenform durch **Rückwärtsrechnen**. An der reduzierten Stufenform lässt sich die Lösungsmenge sogar fast direkt ablesen.

Beispiel 11.26. (Gaußsches Eliminationsverfahren für LGS)

(a) Das lineare Gleichungssystem

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ -x_1 + x_2 &= 2 \\ 2x_1 - 2x_2 + x_3 &= -2 \end{aligned}$$

hat die folgende erweiterte Koeffizientenmatrix:

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ -1 & 1 & 0 & 2 \\ 2 & -2 & 1 & -2 \end{array} \right].$$

Wir bringen diese nun mit elementaren Zeilenoperationen in Stufenform:

$$\begin{aligned} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ -1 & 1 & 0 & 2 \\ 2 & -2 & 1 & -2 \end{array} \right] & \xrightarrow[Z_3 \rightarrow Z_3 + 2 \cdot Z_2]{\Leftrightarrow} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ -1 & 1 & 0 & 2 \\ 0 & 0 & 1 & 2 \end{array} \right] \\ & \xrightarrow[Z_2 \rightarrow Z_2 + Z_1]{\Leftrightarrow} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 3 & 3 & 3 \\ 0 & 0 & 1 & 2 \end{array} \right] \xrightarrow[Z_2 \rightarrow \frac{1}{3} \cdot Z_2]{\Leftrightarrow} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 \end{array} \right] \end{aligned}$$

Als lineares Gleichungssystem haben wir nun:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 & \text{(I)} \\ x_2 + x_3 &= 1 & \text{(II)} \\ x_3 &= 2 & \text{(III)} \end{aligned}$$

Mit „Rückwärtsrechnen“ finden wir also:

Aus (III) : $x_3 = 2$

In (II) einsetzen: $x_2 = 1 - x_3 = 1 - 2 = -1$

In (I) einsetzen: $x_1 = 1 - 2x_2 - 3x_3 = 1 - 2 \cdot (-1) - 3 \cdot 2 = -3$

Also ist die Lösungsmenge des linearen Gleichungssystems

$$\mathbb{L}_{[\mathbf{A}|\vec{\mathbf{b}}]} = \left\{ \begin{bmatrix} -3 \\ -1 \\ 2 \end{bmatrix} \right\}.$$

Das lineare Gleichungssystem

$$\begin{aligned}x_1 + 2x_2 + 3x_3 &= 1 \\ -x_1 + x_2 + x_3 &= 1 \\ 2x_1 - 2x_2 - 2x_3 &= -2\end{aligned}$$

hat die folgende erweiterte Koeffizientenmatrix:

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ -1 & 1 & 1 & 1 \\ 2 & -2 & -2 & -2 \end{array} \right].$$

Wir bringen diese nun mit elementaren Zeilenoperationen in Stufenform:

$$\begin{aligned} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ -1 & 1 & 1 & 1 \\ 2 & -2 & -2 & -2 \end{array} \right] & \xrightarrow[Z_3 \rightarrow Z_3 + 2 \cdot Z_2]{\Leftrightarrow} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ -1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right] \\ & \xrightarrow[Z_2 \rightarrow Z_2 + Z_1]{\Leftrightarrow} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 3 & 4 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right] \xrightarrow[Z_2 \rightarrow \frac{1}{3} \cdot Z_2]{\Leftrightarrow} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & \frac{4}{3} & \frac{2}{3} \\ 0 & 0 & 0 & 0 \end{array} \right] \end{aligned}$$

Die letzte Zeile besteht nur aus Nullen und kann daher ignoriert werden. Wir erhalten aus der Stufenform also die beiden Gleichungen

$$x_1 + 2x_2 + 3x_3 = 1 \quad (\text{I})$$

$$x_2 + \frac{4}{3}x_3 = \frac{2}{3} \quad (\text{II})$$

Wir setzen $x_3 = \lambda$ mit $\lambda \in \mathbb{R}$ beliebig. Dann erhalten wir mit „Rückwärtsrechnen“:

$$\text{Aus (II): } x_2 = \frac{2}{3} - \frac{4}{3}x_3 = \frac{2}{3} - \frac{4}{3}\lambda$$

$$\text{Aus (I): } x_1 = 1 - 2x_2 - 3x_3 = 1 - 2\left(\frac{2}{3} - \frac{4}{3}\lambda\right) - 3\lambda = -\frac{1}{3} - \frac{1}{3}\lambda$$

Also ist die Lösungsmenge des LGS

$$\mathbb{L}_{[\mathbf{A}|\vec{\mathbf{b}}]} = \left\{ \left[\begin{array}{c} -\frac{1}{3} - \frac{1}{3}\lambda \\ \frac{2}{3} - \frac{4}{3}\lambda \\ \lambda \end{array} \right] : \lambda \in \mathbb{R} \right\} = \left\{ \left[\begin{array}{c} -\frac{1}{3} \\ \frac{2}{3} \\ 0 \end{array} \right] + \lambda \left[\begin{array}{c} -\frac{1}{3} \\ -\frac{4}{3} \\ 1 \end{array} \right] : \lambda \in \mathbb{R} \right\}.$$

(c) Das lineare Gleichungssystem

$$\begin{aligned}x_1 + 2x_2 + 3x_3 &= 1 \\ -x_1 + x_2 - x_3 &= 2 \\ 2x_1 - 2x_2 + 2x_3 &= -2\end{aligned}$$

hat die folgende erweiterte Koeffizientenmatrix:

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ -1 & 1 & -1 & 2 \\ 2 & -2 & 2 & -2 \end{array} \right].$$

Mit einer elementaren Zeilenoperation finden wir

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ -1 & 1 & -1 & 2 \\ 2 & -2 & 2 & -2 \end{array} \right] \xrightarrow[Z_3 \rightarrow Z_3 + 2Z_2]{\iff} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ -1 & 1 & -1 & 2 \\ 0 & 0 & 0 & 2 \end{array} \right].$$

Die letzte Zeile ist

$$0x_1 + 0x_2 + 0x_3 = 2 \quad \iff \quad 0 = 2.$$

Da diese Gleichung nicht erfüllbar ist (egal wie wir x_1, x_2 und x_3 wählen), hat das LGS keine Lösung, d.h. es gilt $\mathbb{L}_{[\mathbf{A}|\vec{\mathbf{b}}]} = \emptyset$.

(d) Das lineare Gleichungssystem

$$\begin{aligned}x_1 + x_2 - x_3 + 2x_4 &= 3 \\ 2x_1 - x_2 + x_3 - 2x_4 &= -2 \\ -2x_1 + 2x_2 + 2x_3 + 2x_4 &= 4\end{aligned}$$

hat die erweiterte Koeffizientenmatrix

$$\left[\begin{array}{cccc|c} 1 & 1 & -1 & 2 & 3 \\ 2 & -1 & 1 & -2 & -2 \\ -2 & 2 & 2 & 2 & 4 \end{array} \right].$$

Wir bringen das LGS mit elementaren Zeilenoperationen in reduzierte Stufenform:

$$\left[\begin{array}{cccc|c} 1 & 1 & -1 & 2 & 3 \\ 2 & -1 & 1 & -2 & -2 \\ -2 & 2 & 2 & 2 & 4 \end{array} \right] \xrightarrow[Z_3 \rightarrow Z_3 + Z_2]{\iff} \left[\begin{array}{cccc|c} 1 & 1 & -1 & 2 & 3 \\ 2 & -1 & 1 & -2 & -2 \\ 0 & 1 & 3 & 0 & 2 \end{array} \right]$$

$$\begin{array}{ccc}
\begin{array}{c} Z_2 \rightarrow Z_2 - 2Z_1 \\ \downarrow \\ \Leftrightarrow \end{array} & \left[\begin{array}{cccc|c} 1 & 1 & -1 & 2 & 3 \\ 0 & -3 & 3 & -6 & -8 \\ 0 & 1 & 3 & 0 & 2 \end{array} \right] & \begin{array}{c} Z_2 \rightarrow -\frac{1}{3} \cdot Z_2 \\ \downarrow \\ \Leftrightarrow \end{array} & \left[\begin{array}{cccc|c} 1 & 1 & -1 & 2 & 3 \\ 0 & 1 & -3 & 2 & \frac{8}{3} \\ 0 & 1 & 3 & 0 & 2 \end{array} \right] \\
\begin{array}{c} Z_3 \rightarrow Z_3 - Z_2 \\ \downarrow \\ \Leftrightarrow \end{array} & \left[\begin{array}{cccc|c} 1 & 1 & -1 & 2 & 3 \\ 0 & 1 & -3 & 2 & \frac{8}{3} \\ 0 & 0 & 6 & -2 & -\frac{2}{3} \end{array} \right] & \begin{array}{c} Z_3 \rightarrow \frac{1}{6} \cdot Z_3 \\ \downarrow \\ \Leftrightarrow \end{array} & \left[\begin{array}{cccc|c} 1 & 1 & -1 & 2 & 3 \\ 0 & 1 & -3 & 2 & \frac{8}{3} \\ 0 & 0 & 1 & -\frac{1}{3} & -\frac{1}{9} \end{array} \right] \\
\begin{array}{c} Z_2 \rightarrow Z_2 + 3 \cdot Z_3 \\ Z_1 \rightarrow Z_1 + Z_3 \\ \downarrow \\ \Leftrightarrow \end{array} & \left[\begin{array}{cccc|c} 1 & 1 & 0 & \frac{5}{3} & \frac{26}{9} \\ 0 & 1 & 0 & 1 & \frac{7}{3} \\ 0 & 0 & 1 & -\frac{1}{3} & -\frac{1}{9} \end{array} \right] & \begin{array}{c} Z_1 \rightarrow Z_1 - Z_2 \\ \downarrow \\ \Leftrightarrow \end{array} & \left[\begin{array}{cccc|c} 1 & 0 & 0 & \frac{2}{3} & \frac{5}{9} \\ 0 & 1 & 0 & 1 & \frac{7}{3} \\ 0 & 0 & 1 & -\frac{1}{3} & -\frac{1}{9} \end{array} \right]
\end{array}$$

An der reduzierten Stufenform lesen wir mit $x_4 = \lambda$ mit $\lambda \in \mathbb{R}$ ab:

$$\begin{aligned}
x_1 &= \frac{5}{9} - \frac{2}{3} x_4 = \frac{5}{9} - \frac{2}{3} \lambda, \\
x_2 &= \frac{7}{3} - x_4 = \frac{7}{3} - \lambda, \\
x_3 &= -\frac{1}{9} + \frac{1}{3} x_4 = -\frac{1}{9} + \frac{1}{3} \lambda.
\end{aligned}$$

Also ist die Lösungsmenge des LGS

$$\mathbb{L}_{[\mathbf{A}|\vec{\mathbf{b}}]} = \left\{ \left[\begin{array}{c} \frac{5}{9} - \frac{2}{3} \lambda \\ \frac{7}{3} - \lambda \\ -\frac{1}{9} + \frac{1}{3} \lambda \\ \lambda \end{array} \right] : \lambda \in \mathbb{R} \right\} = \left\{ \left[\begin{array}{c} \frac{5}{9} \\ \frac{7}{3} \\ -\frac{1}{9} \\ 0 \end{array} \right] + \lambda \left[\begin{array}{c} -\frac{2}{3} \\ -1 \\ \frac{1}{3} \\ 1 \end{array} \right] : \lambda \in \mathbb{R} \right\}.$$

Beispiel 11.27. (Gaußsches Eliminationsverfahren für LGS)

Wir betrachten das LGS $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}$ mit der erweiterten Koeffizientenmatrix

$$[\mathbf{A}|\vec{\mathbf{b}}] = \left[\begin{array}{cccccc|c} 0 & 2 & 4 & -2 & 1 & 7 & -1 \\ 1 & 0 & 1 & 3 & 0 & -1 & 1 \\ 1 & 1 & 3 & 2 & 0 & 1 & 1 \\ 0 & 1 & 2 & -1 & -1 & -1 & 1 \\ 3 & 2 & 7 & 7 & -1 & -2 & \alpha \end{array} \right].$$

Mit elementaren Zeilenumformungen finden wir:

$$\begin{array}{ccc}
 \left[\begin{array}{cccccc|c} 0 & 2 & 4 & -2 & 1 & 7 & -1 \\ 1 & 0 & 1 & 3 & 0 & -1 & 1 \\ 1 & 1 & 3 & 2 & 0 & 1 & 1 \\ 0 & 1 & 2 & -1 & -1 & -1 & 1 \\ 3 & 2 & 7 & 7 & -1 & -2 & \alpha \end{array} \right] & \begin{array}{c} Z_1 \leftrightarrow Z_2 \\ \downarrow \\ \iff \end{array} & \left[\begin{array}{cccccc|c} 1 & 0 & 1 & 3 & 0 & -1 & 1 \\ 0 & 2 & 4 & -2 & 1 & 7 & -1 \\ 1 & 1 & 3 & 2 & 0 & 1 & 1 \\ 0 & 1 & 2 & -1 & -1 & -1 & 1 \\ 3 & 2 & 7 & 7 & -1 & -2 & \alpha \end{array} \right] \\
 \\
 & \begin{array}{c} Z_3 \rightarrow Z_3 - Z_1 \\ Z_5 \rightarrow Z_5 - 3Z_1 \\ \downarrow \\ \iff \end{array} & \left[\begin{array}{cccccc|c} 1 & 0 & 1 & 3 & 0 & -1 & 1 \\ 0 & 2 & 4 & -2 & 1 & 7 & -1 \\ 0 & 1 & 2 & -1 & 0 & 2 & 0 \\ 0 & 1 & 2 & -1 & -1 & -1 & 1 \\ 0 & 2 & 4 & -2 & -1 & 1 & \alpha - 3 \end{array} \right] \\
 \\
 & \begin{array}{c} Z_2 \leftrightarrow Z_3 \\ \downarrow \\ \iff \end{array} & \left[\begin{array}{cccccc|c} 1 & 0 & 1 & 3 & 0 & -1 & 1 \\ 0 & 1 & 2 & -1 & 0 & 2 & 0 \\ 0 & 2 & 4 & -2 & 1 & 7 & -1 \\ 0 & 1 & 2 & -1 & -1 & -1 & 1 \\ 0 & 2 & 4 & -2 & -1 & 1 & \alpha - 3 \end{array} \right] \\
 \\
 & \begin{array}{c} Z_3 \rightarrow Z_3 - 2Z_2 \\ Z_4 \rightarrow Z_4 - Z_2 \\ Z_5 \rightarrow Z_5 - 2Z_2 \\ \downarrow \\ \iff \end{array} & \left[\begin{array}{cccccc|c} 1 & 0 & 1 & 3 & 0 & -1 & 1 \\ 0 & 1 & 2 & -1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 & -1 \\ 0 & 0 & 0 & 0 & -1 & -3 & 1 \\ 0 & 0 & 0 & 0 & -1 & -3 & \alpha - 3 \end{array} \right] \\
 \\
 & \begin{array}{c} Z_4 \rightarrow Z_4 + Z_3 \\ Z_5 \rightarrow Z_5 + Z_3 \\ \downarrow \\ \iff \end{array} & \left[\begin{array}{cccccc|c} 1 & 0 & 1 & 3 & 0 & -1 & 1 \\ 0 & 1 & 2 & -1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \alpha - 4 \end{array} \right]
 \end{array}$$

Die erweiterte Koeffizientenmatrix ist nun in **reduzierter Stufenform**. Falls $\alpha \neq 4$, so ist die Lösungsmenge $\mathbb{L}_{[\mathbf{A}|\vec{\mathbf{b}}]} = \emptyset$, denn die letzte Gleichung ist nicht erfüllbar.

Falls $\alpha = 4$ ist, so reduziert sich das lineare Gleichungssystem auf die ersten drei

Gleichungen:

$$x_1 + x_3 + 3x_4 - x_6 = 1 \quad (\text{I})$$

$$x_2 + 2x_3 - x_4 + 2x_6 = 0 \quad (\text{II})$$

$$x_5 + 3x_6 = -1 \quad (\text{III})$$

Setze $x_6 := \lambda_1$, $x_4 := \lambda_2$ und $x_3 := \lambda_3$ mit $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$. Einsetzen in (I) bis (III) und Auflösen nach x_1, x_2 und x_5 liefert

$$\text{aus (III): } x_5 = -1 - 3x_6 = -1 - 3\lambda_1$$

$$\text{aus (II): } x_2 = -2x_6 + x_4 - 2x_3 = -2\lambda_1 + \lambda_2 - 2\lambda_3$$

$$\text{aus (I): } x_1 = 1 + x_6 - 3x_4 - x_3 = 1 + \lambda_1 - 3\lambda_2 - \lambda_3$$

Also finden wir für $\alpha = 4$ die Lösungsmenge

$$\mathbb{L}_{[\mathbf{A}|\vec{\mathbf{b}}]} = \left\{ \begin{bmatrix} 1 + \lambda_1 - 3\lambda_2 - \lambda_3 \\ -2\lambda_1 + \lambda_2 - 2\lambda_3 \\ \lambda_3 \\ \lambda_2 \\ -1 - 3\lambda_1 \\ \lambda_1 \end{bmatrix} : \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R} \right\}$$

$$= \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ -1 \\ 0 \end{bmatrix} + \lambda_1 \begin{bmatrix} 1 \\ -2 \\ 0 \\ 0 \\ -3 \\ 1 \end{bmatrix} + \lambda_2 \begin{bmatrix} -3 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \lambda_3 \begin{bmatrix} -1 \\ -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} : \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R} \right\}.$$

11.4 Lösungstheorie für lineare Gleichungssysteme

Im letzten Teil dieses Kapitels lernen wir erste Resultate zur Lösbarkeit linearer Gleichungssysteme kennen.

Definition 11.28. (homogenes bzw. inhomogenes LGS)

Sei $\mathbf{A} \in \mathbb{K}^{m \times n}$ und $\vec{\mathbf{b}} \in \mathbb{K}^m$.

- (1) Das lineare Gleichungssystem $\mathbf{A} \vec{x} = \vec{b}$ heißt **homogen**, falls $\vec{b} = \vec{0}$ ist. Sonst heißt das lineare Gleichungssystem **inhomogen**.
- (2) Ist $\mathbf{A} \vec{x} = \vec{b}$ ein inhomogenes lineares Gleichungssystem, so heißt $\mathbf{A} \vec{x} = \vec{0}$ das **zugehörige homogene LGS**.

Betrachten wir ein Beispiel.

Beispiel 11.29. (homogenes und inhomogenes LGS)

Das lineare Gleichungssystem

$$\begin{bmatrix} x_1 + x_2 & = & 1 \\ 2x_1 & - & x_3 = 2 \end{bmatrix} \iff \begin{bmatrix} 1 & 1 & 0 \\ 2 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

ist inhomogen. Das zugehörige homogene lineare Gleichungssystem ist

$$\begin{bmatrix} x_1 + x_2 & = & 0 \\ 2x_1 & - & x_3 = 0 \end{bmatrix} \iff \begin{bmatrix} 1 & 1 & 0 \\ 2 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

In den nächsten beiden Sätzen, die wir auch beweisen werden, lernen wir mehr Informationen über die Lösungen von homogenen bzw. inhomogenen Gleichungssystemen.

Satz 11.30. (Lösung des homogenen LGS)

Seien $\mathbf{A} \in \mathbb{K}^{m \times n}$ und $\vec{0} \in \mathbb{K}^m$. Wir betrachten das **homogene lineare Gleichungssystem**

$$\mathbf{A} \vec{x} = \vec{0}. \quad (\text{H})$$

- (1) Das LGS (H) hat immer (mindestens) die **triviale Lösung** $\vec{x} = \vec{0}$, d.h. $\vec{0} \in \mathbb{L}_{[\mathbf{A}|\vec{0}]}$.
- (2) Sind \vec{y} und \vec{z} beide Lösungen von (H), so ist auch $\vec{y} + \vec{z}$ eine Lösung von (H).
(Kurz: $\vec{y}, \vec{z} \in \mathbb{L}_{[\mathbf{A}|\vec{0}]} \implies \vec{y} + \vec{z} \in \mathbb{L}_{[\mathbf{A}|\vec{0}]}$)
- (3) Ist \vec{y} eine Lösung von (H) und $\lambda \in \mathbb{K}$, so ist $\lambda \vec{y}$ Lösung von (H).
(Kurz: $\vec{y} \in \mathbb{L}_{[\mathbf{A}|\vec{0}]}, \lambda \in \mathbb{K} \implies \lambda \vec{y} \in \mathbb{L}_{[\mathbf{A}|\vec{0}]}$)

Beweis von Satz 11.30:

$$(1) \mathbf{A} \vec{\mathbf{0}} = \vec{\mathbf{0}} \implies \vec{\mathbf{0}} \in \mathbb{L}_{[\mathbf{A}|\vec{\mathbf{0}}]}$$

$$(2) \text{ Seien } \vec{\mathbf{y}}, \vec{\mathbf{z}} \in \mathbb{L}_{[\mathbf{A}|\vec{\mathbf{0}}]}. \implies \mathbf{A} \vec{\mathbf{y}} = \vec{\mathbf{0}} \text{ und } \mathbf{A} \vec{\mathbf{z}} = \vec{\mathbf{0}}$$

Nach dem Distributivgesetz für die Matrizenmultiplikation gilt:

$$\mathbf{A} (\vec{\mathbf{y}} + \vec{\mathbf{z}}) = \mathbf{A} \vec{\mathbf{y}} + \mathbf{A} \vec{\mathbf{z}} = \vec{\mathbf{0}} + \vec{\mathbf{0}} = \vec{\mathbf{0}} \implies \vec{\mathbf{y}} + \vec{\mathbf{z}} \in \mathbb{L}_{[\mathbf{A}|\vec{\mathbf{0}}]}$$

$$(3) \text{ Sei } \vec{\mathbf{y}} \in \mathbb{L}_{[\mathbf{A}|\vec{\mathbf{0}}]}. \implies \mathbf{A} \vec{\mathbf{y}} = \vec{\mathbf{0}} \implies \mathbf{A} (\lambda \vec{\mathbf{y}}) = \lambda \mathbf{A} \vec{\mathbf{y}} = \lambda \vec{\mathbf{0}} = \vec{\mathbf{0}} \\ \implies \lambda \vec{\mathbf{y}} \in \mathbb{L}_{[\mathbf{A}|\vec{\mathbf{0}}]} \quad \square$$

Satz 11.31. (Lösung des inhomogenen LGS)

Seien $\mathbf{A} \in \mathbb{K}^{m \times n}$ und $\vec{\mathbf{b}} \in \mathbb{K}^m$. Gegeben sei das *inhomogene* lineare Gleichungssystem

$$\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}, \quad (\text{IH})$$

und das *zugehörige homogene* lineare Gleichungssystem ist

$$\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{0}}. \quad (\text{H})$$

(1) Sind $\vec{\mathbf{y}}$ und $\vec{\mathbf{z}}$ beide Lösungen von (IH), so ist $\vec{\mathbf{y}} - \vec{\mathbf{z}}$ eine Lösung von (H).

$$(\text{Kurz: } \vec{\mathbf{y}}, \vec{\mathbf{z}} \in \mathbb{L}_{[\mathbf{A}|\vec{\mathbf{b}}]} \implies \vec{\mathbf{y}} - \vec{\mathbf{z}} \in \mathbb{L}_{[\mathbf{A}|\vec{\mathbf{0}}]})$$

(2) Kennt man die komplette Lösungsmenge $\mathbb{L}_{[\mathbf{A}|\vec{\mathbf{0}}]}$ von (H) und irgendeine Lösung $\vec{\mathbf{x}}_s$ von (IH), so kennt man die komplette Lösungsmenge $\mathbb{L}_{[\mathbf{A}|\vec{\mathbf{b}}]}$ von (IH):

$$\mathbb{L}_{[\mathbf{A}|\vec{\mathbf{b}}]} = \left\{ \vec{\mathbf{x}}_s + \vec{\mathbf{y}} : \vec{\mathbf{y}} \in \mathbb{L}_{[\mathbf{A}|\vec{\mathbf{0}}]} \right\}.$$

Beweis von Satz 11.31:

$$(1) \text{ Seien } \vec{\mathbf{y}}, \vec{\mathbf{z}} \in \mathbb{L}_{[\mathbf{A}|\vec{\mathbf{b}}]}. \implies \mathbf{A} \vec{\mathbf{y}} = \vec{\mathbf{b}} \text{ und } \mathbf{A} \vec{\mathbf{z}} = \vec{\mathbf{b}} \implies$$

$$\mathbf{A} (\vec{\mathbf{y}} - \vec{\mathbf{z}}) = \mathbf{A} \vec{\mathbf{y}} - \mathbf{A} \vec{\mathbf{z}} = \vec{\mathbf{b}} - \vec{\mathbf{b}} = \vec{\mathbf{0}} \implies \vec{\mathbf{y}} - \vec{\mathbf{z}} \in \mathbb{L}_{[\mathbf{A}|\vec{\mathbf{0}}]}$$

(2) Hier müssen wir die folgenden zwei Teilmengenbeziehungen zeigen

$$\left\{ \vec{\mathbf{x}}_s + \vec{\mathbf{y}} : \vec{\mathbf{y}} \in \mathbb{L}_{[\mathbf{A}|\vec{\mathbf{0}}]} \right\} \subseteq \mathbb{L}_{[\mathbf{A}|\vec{\mathbf{b}}]} \text{ und } \mathbb{L}_{[\mathbf{A}|\vec{\mathbf{b}}]} \subseteq \left\{ \vec{\mathbf{x}}_s + \vec{\mathbf{y}} : \vec{\mathbf{y}} \in \mathbb{L}_{[\mathbf{A}|\vec{\mathbf{0}}]} \right\},$$

$$\text{denn aus diesen folgt } \mathbb{L}_{[\mathbf{A}|\vec{\mathbf{b}}]} = \left\{ \vec{\mathbf{x}}_s + \vec{\mathbf{y}} : \vec{\mathbf{y}} \in \mathbb{L}_{[\mathbf{A}|\vec{\mathbf{0}}]} \right\}.$$

- Sei $\vec{y} \in \mathbb{L}_{[\mathbf{A}|\vec{0}]}$. $\implies \mathbf{A}(\vec{x}_s + \vec{y}) = \mathbf{A}\vec{x}_s + \mathbf{A}\vec{y} = \vec{b} + \vec{0} = \vec{b}$
 $\implies \vec{x}_s + \vec{y} \in \mathbb{L}_{[\mathbf{A}|\vec{b}]}$
 Damit ist $\{\vec{x}_s + \vec{y} : \vec{y} \in \mathbb{L}_{[\mathbf{A}|\vec{0}]}\} \subseteq \mathbb{L}_{[\mathbf{A}|\vec{b}]}$ gezeigt.
- Sei $\vec{x} \in \mathbb{L}_{[\mathbf{A}|\vec{b}]}$. $\xrightarrow{\text{nach (1)}} \vec{y} := \vec{x} - \vec{x}_s \in \mathbb{L}_{[\mathbf{A}|\vec{0}]}$
 $\implies \vec{x} = \vec{x}_s + \vec{y}$ mit $\vec{y} \in \mathbb{L}_{[\mathbf{A}|\vec{0}]}$
 Damit ist $\mathbb{L}_{[\mathbf{A}|\vec{b}]} \subseteq \{\vec{x}_s + \vec{y} : \vec{y} \in \mathbb{L}_{[\mathbf{A}|\vec{0}]}\}$ gezeigt. \square

Als Letztes betrachten wir ein Beispiel für die Anwendung von Satz 11.31.

Beispiel 11.32. (Anwendung von Satz 11.31)

Gesucht sind alle reellen Lösungen des inhomogenen linearen Gleichungssystems

$$\begin{cases} x_1 + 2x_3 = -1 \\ x_2 + x_3 = 2 \end{cases} \iff \left[\begin{array}{ccc|c} 1 & 0 & 2 & -1 \\ 0 & 1 & 1 & 2 \end{array} \right].$$

Durch Inspizieren des linearen Gleichungssystems sieht man, dass

$$\vec{x}_s = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$$

eine Lösung ist. (In der Tat: $1 \cdot (-1) + 0 \cdot 2 + 2 \cdot 0 = -1$ und $0 \cdot (-1) + 1 \cdot 2 + 1 \cdot 0 = 2$.)

Um alle Lösungen des inhomogenen linearen Gleichungssystems zu finden, lösen wir nun das zugehörige homogene lineare Gleichungssystem und nutzen dann Satz 11.31 (2). Das zugehörige homogene lineare Gleichungssystem befindet sich bereits in reduzierter Stufenform, und wir lesen ab:

$$\left[\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 1 & 0 \end{array} \right] \iff \begin{cases} x_1 = -2x_3 \\ x_2 = -x_3 \end{cases}$$

Mit $x_3 = \lambda$ finden wir also

$$\mathbb{L}_{[\mathbf{A}|\vec{0}]} = \left\{ \begin{bmatrix} -2\lambda \\ -\lambda \\ \lambda \end{bmatrix} : \lambda \in \mathbb{R} \right\} = \left\{ \lambda \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} : \lambda \in \mathbb{R} \right\}.$$

Nach Satz 11.31 (2) ist die Lösungsmenge des inhomogenen linearen Gleichungssystems

$$\mathbb{L}_{[\mathbf{A}|\vec{b}]} = \left\{ \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} + \lambda \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} : \lambda \in \mathbb{R} \right\}.$$

KAPITEL 12

Vektorräume

In diesem Kapitel lernen wir den für die lineare Algebra zentralen Begriff eines Vektorraumes kennen. Wir haben bereits einige ganz wichtige Beispiele von Vektorräumen kennengelernt, nämlich \mathbb{R}^n und \mathbb{C}^n , aber auch die Mengen $\mathbb{R}^{m \times n}$ und $\mathbb{C}^{m \times n}$ von reellen bzw. komplexen $m \times n$ -Matrizen sind Vektorräume.

In Teilkapitel 12.1 führen wir zunächst den Begriff eines Vektorraums ein und lernen dann verschiedene Beispiele kennen. In Teilkapitel 12.2 betrachten wir Teilmengen von Vektorräumen, welche die Vektorraumeigenschaften „erben“ – es handelt sich um sogenannte Untervektorräume. Beispielsweise ist die Lösungsmenge eines homogenen linearen Gleichungssystems $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{0}}$ mit $\mathbf{A} \in \mathbb{K}^{m \times n}$, $\vec{\mathbf{0}} \in \mathbb{K}^m$ und dem Vektor der Unbekannten $\vec{\mathbf{x}} \in \mathbb{K}^n$ immer ein Untervektorraum von \mathbb{K}^n . In Teilkapitel 12.3 lernen wir das Konzept von Linearkombinationen kennen und in den Teilkapiteln 12.4 und 12.5 die zentralen Begriffe der linearen Unabhängigkeit von Vektoren und der Basis und Dimension eines Vektorraums.

In Teilkapitel 12.6 nutzen wir eine Basis eines Vektorraums, um Koordinaten einzuführen. Wichtige Koordinatensysteme, die Ihnen vermutlich schon in der Elektrotechnik begegnet sind, sind kartesische, Polar-, Zylinder- und Kugelkoordinaten. Bei diesen Koordinatensystemen sind die Basen allerdings ortsabhängig, und wir werden diese Beispiele erst besprechen können, wenn wir die Differentialrechnung in mehreren Variablen später in dieser Vorlesung kennengelernt haben.

In Teilkapitel 12.7 kommen wir auf Matrizen zurück und lernen den Begriff des Rangs einer Matrix kennen. Mit Hilfe dieses neuen Konzepts gewinnen wir dann weitere wichtige Informationen über die Lösbarkeit linearer Gleichungssysteme.

In Teilkapiteln 12.8 und 12.9 lernen wir den Begriff eines Skalarprodukts und

der Orthogonalität von Vektoren kennen. Wir haben diese Konzepte bereits in Kapitel 2 der HM A für den Spezialfall von \mathbb{R}^n mit dem Standardskalarprodukt kennengelernt. Nun kommen allerdings neue darauf aufbauende Begriffe hinzu.

Auch in diesem Kapitel ist \mathbb{K} wieder \mathbb{R} oder \mathbb{C} , und n ist aus \mathbb{N} .

12.1 Vektorräume: Definition und Beispiele

Wir starten mit der Definition eines Vektorraumes. Dabei sollten Sie als „Standardbeispiel“ immer \mathbb{R}^n mit der klassischen Vektoraddition und der skalaren Multiplikation vor Augen haben.

Definition 12.1. (Vektorraum)

Eine nichtleere Menge V heißt ein \mathbb{K} -Vektorraum (\mathbb{K} -VR), wenn auf V

- eine **Addition** \oplus mit $x \oplus y \in V$ für alle $x, y \in V$, und
- eine **Multiplikation** \odot mit **Skalaren aus** \mathbb{K} (einer **skalaren Multiplikation** „Skalar \odot Vektor“) mit $\lambda \odot x \in V$ für alle $x \in V$ und $\lambda \in \mathbb{K}$

definiert sind mit den folgenden Eigenschaften:

(1) **Assoziativgesetz für \oplus :**

$$x \oplus (y \oplus z) = (x \oplus y) \oplus z \quad \text{für alle } x, y, z \in V.$$

(2) **Kommutativgesetz für \oplus :**

$$x \oplus y = y \oplus x \quad \text{für alle } x, y \in V.$$

(3) **Assoziativgesetz für \odot :**

$$\lambda \odot (\mu \odot x) = (\lambda \cdot \mu) \odot x \quad \text{für alle } \lambda, \mu \in \mathbb{K} \text{ und alle } x \in V.$$

(4) **Distributivgesetze:**

$$(\lambda + \mu) \odot x = (\lambda \odot x) \oplus (\mu \odot x) \quad \text{für alle } \lambda, \mu \in \mathbb{K} \text{ und alle } x \in V, \tag{12.1}$$

$$\lambda \odot (x \oplus y) = (\lambda \odot x) \oplus (\lambda \odot y) \quad \text{für alle } \lambda \in \mathbb{K} \text{ und alle } x, y \in V. \tag{12.2}$$

(5) $1 \odot x = x$ für alle $x \in V$.

(6) **Existenz eines Nullvektors** o_V : Es existiert ein **eindeutig bestimmtes** $o_V \in V$ mit

$$x \oplus o_V = o_V \oplus x = x \quad \text{für alle } x \in V.$$

o_V heißt der **Nullvektor** von V .

(7) **Existenz des Inversen**:

Zu jedem $x \in V$ existiert ein **eindeutig bestimmtes** y mit

$$x \oplus y = y \oplus x = o_V.$$

Wir nennen dieses y dann das **Inverse von** x und wir schreiben $-x$ statt y .

Es gilt immer **Punktrechnung** \odot **vor Strichrechnung** \oplus . Die Klammern auf der rechten Seite von (12.1) und (12.2) darf man also weglassen.

Betrachten wir zunächst ein Beispiel, um uns einige „ungewöhnlich“ aussehende Eigenschaften eines Vektorraumes klar zu machen.

Beispiel 12.2. (Standardbeispiel \mathbb{R}^n)

\mathbb{R}^n mit \oplus als der üblichen Vektoraddition

$$\vec{x} \oplus \vec{y} := \vec{x} + \vec{y} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix}$$

und der üblichen skalaren Multiplikation

$$\lambda \odot \vec{x} := \lambda \vec{x} = \lambda \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{bmatrix}$$

ist ein Vektorraum. Dabei sind die Eigenschaften (1) bis (5) intuitiv klar und leicht nachzurechnen, wenn man das Assoziativgesetz, das Kommutativgesetz und das Distributivgesetz der reellen Zahlen ausnutzt.

Der Nullvektor ist

$$o_{\mathbb{R}^n} = \vec{\mathbf{0}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^n,$$

denn es gilt

$$\vec{x} + \vec{0} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 + 0 \\ \vdots \\ x_n + 0 \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \vec{x} \quad \text{für alle } \vec{x} \in \mathbb{R}^n$$

und analog $\vec{0} + \vec{x} = \vec{x}$ für alle $\vec{x} \in \mathbb{R}^n$.

Das Inverse von $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ ist $-\vec{x} = \begin{bmatrix} -x_1 \\ \vdots \\ -x_n \end{bmatrix}$, denn

$$\vec{x} + (-\vec{x}) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} -x_1 \\ \vdots \\ -x_n \end{bmatrix} = \begin{bmatrix} x_1 + (-x_1) \\ \vdots \\ x_n + (-x_n) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = \vec{0},$$

und wegen (2) folgt damit $(-\vec{x}) + \vec{x} = \vec{0}$.

Bemerkung 12.3. (Notation bei der skalaren Multiplikation)

Es ist durchaus üblich bei der **skalaren Multiplikation** („Skalar \odot Vektor“) den Punkt für die Multiplikation wegzulassen, sofern dieses keine Verwirrung stiften kann.

In Definition 12.1 (3) finden wir plötzlich verschiedene Multiplikationszeichen; wieso macht das Sinn?

$$\lambda \odot (\mu \odot x) = (\lambda \cdot \mu) \odot x \quad \text{für alle } \lambda, \mu \in \mathbb{K} \text{ und alle } x \in V.$$

In $\lambda \odot (\mu \odot x)$ haben wir innerhalb der Klammern eine „Skalar \odot Vektor“-Multiplikation, die wieder einen Vektor ergibt. Daher ist auch in $\lambda \odot (\mu \odot x)$ die Multiplikation außerhalb der Klammern wieder eine „Skalar \odot Vektor“-Multiplikation. In $(\lambda \cdot \mu) \odot x$ werden dagegen zunächst in den Klammern zwei Zahlen $\mu, \lambda \in \mathbb{K}$ multipliziert. Hier haben wir also die normale Multiplikation \cdot von reellen bzw. komplexen Zahlen, die wieder eine reelle bzw. komplexe Zahl liefert. Diese wird nun mit dem Vektor x multipliziert, d.h. in $(\lambda \cdot \mu) \odot x$ liegt außerhalb der Klammern eine „Skalar \odot Vektor“-Multiplikation vor.

Analog kann man sich klarmachen, warum bei dem ersten Distributivgesetz in Definition 12.1 (4) sowohl die Addition \oplus für die Vektoren als auch die Addition $+$ für reelle bzw. komplexe Zahlen auftauchen.

Bevor wir verschiedene Beispiele betrachten, schauen, wir die Eigenschaften eines Vektorraumes noch einmal genauer an und halten noch einige abgeleitete Rechenregeln fest:

Bemerkung 12.4. (Rechenregeln für Vektorräume)

Es gelten die folgenden abgeleiteten Rechenregeln:

$$(8) \quad 0 \odot x = o_V \text{ für alle } x \in V,$$

denn es gilt

$$x \stackrel{(5)}{=} 1 \odot x = (0 + 1) \odot x \stackrel{(4)}{=} 0 \odot x \oplus \underbrace{1 \odot x}_{=x} \stackrel{(5)}{=} 0 \odot x \oplus x.$$

Nun addieren wir auf beiden Seiten das Inverse von x , also $-x$, und nutzen (7) aus:

$$\underbrace{x \oplus (-x)}_{=o_V} = 0 \odot x \oplus \underbrace{x \oplus (-x)}_{=o_V} \iff o_V = 0 \odot x \oplus o_V \stackrel{(6)}{=} 0 \odot x.$$

$$(9) \quad \lambda \odot o_V = o_V \text{ für alle } \lambda \in \mathbb{K},$$

denn es gilt

$$\lambda \odot x \stackrel{(6)}{=} \lambda \odot (o_V \oplus x) \stackrel{(4)}{=} \lambda \odot o_V \oplus \lambda \odot x.$$

Nun addieren wir auf beiden Seiten das Inverse von $\lambda \odot x$, also $-(\lambda \odot x)$, und nutzen (7) aus:

$$\begin{aligned} \underbrace{\lambda \odot x \oplus (-(\lambda \odot x))}_{=o_V} &= \lambda \odot o_V \oplus \underbrace{\lambda \odot x \oplus (-(\lambda \odot x))}_{=o_V} \\ \iff o_V &= \lambda \odot o_V \oplus o_V \stackrel{(6)}{\iff} o_V = \lambda \odot o_V. \end{aligned}$$

$$(10) \quad \text{Es gilt: } -x = (-1) \odot x,$$

denn

$$x \oplus ((-1) \odot x) \stackrel{(5)}{=} 1 \odot x \oplus (-1) \odot x \stackrel{(4)}{=} (1 + (-1)) \odot x = 0 \odot x \stackrel{(8)}{=} o_V.$$

Mit dem Kommutativgesetz (2) folgt daraus direkt $((-1) \odot x) \oplus x = o_V$.

$$(11) \quad \text{Wir haben eine Subtraktion durch: } x \ominus y := x \oplus (-y)$$

Betrachten wir nun einige vertraute und einige neue Beispiele für Vektorräume.

Beispiel 12.5. (Vektorräume)

- (a) Die reellen Zahlen \mathbb{R} mit der üblichen Addition $+$ reeller Zahlen und der skalaren Multiplikation $\lambda \cdot x$ eines Skalars $\lambda \in \mathbb{R}$ mit einer reellen Zahl x (als Vektor in \mathbb{R}) ist ein \mathbb{R} -Vektorraum. Hier sind also die Vektoraddition \oplus bzw. die Multiplikation mit Skalaren \odot gerade die übliche Addition $+$ bzw. die übliche Multiplikation \cdot reeller Zahlen. Der Nullvektor ist die reelle Zahl 0 . Die Inverse von $x \in \mathbb{R}$ ist die reelle Zahl $-x$.
- (b) Die im Kapitel 2 der HM A definierten Pfeile im Anschauungsraum bilden einen \mathbb{R} -Vektorraum bzgl. der dort definierten Addition $+$ und Multiplikation \cdot mit Skalaren. Der Nullvektor ist der Nullpfeil. Hier ist \oplus die übliche Vektoraddition $+$, und \odot ist die skalare Multiplikation \cdot (und wir lassen häufig den Multiplikationspunkt weg).
- (c) Die Menge $\mathbb{K}^n = \mathbb{K}^{n \times 1}$ der Spaltenvektoren mit der Multiplikation mit Skalaren aus \mathbb{K} bilden einen \mathbb{K} -Vektorraum:

$$\text{Addition:} \quad \vec{x} \oplus \vec{y} := \vec{x} + \vec{y} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix}$$

$$\text{Multiplikation mit Skalaren:} \quad \lambda \odot \vec{x} := \lambda \vec{x} = \lambda \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{bmatrix}$$

Hier ist \oplus die Vektoraddition $+$, und \odot ist die Multiplikation \cdot mit Skalaren (und wir lassen häufig den Multiplikationspunkt weg). Der Nullvektor ist der Nullvektor $\vec{0} \in \mathbb{K}^n$.

$$\text{Das Inverse von } \vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \text{ ist } -\vec{x} = \begin{bmatrix} -x_1 \\ \vdots \\ -x_n \end{bmatrix}.$$

- (d) Die Menge $\mathbb{K}^{m \times n}$ der $m \times n$ -Matrizen ist ein \mathbb{K} -Vektorraum:

$$\begin{aligned} \text{Addition:} \quad & \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix} + \begin{bmatrix} b_{1,1} & \cdots & b_{1,n} \\ \vdots & & \vdots \\ b_{m,1} & \cdots & b_{m,n} \end{bmatrix} \\ &= \begin{bmatrix} a_{1,1} + b_{1,1} & \cdots & a_{1,n} + b_{1,n} \\ \vdots & & \vdots \\ a_{m,1} + b_{m,1} & \cdots & a_{m,n} + b_{m,n} \end{bmatrix} \end{aligned}$$

$$\text{Multiplikation mit Skalaren:} \quad \lambda \cdot \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix} = \begin{bmatrix} \lambda a_{1,1} & \cdots & \lambda a_{1,n} \\ \vdots & & \vdots \\ \lambda a_{m,1} & \cdots & \lambda a_{m,n} \end{bmatrix}$$

Hier ist \oplus die Matrizenaddition $+$, und \odot die Multiplikation \cdot einer Zahl in \mathbb{K} mit einer Matrix in $\mathbb{K}^{m \times n}$ (und wir lassen häufig den Multiplikationspunkt weg). Der Nullvektor ist hier die Nullmatrix $\mathbf{O}_{m \times n}$, und die Inverse von

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix}$$

ist

$$-\mathbf{A} = (-1) \cdot \mathbf{A} = \begin{bmatrix} -a_{1,1} & \cdots & -a_{1,n} \\ \vdots & & \vdots \\ -a_{m,1} & \cdots & -a_{m,n} \end{bmatrix}.$$

(e) Die Menge \mathcal{P}_n aller Polynomfunktionen

$$p : \mathbb{K} \rightarrow \mathbb{K}, \quad p(t) = a_0 + a_1 t + a_2 t^2 + \dots + a_n t^n, \quad \text{mit } t \in \mathbb{K},$$

vom Grad $\leq n$, mit Koeffizienten $a_0, a_1, a_2, \dots, a_n \in \mathbb{K}$, ist ein \mathbb{K} -Vektorraum bzgl. der

$$\begin{aligned} \text{Addition:} \quad & (a_0 + a_1 t + a_2 t^2 + \dots + a_n t^n) + (b_0 + b_1 t + b_2 t^2 + \dots + b_n t^n) \\ & = (a_0 + b_0) + (a_1 + b_1) t + (a_2 + b_2) t^2 + \dots + (a_n + b_n) t^n, \end{aligned}$$

$$\begin{aligned} \text{Multiplikation mit Skalaren:} \quad & \lambda \cdot (a_0 + a_1 t + a_2 t^2 + \dots + a_n t^n) \\ & = (\lambda a_0) + (\lambda a_1) t + (\lambda a_2) t^2 + \dots + (\lambda a_n) t^n. \end{aligned}$$

Hier ist \oplus die Addition der Polynomfunktionen, die erfolgt, indem man die Koeffizienten der entsprechenden Potenzen addiert, und die skalare Multiplikation \odot ist die Multiplikation, die man erhält, indem man alle Koeffizienten mit dem Skalar multipliziert. Der Nullvektor ist das Nullpolynom, und das Inverse von $p \in \mathcal{P}_n$ ist $-p = (-1) \cdot p$. Auch hier schreiben wir der Einfachheit halber $+$ und \cdot statt \oplus und \odot .

(f) Sei M eine nichtleere Menge. Die Menge $\mathcal{F}(M)$ aller Funktionen $f : M \rightarrow \mathbb{K}$ bildet einen \mathbb{K} -Vektorraum bzgl. der Addition

$$f + g : M \rightarrow \mathbb{K}, \quad (f + g)(x) := f(x) + g(x),$$

und der skalaren Multiplikation

$$\lambda f : M \rightarrow \mathbb{K}, \quad (\lambda f)(x) := \lambda \cdot f(x) \quad \text{für } \lambda \in \mathbb{K}.$$

Der Nullvektor ist hier die Nullfunktion $o : M \rightarrow \mathbb{K}$, $o(x) := 0$ für alle $x \in M$. Die Inverse von $f \in \mathcal{F}(M)$ ist $-f = (-1) \cdot f$. Auch hier schreiben wir der Einfachheit halber $+$ und \cdot statt \oplus und \odot .

- (g) Die Menge $\mathcal{F}(\mathbb{N})$ aller Folgen $(a_k)_{k \geq 1}$ in \mathbb{K} bildet einen \mathbb{K} -Vektorraum bzgl. der

$$\text{Addition:} \quad (a_k)_{k \geq 1} + (b_k)_{k \geq 1} = (a_k + b_k)_{k \geq 1}$$

$$\text{Multiplikation mit Skalaren:} \quad \lambda \cdot (a_k)_{k \geq 1} = (\lambda a_k)_{k \geq 1}$$

Der Nullvektor ist die Nullfolge $(0)_{k \geq 1} = (0, 0, 0, \dots)$. Das Inverse von $(a_k)_{k \geq 1}$ ist die Folge $(-a_k)_{k \geq 1}$. Auch hier schreiben wir der Einfachheit halber $+$ und \cdot statt \oplus und \odot .

- (h) Als Letztes betrachten wir noch ein „obskures“ Beispiel, bei dem man sieht, dass die Vektoraddition \oplus und die skalare Multiplikation \odot ungewöhnlich aussehen können: Die Menge

$$V := \{(x, y) : x, y \in \mathbb{R} \text{ mit } x > 0\}$$

mit der Addition

$$(x, y) \oplus (u, v) = (x \cdot u, y + v) = (x u, y + v)$$

und der skalaren Multiplikation

$$\lambda \odot (x, y) = (x^\lambda, \lambda \cdot y) = (x^\lambda, \lambda y)$$

ist ein \mathbb{R} -Vektorraum.

Da für $(x, y), (u, v) \in V$ gilt $x > 0$ und $u > 0$, folgt $x u > 0$ und damit

$$(x, y) \oplus (u, v) = (x u, y + v) \in V.$$

Da für $(x, y) \in V$ gilt $x > 0$, folgt für jedes $\lambda \in \mathbb{R}$, dass $x^\lambda > 0$ und damit

$$\lambda \odot (x, y) = (x^\lambda, \lambda y) \in V.$$

Nun überprüfen wir die Vektorraumeigenschaften:

- (1) Für alle $(x, y), (u, v), (z, w) \in V$ gilt

$$\begin{aligned} (x, y) \oplus ((u, v) \oplus (z, w)) &= (x, y) \oplus (u z, v + w) = (x u z, y + v + w) \\ &= (x u, y + v) \oplus (z, w) = ((x, y) \oplus (u, v)) \oplus (z, w). \end{aligned}$$

- (2) Für alle $(x, y), (u, v) \in V$ gilt

$$(x, y) \oplus (u, v) = (x u, y + v) = (u x, v + y) = (u, v) \oplus (x, y).$$

(3) Für alle $(x, y) \in V$ und alle $\lambda, \mu \in \mathbb{R}$ gilt

$$\begin{aligned}\lambda \odot (\mu \odot (x, y)) &= \lambda \odot (x^\mu, \mu y) = ((x^\mu)^\lambda, \lambda \mu y) \\ &= (x^{\lambda\mu}, (\lambda \mu) y) = (\lambda \mu) \odot (x, y).\end{aligned}$$

(4) Für alle $(x, y), (u, v) \in V$ und alle $\lambda, \mu \in \mathbb{R}$ gelten

$$\begin{aligned}(\lambda + \mu) \odot (x, y) &= (x^{\lambda+\mu}, (\lambda + \mu) y) = (x^\lambda x^\mu, \lambda y + \mu y) \\ &= (x^\lambda, \lambda y) \oplus (x^\mu, \mu y) = \lambda \odot (x, y) \oplus \mu \odot (x, y),\end{aligned}$$

$$\begin{aligned}\lambda \odot ((x, y) \oplus (u, v)) &= \lambda \odot (xu, y + v) = ((xu)^\lambda, \lambda(y + v)) \\ &= (x^\lambda u^\lambda, \lambda y + \lambda v) = (x^\lambda, \lambda y) \oplus (u^\lambda, \lambda v) \\ &= \lambda \odot (x, y) \oplus \lambda \odot (u, v).\end{aligned}$$

(5) $1 \odot (x, y) = (x^1, 1 \cdot y) = (x, y)$ für alle $(x, y) \in V$.

(6) Der Vektor $(1, 0)$ ist der Nullvektor, denn

$$(1, 0) \oplus (x, y) = (1 \cdot x, 0 + y) = (x, y) \quad \text{für alle } (x, y) \in V,$$

und wegen (2) folgt auch $(x, y) \oplus (1, 0) = (x, y)$ für alle $(x, y) \in V$.

(7) Wir behaupten, dass $(\frac{1}{x}, -y) \in V$ das zu $(x, y) \in V$ Inverse ist. In der Tat finden wir

$$\begin{aligned}(x, y) \oplus \left(\frac{1}{x}, -y\right) &= \left(x \cdot \frac{1}{x}, y - y\right) = (1, 0), \\ \left(\frac{1}{x}, -y\right) \oplus (x, y) &= \left(\frac{1}{x} \cdot x, -y + y\right) = (1, 0).\end{aligned}$$

Alle Vektorraumeigenschaften sind erfüllt. Also ist V mit \oplus und \odot ist in der Tat ein \mathbb{R} -Vektorraum.

Ab jetzt schreiben wir der Einfachheit halber in beliebigen \mathbb{K} -Vektorräumen V wieder $+$ und \cdot statt \oplus und \odot . Wir sollten uns aber darüber klar sein, dass, je nach Kontext, ein $+$ bzw. ein \cdot eine Vektoraddition in V oder eine Addition (reeller oder komplexer Zahlen) in \mathbb{K} bzw. eine skalare Multiplikation von einem Skalar mit einem Vektor oder eine Multiplikation (reeller oder komplexer Zahlen) in \mathbb{K} sein kann. Das Multiplikationszeichen \cdot wird auch häufig weggelassen, z.B. bedeutet λx für $\lambda \in \mathbb{K}$ und $x \in V$ die „Skalar \cdot Vektor“-Multiplikation $\lambda \cdot x$.

Abschließend betrachten wir noch einige Beispiele, die **keine** Vektorräume sind.

Beispiel 12.6. (keine Vektorräume)

- (a) Die ganzen Zahlen \mathbb{Z} mit der üblichen Addition $+$ und der üblichen Multiplikation \cdot als skalarer Multiplikation sind kein \mathbb{R} -Vektorraum, denn für $\frac{1}{3} \in \mathbb{R}$ und $1 \in \mathbb{Z}$ gilt $\frac{1}{3} \cdot 1 = \frac{1}{3} \notin \mathbb{Z}$.
- (b) Die rationalen Zahlen \mathbb{Q} mit der üblichen Addition $+$ und der üblichen Multiplikation \cdot als skalarer Multiplikation sind kein \mathbb{R} -Vektorraum, denn für $\pi \in \mathbb{R}$ und $1 \in \mathbb{Q}$ gilt $\pi \cdot 1 = \pi \notin \mathbb{Q}$.
- (c) Die Kugeloberfläche

$$S := \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$$

der Kugel um $(0, 0, 0)$ mit Radius 1 mit der üblichen Addition und skalaren Multiplikation ist kein \mathbb{R} -Vektorraum, denn für $(1, 0, 0) \in S$ gilt

$$(1, 0, 0) + (1, 0, 0) = (2, 0, 0) \notin S.$$

- (d) Die Menge M aller streng monoton wachsenden Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$ mit der üblichen Addition von Funktionen und der üblichen Multiplikation von Funktionen mit reellen Zahlen ist kein \mathbb{R} -Vektorraum, denn für $-1 \in \mathbb{R}$ und die streng monoton wachsende Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) := x$, gilt

$$(-1) \cdot f : \mathbb{R} \rightarrow \mathbb{R}, \quad ((-1) \cdot f)(x) = -f(x) = -x,$$

und diese Funktion ist streng monoton fallend. Somit ist sie nicht streng monoton wachsend, also $(-1) \cdot f \notin M$.

12.2 Untervektorräume

Wir führen zunächst den wichtigen Begriff eines Untervektorraums oder linearen Teilraums ein. Bei einem Untervektorraum eines Vektorraums handelt es sich anschaulich um eine Teilmenge, die den Nullvektor enthält und die Vektoraddition und die skalare Multiplikation „erbt“, wobei das Ergebnis dieser Operationen immer wieder in der Teilmenge liegt. Diese Eigenschaften liegen natürlich nur für bestimmte Teilmengen vor.

Satz 12.7. (Untervektorraum oder linearer Teilraum)

Sei V ein \mathbb{K} -Vektorraum und $U \subseteq V$. Falls

- (i) $0_V \in U$ (und damit $U \neq \emptyset$),

(ii) $x, y \in U \implies x + y \in U$, und

(iii) $x \in U, \lambda \in \mathbb{K} \implies \lambda x \in U$

gelten, dann ist U **selbst ein \mathbb{K} -Vektorraum**. Man nennt U einen **Untervektorraum (UVR) von V** (oder einen **linearen Teilraum von V**).

Beweis von Satz 12.7: Wir müssen überprüfen, dass U die Vektorraumeigenschaften erfüllt. Eigenschaften (ii) und (iii) garantieren, dass die Addition $+$ und die skalare Multiplikation \cdot von V für Elemente aus U jeweils ein Ergebnis in U besitzen. Die Eigenschaften (1) bis (5) gelten in V und damit auch in der Teilmenge U von V . Die Eigenschaft (i) garantiert, dass der Nullvektor o_V von V in U liegt. Damit gilt auch $u + o_V = o_V + u = u$ für alle $u \in U$, d.h. o_V ist auch der Nullvektor von U , und (6) ist erfüllt. Nach (iii) gilt $(-1) \cdot u \in U$, und nach Bemerkung 12.4 (10) ist $(-1) \cdot u = -u$ das Inverse zu $u \in U$ in V . Also gilt $u + (-u) = (-u) + u = o_V$. Weil o_V der Nullvektor von U ist, ist $(-1) \cdot u = -u$ auch das Inverse von $u \in U$ in der Teilmenge U . Da U alle Eigenschaften eines \mathbb{K} -Vektorraums erfüllt, ist U ebenfalls ein \mathbb{K} -Vektorraum. \square

Wir sehen an dem vorigen Beweis, dass die Eigenschaften (i) bis (iii) in Satz 12.7 garantieren, dass die Teilmenge U von V die Vektorraumeigenschaften „erbt“.

Betrachten wir einige Beispiele.

Beispiel 12.8. (Untervektorräume)

(a) Die Teilmenge

$$U := \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 : 2x_1 + x_2 = 0 \right\}$$

ist ein Untervektorraum von \mathbb{R}^2 , denn:

(i) $\begin{bmatrix} 0 \\ 0 \end{bmatrix} \in U$, da $2 \cdot 0 + 0 = 0$

(ii) Seien $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \in U$. Dann müssen wir zeigen, dass auch

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix}$$

in U liegt. Wir überprüfen die Bedingung an Elemente in U :

$$2(x_1 + y_1) + (x_2 + y_2) = \underbrace{(2x_1 + x_2)}_{=0} + \underbrace{(2y_1 + y_2)}_{=0} = 0,$$

wobei die Terme in den Klammern jeweils null sind, weil $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ und $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ in U sind. Also folgt $\begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix} \in U$.

(iii) Seien $\lambda \in \mathbb{R}$ und $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in U$. Wir müssen zeigen, dass auch

$$\lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \end{bmatrix}$$

in U liegt. Wir überprüfen die Bedingung an Elemente in U :

$$2(\lambda x_1) + (\lambda x_2) = \lambda \underbrace{(2x_1 + x_2)}_{=0} = \lambda \cdot 0 = 0,$$

wobei der Term in den Klammern null ist, weil $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in U$ ist. Also

$$\text{folgt } \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \end{bmatrix} \in U.$$

Da alle drei Eigenschaften eines Untervektorraums erfüllt sind, ist U ein Untervektorraum von \mathbb{R}^2 .

(b) Die Teilmenge

$$U := \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 : 2x_1 + x_2 = 3 \right\}$$

ist **kein** Untervektorraum von \mathbb{R}^2 , denn $\begin{bmatrix} 0 \\ 0 \end{bmatrix} \notin U$, da $2 \cdot 0 + 0 = 0 \neq 3$.

(c) Für jedes $n \in \mathbb{N}_0$ ist der \mathbb{K} -Vektorraum \mathcal{P}_n aller Polynomfunktionen auf \mathbb{K} vom Grad $\leq n$ (aus Beispiel 12.5 (e)) ein Untervektorraum des \mathbb{K} -Vektorraums $\mathcal{F}(\mathbb{K})$ der \mathbb{K} -wertigen Funktionen auf \mathbb{K} (vgl. Beispiel 12.5 (f) mit $M = \mathbb{K}$). Dieses sieht man direkt, denn die Vektoraddition und die skalare Multiplikation sind in beiden Räumen gleich definiert, und wir wissen bereits, dass:

- (i) die Nullfunktion (also das Nullpolynom) in \mathcal{P}_n liegt,
- (ii) die Addition zweier Polynomfunktionen vom Grad $\leq n$ wieder eine Polynomfunktion vom Grad $\leq n$ ergibt, und
- (iii) die skalare Multiplikation einer Polynomfunktion vom Grad $\leq n$ mit einer Zahl $\lambda \in \mathbb{K}$ wieder eine Polynomfunktion vom Grad $\leq n$ ergibt.

Da alle drei Eigenschaften eines Untervektorraums erfüllt sind, ist \mathcal{P}_n ein Untervektorraum von $\mathcal{F}(\mathbb{K})$.

Der nächste Satz zeigt, dass Untervektorräume für das Verständnis der Lösungsmengen linearer Gleichungssysteme wichtig sind.

Satz 12.9. (Lösungsmenge eines homogenen LGS ist ein UVR)

Sei $\mathbf{A} \in \mathbb{K}^{m \times n}$. Die Lösungsmenge

$$\mathbb{L}_{[\mathbf{A}|\mathbf{0}]} = \{ \vec{x} \in \mathbb{K}^n : \mathbf{A} \vec{x} = \vec{\mathbf{0}} \}$$

des *homogenen* linearen Gleichungssystems $\mathbf{A} \vec{x} = \vec{\mathbf{0}}$ ist ein **Untervektorraum** von \mathbb{K}^n .

Beweis von Satz 12.9: Dieses folgt aus Satz 11.30 und Satz 12.7: Die Aussagen (1) bis (3) in Satz 11.30 garantieren, dass $\mathbb{L}_{[\mathbf{A}|\mathbf{0}]}$ die Eigenschaften (i) bis (iii) in Satz 12.7 erfüllt und somit ein Untervektorraum von \mathbb{K}^n ist. \square

Betrachten wir noch einige weitere Beispiele für Untervektorräume.

Beispiel 12.10. (Untervektorräume)

- (a) Sei $I \subseteq \mathbb{R}$ ein Intervall und $\mathbb{K} = \mathbb{R}$. Dann ist die Menge aller auf I stetigen Funktionen

$$\mathcal{C}(I) = \{ f : I \rightarrow \mathbb{R} : f \text{ ist stetig auf } I \}$$

ein Untervektorraum von dem Raum $\mathcal{F}(I)$ aller reellwertigen Funktionen auf I (vgl. Beispiel 12.5 (f) mit $M = I$), denn

- (i) die Nullfunktion ist stetig und liegt somit in $\mathcal{C}(I)$,
- (ii) die Summe zweier stetiger Funktionen ist ebenfalls stetig, und
- (iii) das λ -fache einer stetigen Funktion ist stetig.

- (b) Sei $I \subseteq \mathbb{R}$ ein offenes Intervall. Für jedes $k \in \mathbb{N}$ ist

$$\mathcal{C}^k(I) = \{ f : I \rightarrow \mathbb{R} : f \text{ ist } k\text{-mal stetig differenzierbar auf } I \}$$

ein Untervektorraum von $\mathcal{C}(I)$. Die Begründung ist analog zu Beispiel (a).

- (c) Sei I ein offenes Intervall und $a \in \mathcal{C}(I)$. Die Menge aller Lösungen der homogenen linearen Differentialgleichung

$$y' + a(t)y = 0$$

ist ein Untervektorraum von $\mathcal{C}^1(I)$. Diese folgt mit unserem Wissen aus Kapitel 9 der HM A.

(d) Die Menge der

- beschränkten Folgen,
- konvergenten Folgen,
- Nullfolgen

ist jeweils ein Untervektorraum des Vektorraums $\mathcal{F}(\mathbb{N})$ der Folgen (vgl. Beispiel 12.5 (g)).

(e) Sei V ein \mathbb{K} -Vektorraum. Dann sind V selber und $\{o_V\}$ jeweils Untervektorräume von V .

12.3 Linearkombinationen

Wir lernen nun Linearkombinationen von Vektoren kennen. Der Begriff der Linearkombination zusammen mit dem Begriff der linearen Unabhängigkeit (siehe Teilkapitel 12.4) sind ganz zentral für die lineare Algebra.

Definition 12.11. (Linearkombination)

Seien V ein \mathbb{K} -Vektorraum und $k \in \mathbb{N}$, und seien $v_1, v_2, \dots, v_k \in V$. Für $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{K}$ heißt

$$x = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k = \sum_{i=1}^k \lambda_i v_i$$

eine **Linearkombination (LK)** von v_1, v_2, \dots, v_k .

Wir halten noch zwei hilfreiche Beobachtungen über Linearkombinationen fest.

Bemerkung 12.12. (Linearkombination)

(1) Ist U ein Untervektorraum eines \mathbb{K} -Vektorraums V und sind die Vektoren $u_1, u_2, \dots, u_k \in U$, so gilt

$$\lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_k u_k \in U \quad \text{für alle } \lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{K}.$$

Dieses folgt durch wiederholte Anwendung der Eigenschaften (ii) und (iii) in Satz 12.7.

(2) $o_V = 0 \cdot v_1 + 0 \cdot v_2 + \dots + 0 \cdot v_k$, d.h. der Nullvektor o_V ist immer eine Linearkombination von beliebigen Vektoren v_1, v_2, \dots, v_k .

Betrachten wir nun zwei Beispiele.

Beispiel 12.13. (Linearkombination)

(a) Seien $V = \mathbb{R}^2$ und $\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ und $\vec{v}_3 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$.

Ist $\begin{bmatrix} 1 \\ 6 \end{bmatrix}$ eine Linearkombination von $\vec{v}_1, \vec{v}_2, \vec{v}_3$?

Wir suchen $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ mit

$$\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \lambda_3 \vec{v}_3 = \lambda_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \lambda_3 \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ 2\lambda_1 + \lambda_2 + 2\lambda_3 \end{bmatrix} \stackrel{!}{=} \begin{bmatrix} 1 \\ 6 \end{bmatrix},$$

d.h. wir suchen eine Lösung des linearen Gleichungssystems

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 2 & 1 & 2 & 6 \end{array} \right]$$

Dieses lineare Gleichungssystem ist lösbar, beispielsweise durch $\lambda_1 = 1$, $\lambda_2 = 4$, $\lambda_3 = 0$ oder durch $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 1$. Also ist die Antwort „ja“. Wir sehen auch, dass ein Vektor unter Umständen auf mehrere Weisen als Linearkombination von vorgegebenen Vektoren dargestellt werden kann.

(b) Sei $V = \mathbb{C}^3$ und $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ und $\vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$.

Ist $\begin{bmatrix} 1 \\ j \\ 0 \end{bmatrix}$ eine Linearkombination von \vec{v}_1, \vec{v}_2 ?

Die Antwort ist „nein“, denn man sieht direkt an

$$\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 = \lambda_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} \lambda_1 + \lambda_2 \\ 0 \\ 2\lambda_2 \end{bmatrix} \stackrel{!}{=} \begin{bmatrix} 1 \\ j \\ 0 \end{bmatrix},$$

dass $0 = j$ nicht erfüllbar ist.

Methode 12.14. (Feststellen, ob ein Vektor $\vec{x} \in \mathbb{K}^n$ eine Linearkombination der Vektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k \in \mathbb{K}^n$ ist)

Die *Linearkombination-Bedingung*

$$\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_k \vec{v}_k \stackrel{!}{=} \vec{x} \quad (12.3)$$

ist äquivalent zu dem linearen Gleichungssystem

$$\underbrace{\begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_k \end{bmatrix}}_{n \times k\text{-Matrix}} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{bmatrix} \stackrel{!}{=} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

wobei die i -te Spalte der Matrix gerade von dem Spaltenvektor \vec{v}_i gebildet wird. Schreiben wir dieses lineare Gleichungssystem mit der erweiterten Koeffizientenmatrix, so finden wir

$$\left[\vec{v}_1 \quad \vec{v}_2 \quad \dots \quad \vec{v}_k \mid \vec{x} \right].$$

Wenn dieses lineare Gleichungssystem lösbar ist, so sind alle Lösungen $\lambda_1, \lambda_2, \dots, \lambda_k$ gültige Koeffizienten $\lambda_1, \lambda_2, \dots, \lambda_k$ für (12.3).

Der nächste Hilfssatz untersucht die Natur der Menge aller Linearkombinationen von fest vorgegebenen Vektoren v_1, v_2, \dots, v_k .

Hilfssatz 12.15. (Die lineare Hülle ist ein Untervektorraum.)

Seien V ein \mathbb{K} -Vektorraum und $v_1, v_2, \dots, v_k \in V$. Die **Menge**

$$\text{LH}(v_1, v_2, \dots, v_k) := \left\{ \sum_{i=1}^k \lambda_i v_i = \lambda_1 v_1 + \dots + \lambda_k v_k : \lambda_1, \dots, \lambda_k \in \mathbb{K} \right\}$$

aller **Linearkombinationen** von v_1, v_2, \dots, v_k ist ein **Untervektorraum** von V und wird die **lineare Hülle** von v_1, v_2, \dots, v_k genannt. (Die lineare Hülle von v_1, v_2, \dots, v_k wird in manchen Büchern auch als der **Span** von v_1, v_2, \dots, v_k bezeichnet.)

Beweis von Hilfssatz 12.15: Wir überprüfen die drei Eigenschaften eines Untervektorraums:

(i) $o_V = 0 \cdot v_1 + 0 \cdot v_2 + \dots + 0 \cdot v_k \in \text{LH}(v_1, v_2, \dots, v_k)$.

(ii) Seien $x, y \in \text{LH}(v_1, v_2, \dots, v_k)$, also

$$x = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k, \quad y = \mu_1 v_1 + \mu_2 v_2 + \dots + \mu_k v_k.$$

Dann gilt nach den Distributivgesetzen

$$\begin{aligned} x + y &= (\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k) + (\mu_1 v_1 + \mu_2 v_2 + \dots + \mu_k v_k) \\ &= (\lambda_1 + \mu_1) v_1 + (\lambda_2 + \mu_2) v_2 + \dots + (\lambda_k + \mu_k) v_k \in \text{LH}(v_1, v_2, \dots, v_k). \end{aligned}$$

(iii) Seien $x = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k$ und $\alpha \in \mathbb{K}$. Dann gilt nach den Distributivgesetzen

$$\begin{aligned} \alpha x &= \alpha(\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k) \\ &= (\alpha \lambda_1) v_1 + (\alpha \lambda_2) v_2 + \dots + (\alpha \lambda_k) v_k \in \text{LH}(v_1, v_2, \dots, v_k). \end{aligned}$$

Also ist $\text{LH}(v_1, v_2, \dots, v_k)$ ein Untervektorraum von V . □

Nun untersuchen wir die lineare Hülle für einige Beispiele.

Beispiel 12.16. (lineare Hülle)

(a) In jedem \mathbb{K} -Vektorraum gilt $\text{LH}(o_V) = \{o_V\}$.

(b) Sei $V = \mathbb{R}^2$. Es gilt

$$\text{LH} \left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix} \right) = \mathbb{R}^2,$$

da sich jedes $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$ als Linearkombination von $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ und $\begin{bmatrix} 0 \\ 2 \end{bmatrix}$ schreiben lässt. Das lineare Gleichungssystem (vgl. auch Beispiel 12.13 (a))

$$\lambda_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \lambda_3 \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \Longleftrightarrow \quad \left[\begin{array}{ccc|c} 1 & 0 & 0 & x_1 \\ 2 & 1 & 2 & x_2 \end{array} \right]$$

ist nämlich für jedes $x_1, x_2 \in \mathbb{R}$ lösbar, z.B. durch $\lambda_1 = x_1$, $\lambda_2 = x_2 - 2x_1$, $\lambda_3 = 0$.

(c) Sei $V = \mathbb{C}^3$. Es gilt

$$\text{LH} \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \right) = \left\{ \begin{bmatrix} z_1 \\ 0 \\ z_3 \end{bmatrix} : z_1, z_3 \in \mathbb{C} \right\},$$

da das lineare Gleichungssystem

$$\lambda_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \iff \left[\begin{array}{cc|c} 1 & 1 & z_1 \\ 0 & 0 & z_2 \\ 0 & 2 & z_3 \end{array} \right] \xrightarrow{\substack{Z_1 \rightarrow Z_1 - \frac{1}{2}Z_3 \\ \text{dann: } Z_3 \rightarrow \frac{1}{2}Z_3 \\ Z_2 \leftrightarrow Z_3}} \left[\begin{array}{cc|c} 1 & 0 & z_1 - \frac{1}{2}z_3 \\ 0 & 1 & \frac{1}{2}z_3 \\ 0 & 0 & z_2 \end{array} \right]$$

genau dann lösbar ist, wenn $z_2 = 0$ ist.

(d) Seien $V = \mathbb{K}^n$ und

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \vec{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Dann gilt $\text{LH}(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n) = \mathbb{K}^n$, denn für jedes $\vec{x} \in \mathbb{K}^n$ gilt

$$x_1 \vec{e}_1 + x_2 \vec{e}_2 + \dots + x_n \vec{e}_n = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \vec{x}.$$

Mit unserem neuen Verständnis von Linearkombinationen und der linearen Hülle können wir nun auch die Lösung eines homogenen linearen Gleichungssystems als lineare Hülle geeigneter Vektoren schreiben. Betrachten wir dazu ein Beispiel.

Beispiel 12.17. (Lösungsmenge eines homogenen LGS als lineare Hülle)

Gegeben sei das reelle homogene lineare Gleichungssystem

$$\left[\begin{array}{cccc|c} 1 & 3 & 0 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \end{array} \right],$$

welches bereits in reduzierter Stufenform vorliegt. Wir setzen $x_2 = \lambda_1$ und $x_4 = \lambda_2$ und lesen ab

$$x_1 = -3x_2 - 2x_4 = -3\lambda_1 - 2\lambda_2,$$

$$x_2 = \lambda_1,$$

$$x_3 = x_4 = \lambda_2,$$

$$x_4 = \lambda_2,$$

d.h. die Lösungsmenge ist

$$\begin{aligned} \mathbb{L}_{[\mathbf{A}|\vec{0}]} &= \left\{ \begin{bmatrix} -3\lambda_1 - 2\lambda_2 \\ \lambda_1 \\ \lambda_2 \\ \lambda_2 \end{bmatrix} : \lambda_1, \lambda_2 \in \mathbb{R} \right\} \\ &= \left\{ \lambda_1 \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} -2 \\ 0 \\ 1 \\ 1 \end{bmatrix} : \lambda_1, \lambda_2 \in \mathbb{R} \right\} = \text{LH} \left(\begin{pmatrix} \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \\ 1 \end{bmatrix} \end{pmatrix} \right). \end{aligned}$$

Der nächste Satz liefert nützliche Informationen über die lineare Hülle.

Satz 12.18. (Invarianz der linearen Hülle)

Seien V ein \mathbb{K} -Vektorraum und $v_1, v_2, \dots, v_k \in V$. Die lineare Hülle $\text{LH}(v_1, v_2, \dots, v_k)$ ändert sich nicht bei

- (1) dem Vertauschen zweier Vektoren,
- (2) der Multiplikation eines Vektors mit einem $\lambda \in \mathbb{K} \setminus \{0\}$,
- (3) der Addition des μ -fachen eines Vektors zu einem anderen Vektor (wobei $\mu \in \mathbb{K}$).

Beweis von Satz 12.18: Diesen Satz beweisen wir in einer Übungsaufgabe. \square

Wir können diesen Satz nutzen, um die Darstellung der linearen Hülle zu vereinfachen. Dieses ist in dem nachfolgenden Beispiel demonstriert. Dabei verwenden wir eine **analoge Notation zu den elementaren Zeilenoperationen einer Matrix**: $v_i \leftrightarrow v_j$ deutet das Vertauschen der Vektoren v_i und v_j an, und $v_i \rightarrow \lambda v_i$ (mit $\lambda \neq 0$) bzw. $v_i \rightarrow v_i + \mu v_j$ (mit $j \neq i$) bedeuten, dass v_i mit $\lambda \neq 0$ multipliziert wird bzw. dass zu v_i das μ -fache des Vektors v_j addiert wird.

Beispiel 12.19. (Invarianz der linearen Hülle)

(a) Sei $V = \mathbb{R}^2$. Dann gilt:

$$\text{LH} \left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix} \right)$$

$$\begin{aligned}
&= \text{LH} \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix} \right) && \left(\text{wegen: } \begin{bmatrix} 1 \\ 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 \\ 2 \end{bmatrix} + (-1) \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) \\
&= \text{LH} \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) && \left(\text{wegen: } \begin{bmatrix} 0 \\ 2 \end{bmatrix} \rightarrow \frac{1}{2} \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) \\
&= \text{LH} \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) && \left(\text{doppelter Vektor wird weggelassen} \right) \\
&= \mathbb{R}^2 && \left(\text{vgl. Beispiel 12.16 (d)} \right)
\end{aligned}$$

(b) Sei $V = \mathcal{P}_3$. Dann gilt:

$$\begin{aligned}
&\text{LH}(t^2, 1 + t^2, 2 + 2t + 2t^3) \\
&= \text{LH}(t^2, 1, 2 + 2t + 2t^3) && \left(\text{da: } 1 + t^2 \rightarrow (1 + t^2) + (-1) \cdot t^2 \right) \\
&= \text{LH}(t^2, 1, 1 + t + t^3) && \left(\text{da: } 2 + 2t + 2t^3 \rightarrow \frac{1}{2} \cdot (2 + 2t + 2t^3) \right) \\
&= \text{LH}(t^2, 1, t + t^3) && \left(\text{da: } 1 + t + t^3 \rightarrow (1 + t + t^3) + (-1) \cdot 1 \right) \\
&= \text{LH}(1, t^2, t + t^3). && \left(\text{da: } 1 \leftrightarrow t^2 \right)
\end{aligned}$$

12.4 Lineare Unabhängigkeit

Wir beginnen mit der Definition des Begriffes „linear unabhängig“.

Definition 12.20. (linear unabhängig bzw. linear abhängig)

Sei V ein \mathbb{K} -Vektorraum. Die Vektoren $v_1, v_2, \dots, v_k \in V$ heißen **linear unabhängig** wenn gilt:

$$\sum_{i=1}^k \lambda_i v_i = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k = o_V \quad \implies \quad \lambda_1 = \lambda_2 = \dots = \lambda_k = 0.$$

Andernfalls heißen v_1, v_2, \dots, v_k **linear abhängig**.

In der nachfolgenden Bemerkung halten wir einige wichtige Beobachtungen fest. Danach betrachten wir verschiedene Beispiele.

Bemerkung 12.21. (linear unabhängig bzw. linear abhängig)

- (1) Ist mindestens einer der Vektoren v_1, v_2, \dots, v_k der Nullvektor o_V , so sind v_1, v_2, \dots, v_k **linear abhängig**.
- (2) Sonderfall $k = 1$: v_1 ist **linear abhängig**. $\implies v_1 = o_V$.
- (3) Sonderfall $k = 2$: v_1, v_2 sind genau dann **linear abhängig**, wenn gilt: Es existiert $\lambda \in \mathbb{K}$ mit $v_1 = \lambda v_2$, oder es existiert $\mu \in \mathbb{K}$ mit $v_2 = \mu v_1$.
- (4) v_1, v_2, \dots, v_k sind genau dann **linear abhängig**, wenn es Koeffizienten $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{K}$ gibt, die nicht alle gleich null sind und

$$\sum_{i=1}^k \lambda_i v_i = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k = o_V$$

erfüllen.

- (5) Lässt sich für Vektoren v_1, v_2, \dots, v_k einer dieser Vektoren als Linearkombination der anderen darstellen, so sind die Vektoren **linear abhängig**.

Erklärung: Ist beispielsweise der Vektor v_k ein Linearkombination von v_1, v_2, \dots, v_{k-1} , so gibt es Koeffizienten $\lambda_1, \lambda_2, \dots, \lambda_{k-1} \in \mathbb{K}$ mit

$$v_k = \sum_{i=1}^{k-1} \lambda_i v_i \iff \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_{k-1} v_{k-1} + (-1) \cdot v_k = o_V,$$

und wir sehen, dass $v_1, v_2, \dots, v_{k-1}, v_k$ linear abhängig sind.

- (6) Sind v_1, v_2, \dots, v_k **linear unabhängig** und gilt $x \notin \text{LH}(v_1, v_2, \dots, v_k)$, so sind die Vektoren v_1, v_2, \dots, v_k, x **linear unabhängig**.

Gilt $x \in \text{LH}(v_1, v_2, \dots, v_k)$, so sind v_1, v_2, \dots, v_k, x **linear abhängig**. (Hierbei können v_1, v_2, \dots, v_k linear unabhängig oder linear abhängig sein.)

Beweis: Die zweite Aussage folgt direkt aus (5). – Betrachten wir nun den Fall, dass v_1, v_2, \dots, v_k linear unabhängig sind und dass gilt $x \notin \text{LH}(v_1, v_2, \dots, v_k)$. Aus

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k + \lambda_{k+1} x = o_V \tag{12.4}$$

folgt dann $\lambda_{k+1} = 0$, denn sonst wäre (durch Auflösen nach x) nämlich $x \in \text{LH}(v_1, v_2, \dots, v_k)$. Damit reduziert sich (12.4) zu

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k = o_V,$$

und wegen der linearen Unabhängigkeit von v_1, v_2, \dots, v_k folgt $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$. Also gilt $\lambda_1 = \lambda_2 = \dots = \lambda_k = \lambda_{k+1} = 0$, d.h. v_1, v_2, \dots, v_k, x sind linear unabhängig. \square

- (7) Sind die Vektoren v_1, v_2, \dots, v_k **linear unabhängig**, so besteht jede Teilmenge dieser Vektoren ebenfalls aus **linear unabhängigen** Vektoren. – Fügt man dagegen zu linear unabhängigen Vektoren v_1, v_2, \dots, v_k weitere Vektoren hinzu, so kann das neue System aus Vektoren linear unabhängig oder linear abhängig sein.

Erklärung: Betrachten wir beispielsweise v_1, v_2, \dots, v_m mit $m < k$. Wenn v_1, v_2, \dots, v_m linear abhängig wären, würde die Gleichung

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_m v_m = o_V$$

mit Koeffizienten $\lambda_1, \lambda_2, \dots, \lambda_m$, die nicht alle gleich null sind, gelten. Dann würde aber mit dieser Wahl der Koeffizienten $\lambda_1, \lambda_2, \dots, \lambda_m$ und mit $\lambda_{m+1} = \dots = \lambda_k = 0$ auch

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_m v_m + \dots + \lambda_k v_k = o_V$$

gelten. Dieses steht im Widerspruch zur linearen Unabhängigkeit von v_1, v_2, \dots, v_k . Also müssen auch v_1, v_2, \dots, v_m linear unabhängig sein.

- (8) Sind die Vektoren $v_1, v_2, \dots, v_k \in V$ **linear abhängig** und fügt man weitere Vektoren $w_1, w_2, \dots, w_m \in V$ hinzu, so sind die Vektoren $v_1, v_2, \dots, v_k, w_1, w_2, \dots, w_m$ ebenfalls **linear abhängig**.

Erklärung: Weil v_1, v_2, \dots, v_k linear abhängig sind, gibt es Koeffizienten $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{K}$, die nicht alle gleich null sind, mit

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k = o_V.$$

Mit dieser Wahl von $\lambda_1, \lambda_2, \dots, \lambda_k$ nicht alle gleich null und mit $\mu_1 = \mu_2 = \dots = \mu_m = 0$ gilt dann

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k + \mu_1 w_1 + \mu_2 w_2 + \dots + \mu_m w_m = o_V.$$

Also sind $v_1, v_2, \dots, v_k, w_1, w_2, \dots, w_m$ linear abhängig.

Beispiel 12.22. (linear unabhängig bzw. linear abhängig)

- (a) Seien $V = \mathbb{R}^2$, $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ und $\vec{v}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$.

Sind \vec{v}_1, \vec{v}_2 linear unabhängig?

Wir betrachten

$$\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 = \lambda_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} \lambda_1 + 2\lambda_2 \\ -\lambda_2 \end{bmatrix} = \vec{0}.$$

Das homogene lineare Gleichungssystem ist

$$\left[\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & -1 & 0 \end{array} \right] \xrightarrow{Z_2 \leftrightarrow -Z_2} \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 1 & 0 \end{array} \right],$$

welches nun in Stufenform ist, und wir lesen ab: $\lambda_2 = 0$ und $\lambda_1 = -2\lambda_2 = 0$. Also sind \vec{v}_1, \vec{v}_2 linear unabhängig.

(b) Seien $V = \mathbb{C}^3$ und $\vec{v}_1 = \begin{bmatrix} 1 \\ -j \\ 0 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} j \\ -1 \\ 0 \end{bmatrix}$, $\vec{v}_3 = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$.

Sind die Vektoren $\vec{v}_1, \vec{v}_2, \vec{v}_3$ linear unabhängig oder linear abhängig?

Wir betrachten

$$\begin{aligned} \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \lambda_3 \vec{v}_3 &= \lambda_1 \begin{bmatrix} 1 \\ -j \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} j \\ -1 \\ 0 \end{bmatrix} + \lambda_3 \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 + j\lambda_2 + 2\lambda_3 \\ -j\lambda_1 - \lambda_2 \\ 0 \end{bmatrix} = \vec{0}, \end{aligned}$$

d.h. wir erhalten das homogene lineare Gleichungssystem

$$\left[\begin{array}{ccc|c} 1 & j & 2 & 0 \\ -j & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right].$$

Dieses hat nur zwei nicht-triviale Gleichungen aber drei Unbekannte. Daher muss es unendlich viele Lösungen haben, und insbesondere gibt es Lösungen $\lambda_1, \lambda_2, \lambda_3$, die nicht alle gleich null sind. Also sind die Vektoren $\vec{v}_1, \vec{v}_2, \vec{v}_3$ in \mathbb{C}^3 linear abhängig.

Alternative Antwort: Wir beobachten, dass

$$1 \vec{v}_1 + (-j) \vec{v}_2 + (-1) \vec{v}_3 = \begin{bmatrix} 1 \\ -j \\ 0 \end{bmatrix} + \begin{bmatrix} -j^2 \\ j \\ 0 \end{bmatrix} + \begin{bmatrix} -2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 + 1 - 2 \\ -j + j \\ 0 \end{bmatrix} = \vec{0}$$

gilt. Also sind die Vektoren $\vec{v}_1, \vec{v}_2, \vec{v}_3$ in \mathbb{C}^3 linear abhängig.

Methode 12.23. (Vorgehensweise beim Untersuchen von Vektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ in \mathbb{K}^n auf lineare Unabhängigkeit)

Wir halten noch einmal allgemein fest, wir **im Fall** $V = \mathbb{K}^n$ **vorgehen**, wenn wir Vektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ in \mathbb{K}^n **auf lineare Unabhängigkeit bzw. lineare Abhängigkeit überprüfen** wollen:

Wir müssen die Gleichung

$$\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_k \vec{v}_k = \vec{0} \quad (12.5)$$

untersuchen und alle ihre Lösungen $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{K}$ bestimmen.

- Finden wir nur die einzige Lösung $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$, so sind $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ **linear unabhängig**.
- Gibt es dagegen mindestens eine Lösung, bei der nicht alle $\lambda_1, \lambda_2, \dots, \lambda_k$ gleich null sind, so sind $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ **linear abhängig**.

Die Gleichung (12.5) ist ein lineares Gleichungssystem mit den Unbekannten $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{K}$, dessen erweiterte Koeffizientenmatrix durch

$$[\vec{v}_1 \ \vec{v}_2 \ \dots \ \vec{v}_k \mid \vec{0}]$$

gegeben ist. Dabei ist die erweiterte Koeffizientenmatrix so zu lesen, dass ihre Spaltenvektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ sind und $\vec{0} \in \mathbb{K}^n$ auf der rechten Seite steht.

Wir betrachten nun noch einige weitere Beispiele.

Beispiel 12.24. (linear unabhängig bzw. linear abhängig)

(a) Seien $V = \mathbb{K}^n$ und

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \vec{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Dann sind $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ linear unabhängig, denn

$$\sum_{i=1}^n \lambda_i \vec{e}_i = \vec{0} \quad \iff \quad [\vec{e}_1 \ \vec{e}_2 \ \dots \ \vec{e}_n \mid \vec{0}]$$

$$\iff \left[\begin{array}{cccc|cc} 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & \ddots & \vdots & \vdots & 0 \\ 0 & 0 & \ddots & 0 & 0 & \vdots \\ \vdots & \vdots & \ddots & 1 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 \end{array} \right],$$

also $\lambda_n = 0, \lambda_{n-1} = 0, \dots, \lambda_2 = 0$ und $\lambda_1 = 0$.

- (b) Sie $V = \mathcal{P}_n$. Dann sind die Monome $p_0(t) := 1, p_1(t) := t, p_2(t) := t^2, \dots, p_n(t) := t^n$ linear unabhängig, denn die Gleichung

$$\begin{aligned} & \lambda_0 p_0(t) + \lambda_1 p_1(t) + \lambda_2 p_2(t) + \dots + \lambda_n p_n(t) \\ & = \lambda_0 \cdot 1 + \lambda_1 \cdot t + \lambda_2 \cdot t^2 + \dots + \lambda_n \cdot t^n = o(t) = 0 \quad \text{für alle } t \in \mathbb{K}, \end{aligned} \tag{12.6}$$

wobei $o : \mathbb{K} \rightarrow \mathbb{K}, o(t) := 0$, die Nullfunktion ist, hat nur die Lösung $\lambda_0 = \lambda_1 = \lambda_2 = \dots = \lambda_n = 0$. (*Erklärung:* Setze in (12.6) $t = 0$ ein. Dann folgt $\lambda_0 = 0$. Ableiten von (12.6) nach t liefert $\lambda_1 + 2\lambda_2 t + \dots + n\lambda_n t^{n-1} = 0$. Danach $t = 0$ Setzen liefert $\lambda_1 = 0$. Erneutes (also insgesamt zweifaches) Ableiten von (12.6) nach t liefert $2\lambda_2 + 6\lambda_3 t + \dots + n(n-1)\lambda_n t^{n-2} = 0$. Danach $t = 0$ Setzen liefert $\lambda_2 = 0$. Führt man so fort, so findet man, dass alle $\lambda_i, i = 0, 1, 2, \dots, n$, den Wert null haben.)

- (c) Sei $V = \mathcal{C}(\mathbb{R})$. Die Funktionen $f_1(t) := \sin(t)$ und $f_2(t) := \cos(t)$ sind linear unabhängig, denn:

Wir betrachten die Gleichung

$$\lambda_1 \sin(t) + \lambda_2 \cos(t) = 0 \quad \text{für alle } t \in \mathbb{R}.$$

Wir setzen nun $t = 0$ bzw. $t = \pi/2$ ein und erhalten:

$$\begin{aligned} \lambda_1 \sin(0) + \lambda_2 \cos(0) = 0 & \iff \lambda_2 = 0, \\ \lambda_1 \sin\left(\frac{\pi}{2}\right) + \lambda_2 \cos\left(\frac{\pi}{2}\right) = 0 & \iff \lambda_1 = 0. \end{aligned}$$

Also muss gelten $\lambda_1 = \lambda_2 = 0$, und $f_1(t) = \sin(t)$ und $f_2(t) = \cos(t)$ sind linear unabhängig.

12.5 Basis und Dimension

Aufbauend auf die wichtigen Begriffe der linearen Unabhängigkeit und der Dimension lernen wir nun die Begriffe einer Basis eines Vektorraums und der Dimension eines Vektorraums kennen.

Definition 12.25. (Basis)

Sei V ein \mathbb{K} -Vektorraum, der nicht nur aus o_V besteht. Ein System $B = (b_1, b_2, \dots, b_n)$ von Vektoren $b_1, b_2, \dots, b_n \in V$ heißt eine **Basis von V** , wenn gilt:

- (i) b_1, b_2, \dots, b_n sind linear unabhängig, und
- (ii) $V = \text{LH}(b_1, b_2, \dots, b_n)$.

Betrachten wir zunächst einige Beispiele.

Beispiel 12.26. (Basis)

- (a) Seien $V = \mathbb{K}^n$ und

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \vec{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Aus Beispiel 12.24 (a) wissen wir, dass $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ linear unabhängig sind, und aus Beispiel 12.16 (d) wissen wir, dass $\text{LH}(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n) = \mathbb{K}^n$. Also ist $E := (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ eine Basis von \mathbb{K}^n . Wir nennen $E := (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ die **Standardbasis von \mathbb{K}^n** .

- (b) Seien $V = \mathbb{R}^2$ und $B = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right)$ und $\tilde{B} = \left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)$.

Aus dem vorigen Beispiel wissen wir bereits, dass B eine Basis für \mathbb{R}^2 ist.

Wir zeigen nun, dass auch \tilde{B} eine Basis für \mathbb{R}^2 ist:

$$\begin{aligned} \lambda_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} &\iff \begin{bmatrix} \lambda_1 + \lambda_2 \\ -\lambda_1 + \lambda_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \iff \begin{bmatrix} 1 & 1 & | & 0 \\ -1 & 1 & | & 0 \end{bmatrix} &\xrightarrow{Z_2 \rightarrow Z_2 + Z_1} \begin{bmatrix} 1 & 1 & | & 0 \\ 0 & 2 & | & 0 \end{bmatrix} \\ \iff \begin{bmatrix} 1 & 1 & | & 0 \\ 0 & 1 & | & 0 \end{bmatrix} &\xrightarrow{Z_1 \rightarrow Z_1 - Z_2} \begin{bmatrix} 1 & 0 & | & 0 \\ 0 & 1 & | & 0 \end{bmatrix}. \end{aligned}$$

Also folgt $\lambda_1 = \lambda_2 = 0$, d.h. $\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ sind linear unabhängig.

Mit einer analogen Rechnung folgt

$$\begin{aligned} \lambda_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &\iff \begin{bmatrix} \lambda_1 + \lambda_2 \\ -\lambda_1 + \lambda_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ \iff \left[\begin{array}{cc|c} 1 & 1 & x_1 \\ -1 & 1 & x_2 \end{array} \right] &\xrightarrow{Z_2 \rightarrow Z_2 + Z_1} \left[\begin{array}{cc|c} 1 & 1 & x_1 \\ 0 & 2 & x_1 + x_2 \end{array} \right] \\ \xrightarrow{Z_2 \rightarrow \frac{1}{2}Z_2} \left[\begin{array}{cc|c} 1 & 1 & x_1 \\ 0 & 1 & \frac{1}{2}(x_1 + x_2) \end{array} \right] &\xrightarrow{Z_1 \rightarrow Z_1 - Z_2} \left[\begin{array}{cc|c} 1 & 0 & \frac{1}{2}(x_1 - x_2) \\ 0 & 1 & \frac{1}{2}(x_1 + x_2) \end{array} \right], \end{aligned}$$

d.h. $\lambda_1 = \frac{1}{2}(x_1 - x_2)$ und $\lambda_2 = \frac{1}{2}(x_1 + x_2)$.

Also lässt sich jeder Vektor $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$ als Linearkombination von $\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

schreiben, d.h. $\text{LH} \left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) = \mathbb{R}^2$. Also ist \tilde{B} ebenfalls eine Basis von \mathbb{R}^2 .

(c) Sei $V = \mathbb{R}^2$. $\left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix} \right)$ ist **keine** Basis für \mathbb{R}^2 , denn:

Die Vektoren $\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix}$ sind linear abhängig, da gilt

$$0 \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} + (-2) \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} + 1 \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

(d) Sei $V = \mathbb{C}^3$. $\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \right)$ ist **keine** Basis für \mathbb{C}^3 , denn aus Beispiel

12.16 (c)

wissen wir, dass gilt

$$\text{LH} \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \right) = \left\{ \begin{bmatrix} z_1 \\ 0 \\ z_3 \end{bmatrix} : z_1, z_3 \in \mathbb{C} \right\} \neq \mathbb{C}^3.$$

(e) Sei $V = \mathcal{P}_3$. $(1, t, t^2, t^3)$ ist eine Basis von \mathcal{P}_3 , denn:

Aus Beispiel 12.24 (b) wissen wir, dass die Polynomfunktionen $1, t, t^2, t^3$ linear unabhängig sind.

Weiter gilt $\text{LH}(1, t, t^2, t^3) = \mathcal{P}_3$, da sich jedes Polynom p vom Grad ≤ 3 als $p(t) = \lambda_0 + \lambda_1 t + \lambda_2 t^2 + \lambda_3 t^3$ schreiben lässt.

(f) Sei $V = \mathbb{K}^{2 \times 3}$. Dann ist

$$B = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right)$$

eine Basis von $\mathbb{K}^{2 \times 3}$. Zur Übung sollten Sie dieses nachweisen.

An Beispiel 12.26 (b) sehen wir bereits, dass ein Vektorraum mehrere Basen haben kann.

Wir lernen nun diverse Sätze über Basen („Basen“ ist der Plural von „Basis“) kennen, die uns erlauben das Konzept der Basis eines Vektorraums besser zu verstehen. Diese Sätze werden uns auch dabei helfen, für einen gegebenen Vektorraum eine Basis zu finden.

Folgerung 12.27. (eindeutige Darstellung von Vektoren bzgl. Basis)

Sei V ein \mathbb{K} -Vektorraum, der nicht nur aus o_V besteht. Sei (b_1, b_2, \dots, b_k) eine Basis von V . Dann hat jeder Vektor $x \in V$ eine Darstellung

$$x = \sum_{j=1}^k c_j b_j = c_1 b_1 + c_2 b_2 + \dots + c_k b_k$$

mit **eindeutig bestimmten Koeffizienten** c_1, c_2, \dots, c_k .

Beweis von Folgerung 12.27: Weil (b_1, b_2, \dots, b_k) eine Basis ist, gilt (nach der Definition einer Basis) $\text{LH}(b_1, b_2, \dots, b_k) = V$, d.h. jedes $x \in V$ hat eine Darstellung der Form

$$x = c_1 b_1 + c_2 b_2 + \dots + c_k b_k. \quad (12.7)$$

Um zu zeigen, dass diese Darstellung eindeutig ist, nehmen wir an, dass

$$x = d_1 b_1 + d_2 b_2 + \dots + d_k b_k$$

eine weitere Darstellung dieser Form ist. Subtrahieren der beiden Gleichungen liefert

$$o_V = (d_1 - c_1) b_1 + (d_2 - c_2) b_2 + \dots + (d_k - c_k) b_k. \quad (12.8)$$

Weil (b_1, b_2, \dots, b_k) eine Basis ist, sind die Vektoren b_1, b_2, \dots, b_k linear unabhängig. Also folgt aus (12.8), dass

$$\begin{aligned} d_1 - c_1 = 0, \quad d_2 - c_2 = 0, \quad \dots, \quad d_k - c_k = 0 \\ \iff d_1 = c_1, \quad d_2 = c_2, \quad \dots, \quad d_k = c_k, \end{aligned}$$

d.h. die Koeffizienten in (12.7) sind eindeutig bestimmt. \square

Satz 12.28. (Existenz einer Basis)

Sei V ein \mathbb{K} -Vektorraum, der nicht nur aus o_V besteht. Es seien $W = \{w_1, w_2, \dots, w_N\} \subseteq V$ und $V = \text{LH}(w_1, w_2, \dots, w_N)$. Sei U eine nicht-leere Teilmenge von W , die aus **linear unabhängigen** Vektoren besteht. Dann gibt es eine **Basis** $B = (b_1, b_2, \dots, b_n)$ von V mit

$$U \subseteq \{b_1, b_2, \dots, b_n\} \subseteq W.$$

Unter den obigen Voraussetzungen an V und W kann jede Teilmenge von W , die aus linear unabhängigen Vektoren besteht, durch Hinzunahme weiterer geeigneter Vektoren aus W zu einer Basis von V ergänzt werden.

Beweis von Satz 12.28: Unter allen Mengen X mit den Eigenschaften

- (i) $U \subseteq X \subseteq W$, und
- (ii) X besteht aus linear unabhängigen Vektoren,

wählen wir eine mit maximaler Elementezahl aus und nennen ihre Elemente b_1, b_2, \dots, b_n . Dann sind b_1, b_2, \dots, b_n nach Konstruktion linear unabhängig. Wir wollen nun zeigen, dass $\text{LH}(b_1, b_2, \dots, b_n) = V$, so, dass $B := (b_1, b_2, \dots, b_n)$ die gesuchte Basis aus dem Satz ist.

Als Vorbereitungsschritt zum Nachweis von $\text{LH}(b_1, b_2, \dots, b_n) = V$ zeigen wir mit einem *Widerspruchsbeweis*, dass $W \subseteq \text{LH}(b_1, b_2, \dots, b_n)$ gilt:

Angenommen $W \subseteq \text{LH}(b_1, b_2, \dots, b_n)$ gelte **nicht**. Dann gibt es ein $b_0 \in W$ mit $b_0 \notin \text{LH}(b_1, b_2, \dots, b_n)$. Dann gilt $X := \{b_0, b_1, \dots, b_n\} \subseteq W$. Da die Menge $X = \{b_0, b_1, \dots, b_n\}$ aber $n + 1$ Elemente hat, kann sie nicht zu den Mengen X mit Eigenschaften (i) und (ii) gehören. Da $X = \{b_0, b_1, \dots, b_n\}$ (i) erfüllt, muss (ii) verletzt sein, d.h. b_0, b_1, \dots, b_n sind linear abhängig.

Betrachten wir nun Gleichung

$$\lambda_0 b_0 + \lambda_1 b_1 + \dots + \lambda_n b_n = o_V, \tag{12.9}$$

so folgt $\lambda_0 = 0$, denn sonst wäre $b_0 \in \text{LH}(b_1, b_2, \dots, b_n)$. Damit reduziert sich (12.9) zu

$$\lambda_1 b_1 + \dots + \lambda_n b_n = o_V,$$

und wegen der linearen Unabhängigkeit von b_1, b_2, \dots, b_n folgt, dass $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$ ist. Also sind $\lambda_0 = \lambda_1 = \lambda_2 = \dots = \lambda_n = 0$, und b_0, b_1, \dots, b_n sind linear unabhängig. ∇ Dieses ist ein Widerspruch zu der linearen Abhängigkeit von b_0, b_1, \dots, b_n .

Also war die Annahme, dass $W \subseteq \text{LH}(b_1, b_2, \dots, b_n)$ nicht gelte, falsch, und wir haben gezeigt, dass $W = \{w_1, w_2, \dots, w_N\} \subseteq \text{LH}(b_1, b_2, \dots, b_n)$ gilt. Daraus folgt

$$V = \text{LH}(w_1, w_2, \dots, w_N) \subseteq \text{LH}(b_1, b_2, \dots, b_n).$$

Da $\text{LH}(b_1, b_2, \dots, b_n)$ ein Untervektorraum von V ist, gilt $\text{LH}(b_1, b_2, \dots, b_n) \subseteq V$. Aus den beiden Teilmengenbeziehungen $V \subseteq \text{LH}(b_1, b_2, \dots, b_n)$ und $\text{LH}(b_1, b_2, \dots, b_n) \subseteq V$ folgt $\text{LH}(b_1, b_2, \dots, b_n) = V$. \square

Die folgende Bemerkung erläutert Satz 12.28.

Bemerkung 12.29. (endlich erzeugt, Basis, und lineare Hülle)

- (1) Falls endlich viele Vektoren $w_1, w_2, \dots, w_N \in V$ existieren mit $\text{LH}(w_1, w_2, \dots, w_N) = V$, so nennen wir V **endlich erzeugt**.
- (2) Hat V eine Basis (b_1, b_2, \dots, b_n) , so ist V endlich erzeugt, denn es gilt $V = \text{LH}(b_1, b_2, \dots, b_n)$.
- (3) Ist umgekehrt V endlich erzeugt, so folgt aus Satz 12.28, dass V eine Basis (b_1, b_2, \dots, b_n) besitzt.

Beispiele: \mathbb{K}^n und \mathcal{P}_n sind endlich erzeugt. Der \mathbb{R} -Vektorraum $\mathcal{C}(\mathbb{R})$ der stetigen Funktionen und der \mathbb{R} -Vektorraum \mathcal{P} der reellen Polynomfunktionen sind **nicht** endlich erzeugt.

Folgerung 12.30. (Basisergänzungssatz)

Sei V ein \mathbb{K} -Vektorraum, der nicht nur aus o_V besteht. Ist V **endlich erzeugt** und sind $b_1, b_2, \dots, b_k \in V$ **linear unabhängig**, so lässt sich (b_1, b_2, \dots, b_k) zu einer Basis (b_1, b_2, \dots, b_n) von V ergänzen.

Beweis von Folgerung 12.30: Da V endlich erzeugt ist, existieren w_1, w_2, \dots, w_n

mit $V = \text{LH}(w_1, w_2, \dots, w_n)$. Seien $U := \{b_1, b_2, \dots, b_k\}$ und

$$W := U \cup \{w_1, w_2, \dots, w_n\} = \{b_1, b_2, \dots, b_k, w_1, w_2, \dots, w_n\}.$$

Dann gelten $\text{LH}(b_1, b_2, \dots, b_k, w_1, w_2, \dots, w_n) = V$ und $U \subseteq W$. Aus Satz 12.28 folgt die Behauptung. \square

Satz 12.31. (mehr Vektoren als in einer Basis sind linear abhängig)

Sei V ein \mathbb{K} -Vektorraum, der nicht nur aus o_V besteht. Seien (b_1, b_2, \dots, b_n) eine Basis von V , $N > n$ und $v_1, v_2, \dots, v_N \in V$. Dann sind v_1, v_2, \dots, v_N linear abhängig.

Beweis von Satz 12.31: Da $V = \text{LH}(b_1, b_2, \dots, b_n)$ gilt und $v_1, v_2, \dots, v_N \in V$ sind, gibt es Koeffizienten $\alpha_{i,k}$, $i = 1, 2, \dots, n$; $k = 1, 2, \dots, N$, mit

$$v_k = \sum_{i=1}^n \alpha_{i,k} b_i, \quad k = 1, 2, \dots, N. \quad (12.10)$$

Seien nun $\lambda_1, \lambda_2, \dots, \lambda_N \in \mathbb{K}$ mit

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_N v_N = o_V. \quad (12.11)$$

Wir setzen nun (12.10) in (12.11) ein:

$$\begin{aligned} o_V &= \lambda_1 \sum_{i=1}^n \alpha_{i,1} b_i + \lambda_2 \sum_{i=1}^n \alpha_{i,2} b_i + \dots + \lambda_N \sum_{i=1}^n \alpha_{i,N} b_i \\ &= \sum_{i=1}^n (\alpha_{i,1} \lambda_1 + \alpha_{i,2} \lambda_2 + \dots + \alpha_{i,N} \lambda_N) b_i. \end{aligned}$$

Da b_1, b_2, \dots, b_n linear unabhängig sind, folgt daraus

$$\alpha_{i,1} \lambda_1 + \alpha_{i,2} \lambda_2 + \dots + \alpha_{i,N} \lambda_N = 0 \quad \text{für alle } i = 1, 2, \dots, n.$$

Dieses ist ein homogenes lineares Gleichungssystem mit n Gleichungen und den N Unbekannten $\lambda_1, \lambda_2, \dots, \lambda_N$. Da $N > n$, hat das lineare Gleichungssystem unendlich viele Lösungen, d.h. die Gleichung (12.11) hat unendlich viele Lösungen. Es folgt, dass v_1, v_2, \dots, v_N linear abhängig sind. \square

Aus den vorhergegangenen Sätzen ziehen wir zwei wichtige Folgerungen.

Folgerung 12.32. (Dimension eines Vektorraums)

Sei V ein \mathbb{K} -Vektorraum, der nicht nur aus o_V besteht. V sei endlich erzeugt, und es seien (b_1, b_2, \dots, b_n) und $(\widetilde{b}_1, \widetilde{b}_2, \dots, \widetilde{b}_m)$ Basen von V . Dann gilt $m = n$. n heißt die **Dimension von V** , und wir schreiben $\dim(V) := n$.

Beweis von Folgerung 12.32: Wir nehmen an, dass gelte $m > n$ und führen dieses zu einem Widerspruch: Ist $m > n$, so sind nach Satz 12.31 die m Vektoren $\widetilde{b}_1, \widetilde{b}_2, \dots, \widetilde{b}_m$ linear abhängig. Dies ist ein Widerspruch dazu, dass $(\widetilde{b}_1, \widetilde{b}_2, \dots, \widetilde{b}_m)$ eine Basis von V ist. Analog führt man die Annahme $m < n$ zu einem Widerspruch. Also muss gelten $m = n$. \square

Bemerkung 12.33. (Dimension)

- (1) Um Fallunterscheidungen zu vermeiden, definieren wir zusätzlich $\dim(\{o_V\}) := 0$ für den Fall $V = \{o_V\}$.
- (2) Ist ein Vektorraum V endlich erzeugt, so nennen wir V **endlich dimensional**.
- (3) Gilt für einen Vektorraum V , dass $\dim(V) = n$ ist, so nennen wir V **n -dimensional**. Insbesondere ist jeder n -dimensionale Vektorraum auch **endlich dimensional**.
- (4) Ist ein Vektorraum V **nicht** endlich erzeugt, so sagen wir, V ist **unendlich dimensional**, und schreiben $\dim(V) = \infty$.

Folgerung 12.34. ($\dim(V) = n \Rightarrow n$ lin. unabh. Vektoren bilden Basis)

Ist V ein endlich dimensionaler \mathbb{K} -Vektorraum mit $\dim(V) = n$, so bilden n **linear unabhängige Vektoren immer eine Basis** von V .

Beweis von Folgerung 12.34: Seien v_1, v_2, \dots, v_n beliebige linear unabhängige Vektoren in V . Angenommen (v_1, v_2, \dots, v_n) wäre keine Basis von V . Dann gäbe es einen Vektor w in V , der nicht als Linearkombination von v_1, v_2, \dots, v_n darstellbar ist. Damit hätten wir $n + 1$ linear unabhängige Vektoren v_1, v_2, \dots, v_n, w in einem n -dimensionalen \mathbb{K} -Vektorraum. Dies ist ein Widerspruch zu Satz 12.31. Also war die Annahme falsch, und (v_1, v_2, \dots, v_n) muss eine Basis von V sein. \square

Betrachten wir nun einige weitere Beispiele.

Beispiel 12.35. (Dimension)

- (a) In Beispiel 12.26 (a) haben wir gesehen, dass die Standardbasis $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ eine Basis von \mathbb{K}^n ist. Also gilt $\dim(\mathbb{K}^n) = n$.
- (b) Analog zu Beispiel 12.26 (e) zeigt man, dass $(1, t, t^2, \dots, t^n)$ eine Basis für den Vektorraum \mathcal{P}_n der Polynomfunktionen vom Grad $\leq n$ ist. Damit folgt $\dim(\mathcal{P}_n) = n + 1$.
- (c) In Beispiel 12.26 (f) haben wir eine Basis für $\mathbb{K}^{2 \times 3}$ angegeben und sehen somit, dass $\dim(\mathbb{K}^{2 \times 3}) = 6 = 2 \cdot 3$.

In einer Verallgemeinerung von Beispiel 12.26 (f) überlegt man sich leicht, dass $\dim(\mathbb{K}^{m \times n}) = m \cdot n$ gilt.

- (d) In einer Verallgemeinerung von (b) sieht man, dass $p_i(t) = t^i, i = 1, 2, \dots, n$, für jedes $n \in \mathbb{N}$ in dem Vektorraum \mathcal{P} aller Polynomfunktionen linear unabhängig sind, d.h es gibt beliebig viele linear unabhängige Polynomfunktionen. Also ist der Vektorraum \mathcal{P} nicht endlich erzeugt, und es gilt $\dim(\mathcal{P}) = \infty$.
- (e) Da alle (reellen) Polynomfunktionen stetig sind gilt $\mathcal{P} \subseteq \mathcal{C}(\mathbb{R})$ und \mathcal{P} ist ein Untervektorraum von $\mathcal{C}(\mathbb{R})$. Da nach (d) $\dim(\mathcal{P}) = \infty$ ist, muss auch $\dim(\mathcal{C}(\mathbb{R})) = \infty$ gelten.

12.6 Koordinaten

In diesem Teilkapitel interessieren wir uns dafür, wie sich Vektoren in einem endlich erzeugten Vektorraum als Linearkombination der Vektoren einer Basis darstellen lassen.

Satz 12.36. (Darstellung eines Vektors bzgl. Basis: Koordinaten)

Sei V ein \mathbb{K} -Vektorraum mit $\dim(V) = n$, und sei $B = (b_1, b_2, \dots, b_n)$ eine Basis von V . Dann hat jeder Vektor $v \in V$ eine **eindeutige Darstellung**

$$v = v_1 b_1 + v_2 b_2 + \dots + v_n b_n = \sum_{i=1}^n v_i b_i \quad (12.12)$$

mit **eindeutig bestimmten** $v_1, v_2, \dots, v_n \in \mathbb{K}$. Man nennt v_1, v_2, \dots, v_n die

Koordinaten von v bzgl. der Basis B . Der Spaltenvektor

$$\vec{v}_B := \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

(in \mathbb{K}^n) heißt der **Koordinatenvektor von v bzgl. B** .

Beweis von Satz 12.36: Die Existenz von $v_1, v_2, \dots, v_n \in \mathbb{K}$ folgt aus der Tatsache, dass (b_1, b_2, \dots, b_n) eine Basis von V ist und somit $\text{LH}(b_1, b_2, \dots, b_n) = V$ gilt.

Die Eindeutigkeit der Koeffizienten v_1, v_2, \dots, v_n folgt aus der linearen Unabhängigkeit von b_1, b_2, \dots, b_n , wie jetzt genauer ausgeführt wird:

Wir nehmen an, es gebe neben (12.12) eine weitere Darstellung

$$v = \tilde{v}_1 b_1 + \tilde{v}_2 b_2 + \dots + \tilde{v}_n b_n. \quad (12.13)$$

Gleichsetzen von (12.13) und (12.12) liefert dann

$$\tilde{v}_1 b_1 + \tilde{v}_2 b_2 + \dots + \tilde{v}_n b_n = v_1 b_1 + v_2 b_2 + \dots + v_n b_n,$$

und auf eine Seite Bringen und Sortieren ergibt:

$$\begin{aligned} & (v_1 - \tilde{v}_1) b_1 + (v_2 - \tilde{v}_2) b_2 + \dots + (v_n - \tilde{v}_n) b_n = 0 \\ \implies & v_1 - \tilde{v}_1 = v_2 - \tilde{v}_2 = \dots = v_n - \tilde{v}_n = 0 \\ & \text{(da } b_1, b_2, \dots, b_n \text{ linear unabhängig)} \\ \implies & v_1 = \tilde{v}_1, \quad v_2 = \tilde{v}_2, \quad \dots, \quad v_n = \tilde{v}_n. \end{aligned}$$

Also sind die Darstellungen (12.13) und (12.12) identisch, und wir haben gezeigt, dass die Darstellung (12.12) eindeutig ist. \square

Betrachten wir einige Beispiele.

Beispiel 12.37. (Koordinaten)

(a) Sei $V = \mathbb{K}^n$ mit der Standardbasis $E = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$.

$$\text{Für } \vec{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{K}^n \text{ gilt: } \vec{v} = v_1 \vec{e}_1 + v_2 \vec{e}_2 + \dots + v_n \vec{e}_n.$$

Also ist $\vec{v}_E = \vec{v}$, d.h. \vec{v} ist gleich seinem Koordinatenvektor \vec{v}_E bzgl. der Standardbasis E .

(b) Sei $V = \mathcal{P}_n$ mit der Basis $B = (1, t, t^2, \dots, t^n)$. Der zu der Polynomfunktion

$$p(t) = a_0 + a_1 t + a_2 t^2 + \dots + a_n t^n$$

gehörende Koordinatenvektor ist dann $\vec{p}_B = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{K}^{n+1}$.

(c) Sei $V = \mathbb{R}^2$ mit der Basis $B = (\vec{b}_1, \vec{b}_2)$ mit $\vec{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\vec{b}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Für $\vec{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ gilt $\vec{v} = \vec{b}_1 + \vec{b}_2$, d.h. $\vec{v}_B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Für $\vec{w} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ gilt $\vec{w} = -\vec{b}_1 + 2\vec{b}_2$, d.h. $\vec{w}_B = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$.

Wir untersuchen nun das **Basiswechsel-Problem**: Sei V ein \mathbb{K} -Vektorraum mit $\dim(V) = n$, und seien $B = (b_1, b_2, \dots, b_n)$ und $\tilde{B} = (\tilde{b}_1, \tilde{b}_2, \dots, \tilde{b}_n)$ zwei Basen von V . Ein Vektor $v \in V$ hat bzgl. jeder dieser beiden Basen einen Koordinatenvektor, also \vec{v}_B und $\vec{v}_{\tilde{B}}$ sind die beiden Koordinatenvektoren von v bzgl. der Basen B und \tilde{B} . **Wie berechnet man $\vec{v}_{\tilde{B}}$ aus \vec{v}_B ?**

Seien

$$\begin{aligned} v = v_1 b_1 + v_2 b_2 + \dots + v_n b_n, & \quad \text{also} \quad \vec{v}_B = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}, \\ v = \tilde{v}_1 \tilde{b}_1 + \tilde{v}_2 \tilde{b}_2 + \dots + \tilde{v}_n \tilde{b}_n, & \quad \text{also} \quad \vec{v}_{\tilde{B}} = \begin{bmatrix} \tilde{v}_1 \\ \vdots \\ \tilde{v}_n \end{bmatrix}. \end{aligned} \quad (12.14)$$

Da \tilde{B} eine Basis von V ist, existieren eindeutig bestimmte $\alpha_{i,k} \in \mathbb{K}$ mit

$$b_k = \sum_{i=1}^n \alpha_{i,k} \tilde{b}_i, \quad k = 1, 2, \dots, n. \quad (12.15)$$

Einsetzen in die Darstellung von v bzgl. der Basis B liefert

$$v = \sum_{k=1}^n v_k b_k = \sum_{k=1}^n v_k \left(\sum_{i=1}^n \alpha_{i,k} \tilde{b}_i \right) = \sum_{i=1}^n \left(\sum_{k=1}^n \alpha_{i,k} v_k \right) \tilde{b}_i. \quad (12.16)$$

Links steht nun eine Darstellung von v bzgl. der Basis \tilde{B} . Da die Darstellung bzgl. einer Basis eindeutig ist, folgt aus (12.14) und (12.16), dass

$$\tilde{v}_i = \sum_{k=1}^n \alpha_{i,k} v_k, \quad i = 1, 2, \dots, n, \quad (12.17)$$

gelten muss. Mit der Matrix $\mathbf{T}_{\tilde{B},B} := [\alpha_{i,k}] \in \mathbb{K}^{n \times n}$ können wir (12.17) schreiben als

$$\vec{v}_{\tilde{B}} = \mathbf{T}_{\tilde{B},B} \vec{v}_B.$$

Die Matrix $\mathbf{T}_{\tilde{B},B}$ heißt die **Transformationsmatrix** zum Basiswechsel von B nach \tilde{B} .

Achtung: Damit man die Matrix $\mathbf{T}_{\tilde{B},B} := [\alpha_{i,k}] \in \mathbb{K}^{n \times n}$ (und nicht ihre transponierte Matrix) erhält, müssen in (12.15) in der Darstellung von b_k bei den $\alpha_{i,k}$ der zweite Index für den Vektor b_k stehen und der erste Index muss der Summationsindex sein.

Beispiel 12.38. (Basiswechsel)

Seien $V = \mathbb{R}^2$, $E = (\vec{e}_1, \vec{e}_2)$ die Standardbasis und B die Basis $B := \left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \end{bmatrix} \right)$ von \mathbb{R}^2 .

Gesucht ist die Transformationsmatrix $\mathbf{T}_{B,E}$ für den Basiswechsel von E nach B .

Analog zu dem allgemeinen Fall stellen wir die Vektoren in (der alten Basis) E als Linearkombinationen der Vektoren in (der neuen Basis) B dar. Dabei bezeichnen wir die Koeffizienten analog zu dem allgemeinen Fall und bestimmen ihre Zahlenwerte durch das Lösen von linearen Gleichungssystemen:

Gesucht sind also $\alpha_{1,1}, \alpha_{2,1}, \alpha_{1,2}, \alpha_{2,2} \in \mathbb{R}$ mit

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \alpha_{1,1} \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \alpha_{2,1} \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \alpha_{1,2} \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \alpha_{2,2} \begin{bmatrix} 3 \\ 4 \end{bmatrix}.$$

Wir erhalten die beiden linearen Gleichungssysteme

$$\left[\begin{array}{cc|c} 1 & 3 & 1 \\ 2 & 4 & 0 \end{array} \right] \quad \text{und} \quad \left[\begin{array}{cc|c} 1 & 3 & 0 \\ 2 & 4 & 1 \end{array} \right].$$

Da diese beide die gleiche Koeffizientenmatrix haben und sich nur durch die rechten Seiten unterscheiden, können wir beide lineare Gleichungssysteme wie folgt „in einem Aufwasch“ erledigen:

$$\begin{array}{ccc}
 \left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 2 & 4 & 0 & 1 \end{array} \right] & \begin{array}{c} Z_2 \rightarrow Z_2 - 2Z_1 \\ \downarrow \\ \Leftrightarrow \end{array} & \left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & -2 & -2 & 1 \end{array} \right] \\
 \\
 \begin{array}{c} Z_2 \rightarrow -\frac{1}{2}Z_2 \\ \downarrow \\ \Leftrightarrow \end{array} & \left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & 1 & 1 & -\frac{1}{2} \end{array} \right] & \begin{array}{c} Z_1 \rightarrow Z_1 - 3Z_2 \\ \downarrow \\ \Leftrightarrow \end{array} & \left[\begin{array}{cc|cc} 1 & 0 & -2 & \frac{3}{2} \\ 0 & 1 & 1 & -\frac{1}{2} \end{array} \right] & (12.18)
 \end{array}$$

Wir lesen ab, dass gilt

$$\alpha_{1,1} = -2, \quad \alpha_{2,1} = 1, \quad \alpha_{1,2} = \frac{3}{2}, \quad \alpha_{2,2} = -\frac{1}{2}.$$

Also ist die Transformationsmatrix

$$\mathbf{T}_{B,E} = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} \\ \alpha_{2,1} & \alpha_{2,2} \end{bmatrix} = \begin{bmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{bmatrix}.$$

Dieses ist gerade die Matrix, die wir in (12.18) in der letzten erweiterten Koeffizientenmatrix auf der rechten Seite vorfinden.

Wollen wir beispielsweise die Darstellung des Vektor $\vec{\mathbf{v}} = \begin{bmatrix} 4 \\ 6 \end{bmatrix} = \vec{\mathbf{v}}_E$ bzgl. der Basis B bestimmen, so rechnen wir

$$\vec{\mathbf{v}}_B = \mathbf{T}_{B,E} \vec{\mathbf{v}}_E = \begin{bmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Also gilt

$$\vec{\mathbf{v}} = 1 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 1 \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix}.$$

Wir halten in einer Bemerkung fest, was wir über die Berechnung der Transformationsmatrix eines Basiswechsels gelernt haben.

Methode 12.39. (Berechnung der Transformationsmatrix für einen Basiswechsel in \mathbb{K}^n)

Seien $B = (\vec{\mathbf{b}}_1, \vec{\mathbf{b}}_2, \dots, \vec{\mathbf{b}}_n)$ und $D = (\vec{\mathbf{d}}_1, \vec{\mathbf{d}}_2, \dots, \vec{\mathbf{d}}_n)$ zwei Basen von \mathbb{K}^n . Die

Transformationsmatrix $\mathbf{T}_{D,B}$ für den Basiswechsel von der Basis B zur Basis D berechnet man wie folgt: Wir schreiben

$$\left[\vec{d}_1 \quad \vec{d}_2 \quad \cdots \quad \vec{d}_n \mid \vec{b}_1 \quad \vec{b}_2 \quad \cdots \quad \vec{b}_n \right]$$

und formen diese erweiterte Koeffizientenmatrix mit geeigneten elementaren Zeilenumformungen um, bis wir eine erweiterte Koeffizientenmatrix der Form $[\mathbf{E}_n \mid \mathbf{C}]$ erhalten, wobei \mathbf{E}_n die $n \times n$ -**Einheitsmatrix** ist mit den Einträgen $(\mathbf{E}_n)_{i,k} = 1$ für $i = k$ und $(\mathbf{E}_n)_{i,k} = 0$ für $i \neq k$. Die Matrix \mathbf{C} ist dann die Transformationsmatrix $\mathbf{T}_{D,B}$ für den Basiswechsel von der Basis B zur Basis D , also $\mathbf{T}_{D,B} := \mathbf{C}$.

12.7 Der Rang einer Matrix

Zunächst führen wir einige neue Bezeichnungen ein.

Definition 12.40. (Spaltenraum und Zeilenraum, Spaltenrang und Zeilenrang)

Sei $\mathbf{A} = [a_{i,k}] \in \mathbb{K}^{m \times n}$, also

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}$$

(1) **Spaltenvektoren der Matrix \mathbf{A} :** Wir bezeichnen die n Spaltenvektoren der Matrix \mathbf{A} mit $\vec{s}_1, \vec{s}_2, \dots, \vec{s}_n \in \mathbb{K}^{m \times 1} = \mathbb{K}^m$. Es gilt also

$$\vec{s}_1 = \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{m,1} \end{bmatrix}, \quad \vec{s}_2 = \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{m,2} \end{bmatrix}, \quad \dots, \quad \vec{s}_n = \begin{bmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{m,n} \end{bmatrix}.$$

Weiter bezeichnen wir mit

$$S(\mathbf{A}) = \text{LH}(\vec{s}_1, \vec{s}_2, \dots, \vec{s}_n)$$

den **Spaltenraum der Matrix \mathbf{A}** . Dieser ist nach Hilfssatz 12.15 ein Untervektorraum von \mathbb{K}^m . Die Dimension $\dim(S(\mathbf{A}))$ bezeichnen wir als **Spaltenrang**.

- (2) **Zeilenvektoren der Matrix \mathbf{A}** : Wir bezeichnen die m Zeilenvektoren der Matrix \mathbf{A} mit $\vec{\mathbf{z}}_1^T, \vec{\mathbf{z}}_2^T, \dots, \vec{\mathbf{z}}_m^T \in \mathbb{K}^{1 \times n}$. Es gilt also

$$\begin{aligned}\vec{\mathbf{z}}_1^T &= [a_{1,1} \ a_{1,2} \ \cdots \ a_{1,n}] \\ \vec{\mathbf{z}}_2^T &= [a_{2,1} \ a_{2,2} \ \cdots \ a_{2,n}] \\ &\vdots \\ \vec{\mathbf{z}}_m^T &= [a_{m,1} \ a_{m,2} \ \cdots \ a_{m,n}]\end{aligned}$$

Weiter bezeichnen wir mit

$$Z(\mathbf{A}) = \text{LH}(\vec{\mathbf{z}}_1^T, \vec{\mathbf{z}}_2^T, \dots, \vec{\mathbf{z}}_m^T)$$

den **Zeilenraum der Matrix \mathbf{A}** . Dieser ist nach Hilfssatz 12.15 ein Untervektorraum von $\mathbb{K}^{1 \times n}$. Die Dimension $\dim(Z(\mathbf{A}))$ bezeichnen wir als **Zeilenrang**.

(**Notation:** Ein Vektor $\vec{\mathbf{x}} \in \mathbb{K}^n$ ist immer ein Spaltenvektor. Daher müssen wir die Zeilenvektoren als $\vec{\mathbf{z}}_i^T \in \mathbb{K}^{1 \times n}$ schreiben, also $\vec{\mathbf{z}}_i \in \mathbb{K}^n$ ist ein Spaltenvektor, und durch die Transposition $\vec{\mathbf{z}}_i^T$ wird er ein Zeilenvektor.)

In der nächsten Bemerkung halten wir die sehr wichtige Anschauung des Spalten- und Zeilenranges fest:

Bemerkung 12.41. (Spalten- und Zeilenrang)

- (1) Der Spaltenrang ist die **maximale Anzahl an linear unabhängigen Spaltenvektoren**. Insbesondere gilt $\dim(S(\mathbf{A})) \leq \min\{m, n\}$ für jede Matrix $\mathbf{A} \in \mathbb{K}^{m \times n}$.
- (2) Der Zeilenrang ist die **maximale Anzahl an linear unabhängigen Zeilenvektoren**. Insbesondere gilt $\dim(Z(\mathbf{A})) \leq \min\{m, n\}$ für jede Matrix $\mathbf{A} \in \mathbb{K}^{m \times n}$.

Warum gelten $\dim(S(\mathbf{A})) \leq \min\{m, n\}$ und $\dim(Z(\mathbf{A})) \leq \min\{m, n\}$?

Erklärung: Es gibt in \mathbb{K}^m und damit im Untervektorraum $S(\mathbf{A})$ maximal m linear unabhängige Vektoren, d.h. $\dim(S(\mathbf{A})) \leq m$. Wegen $S(\mathbf{A}) =$

$\text{LH}(\vec{s}_1, \vec{s}_2, \dots, \vec{s}_n)$ folgt weiter $\dim(S(\mathbf{A})) \leq n$. Wir finden also insgesamt $\dim(S(\mathbf{A})) \leq \min\{m, n\}$. – Dass $\dim(Z(\mathbf{A})) \leq \min\{m, n\}$ gilt, zeigt man mit einer analogen Argumentation.

Betrachten wir ein Beispiel, um uns an diese einfache aber etwas gewöhnungsbedürftige Notation und Terminologie zu gewöhnen.

Beispiel 12.42. (Spaltenraum und Zeilenraum)

(a) Sei
$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 3 & 0 & 2 \\ 0 & 1 & 2 & 0 & 7 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{4 \times 5}.$$

Dann sind die Spaltenvektoren

$$\vec{s}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{s}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{s}_3 = \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{s}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{s}_5 = \begin{bmatrix} 2 \\ 7 \\ 4 \\ 0 \end{bmatrix},$$

und der Spaltenraum ist

$$S(\mathbf{A}) = \text{LH} \left(\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 7 \\ 4 \\ 0 \end{bmatrix} \right) = \text{LH} \left(\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right).$$

(*Erklärung:* An der von den Spaltenvektoren gebildeten Matrix sieht man, dass man alle Vektoren in $S(\mathbf{A}) = \text{LH}(\vec{s}_1, \vec{s}_2, \vec{s}_3, \vec{s}_4, \vec{s}_5)$ als Linearkombinationen der ersten drei Basisvektoren $\vec{s}_1 = \vec{e}_1$, $\vec{s}_2 = \vec{e}_2$, $\vec{s}_4 = \vec{e}_3$ der Standardbasis von \mathbb{R}^4 bilden kann.) Der Spaltenrang ist also $\dim(S(\mathbf{A})) = 3$.

Die Zeilenvektoren sind

$$\begin{aligned} \vec{z}_1^T &= [1 \ 0 \ 3 \ 0 \ 2], & \vec{z}_2^T &= [0 \ 1 \ 2 \ 0 \ 7], \\ \vec{z}_3^T &= [0 \ 0 \ 0 \ 1 \ 4], & \vec{z}_4^T &= [0 \ 0 \ 0 \ 0 \ 0], \end{aligned}$$

und der Zeilenraum ist

$$\begin{aligned}
Z(\mathbf{A}) &= \text{LH}\left([1 \ 0 \ 3 \ 0 \ 2], [0 \ 1 \ 2 \ 0 \ 7], [0 \ 0 \ 0 \ 1 \ 4], [0 \ 0 \ 0 \ 0 \ 0] \right) \\
&= \text{LH}\left([1 \ 0 \ 3 \ 0 \ 2], [0 \ 1 \ 2 \ 0 \ 7], [0 \ 0 \ 0 \ 1 \ 4] \right).
\end{aligned}$$

Man kann leicht ausrechnen, dass der Zeilenrang $\dim(Z(\mathbf{A})) = 3$ ist.

(b) Sei $\mathbf{A} = \mathbf{O}_{m \times n}$ die $m \times n$ -Nullmatrix. Dann gilt

$$\begin{aligned}
S(\mathbf{A}) &= \text{LH}\left(\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \right) = \left\{ \underbrace{\vec{\mathbf{0}}}_{\in \mathbb{K}^m} \right\}, \\
Z(\mathbf{A}) &= \text{LH}\left([0 \ \dots \ 0] \right) = \left\{ \underbrace{\vec{\mathbf{0}}^T}_{\in \mathbb{K}^{1 \times n}} \right\},
\end{aligned}$$

und es gilt $\dim(S(\mathbf{A})) = 0$ und $\dim(Z(\mathbf{A})) = 0$.

In dem vorigen Beispiel galt $\dim(S(\mathbf{A})) = \dim(Z(\mathbf{A}))$. Dieses ist immer der Fall, wie der nächste Satz zeigt.

Satz 12.43. (Spaltenrang = Zeilenrang)

Für jede Matrix $\mathbf{A} \in \mathbb{K}^{m \times n}$ gilt

$$\dim(S(\mathbf{A})) = \dim(Z(\mathbf{A})) \leq \min\{m, n\}.$$

Beweisskizze für Satz 12.43:

Schritt 1: Geht $\tilde{\mathbf{A}}$ aus \mathbf{A} durch elementare Zeilenoperationen hervor, so gilt $Z(\tilde{\mathbf{A}}) = Z(\mathbf{A})$, also insbesondere

$$\dim(Z(\tilde{\mathbf{A}})) = \dim(Z(\mathbf{A})). \quad (12.19)$$

(Dieses sieht man leicht, wenn man sich klar macht, was die elementaren Zeilenoperationen auf den Zeilen für die Zeilenvektoren bewirken.)

Schritt 2: Geht $\tilde{\mathbf{A}}$ aus \mathbf{A} durch elementare Zeilenoperationen hervor, so gilt im Allgemeinen nicht $S(\tilde{\mathbf{A}}) = S(\mathbf{A})$, aber es gilt trotzdem

$$\dim(S(\tilde{\mathbf{A}})) = \dim(S(\mathbf{A})). \quad (12.20)$$

Dieses sieht man wie folgt:

Sei $p := \dim(S(\mathbf{A}))$, und seien $\vec{s}_{j_1}, \vec{s}_{j_2}, \dots, \vec{s}_{j_p}$ linear unabhängige Spaltenvektoren von \mathbf{A} . Dann hat das lineare Gleichungssystem

$$\left[\begin{array}{cccc|c} \vec{s}_{j_1} & \vec{s}_{j_2} & \cdots & \vec{s}_{j_p} & \vec{0} \end{array} \right] \quad (12.21)$$

nur die triviale Lösung $\lambda_1 = \lambda_2 = \dots = \lambda_p = 0$. Elementare Zeilenoperationen führen \mathbf{A} in $\tilde{\mathbf{A}}$ über, und durch die gleichen elementaren Zeilenumformungen geht (12.21) in

$$\left[\begin{array}{cccc|c} \vec{s}_{j_1} & \vec{s}_{j_2} & \cdots & \vec{s}_{j_p} & \vec{0} \end{array} \right] \quad (12.22)$$

über. Dabei hat (12.22) genau wie (12.21) nur die triviale Lösung $\lambda_1 = \lambda_2 = \dots = \lambda_p = 0$, und die neuen Spaltenvektoren $\vec{s}_{j_1}, \vec{s}_{j_2}, \dots, \vec{s}_{j_p}$ sind linear unabhängig. Da die Vektoren $\vec{s}_{j_1}, \vec{s}_{j_2}, \dots, \vec{s}_{j_p}$ Spaltenvektoren von $\tilde{\mathbf{A}}$ sind folgt

$$\dim(S(\mathbf{A})) = p \leq \dim(S(\tilde{\mathbf{A}})), \quad (12.23)$$

denn $\tilde{\mathbf{A}}$ hat mindestens so viele linear unabhängige Spaltenvektoren wie \mathbf{A} . Analog geht auch \mathbf{A} aus $\tilde{\mathbf{A}}$ durch elementare Zeilenoperationen hervor, und mit der gleichen Argumentation folgert man

$$\dim(S(\tilde{\mathbf{A}})) \leq \dim(S(\mathbf{A})). \quad (12.24)$$

Aus (12.23) und (12.24) folgt dann

$$\dim(S(\mathbf{A})) = \dim(S(\tilde{\mathbf{A}})).$$

Schritt 3: Sei nun \mathbf{C} die reduzierte Stufenform von \mathbf{A} . Dann sieht man an der reduzierten Stufenform direkt, dass gilt

$$\dim(Z(\mathbf{C})) = \dim(S(\mathbf{C})). \quad (12.25)$$

Kombiniert man (12.19) und (12.20) mit $\tilde{\mathbf{A}} = \mathbf{C}$ mit (12.25), so findet man

$$\dim(Z(\mathbf{A})) = \dim(Z(\mathbf{C})) = \dim(S(\mathbf{C})) = \dim(S(\mathbf{A})).$$

Da $\dim(S(\mathbf{A}))$ gemäß Definition 12.40 (vgl. auch Bemerkung 12.41) die maximale Anzahl der linear unabhängigen Spaltenvektoren ist, gilt $\dim(S(\mathbf{A})) \leq n$. Da $\dim(Z(\mathbf{A}))$ gemäß Definition die maximale Anzahl der linear unabhängigen Zeilenvektoren ist, gilt $\dim(Z(\mathbf{A})) \leq m$. Damit folgt auch

$$\dim(S(\mathbf{A})) = \dim(Z(\mathbf{A})) \leq \min\{m, n\},$$

und der Satz ist bewiesen. □

Definition 12.44. (Rang einer Matrix)

Sei $\mathbf{A} \in \mathbb{K}^{m \times n}$. Dann definieren wir den **Rang** von \mathbf{A} durch

$$\text{Rang}(\mathbf{A}) := \dim(S(\mathbf{A})) = \dim(Z(\mathbf{A})).$$

Wir halten noch fest, was wir im Beweis von Satz 12.43 gesehen haben.

Hilfssatz 12.45. (Rang ist invariant unter elem. Zeilenoperationen)

Der Rang einer Matrix $\mathbf{A} \in \mathbb{K}^{m \times n}$ ändert sich durch elementare Zeilenoperationen nicht.

Satz 12.46. (Rang von Matrix und Transponierter sind gleich)

Sei $\mathbf{A} \in \mathbb{K}^{m \times n}$. Dann gilt $\text{Rang}(\mathbf{A}^T) = \text{Rang}(\mathbf{A})$.

Beweis von Satz 12.46: Es gilt

$$\text{Rang}(\mathbf{A}^T) = \dim(Z(\mathbf{A}^T)) = \dim(S(\mathbf{A})) = \text{Rang}(\mathbf{A}). \quad \square$$

Betrachten wir nun einige Beispiele, um den Umgang mit dem Begriff des Ranges einer Matrix zu üben.

Beispiel 12.47. (Rang einer Matrix)

(a) Es gilt

$$\text{Rang} \left(\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \end{bmatrix} \right) = 3,$$

denn die ersten drei Spaltenvektoren sind offensichtlich linear unabhängig (wegen der Stufenform der Matrix), und mehr als drei linear unabhängige Vektoren kann es in \mathbb{R}^3 nicht geben.

(b) Da sich der Rang einer Matrix unter elementaren Zeilenoperationen nicht ändert (siehe Hilfssatz 12.45), können wir elementare Zeilenoperationen nut-

zen, um die Matrix zu „vereinfachen“, bevor wir den Rang berechnen:

$$\text{Rang} \left(\begin{bmatrix} 1 & 2 & 7 & -8 & 4 \\ -11 & 3 & -1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 4 & 14 & -16 & 8 \end{bmatrix} \right)$$

$$\stackrel{\substack{Z_4 \rightarrow Z_4 - 2Z_1 \\ \Downarrow \\ \cong}}{\text{Rang}} \left(\begin{bmatrix} 1 & 2 & 7 & -8 & 4 \\ -11 & 3 & -1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right) = 2.$$

Hier haben wir im letzten Schritt den Rang über den Zeilenrang bestimmt. Dieser ist ≤ 2 , da es nur zwei Zeilen gibt, die keine Nullzeilen sind. Weil keiner der ersten beiden Zeilenvektoren der Matrix ein Vielfaches des anderen ist, muss der Zeilenrang gleich 2 sein.

(c) Es gilt

$$\text{Rang} \left(\begin{bmatrix} 1 & 1 & 1 \\ 3 & 2 & 1 \\ 2 & 1 & \alpha \end{bmatrix} \right) \stackrel{\substack{Z_2 \rightarrow Z_2 - 3Z_1 \\ Z_3 \rightarrow Z_3 - 2Z_1 \\ \Downarrow \\ \cong}}{\text{Rang}} \left(\begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & -1 & \alpha - 2 \end{bmatrix} \right)$$

$$\stackrel{\substack{Z_3 \rightarrow Z_3 - Z_2 \\ \Downarrow \\ \cong}}{\text{Rang}} \left(\begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & \alpha \end{bmatrix} \right) \stackrel{\substack{Z_2 \rightarrow -Z_2 \\ \Downarrow \\ \cong}}{\text{Rang}} \left(\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & \alpha \end{bmatrix} \right)$$

$$= \begin{cases} 2, & \text{falls } \alpha = 0, \\ 3, & \text{falls } \alpha \neq 0. \end{cases}$$

Als letzte Anwendung des Ranges einer Matrix gewinnen wir zwei weitere Informationen über die Lösbarkeit eines linearen Gleichungssystems.

Satz 12.48. (Lösbarkeit eines LGS)

Sei $\mathbf{A} \in \mathbb{K}^{m \times n}$. Dann gelten:

- (1) Für die Lösungsmenge $\mathbb{L}_{[\mathbf{A}|\vec{\mathbf{0}}]}$ des homogenen linearen Gleichungssystems $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{0}}$ mit $\vec{\mathbf{0}} \in \mathbb{K}^m$ und $\vec{\mathbf{x}} \in \mathbb{K}^n$ gilt

$$\dim(\mathbb{L}_{[\mathbf{A}|\vec{\mathbf{0}}]}) = n - \text{Rang}(\mathbf{A}).$$

(2) Für $\vec{\mathbf{b}} \in \mathbb{K}^m$ ist das lineare Gleichungssystem $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}$ genau dann lösbar, wenn gilt

$$\text{Rang}([\mathbf{A} | \vec{\mathbf{b}}]) = \text{Rang}(\mathbf{A}).$$

Beweis von Satz 12.48:

- (1) Betrachten wir $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{0}} \iff [\mathbf{A} | \vec{\mathbf{0}}]$. Dann gilt $\text{Rang}(\mathbf{A}) = ([\mathbf{A} | \vec{\mathbf{0}}])$, da der Nullvektor $\vec{\mathbf{0}}$ und die Spaltenvektoren der Matrix \mathbf{A} linear abhängig sind. Durch elementare Zeilenumformungen können wir $[\mathbf{A} | \vec{\mathbf{0}}]$ in Stufenform bringen, und der Rang ändert sich dabei nicht. In der Stufenform finden wir genau $\text{Rang}(\mathbf{A})$ linear unabhängige Spaltenvektoren vor, und man sieht an der Stufenform, dass die Lösung von $n - \text{Rang}(\mathbf{A})$ frei wählbaren Parametern abhängt. Also gilt

$$\dim(\mathbb{L}_{[\mathbf{A} | \vec{\mathbf{0}}]}) = n - \text{Rang}(\mathbf{A}).$$

- (2) Bezeichnen wir die Spaltenvektoren von \mathbf{A} mit $\vec{\mathbf{s}}_1, \vec{\mathbf{s}}_2, \dots, \vec{\mathbf{s}}_n$, so können wir $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}$ schreiben als

$$x_1 \vec{\mathbf{s}}_1 + x_2 \vec{\mathbf{s}}_2 + \dots + x_n \vec{\mathbf{s}}_n = \vec{\mathbf{b}},$$

d.h. das lineare Gleichungssystem $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}$ ist genau dann lösbar, wenn gilt $\vec{\mathbf{b}} \in \text{LH}(\vec{\mathbf{s}}_1, \vec{\mathbf{s}}_2, \dots, \vec{\mathbf{s}}_n)$. Da $\text{Rang}(\mathbf{A})$ bzw. $\text{Rang}([\mathbf{A} | \vec{\mathbf{b}}])$ die Anzahl der linear unabhängigen Spaltenvektoren in \mathbf{A} bzw. in $[\mathbf{A} | \vec{\mathbf{b}}]$ ist, gilt $\vec{\mathbf{b}} \in \text{LH}(\vec{\mathbf{s}}_1, \vec{\mathbf{s}}_2, \dots, \vec{\mathbf{s}}_n)$ genau dann, wenn $\text{Rang}([\mathbf{A} | \vec{\mathbf{b}}]) = \text{Rang}(\mathbf{A})$ gilt (vgl. auch Bemerkung 12.21 (6)). \square

Für den Sonderfall, dass $m = n$ ist und $\text{Rang}(\mathbf{A}) = n$ folgt mit Hilfe von Satz 12.48 der folgende Satz.

Satz 12.49. (äquivalente Aussagen über ein LGS mit $m = n$)

Sei $\mathbf{A} \in \mathbb{K}^{n \times n}$. Dann sind die folgenden Aussagen **äquivalent**:

- (i) $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}$ ist für jedes $\vec{\mathbf{b}} \in \mathbb{K}^n$ **eindeutig lösbar**.
- (ii) $\text{Rang}(\mathbf{A}) = n$

(iii) $\mathbb{L}_{[\mathbf{A}|\vec{\mathbf{0}}]} = \{\vec{\mathbf{0}}\}$, d.h. $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{0}}$ hat **nur die Lösung** $\vec{\mathbf{x}} = \vec{\mathbf{0}}$.
(Es gilt dann auch $\dim(\mathbb{L}_{[\mathbf{A}|\vec{\mathbf{0}}]}) = 0$.)

Wir wollen auch diesen Satz beweisen, weil der Beweis die Aussagen verständlicher macht.

Beweis von Satz 12.49: Das Kernstück des Beweises bildet die folgende Vorüberlegung: Bezeichnen wir die Spaltenvektoren von \mathbf{A} mit $\vec{\mathbf{s}}_1, \vec{\mathbf{s}}_2, \dots, \vec{\mathbf{s}}_n$, so können wir $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}$ schreiben als

$$x_1 \vec{\mathbf{s}}_1 + x_2 \vec{\mathbf{s}}_2 + \dots + x_n \vec{\mathbf{s}}_n = \vec{\mathbf{b}}, \quad (12.26)$$

d.h. das lineare Gleichungssystem $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}$ ist genau dann lösbar, wenn $\vec{\mathbf{b}} \in \text{LH}(\vec{\mathbf{s}}_1, \vec{\mathbf{s}}_2, \dots, \vec{\mathbf{s}}_n)$ gilt.

Wir zeigen nun mit Hilfe dieser Vorüberlegung die folgenden drei Aussagen: (i) \Rightarrow (ii), (ii) \Rightarrow (iii) und (iii) \Rightarrow (i). Mit diesen kann man dann „zyklisch“ aus jeder der drei Aussagen jede andere folgern und hat damit die Äquivalenzen bewiesen.

- *Beweis von (i) \Rightarrow (ii):* Es sei $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}$ für jedes $\vec{\mathbf{b}} \in \mathbb{K}^n$ eindeutig lösbar. Dieses bedeutet nach der Vorüberlegung, dass jeder Vektor $\vec{\mathbf{b}} \in \mathbb{K}^n$ in $\text{LH}(\vec{\mathbf{s}}_1, \vec{\mathbf{s}}_2, \dots, \vec{\mathbf{s}}_n)$ liegt. Also gilt $\text{LH}(\vec{\mathbf{s}}_1, \vec{\mathbf{s}}_2, \dots, \vec{\mathbf{s}}_n) = \mathbb{K}^n$. Wegen $\dim(\mathbb{K}^n) = n$, kann dieses nur dann der Fall sein, wenn $\vec{\mathbf{s}}_1, \vec{\mathbf{s}}_2, \dots, \vec{\mathbf{s}}_n$ eine Basis von \mathbb{K}^n bilden und somit linear unabhängig sind. Dann ist $\text{Rang}(\mathbf{A}) = n$.
- *Beweis von (ii) \Rightarrow (iii):* Sei $\text{Rang}(\mathbf{A}) = n$. Dann sind die n Spaltenvektoren von \mathbf{A} linear unabhängig (da $\text{Rang}(\mathbf{A})$ die Anzahl der linear unabhängigen Spaltenvektoren von \mathbf{A} ist). Dann hat (12.26) mit $\vec{\mathbf{b}} = \vec{\mathbf{0}}$ nur die Lösung $x_1 = x_2 = \dots = x_n = 0$, und somit hat $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{0}}$ nach der Vorüberlegung nur die Lösung $\vec{\mathbf{x}} = \vec{\mathbf{0}}$.
- *Beweis von (iii) \Rightarrow (i):* Das lineare Gleichungssystem $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{0}}$ habe nur die Lösung $\vec{\mathbf{x}} = \vec{\mathbf{0}}$. Aus (12.26) mit $\vec{\mathbf{b}} = \vec{\mathbf{0}}$ folgt dann, dass $x_1 = x_2 = \dots = x_n = 0$ sind. Die Spaltenvektoren $\vec{\mathbf{s}}_1, \vec{\mathbf{s}}_2, \dots, \vec{\mathbf{s}}_n$ von \mathbf{A} sind also n linear unabhängige Vektoren in \mathbb{K}^n . Also ist $(\vec{\mathbf{s}}_1, \vec{\mathbf{s}}_2, \dots, \vec{\mathbf{s}}_n)$ eine Basis von \mathbb{K}^n , und jedes $\vec{\mathbf{b}} \in \mathbb{K}^n$ lässt sich mit (12.26) eindeutig als Linearkombination vom $\vec{\mathbf{s}}_1, \vec{\mathbf{s}}_2, \dots, \vec{\mathbf{s}}_n$ darstellen. Also ist $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}$ nach der Vorüberlegung für jedes $\vec{\mathbf{b}} \in \mathbb{K}^n$ eindeutig lösbar. \square

Betrachten wir zwei Beispiele zu Satz 12.48 (und Satz 12.49).

Beispiel 12.50. (Lösbarkeit eines LGS)

(a) Für das lineare Gleichungssystem

$$[\mathbf{A}|\vec{\mathbf{0}}] = \left[\begin{array}{ccccc|c} 2 & -1 & 3 & 5 & 7 & 0 \\ 4 & 1 & -5 & 3 & 0 & 0 \\ 6 & 0 & 3 & -9 & 0 & 0 \end{array} \right]$$

ist die Spaltenzahl $n = 5$, und die Zeilenzahl ist $m = 3$. Es gilt somit $\text{Rang}(\mathbf{A}) \leq 3 = \min\{3, 5\}$. Wir finden drei linear unabhängige Spaltenvektoren

$$\vec{\mathbf{s}}_5 = \begin{bmatrix} 7 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{\mathbf{s}}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{\mathbf{s}}_1 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix},$$

d.h. es gilt $\text{Rang}(\mathbf{A}) = 3$. Die lineare Unabhängigkeit der Spaltenvektoren folgt direkt aus der von diesen gebildeten Matrix (wegen deren Stufenform):

$$\left[\begin{array}{ccc} 7 & -1 & 2 \\ 0 & 1 & 4 \\ 0 & 0 & 6 \end{array} \right] \begin{array}{l} Z_1 \rightarrow \frac{1}{7}Z_1 \\ Z_3 \rightarrow \frac{1}{6}Z_3 \\ \downarrow \\ \iff \end{array} \left[\begin{array}{ccc} 1 & -\frac{1}{7} & \frac{2}{7} \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{array} \right].$$

Also gilt nach Satz 12.48 (1)

$$\dim(\mathbb{L}_{[\mathbf{A}|\vec{\mathbf{0}}]}) = n - \text{Rang}(\mathbf{A}) = 5 - 3 = 2.$$

(b) In Beispiel 11.24 wurde das lineare Gleichungssystem

$$[\mathbf{A}|\vec{\mathbf{b}}] = \left[\begin{array}{cccc|c} 1 & 1 & -3 & 1 & 1 \\ 2 & 1 & 1 & -1 & 0 \\ 0 & 2 & -13 & 1 & -1 \\ 2 & -1 & 14 & -2 & \alpha \end{array} \right]$$

betrachtet und mit den Gaußschen Eliminationsverfahren in die Stufenform

$$\left[\begin{array}{cccc|c} 1 & 1 & -3 & 1 & 1 \\ 0 & 1 & -7 & 3 & 2 \\ 0 & 0 & 1 & -5 & -5 \\ 0 & 0 & 0 & 0 & \alpha - 1 \end{array} \right]$$

gebracht. Nun können wir an der Stufenform direkt ablesen

$$\text{Rang}(\mathbf{A}) = 3 \quad \text{und} \quad \text{Rang}([\mathbf{A}|\vec{\mathbf{b}}]) = \begin{cases} 3 & \text{wenn } \alpha = 1, \\ 4 & \text{wenn } \alpha \neq 1. \end{cases}$$

Also ist das lineare Gleichungssystem nach Satz 12.48 (2) genau dann lösbar, wenn $\alpha = 1$ ist.

12.8 Skalarprodukt und Norm

In Verallgemeinerung der Standardskalarproduktes, welches wir aus Kapitel 2 der HM A kennen, führen wir nun für beliebige \mathbb{K} -Vektorräume ein Skalarprodukt und eine Norm ein.

Definition 12.51. (Skalarprodukt)

Sei V ein \mathbb{K} -Vektorraum. $\langle x|y \rangle$ heißt ein **Skalarprodukt** von $x, y \in V$, wenn gelten:

$$(S1) \quad \langle x|y \rangle \in \mathbb{K} \quad \text{für alle } x, y \in V.$$

$$(S2) \quad \langle x|y \rangle = \overline{\langle y|x \rangle} \quad \text{für alle } x, y \in V.$$

$$(S3) \quad \langle x|y+z \rangle = \langle x|y \rangle + \langle x|z \rangle \quad \text{für alle } x, y, z \in V.$$

$$(S4) \quad \langle x|\lambda y \rangle = \lambda \langle x|y \rangle \quad \text{für alle } x, y \in V \text{ und alle } \lambda \in \mathbb{K}.$$

$$(S5) \quad \langle x|x \rangle \geq 0 \quad \text{für alle } x \in V.$$

$$(S6) \quad \langle x|x \rangle = 0 \iff x = o_V$$

V heißt dann ein **Vektorraum mit dem Skalarprodukt (VR mit SP)** $\langle x|y \rangle$. (Anmerkung: Der „Überstrich“ in $\overline{\langle y|x \rangle}$ steht dabei für das Bilden der zu $\langle y|x \rangle$ konjugiert komplexen Zahl.)

Wir halten zunächst einige abgeleitete Eigenschaften des Skalarprodukts fest.

Bemerkung 12.52. (weitere Eigenschaften des Skalarprodukts)

(S7) Aus (S2) und (S3) folgt

$$\langle x+y|z \rangle = \overline{\langle z|x+y \rangle} = \overline{\langle z|x \rangle + \langle z|y \rangle} = \overline{\langle z|x \rangle} + \overline{\langle z|y \rangle} = \langle x|z \rangle + \langle y|z \rangle.$$

Also gilt

$$\langle x+y|z \rangle = \langle x|z \rangle + \langle y|z \rangle \quad \text{für alle } x, y, z \in V.$$

(S8) Aus (S2) und (S4) folgt für alle $x, y \in V$ und alle $\lambda \in \mathbb{K}$

$$\langle \lambda x|y \rangle = \overline{\langle y|\lambda x \rangle} = \overline{\lambda \langle y|x \rangle} = \bar{\lambda} \overline{\langle y|x \rangle} = \bar{\lambda} \langle x|y \rangle.$$

(S9) Aus (S4) und (S8) folgt leicht

$$\langle x|o_V \rangle = \langle o_V|x \rangle = 0 \quad \text{für alle } x \in V.$$

Wenn $\mathbb{K} = \mathbb{R}$ ist, dann entfällt das Bilden der konjugiert komplexen Zahlen natürlich, denn $\bar{a} = a$ für jedes $a \in \mathbb{R}$.

Betrachten wir einige Beispiele.

Beispiel 12.53. (Skalarprodukt)

(a) Sei $V = \mathbb{R}^n$. Für $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^n$ definieren wir

$$\langle \vec{x} | \vec{y} \rangle := \vec{x} \cdot \vec{y} = x_1 y_1 + x_2 y_2 + \dots + x_n y_n = \sum_{k=1}^n x_k y_k.$$

Dieses ist das **Standardskalarprodukt auf \mathbb{R}^n** (vgl. Kapitel 2 der HM A).

(b) Sei $V = \mathbb{C}^n$. Für $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{C}^n$ definieren wir

$$\langle \vec{x} | \vec{y} \rangle := \vec{x} \cdot \vec{y} := \bar{x}_1 y_1 + \bar{x}_2 y_2 + \dots + \bar{x}_n y_n = \sum_{k=1}^n \bar{x}_k y_k.$$

Dieses ist das sogenannte **Standardskalarprodukt auf \mathbb{C}^n** .

(c) Sei $V = \mathcal{C}([a, b])$ der Vektorraum der auf dem Intervall $[a, b]$ stetigen reellwertigen Funktionen. Für f, g in $\mathcal{C}([a, b])$ definiert

$$\langle f | g \rangle := \int_a^b f(x) g(x) dx \quad (12.27)$$

ein Skalarprodukt.

Wir werden jeweils in einer Übungsaufgabe überprüfen, dass die Skalarprodukte aus dem vorigem Beispiel wirklich die Eigenschaften (S1) bis (S6) haben.

Nun führen wir die zu dem Skalarprodukt gehörige Norm ein.

Definition 12.54. (Norm)

Sei V ein \mathbb{K} -Vektorraum mit einem Skalarprodukt $\langle x | y \rangle$, $x, y \in V$. Dann definieren wir die (**zum Skalarprodukt gehörige**) **Norm** durch

$$\|x\| := \sqrt{\langle x | x \rangle}, \quad x \in V.$$

Betrachten wir auch hierzu einige Beispiele.

Beispiel 12.55. (Norm)

(a) Sei $V = \mathbb{R}^n$ mit dem Standardskalarprodukt.

Für $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ ist die (zugehörige) Standardnorm

$$\|\vec{x}\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} = |\vec{x}|.$$

Dieses ist die Länge des Vektors \vec{x} (vgl. Kapitel 2 der HM A).

(b) Sei $V = \mathbb{C}^n$ mit dem Standardskalarprodukt.

Für $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{C}^n$ ist die (zugehörige) Standardnorm

$$\|\vec{x}\| = \sqrt{\overline{x_1}x_1 + \overline{x_2}x_2 + \dots + \overline{x_n}x_n} = \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2}.$$

(c) Für $V = \mathcal{C}([a, b])$ mit dem Skalarprodukt (12.27) ist die (zugehörige) Norm

$$\|f\| = \left(\int_a^b f(x) \cdot f(x) \, dx \right)^{1/2} = \left(\int_a^b [f(x)]^2 \, dx \right)^{1/2}, \quad f \in \mathcal{C}([a, b]).$$

Im nächsten Satz sind die wichtigsten Eigenschaften einer (zu einem Skalarprodukt gehörigen) Norm zusammengestellt.

Satz 12.56. (Eigenschaften der zum Skalarprodukt gehörigen Norm)

Sei V ein \mathbb{K} -Vektorraum mit einem Skalarprodukt $\langle x|y \rangle$, $x, y \in V$, und sei $\|x\| := \sqrt{\langle x|x \rangle}$ die zum Skalarprodukt gehörige Norm. Für alle $x, y \in V$ und alle $\lambda \in \mathbb{K}$ gelten:

(1) $\|x\| \geq 0$

(2) $\|x\| = 0 \iff x = o_V$

(3) $\|\lambda x\| = |\lambda| \cdot \|x\|$

(4) $|\langle x|y \rangle| \leq \|x\| \cdot \|y\|$ (**Cauchy-Schwarz-Ungleichung**)

$$(5) |\langle x|y \rangle| = \|x\| \cdot \|y\| \iff x, y \text{ sind linear abhängig.}$$

$$(6) \|x + y\| \leq \|x\| + \|y\| \quad (\text{Dreiecksungleichung})$$

Wir beweisen die Eigenschaften der (zu einem Skalarprodukt gehörigen) Norm.

Beweis von Satz 12.56:

- (1) $\|x\| \geq 0$ ist klar, weil die Wurzel in $\|x\| = \sqrt{\langle x|x \rangle}$ per Definition nicht-negativ ist.
- (2) Diese Eigenschaft folgt direkt aus Eigenschaft (S6) in Definition 12.51.
- (3) Für alle $x \in V$ und alle $\lambda \in \mathbb{K}$ gilt

$$\|\lambda x\|^2 = \langle \lambda x | \lambda x \rangle = \lambda \cdot \bar{\lambda} \cdot \langle x|x \rangle = |\lambda| \cdot \|x\|^2,$$

und Wurzelziehen liefert die Aussage (3).

- (4) Ist $y = o_V$, so sind beide Seiten der zu zeigenden Ungleichung $|\langle x|y \rangle| \leq \|x\| \cdot \|y\|$ nach (S9) aus Bemerkung 12.52 und Satz 12.56 (2) gleich null, und es gilt Gleichheit. Also sei nun $y \neq o_V$ und damit nach Satz 12.56 (2) $\|y\| \neq 0$. Wir setzen

$$\lambda := \frac{\langle y|x \rangle}{\|y\|^2}$$

und berechnen

$$\begin{aligned} 0 \leq \langle x - \lambda y | x - \lambda y \rangle &= \langle x|x \rangle - \lambda \langle x|y \rangle - \bar{\lambda} \langle y|x \rangle + |\lambda|^2 \langle y|y \rangle \\ &= \|x\|^2 - \lambda \overline{\langle y|x \rangle} - \bar{\lambda} \langle y|x \rangle + |\lambda|^2 \|y\|^2 \\ &= \|x\|^2 - \frac{\langle y|x \rangle}{\|y\|^2} \overline{\langle y|x \rangle} - \frac{\overline{\langle y|x \rangle}}{\|y\|^2} \langle y|x \rangle + \frac{|\langle y|x \rangle|^2}{\|y\|^4} \|y\|^2 \\ &= \|x\|^2 - 2 \frac{|\langle y|x \rangle|^2}{\|y\|^2} + \frac{|\langle y|x \rangle|^2}{\|y\|^2} \\ &= \|x\|^2 - \frac{|\langle y|x \rangle|^2}{\|y\|^2}. \end{aligned}$$

Es gilt somit

$$\begin{aligned} 0 \leq \langle x - \lambda y | x - \lambda y \rangle = \|x\|^2 - \frac{|\langle y|x \rangle|^2}{\|y\|^2} &\iff \frac{|\langle y|x \rangle|^2}{\|y\|^2} \leq \|x\|^2 \\ \iff |\langle y|x \rangle|^2 \leq \|x\|^2 \cdot \|y\|^2 &\iff |\langle y|x \rangle| \leq \|x\| \cdot \|y\|, \quad (12.28) \end{aligned}$$

und wir haben die Cauchy-Schwarz-Ungleichung bewiesen.

- (5) Ist $y = o_V$ oder $x = o_V$, so sind die beiden Vektoren x, y immer linear abhängig, und es gilt nach (S9) aus Bemerkung 12.52 und Satz 12.56 (2)

$$0 = |\langle x|y \rangle| = \|x\| \cdot \|y\| = 0.$$

Nun nehmen wir an, dass x und y beide ungleich dem Nullvektor o_V sind.

Wir zeigen zuerst „ \implies “: Es gelte also $|\langle y|x \rangle| = \|x\| \cdot \|y\|$. Dann folgt aus der Rechnung im Beweis von Satz 12.56 (4), wenn wir alle Schritte rückwärts durchführen, dass

$$0 = \langle x - \lambda y | x - \lambda y \rangle.$$

Nach (S5) in Definition 12.51 folgt dann $x - \lambda y = 0$, und wir sehen, dass x, y linear abhängig sind.

Nun zeigen wir „ \impliedby “: Sind $x \neq 0$ und $y \neq 0$ linear abhängig, so gibt es ein $\mu \in \mathbb{K}$ mit $x = \mu y \Leftrightarrow x - \mu y = 0$. Wir wollen die Rechenschritte aus dem Beweis von Satz 12.56 (4) nutzen und berechnen dazu λ :

$$\lambda := \frac{\langle y|x \rangle}{\|y\|^2} = \frac{\langle y|\mu y \rangle}{\|y\|^2} = \frac{\mu \cdot \langle y|y \rangle}{\|y\|^2} = \mu \cdot \frac{\|y\|^2}{\|y\|^2} = \mu.$$

Also gilt in (12.28), wobei wir nun wegen $x - \lambda y = 0$ nach (S6) in Satz 12.51 statt dem \leq ein $=$ haben:

$$\begin{aligned} 0 = \langle x - \lambda y | x - \lambda y \rangle &= \|x\|^2 - \frac{|\langle y|x \rangle|^2}{\|y\|^2} && \iff && \frac{|\langle y|x \rangle|^2}{\|y\|^2} = \|x\|^2 \\ &\iff |\langle y|x \rangle|^2 = \|x\|^2 \cdot \|y\|^2 && \iff && |\langle y|x \rangle| = \|x\| \cdot \|y\|. \end{aligned}$$

Damit haben wir beide Richtungen gewiesen.

- (6) Wir nutzen Satz 12.56 (5) aus:

$$\begin{aligned} \|x + y\|^2 &= \langle x + y | x + y \rangle = \langle x|x \rangle + \langle x|y \rangle + \langle y|x \rangle + \langle y|y \rangle \\ &\leq \underbrace{\langle x|x \rangle}_{=\|x\|^2} + \underbrace{|\langle x|y \rangle|}_{\leq \|x\| \cdot \|y\|} + \underbrace{|\langle y|x \rangle|}_{\leq \|x\| \cdot \|y\|} + \underbrace{\langle y|y \rangle}_{=\|y\|^2} \\ &\leq \|x\|^2 + 2 \cdot \|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \end{aligned}$$

Wurzelziehen liefert $\|x + y\| \leq \|x\| + \|y\|$. □

12.9 Orthogonalität

Aufbauend auf das Konzept eines Skalarprodukts lernen wir in diesem Teilkapitel das Konzept der Orthogonalität von Vektoren kennen und ziehen wichtige Schlussfolgerungen über orthogonale Vektoren.

Definition 12.57. (orthogonale Vektoren)

Sei V ein \mathbb{K} -Vektorraum mit einem **Skalarprodukt** $\langle x|y \rangle$, $x, y \in V$. Zwei Vektoren $x, y \in V$ heißen **orthogonal** (in Symbolen $x \perp y$), falls $\langle x|y \rangle = 0$ ist.

Betrachten wir einige Beispiele.

Beispiel 12.58. (orthogonale Vektoren)

(a) Sei $V = \mathbb{R}^3$ mit dem Standardskalarprodukt.

Dann sind $\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$ und $\begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$ orthogonal, denn

$$\left\langle \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \middle| \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\rangle = 1 \cdot (-2) + (-1) \cdot 0 + 2 \cdot 1 = -2 + 0 + 2 = 0.$$

(b) Sei $V = \mathcal{C}([0, 2\pi])$ der \mathbb{R} -Vektorraum der reellwertigen stetigen Funktionen auf $[0, 2\pi]$ mit dem Skalarprodukt

$$\langle f|g \rangle = \int_0^{2\pi} f(x) g(x) dx. \quad (12.29)$$

Seien $f : [0, 2\pi] \rightarrow \mathbb{R}$, $f(x) := \sin(x)$, und $g : [0, 2\pi] \rightarrow \mathbb{R}$, $g(x) := \cos(x)$, aus $V = \mathcal{C}([0, 2\pi])$. Dann gilt

$$\begin{aligned} \langle f|g \rangle &= \int_0^{2\pi} \sin(x) \cos(x) dx = \int_0^{2\pi} \frac{1}{2} \sin(2x) dx \\ &= \left[-\frac{1}{4} \cos(2x) \right]_{x=0}^{x=2\pi} = -\frac{1}{4} \cos(4\pi) - \left(-\frac{1}{4} \cos(0) \right) = -\frac{1}{4} + \frac{1}{4} = 0, \end{aligned}$$

wobei wir im zweiten Schritt

$$\sin(2x) = \sin(x + x) = \sin(x) \cos(x) + \cos(x) \sin(x) = 2 \sin(x) \cos(x)$$

genutzt haben. Also sind die Funktionen $f : [0, 2\pi] \rightarrow \mathbb{R}$, $f(x) := \sin(x)$, und $g : [0, 2\pi] \rightarrow \mathbb{R}$, $g(x) := \cos(x)$, orthogonal.

Wir betrachten nun Teilmengen von Vektorräumen, die aus sogenannten „paarweise orthogonalen“ Vektoren bestehen.

Definition 12.59. (Orthonormalsystem)

Sei V ein \mathbb{K} -Vektorraum mit einem Skalarprodukt $\langle x|y \rangle$, $x, y \in V$. Ein System (e_1, e_2, \dots, e_m) von Vektoren $e_1, e_2, \dots, e_m \in V$ heißt ein **Orthonormalsystem (ONS)** in V , falls

$$\langle e_i|e_k \rangle = \begin{cases} 1, & \text{wenn } i = k, \\ 0, & \text{wenn } i \neq k. \end{cases}$$

Wir sagen dann, die Vektoren e_1, e_2, \dots, e_m sind **paarweise orthogonal** (da $\langle e_i|e_k \rangle = 0$ wenn $i \neq k$) und **normiert** (da $\|e_i\| = \sqrt{\langle e_i|e_i \rangle} = 1$ für alle $i = 1, 2, \dots, m$).

Betrachten wir einige Beispiele für Orthonormalsysteme.

Beispiel 12.60. (Orthonormalsysteme)

(a) Sei $V = \mathbb{K}^n$ mit dem Standardskalarprodukt, und seien

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \vec{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

die Vektoren der Standardbasis. Die Standardbasis $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ ist ein Orthonormalsystem in \mathbb{K}^n (mit dem Standardskalarprodukt).

(b) Sei $V = \mathcal{C}([0, 2\pi])$ mit dem Skalarprodukt (12.29). Seien

$$e_1 : [0, 2\pi] \rightarrow \mathbb{R}, \quad e_1(x) := \frac{1}{\sqrt{\pi}} \sin(x),$$

$$e_2 : [0, 2\pi] \rightarrow \mathbb{R}, \quad e_2(x) := \frac{1}{\sqrt{\pi}} \cos(x).$$

Aus der Rechnung in Beispiel 12.58 (b) folgt direkt, dass gilt

$$\langle e_1|e_2 \rangle = \int_0^{2\pi} \frac{1}{\sqrt{\pi}} \sin(x) \frac{1}{\sqrt{\pi}} \cos(x) dx = \frac{1}{\pi} \int_0^{2\pi} \sin(x) \cos(x) dx = 0.$$

Also sind e_1 und e_2 orthogonal. Analog zeigt man (z.B. mit partieller Integration), dass gilt

$$\begin{aligned}\langle e_1|e_1\rangle &= \int_0^{2\pi} \frac{1}{\sqrt{\pi}} \sin(x) \frac{1}{\sqrt{\pi}} \sin(x) \, dx = \frac{1}{\pi} \int_0^{2\pi} \sin^2(x) \, dx = 1, \\ \langle e_2|e_2\rangle &= \int_0^{2\pi} \frac{1}{\sqrt{\pi}} \cos(x) \frac{1}{\sqrt{\pi}} \cos(x) \, dx = \frac{1}{\pi} \int_0^{2\pi} \cos^2(x) \, dx = 1.\end{aligned}$$

Damit ist nachgewiesen, dass (e_1, e_2) ein Orthonormalsystem in $\mathcal{C}([0, 2\pi])$ mit dem Skalarprodukt (12.29) ist.

Der nächste Satz stellt einen Zusammenhang von Orthogonalität mit linearer Unabhängigkeit her.

Satz 12.61. (Orthonormalsystem und lineare Unabhängigkeit)

Sei V ein \mathbb{K} -Vektorraum mit einem Skalarprodukt $\langle x|y\rangle$, $x, y \in V$. Ist (e_1, e_2, \dots, e_m) ein **Orthonormalsystem** in V , so sind die Vektoren e_1, e_2, \dots, e_m **linear unabhängig**.

Beweis von Satz 12.61: Seien $\lambda_1, \lambda_2, \dots, \lambda_m \in \mathbb{K}$ mit

$$\lambda_1 e_1 + \lambda_2 e_2 + \dots + \lambda_m e_m = o_V.$$

Für jedes $i = 1, 2, \dots, m$ nehmen wir auf beiden Seiten das Skalarprodukt mit e_i , also

$$\begin{aligned}0 &= \langle e_i|o_V\rangle = \langle e_i|\lambda_1 e_1 + \lambda_2 e_2 + \dots + \lambda_i e_i + \dots + \lambda_m e_m\rangle \\ &= \lambda_1 \underbrace{\langle e_i|e_1\rangle}_{=0} + \lambda_2 \underbrace{\langle e_i|e_2\rangle}_{=0} + \dots + \lambda_i \underbrace{\langle e_i|e_i\rangle}_{=1} + \dots + \lambda_m \underbrace{\langle e_i|e_m\rangle}_{=0} = \lambda_i.\end{aligned}$$

Also folgt $\lambda_1 = \lambda_2 = \dots = \lambda_m = 0$, und wir haben gezeigt, dass e_1, e_2, \dots, e_m linear unabhängig sind. \square

Folgerung 12.62. (Orthonormalbasis)

Sei V ein \mathbb{K} -Vektorraum mit einem Skalarprodukt $\langle x|y\rangle$, $x, y \in V$. Ist $\dim(V) = n$ und (e_1, e_2, \dots, e_n) ein Orthonormalsystem in V , so ist (e_1, e_2, \dots, e_n) eine Basis von V . Wir nennen eine solche Basis eine **Orthonormalbasis (ONB)** von V .

Betrachten wir ein Beispiel für eine Orthonormalbasis.

Beispiel 12.63. (Orthonormalbasis)

Die Standardbasis $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ von \mathbb{R}^n bzw. \mathbb{C}^n mit dem Standardskalarprodukt von \mathbb{R}^n bzw. \mathbb{C}^n ist eine Orthonormalbasis dieser Vektorräume.

Bemerkung 12.64. (Finden einer Orthonormalbasis in einem endlichdimensionalen Vektorraum)

In jedem endlich dimensionalen Vektorraum mit einem Skalarprodukt kann man eine Orthonormalbasis wählen. Mit Hilfe des „**Orthogonalisierungsverfahrens von Gram-Schmidt**“ kann man aus jeder gegebenen Basis eine Orthonormalbasis machen. Wir besprechen dieses auf einem Übungszettel.

Als Letztes lernen wir, wie sich Vektoren als Linearkombination bzgl. einer Orthonormalbasis besonders elegant darstellen lassen.

Satz 12.65. (Darstellung von Vektoren bzgl einer Orthonormalbasis)

Sei V ein n -dimensionaler \mathbb{K} -Vektorraum mit einem Skalarprodukt $\langle x|y \rangle$, $x, y \in V$, und sei $E = (e_1, e_2, \dots, e_n)$ eine **Orthonormalbasis** von V . Seien $x, y \in V$. Dann gelten:

$$(1) \quad x = \sum_{k=1}^n \langle e_k|x \rangle e_k$$

$$(2) \quad \langle x|y \rangle = \sum_{k=1}^n \overline{\langle e_k|x \rangle} \langle e_k|y \rangle$$

$$(3) \quad \|x\|^2 = \sum_{k=1}^n |\langle e_k|x \rangle|^2$$

Schauen wir uns die Aussagen in Satz 12.65 genauer an, um diesen sehr wichtigen Satz zu verstehen:

Satz 12.65 (1) gibt Informationen über die Darstellung von Vektoren bzgl. der Orthonormalbasis $E = (e_1, e_2, \dots, e_n)$. Genauer gilt für $x, y \in V$

$$x = \sum_{k=1}^n \underbrace{\langle e_k|x \rangle}_{=: x_k} e_k = \sum_{k=1}^n x_k e_k \quad \text{und} \quad y = \sum_{k=1}^n \underbrace{\langle e_k|y \rangle}_{=: y_k} e_k = \sum_{k=1}^n y_k e_k, \quad (12.30)$$

d.h. die Koeffizienten $x_k = \langle e_k | x \rangle$, $k = 1, 2, \dots, n$, bzw. $y_k = \langle e_k | y \rangle$, $k = 1, 2, \dots, n$, sind gerade die Koordinaten von x bzw. y bzgl. der Orthonormalbasis $E = (e_1, e_2, \dots, e_n)$. Genauer ist

$$\vec{x}_E = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \langle e_1 | x \rangle \\ \vdots \\ \langle e_n | x \rangle \end{bmatrix} \quad \text{und} \quad \vec{y}_E = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \langle e_1 | y \rangle \\ \vdots \\ \langle e_n | y \rangle \end{bmatrix}.$$

Satz 12.65 (2) besagt, dass für die Vektoren x und y in (12.30) gilt

$$\langle x | y \rangle = \sum_{k=1}^n \underbrace{\langle e_k | x \rangle}_{= \overline{x_k}} \underbrace{\langle e_k | y \rangle}_{= y_k} = \sum_{k=1}^n \overline{x_k} \cdot y_k, \quad (12.31)$$

d.h. das Skalarprodukt $\langle x | y \rangle$ lässt sich einfach berechnen, indem man auf die Koordinatenvektoren \vec{x}_E und \vec{y}_E von x und y das Standardskalarprodukt für \mathbb{K}^n anwendet.

Satz 12.65 (3) ist ein Sonderfall von Satz 12.65 (2) mit $x = y$, und an (12.31) lesen wir ab, dass Satz 12.65 (3) nun

$$\langle x | x \rangle = \sum_{k=1}^n \underbrace{\langle e_k | x \rangle}_{= \overline{x_k}} \underbrace{\langle e_k | x \rangle}_{= x_k} = \sum_{k=1}^n \overline{x_k} \cdot x_k = \sum_{k=1}^n |x_k|^2 \quad (12.32)$$

ist. Auch hier berechnet sich die Norm einfach, indem man die Standardnorm für \mathbb{K}^n auf den Koordinatenvektor \vec{x}_E von x anwendet.

Achtung: Es ist ganz wichtig, zu beachten, dass die einfachen Formeln (12.31) und (12.32) **nur gelten**, wenn die Koordinatendarstellungen bzgl. einer **Orthonormalbasis** vorliegen! Ebenso gilt die einfache Formel $x_k = \langle e_k | x \rangle$ für die Koordinaten in (12.30) **nur** bei Koordinaten bzgl. einer **Orthonormalbasis**.

Wir beweisen nun Satz 12.65.

Beweis von Satz 12.65:

- (1) Da $E = (e_1, e_2, \dots, e_n)$ eine Basis von V ist, hat x eine eindeutige Darstellung

$$x = x_1 e_1 + x_2 e_2 + \dots + x_n e_n \quad (12.33)$$

mit eindeutig bestimmten $x_1, x_2, \dots, x_n \in \mathbb{K}$. Da (e_1, e_2, \dots, e_n) ein Orthonormalsystem ist, können wir x_k direkt berechnen, indem wir in (12.33) auf beiden Seiten das Skalarprodukt mit e_k bilden:

$$\langle e_k | x \rangle = \langle e_k | x_1 e_1 + x_2 e_2 + \dots + x_k e_k + \dots + x_n e_n \rangle \quad \iff$$

$$\langle e_k | x \rangle = x_1 \underbrace{\langle e_k | e_1 \rangle}_{=0} + x_2 \underbrace{\langle e_k | e_2 \rangle}_{=0} + \dots + x_k \underbrace{\langle e_k | e_k \rangle}_{=1} + \dots + x_n \underbrace{\langle e_k | e_n \rangle}_{=0} = x_k,$$

weil $\langle e_i | e_k \rangle = 1$ wenn $i = k$ und $\langle e_i | e_k \rangle = 0$ wenn $i \neq k$. Also gilt $x_k = \langle e_k | x \rangle$ für $k = 1, 2, \dots, n$.

- (2) Wir nutzen Satz 12.65 (1) und die Eigenschaften des Skalarprodukts: Für $x, y \in V$ gilt

$$\begin{aligned} x &= \langle e_1 | x \rangle e_1 + \langle e_2 | x \rangle e_2 + \dots + \langle e_n | x \rangle e_n \quad \implies \\ \langle x | y \rangle &= \left\langle \langle e_1 | x \rangle e_1 + \langle e_2 | x \rangle e_2 + \dots + \langle e_n | x \rangle e_n \mid y \right\rangle \\ &= \overline{\langle e_1 | x \rangle} \langle e_1 | y \rangle + \overline{\langle e_2 | x \rangle} \langle e_2 | y \rangle + \dots + \overline{\langle e_n | x \rangle} \langle e_n | y \rangle \\ &= \sum_{k=1}^n \overline{\langle e_k | x \rangle} \langle e_k | y \rangle. \end{aligned}$$

- (3) Wir nutzen Satz 12.65 (2):

$$\begin{aligned} \|x\|^2 &= \overline{\langle e_1 | x \rangle} \langle e_1 | x \rangle + \overline{\langle e_2 | x \rangle} \langle e_2 | x \rangle + \dots + \overline{\langle e_n | x \rangle} \langle e_n | x \rangle \\ &= |\langle e_1 | x \rangle|^2 + |\langle e_2 | x \rangle|^2 + \dots + |\langle e_n | x \rangle|^2 \\ &= \sum_{k=1}^n |\langle e_k | x \rangle|^2. \end{aligned}$$

Damit ist Satz 12.65 bewiesen. □

Betrachten wir zwei Beispiele zu Satz 12.65.

Beispiel 12.66. (Darstellung von Vektoren bzgl. Orthonormalbasis)

- (a) Sei $V = \mathbb{K}^n$ mit dem Standardskalarprodukt, und seien

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \vec{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

die Vektoren der Standardbasis. Die Standardbasis $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ ist eine Orthonormalbasis in \mathbb{K}^n (mit dem Standardskalarprodukt): Für $\vec{x} \in \mathbb{K}^n$

gilt also

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{k=1}^n x_k \vec{e}_k,$$

und wir haben in der Tat

$$\langle \vec{e}_k | \vec{x} \rangle = 0 \cdot x_1 + \dots + 0 \cdot x_{k-1} + 1 \cdot x_k + 0 \cdot x_{k+1} + \dots + 0 \cdot x_n = x_k$$

für $k = 1, 2, \dots, n$.

(b) In \mathbb{R}^3 mit dem Standardskalarprodukt ist

$$E = \left(\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \right)$$

eine Orthonormalbasis, denn: Es handelt sich um $\dim(\mathbb{R}^3) = 3$ Vektoren, und

$$\left\langle \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \middle| \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \right\rangle = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} = \frac{1}{4} + \frac{1}{4} + \frac{1}{2} = 1,$$

$$\left\langle \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \middle| \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \right\rangle = \left(-\frac{1}{\sqrt{2}}\right) \cdot \left(-\frac{1}{\sqrt{2}}\right) + \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} = \frac{1}{2} + \frac{1}{2} = 1,$$

$$\begin{aligned} \left\langle \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \middle| \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \right\rangle &= \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} + \left(-\frac{1}{\sqrt{2}}\right) \cdot \left(-\frac{1}{\sqrt{2}}\right) \\ &= \frac{1}{4} + \frac{1}{4} + \frac{1}{2} = 1, \end{aligned}$$

d.h. die Vektoren von E sind normiert, und

$$\begin{aligned} \left\langle \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \middle| \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \right\rangle &= \frac{1}{2} \cdot \left(-\frac{1}{\sqrt{2}}\right) + \frac{1}{2} \cdot \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} \cdot 0 \\ &= -\frac{1}{2\sqrt{2}} + \frac{1}{2\sqrt{2}} = 0, \end{aligned}$$

$$\left\langle \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \middle| \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \right\rangle = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{\sqrt{2}} \cdot \left(-\frac{1}{\sqrt{2}}\right) = \frac{1}{4} + \frac{1}{4} - \frac{1}{2} = 0,$$

$$\left\langle \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \middle| \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \right\rangle = \left(-\frac{1}{\sqrt{2}}\right) \cdot \frac{1}{2} + \frac{1}{\sqrt{2}} \cdot \frac{1}{2} + 0 \cdot \left(-\frac{1}{\sqrt{2}}\right)$$

$$= -\frac{1}{2\sqrt{2}} + \frac{1}{2\sqrt{2}} = 0,$$

d.h. die Vektoren von E sind paarweise orthogonal.

Die Darstellung des Vektors

$$\vec{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

als Linearkombination bzgl. der Orthonormalbasis E berechnet sich mit

$$\left\langle \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \middle| \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\rangle = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 2 + \frac{1}{\sqrt{2}} \cdot 3 = \frac{3}{2} + \frac{3}{\sqrt{2}} = \frac{3(1 + \sqrt{2})}{2},$$

$$\left\langle \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \middle| \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\rangle = \left(-\frac{1}{\sqrt{2}}\right) \cdot 1 + \frac{1}{\sqrt{2}} \cdot 2 + 0 \cdot 3 = -\frac{1}{\sqrt{2}} + \frac{2}{\sqrt{2}} = \frac{1}{\sqrt{2}},$$

$$\left\langle \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \middle| \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\rangle = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 2 + \left(-\frac{1}{\sqrt{2}}\right) \cdot 3 = \frac{3}{2} - \frac{3}{\sqrt{2}} = \frac{3(1 - \sqrt{2})}{2}$$

als

$$\frac{3(1 + \sqrt{2})}{2} \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{\sqrt{2}} \end{bmatrix} + \frac{1}{\sqrt{2}} \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} + \frac{3(1 - \sqrt{2})}{2} \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

KAPITEL 13

Quadratische Matrizen

In diesem Kapitel betrachten wir ausschließlich quadratische Matrizen und lernen diverse teils neue Konzepte kennen, die nur für quadratische Matrizen definiert werden können. Als Erstes ist sind hier die inverse Matrix (siehe Teilkapitel 13.2) und die Determinante (siehe Teilkapitel 13.3, 13.4 und 13.5) zu nennen. In Teilkapitel 13.6 lernen wir dann die neuen Begriffe des Eigenwertes und Eigenvektors einer quadratischen Matrix kennen.

In diesem Kapitel sei, wie bereits in den vorigen Kapiteln, \mathbb{K} immer entweder \mathbb{R} oder \mathbb{C} , und es sei $n \in \mathbb{N}$.

13.1 Einführung zu quadratischen Matrizen

In der HM A (siehe Kapitel 2 im Skript der HM A) haben wir bereits die folgende Definition gelernt:

Definition 13.1. (quadratische Matrix)

Gilt bei einer Matrix $\mathbf{A} \in \mathbb{K}^{m \times n}$, dass $m = n$ ist, also wenn die Matrix genauso viele Zeilen wie Spalten hat, dann nennen wir die Matrix eine quadratische Matrix.

Von jetzt an betrachten wir in diesem Kapitel nur noch quadratische Matrizen.

Definition 13.2. (Diagonalmatrix, Einheitsmatrix, untere/obere Dreiecksmatrix und Spur)

Sei $\mathbf{A} = [a_{i,k}] \in \mathbb{K}^{n \times n}$ eine quadratische Matrix.

- (1) $\mathbf{A} = [a_{i,k}]$ heißt eine **Diagonalmatrix**, wenn $a_{i,k} = 0$ für $i \neq k$, also wenn

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & 0 & \cdots & 0 \\ 0 & a_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{n,n} \end{bmatrix}.$$

- (2) $\mathbf{A} = [a_{i,k}]$ heißt die **Einheitsmatrix** (von $\mathbb{K}^{n \times n}$), wenn \mathbf{A} eine Diagonalmatrix ist mit $a_{1,1} = a_{2,2} = \dots = a_{n,n} = 1$, also wenn

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} =: \mathbf{E}_n.$$

- (3) $\mathbf{A} = [a_{i,k}]$ heißt eine **untere Dreiecksmatrix**, wenn $a_{i,k} = 0$ für $i < k$, also wenn

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & 0 & \cdots & 0 \\ a_{2,1} & a_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ a_{n,1} & \cdots & a_{n,n-1} & a_{n,n} \end{bmatrix}.$$

- (4) $\mathbf{A} = [a_{i,k}]$ heißt eine **obere Dreiecksmatrix**, wenn $a_{i,k} = 0$ für $i > k$, also wenn

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ 0 & a_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ 0 & \cdots & 0 & a_{n,n} \end{bmatrix}.$$

- (5) Die **Spur** von $\mathbf{A} = [a_{i,k}]$ ist definiert durch

$$\text{Spur}(\mathbf{A}) = \text{Spur} \left(\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix} \right) := a_{1,1} + a_{2,2} + \dots + a_{n,n}.$$

(Also: Die Spur ist die Summe der Einträge auf der Diagonale der Matrix von links oben nach rechts unten.)

Betrachten wir einige Beispiele.

Beispiel 13.3. (Diagonalmatrix, Einheitsmatrix, untere/obere Dreiecksmatrix und Spur)

(a) Hier sind einige Diagonalmatrizen:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} -17 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & 0 & \pi \end{bmatrix}.$$

Die Matrix $\mathbf{D} = \begin{bmatrix} 0 & 0 & 5 \\ 0 & 3 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ ist **keine** Diagonalmatrix, denn:

Die Diagonale einer Matrix verläuft **von links oben nach rechts unten!**

(b) Die Einheitsmatrix in $\mathbb{R}^{2 \times 2}$ und $\mathbb{C}^{2 \times 2}$ bzw. in $\mathbb{R}^{3 \times 3}$ und $\mathbb{C}^{3 \times 3}$ ist jeweils

$$\mathbf{E}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{bzw.} \quad \mathbf{E}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

(c) Seien

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -3 & 0 & -7 \\ 0 & \sqrt{2} & 1 \\ 0 & 0 & \pi \end{bmatrix},$$

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 3 & 0 & 0 \\ -2 & -7 & 0 \\ 0 & -1 & -2 \end{bmatrix}.$$

Dann sind \mathbf{A} und \mathbf{B} obere Dreiecksmatrizen, und \mathbf{C} und \mathbf{D} sind untere Dreiecksmatrizen.

Jede Diagonalmatrix ist sowohl eine untere als auch eine obere Dreiecksmatrix.

(d) Berechnen wir die Spur für einige quadratische Matrizen

$$\text{Spur} \left(\begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix} \right) = 1 + 5 + 9 = 15,$$

$$\text{Spur} \left(\begin{bmatrix} 2 & -1 & 6 & -7 \\ -3 & 3 & -8 & 13 \\ 4 & -2 & 4 & -3 \\ 7 & 6 & 8 & -9 \end{bmatrix} \right) = 2 + 3 + 4 + (-9) = 0.$$

Für die Einheitsmatrix gilt

$$\text{Spur}(\mathbf{E}_n) = \text{Spur} \left(\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} \right) = \underbrace{1 + 1 + \cdots + 1}_{n \text{ mal}} = n.$$

Wir halten eine wichtige Eigenschaft der Einheitsmatrix als Hilfssatz fest: Diese spielt bei der Multiplikation von Matrizen die Rolle der Zahl 1 bei der Multiplikation von reellen oder komplexen Zahlen.

Hilfssatz 13.4. (Multiplikation mit der Einheitsmatrix)

Sei \mathbf{E}_n die Einheitsmatrix von $\mathbb{K}^{n \times n}$. Dann gilt

$$\mathbf{A} \cdot \mathbf{E}_n = \mathbf{E}_n \cdot \mathbf{A} = \mathbf{A} \quad \text{für alle } \mathbf{A} \in \mathbb{K}^{n \times n}$$

und

$$\mathbf{E}_n \vec{\mathbf{x}} = \vec{\mathbf{x}} \quad \text{für alle } \vec{\mathbf{x}} \in \mathbb{K}^n.$$

Beweis von Hilfssatz 13.4: Diesen Hilfssatz haben wir in einer Übungsaufgabe bewiesen. \square

Als Letztes lernen wir noch einen nützlichen Satz über die Spur des Produkts zweier Matrizen kennen.

Satz 13.5. (Spur des Produkts zweier Matrizen)

Für $\mathbf{B} \in \mathbb{K}^{n \times p}$ und $\mathbf{C} \in \mathbb{K}^{p \times n}$ gilt $\mathbf{B} \cdot \mathbf{C} \in \mathbb{K}^{n \times n}$ und $\mathbf{C} \cdot \mathbf{B} \in \mathbb{K}^{p \times p}$ und

$$\text{Spur}(\mathbf{B} \cdot \mathbf{C}) = \text{Spur}(\mathbf{C} \cdot \mathbf{B}).$$

Beweis von Satz 13.5: Wir beweisen diesen Satz in einer Übungsaufgabe. \square

13.2 Invertierbare Matrizen

Als Vorbereitung für die Definition der **inversen Matrix** starten wir mit einem Hilfssatz, der uns schon bekanntes Wissen neu zusammenstellt.

Hilfssatz 13.6. (äquivalente Aussagen über quadratische Matrizen)

Sei $\mathbf{A} = [a_{i,k}] \in \mathbb{K}^{n \times n}$, und sei $(\vec{\mathbf{e}}_1, \vec{\mathbf{e}}_2, \dots, \vec{\mathbf{e}}_n)$ die Standardbasis von \mathbb{K}^n . Dann sind folgende Aussagen äquivalent:

- (i) Das lineare Gleichungssystem $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}$ ist für jedes $\vec{\mathbf{b}} \in \mathbb{K}^n$ **eindeutig lösbar**.
- (ii) Die n linearen Gleichungssysteme $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{e}}_1$, $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{e}}_2$, \dots , $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{e}}_n$ sind **eindeutig lösbar**.
- (iii) Es existiert **genau ein** $\mathbf{X} \in \mathbb{K}^{n \times n}$ mit $\mathbf{A} \cdot \mathbf{X} = \mathbf{E}_n$.

Beweis von Hilfssatz 13.6: Wir zeigen (i) \Leftrightarrow (ii) und (ii) \Leftrightarrow (iii). Durch Verkettung dieser beiden Äquivalenzen folgt daraus dann auch (i) \Leftrightarrow (iii).

- *Beweis von (i) \Rightarrow (ii):* Ist $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}$ für jedes $\vec{\mathbf{b}} \in \mathbb{K}^n$ eindeutig lösbar, so gilt dieses insbesondere für $\vec{\mathbf{b}} = \vec{\mathbf{e}}_1$, $\vec{\mathbf{b}} = \vec{\mathbf{e}}_2$, \dots , bzw. $\vec{\mathbf{b}} = \vec{\mathbf{e}}_n$.
- *Beweis von (ii) \Rightarrow (i):* Es seien $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{e}}_1$, $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{e}}_2$, \dots , $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{e}}_n$ eindeutig lösbar, und wir bezeichnen deren Lösungen als $\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2, \dots, \vec{\mathbf{x}}_n$, also $\mathbf{A} \vec{\mathbf{x}}_i = \vec{\mathbf{e}}_i$ für $i = 1, 2, \dots, n$.

Jeder Vektor $\vec{\mathbf{b}} \in \mathbb{K}^n$ hat die eindeutige Darstellung

$$\vec{\mathbf{b}} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \sum_{i=1}^n b_i \vec{\mathbf{e}}_i$$

bzgl. der Standardbasis $(\vec{\mathbf{e}}_1, \vec{\mathbf{e}}_2, \dots, \vec{\mathbf{e}}_n)$ von \mathbb{K}^n . Wir zeigen nun, dass

$$\vec{\mathbf{x}} := \sum_{i=1}^n b_i \vec{\mathbf{x}}_i$$

eine Lösung von $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}$ ist: In der Tat gilt

$$\mathbf{A} \cdot \vec{\mathbf{x}} = \mathbf{A} \left(\sum_{i=1}^n b_i \vec{\mathbf{x}}_i \right) = \sum_{i=1}^n \mathbf{A} b_i \vec{\mathbf{x}}_i = \sum_{i=1}^n b_i \underbrace{\mathbf{A} \vec{\mathbf{x}}_i}_{=\vec{\mathbf{e}}_i} = \sum_{i=1}^n b_i \vec{\mathbf{e}}_i = \vec{\mathbf{b}}.$$

Dass wir die Matrix in die Summe schieben dürfen, liegt an dem Distributivgesetz für die Matrizenrechnung.

Wir müssen uns noch überlegen, warum $\vec{\mathbf{x}} = \sum_{i=1}^n b_i \vec{\mathbf{x}}_i$ die einzige Lösung von $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}$ ist. Dazu zeigen wir zunächst, dass die Vektoren $\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2, \dots, \vec{\mathbf{x}}_n$ linear unabhängig sind und somit eine Basis von \mathbb{K}^n bilden: Sei also

$$\begin{aligned} \sum_{i=1}^n \lambda_i \vec{\mathbf{x}}_i = \vec{\mathbf{0}} &\quad \implies \quad \mathbf{A} \left(\sum_{i=1}^n \lambda_i \vec{\mathbf{x}}_i \right) = \mathbf{A} \vec{\mathbf{0}} \\ \implies \sum_{i=1}^n \lambda_i \underbrace{\mathbf{A} \vec{\mathbf{x}}_i}_{=\vec{\mathbf{e}}_i} = \vec{\mathbf{0}} &\quad \implies \quad \sum_{i=1}^n \lambda_i \vec{\mathbf{e}}_i = \vec{\mathbf{0}}, \end{aligned}$$

und es folgt $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$, d.h. $\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2, \dots, \vec{\mathbf{x}}_n$ sind linear unabhängig und bilden somit eine Basis von \mathbb{K}^n .

Wir nehmen nun an, es gebe eine weitere Lösung $\vec{\mathbf{y}}$ mit $\mathbf{A} \vec{\mathbf{y}} = \vec{\mathbf{b}}$. Dann hat $\vec{\mathbf{y}}$ eine eindeutige Darstellung $\vec{\mathbf{y}} = \sum_{i=1}^n c_i \vec{\mathbf{x}}_i$ bzgl. der Basis $(\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2, \dots, \vec{\mathbf{x}}_n)$ von \mathbb{K}^n . Es folgt

$$\sum_{i=1}^n b_i \vec{\mathbf{e}}_i = \vec{\mathbf{b}} = \mathbf{A} \vec{\mathbf{y}} = \mathbf{A} \left(\sum_{i=1}^n c_i \vec{\mathbf{x}}_i \right) = \sum_{i=1}^n c_i \underbrace{\mathbf{A} \vec{\mathbf{x}}_i}_{=\vec{\mathbf{e}}_i} = \sum_{i=1}^n c_i \vec{\mathbf{e}}_i,$$

und da $(\vec{\mathbf{e}}_1, \vec{\mathbf{e}}_2, \dots, \vec{\mathbf{e}}_n)$ eine Basis von \mathbb{K}^n ist folgt $c_i = b_i$, $i = 1, 2, \dots, n$, d.h. $\vec{\mathbf{y}} = \vec{\mathbf{x}}$. Also ist die Lösung von $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}$ eindeutig bestimmt.

- *Beweis von (ii) \Leftrightarrow (iii):* Es gilt $\mathbf{E}_n = [\vec{e}_1 \ \vec{e}_2 \ \cdots \ \vec{e}_n]$. Es sei \mathbf{X} die Matrix deren Spaltenvektoren die Lösungsvektoren $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n$ von $\mathbf{A} \vec{x}_1 = \vec{e}_1$, $\mathbf{A} \vec{x}_2 = \vec{e}_2$, \dots , $\mathbf{A} \vec{x}_n = \vec{e}_n$ sind. Dann gilt:

$$\begin{aligned} \mathbf{A} \cdot \mathbf{X} = \mathbf{E}_n &\iff \mathbf{A} [\vec{x}_1 \ \vec{x}_2 \ \cdots \ \vec{x}_n] = [\vec{e}_1 \ \vec{e}_2 \ \cdots \ \vec{e}_n] \\ &\iff \left(\mathbf{A} \vec{x}_1 = \vec{e}_1, \quad \mathbf{A} \vec{x}_2 = \vec{e}_2, \quad \dots, \quad \mathbf{A} \vec{x}_n = \vec{e}_n \right) \end{aligned}$$

Daran sehen wir direkt, dass $\mathbf{A} \cdot \mathbf{X} = \mathbf{E}_n$ genau dann in $\mathbb{K}^{n \times n}$ eindeutig lösbar ist, wenn $\mathbf{A} \vec{x} = \vec{e}_1$, $\mathbf{A} \vec{x} = \vec{e}_2$, \dots , $\mathbf{A} \vec{x} = \vec{e}_n$ jeweils in \mathbb{K}^n eindeutig lösbar sind. \square

Wir interessieren uns nun für die Matrix \mathbf{X} in Hilfssatz 13.6 (iii), falls diese existiert. Diese werden wir weiter unten als die **inverse Matrix zu \mathbf{A}** einführen.

Zunächst einmal wollen wir uns noch überlegen, dass, falls ein $\mathbf{X} \in \mathbb{K}^{n \times n}$ mit $\mathbf{A} \cdot \mathbf{X} = \mathbf{E}_n$ existiert, die Matrix \mathbf{X} **eindeutig bestimmt** ist und auch

$$\mathbf{X} \cdot \mathbf{A} = \mathbf{E}_n$$

erfüllt. (Da die Matrizenmultiplikation nicht kommutativ ist, ist dieses nicht selbstverständlich!)

Es existiere also $\mathbf{X} \in \mathbb{K}^{n \times n}$ mit $\mathbf{A} \cdot \mathbf{X} = \mathbf{E}_n$.

Nach Hilfssatz 13.6 ist dann das lineare Gleichungssystem $\mathbf{A} \vec{x} = \vec{b}$ für jedes $\vec{b} \in \mathbb{K}^n$ lösbar. (Da wir die Eindeutigkeit von \mathbf{X} nicht verlangen, verlieren wir auch in Hilfssatz 13.6 (i) und (ii) die Eindeutigkeit. Man macht sich dieses klar, wenn man den Beweis des Satzes noch einmal durchgeht.) Mit den Resultaten aus Teilkapitel 12.7 folgt daraus $\text{Rang}(\mathbf{A}) = n$. Mit Satz 12.46 folgt $\text{Rang}(\mathbf{A}^T) = \text{Rang}(\mathbf{A}) = n$.

Nach Satz 12.49 folgt aus $\text{Rang}(\mathbf{A}^T) = n$, dass das lineare Gleichungssystem $\mathbf{A}^T \vec{y} = \vec{c}$ ist für jedes $\vec{c} \in \mathbb{K}^n$ eindeutig lösbar ist. Nach Hilfssatz 13.6 existiert daher ein $\mathbf{Y} \in \mathbb{K}^{n \times n}$ mit $\mathbf{A}^T \cdot \mathbf{Y} = \mathbf{E}_n$. Durch Transponieren finden wir

$$(\mathbf{A}^T \cdot \mathbf{Y})^T = \mathbf{E}_n^T \iff \mathbf{Y}^T \cdot \mathbf{A} = \mathbf{E}_n, \quad (13.1)$$

und Multiplikation mit \mathbf{X} von rechts liefert

$$\mathbf{Y}^T \cdot \underbrace{\mathbf{A} \cdot \mathbf{X}}_{=\mathbf{E}_n} = \underbrace{\mathbf{E}_n \cdot \mathbf{X}}_{=\mathbf{X}}, \iff \mathbf{Y}^T = \mathbf{Y}^T \cdot \mathbf{E}_n = \mathbf{X}.$$

Aus (13.1) folgt wegen $\mathbf{Y}^T = \mathbf{X}$, dass $\mathbf{X} \in \mathbb{K}^{n \times n}$ mit $\mathbf{A} \cdot \mathbf{X} = \mathbf{E}_n$ auch $\mathbf{X} \cdot \mathbf{A} = \mathbf{E}_n$ erfüllt.

Um zu zeigen, dass die Matrix \mathbf{X} eindeutig bestimmt ist, nehmen wir an, es gebe eine weitere Matrix $\tilde{\mathbf{X}}$ mit $\mathbf{A} \cdot \tilde{\mathbf{X}} = \mathbf{E}_n$. Multiplikation von links mit \mathbf{X} liefert

$$\underbrace{\mathbf{X} \cdot \mathbf{A}}_{=\mathbf{E}_n} \cdot \tilde{\mathbf{X}} = \mathbf{X} \cdot \mathbf{E}_n \quad \iff \quad \underbrace{\mathbf{E}_n \cdot \tilde{\mathbf{X}}}_{=\tilde{\mathbf{X}}} = \underbrace{\mathbf{X} \cdot \mathbf{E}_n}_{=\mathbf{X}} \quad \iff \quad \tilde{\mathbf{X}} = \mathbf{X},$$

d.h. die Matrix \mathbf{X} ist eindeutig bestimmt.

Nach diesen Vorbereitungen führen wir die inverse Matrix ein.

Definition 13.7. (invertierbar; inverse Matrix)

Eine Matrix $\mathbf{A} \in \mathbb{K}^{n \times n}$ heißt *invertierbar*, falls eine Matrix $\mathbf{X} \in \mathbb{K}^{n \times n}$ existiert mit

$$\mathbf{A} \cdot \mathbf{X} = \mathbf{E}_n.$$

(Es gilt dann auch $\mathbf{X} \cdot \mathbf{A} = \mathbf{E}_n$. Die Matrix \mathbf{X} ist *eindeutig bestimmt*.) In diesem Fall heißt \mathbf{X} die *zu \mathbf{A} inverse Matrix*. Sie wird mit \mathbf{A}^{-1} bezeichnet. (Man sagt auch, \mathbf{A}^{-1} ist die inverse Matrix von/zu \mathbf{A} .)

Betrachten wir zunächst einige Beispiele.

Beispiel 13.8. (invertierbar; inverse Matrix)

(a) $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ ist invertierbar mit $\mathbf{A}^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$, denn

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{E}_2.$$

(b) $\mathbf{B} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$ ist invertierbar mit $\mathbf{B}^{-1} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$, denn

$$\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{E}_2.$$

(c) $\mathbf{C} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ ist nicht invertierbar, denn für jedes $\mathbf{X} = \begin{bmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \end{bmatrix} \in \mathbb{K}^{2 \times 2}$ gilt

$$\mathbf{C} \cdot \mathbf{X} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \end{bmatrix} = \begin{bmatrix} x_{2,1} & x_{2,2} \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{E}_2.$$

Wir kommen nun noch einmal auf Hilfssatz 13.6 zurück, denn dieser liefert uns Zusammen mit Satz 12.49 den folgenden wichtigen Satz.

Satz 13.9. (äquivalente Aussagen für invertierbare Matrizen)

Sei $\mathbf{A} \in \mathbb{K}^{n \times n}$. Dann sind die folgenden Aussagen äquivalent:

- (i) \mathbf{A} ist *invertierbar*.
- (ii) $\text{Rang}(\mathbf{A}) = n$.
- (iii) $\mathbf{A} \vec{x} = \vec{b}$ ist für jedes $\vec{b} \in \mathbb{K}^n$ *eindeutig lösbar*. (Es gilt dann $\vec{x} = \mathbf{A}^{-1} \vec{b}$.)
- (iv) Die n linearen Gleichungssysteme $\mathbf{A} \vec{x} = \vec{e}_1$, $\mathbf{A} \vec{x} = \vec{e}_2$, \dots , $\mathbf{A} \vec{x} = \vec{e}_n$ sind *eindeutig lösbar*.
- (v) $\mathbf{A} \vec{x} = \vec{0}$ hat *nur die Lösung* $\vec{x} = \vec{0}$.

Beweis von Satz 13.9: Betrachten wir zunächst Hilfssatz 13.6. Aussage (iii) in Hilfssatz 13.6 besagt nach Definition 13.7 gerade, dass \mathbf{A} invertierbar ist. Also sind nach Hilfssatz 13.6 Aussagen (i), (iii) und (iv) äquivalent. Aus Satz 12.49 folgt weiter, dass (iii) zu (ii) und (v) äquivalent ist. Damit ist gezeigt, dass alle Aussagen in Satz 13.9 äquivalent sind. \square

Ist $\mathbf{A} \in \mathbb{K}^{n \times n}$ invertierbar, so folgt:

$$\mathbf{A} \vec{x} = \vec{b} \quad \iff \quad \underbrace{\mathbf{A}^{-1} \mathbf{A}}_{=\mathbf{E}_n} \vec{x} = \mathbf{A}^{-1} \vec{b} \quad \iff \quad \vec{x} = \mathbf{A}^{-1} \vec{b}$$

Was ist, wenn $\mathbf{A} \in \mathbb{K}^{n \times n}$ nicht invertierbar ist? Mittels Kontraposition (vgl. Anhang B im Skript der HM A) erhält man, dass die Verneinungen der Aussagen in Satz 13.9 äquivalent sind.

Folgerung 13.10. (äquiv. Aussagen für nicht invertierbare Matrizen)

Sei $\mathbf{A} \in \mathbb{K}^{n \times n}$. Dann sind die folgenden Aussagen äquivalent:

- (i) \mathbf{A} ist *nicht invertierbar*.
- (ii) $\text{Rang}(\mathbf{A}) < n$.
- (iii) $\mathbf{A} \vec{x} = \vec{b}$ ist *nicht* für jedes $\vec{b} \in \mathbb{K}^n$ *eindeutig lösbar*. (Es gibt dann mindestens ein $\vec{b} \in \mathbb{K}^n$, für das $\mathbf{A} \vec{x} = \vec{b}$ nicht lösbar oder nicht eindeutig lösbar ist.)

(iv) Die n linearen Gleichungssysteme $\mathbf{A} \vec{x} = \vec{e}_1, \mathbf{A} \vec{x} = \vec{e}_2, \dots, \mathbf{A} \vec{x} = \vec{e}_n$ sind **nicht alle eindeutig lösbar**. (Insbesondere gibt es mindestens ein \vec{e}_k , für das $\mathbf{A} \vec{x} = \vec{e}_k$ nicht lösbar ist.)

(v) $\mathbf{A} \vec{x} = \vec{0}$ hat **unendlich viele weitere Lösungen** außer $\vec{x} = \vec{0}$.

Aus Satz 13.9 ergibt sich ein nützliches **Verfahren zur Berechnung der Inversen \mathbf{A}^{-1} einer invertierbaren Matrix \mathbf{A}** : Wir suchen eine Matrix \mathbf{X} mit $\mathbf{A} \cdot \mathbf{X} = \mathbf{E}_n$. Dieses können wir auch als die **Lösung der folgenden n linearen Gleichungssysteme**

$$\mathbf{A} \vec{x}_1 = \vec{e}_1, \quad \mathbf{A} \vec{x}_2 = \vec{e}_2, \quad \dots, \quad \mathbf{A} \vec{x}_n = \vec{e}_n,$$

für die Spaltenvektoren $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n$ von \mathbf{X} auffassen, denn

$$\mathbf{A} \cdot \mathbf{X} = \mathbf{E}_n \quad \iff \quad \mathbf{A} \cdot [\vec{x}_1 \ \vec{x}_2 \ \dots \ \vec{x}_n] = [\vec{e}_1 \ \vec{e}_2 \ \dots \ \vec{e}_n].$$

Wir behandeln alle n linearen Gleichungssysteme gemeinsam und schreiben dazu die erweiterte Koeffizientenmatrix $[\mathbf{A} \mid \mathbf{E}_n]$. Wenn wir diese erweiterte Koeffizientenmatrix mit elementaren Zeilenoperationen in eine Form bringen, bei der links \mathbf{E}_n steht, also $[\mathbf{E}_n \mid \mathbf{X}]$, so können wir die Inverse $\mathbf{A}^{-1} := \mathbf{X}$ ablesen. Wir halten diese Vorgehensweise als Verfahren zur Berechnung der inversen Matrix fest.

Methode 13.11. (Gauß-Jordan-Verfahren)

Falls $\mathbf{A} \in \mathbb{K}^{n \times n}$ **invertierbar** ist, so lässt sich $[\mathbf{A} \mid \mathbf{E}_n]$ mit elementaren Zeilenoperationen in die Form $[\mathbf{E}_n \mid \mathbf{X}]$ bringen. Die **inverse Matrix** von \mathbf{A} ist dann $\mathbf{A}^{-1} := \mathbf{X}$.

(Falls \mathbf{A} nicht invertierbar ist, so ist es **nicht möglich**, $[\mathbf{A} \mid \mathbf{E}_n]$ mit elementaren Zeilenoperationen in die Form $[\mathbf{E}_n \mid \mathbf{X}]$ zu bringen.)

Beispiel 13.12. (Berechnung der Inversen mit Gauß-Jordan-Verfahren)

(a) Sei $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.

Da keiner der beiden Spaltenvektoren von \mathbf{A} ein Vielfaches des anderen ist, sind die beiden Spaltenvektoren linear unabhängig. Es folgt $\text{Rang}(\mathbf{A}) = 2$, und nach Satz 13.9 ist \mathbf{A} invertierbar. Wir berechnen nun \mathbf{A}^{-1} mit den

Gauß-Jordan-Verfahren:

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array} \right] \xrightarrow[Z_2 \rightarrow Z_2 - 3Z_1]{\downarrow} \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array} \right] \xrightarrow[Z_2 \rightarrow -\frac{1}{2}Z_2]{Z_1 \rightarrow Z_1 + Z_2} \left[\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array} \right]$$

Also gilt $\mathbf{A}^{-1} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$.

(b) Sei $\mathbf{A} = \begin{bmatrix} \frac{3}{2} & 0 & \frac{1}{2} \\ 0 & 3 & 0 \\ \frac{1}{2} & 0 & \frac{3}{2} \end{bmatrix}$.

Man kann sehen, dass die Spaltenvektoren von \mathbf{A} linear unabhängig sind (und dieses rechnet man auch leicht nach), und somit gilt $\text{Rang}(\mathbf{A}) = 3$, und nach Satz 13.9 ist \mathbf{A} invertierbar. Wir berechnen nun \mathbf{A}^{-1} mit den Gauß-Jordan-Verfahren:

$$\left[\begin{array}{ccc|ccc} \frac{3}{2} & 0 & \frac{1}{2} & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{3}{2} & 0 & 0 & 1 \end{array} \right] \xrightarrow[Z_3 \rightarrow 3Z_3]{Z_2 \rightarrow \frac{1}{3}Z_2} \left[\begin{array}{ccc|ccc} \frac{3}{2} & 0 & \frac{1}{2} & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{3} & 0 \\ \frac{3}{2} & 0 & \frac{9}{2} & 0 & 0 & 3 \end{array} \right]$$

$$\xrightarrow[Z_3 \rightarrow Z_3 - Z_1]{Z_1 \rightarrow \frac{2}{3}Z_1} \left[\begin{array}{ccc|ccc} \frac{3}{2} & 0 & \frac{1}{2} & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 4 & -1 & 0 & 3 \end{array} \right] \xrightarrow[Z_3 \rightarrow \frac{1}{4}Z_3]{Z_3 \rightarrow \frac{1}{4}Z_3} \left[\begin{array}{ccc|ccc} 1 & 0 & \frac{1}{3} & \frac{2}{3} & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 & -\frac{1}{4} & 0 & \frac{3}{4} \end{array} \right]$$

$$\xrightarrow[Z_1 \rightarrow Z_1 - \frac{1}{3}Z_3]{Z_1 \rightarrow Z_1 - \frac{1}{3}Z_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{3}{4} & 0 & -\frac{1}{4} \\ 0 & 1 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 & -\frac{1}{4} & 0 & \frac{3}{4} \end{array} \right] \implies \mathbf{A}^{-1} = \begin{bmatrix} \frac{3}{4} & 0 & -\frac{1}{4} \\ 0 & \frac{1}{3} & 0 \\ -\frac{1}{4} & 0 & \frac{3}{4} \end{bmatrix}$$

Somit ist die zu \mathbf{A} inverse Matrix $\mathbf{A}^{-1} = \begin{bmatrix} \frac{3}{4} & 0 & -\frac{1}{4} \\ 0 & \frac{1}{3} & 0 \\ -\frac{1}{4} & 0 & \frac{3}{4} \end{bmatrix}$.

Wir halten abschließend noch einige Resultate über inverse Matrizen fest.

Satz 13.13. (Resultate für invertierbare Matrizen)

(1) Ist $\mathbf{A} \in \mathbb{K}^{n \times n}$ invertierbar, so ist auch ihre inverse Matrix \mathbf{A}^{-1} invertierbar, und es gilt $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.

(2) Sind $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{n \times n}$ invertierbar, so ist auch $\mathbf{A} \cdot \mathbf{B}$ invertierbar, und es gilt

$$(\mathbf{A} \cdot \mathbf{B})^{-1} = \mathbf{B}^{-1} \cdot \mathbf{A}^{-1}.$$

(3) Ist $\mathbf{A} \in \mathbb{K}^{n \times n}$ invertierbar so ist auch \mathbf{A}^T invertierbar, und es gilt $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$.

Bemerkung 13.14. (Transformationsmatrizen und ihre Inversen)

Sei V ein n -dimensionaler \mathbb{K} -Vektorraum und seien B und \tilde{B} Basen von V . Für die Transformationsmatrizen $\mathbf{T}_{\tilde{B}, B}$ und $\mathbf{T}_{B, \tilde{B}}$ gilt dann: $\mathbf{T}_{\tilde{B}, B} \cdot \mathbf{T}_{B, \tilde{B}} = \mathbf{E}_n$. Also ist $\mathbf{T}_{B, \tilde{B}} = (\mathbf{T}_{\tilde{B}, B})^{-1}$ und $\mathbf{T}_{\tilde{B}, B} = (\mathbf{T}_{B, \tilde{B}})^{-1}$.

Zum Abschluss beweisen wir Satz 13.13.

Beweis von Satz 13.13:

(1) Nach Definition der inversen Matrix \mathbf{A}^{-1} von \mathbf{A} gilt

$$\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{E}_n.$$

Daran kann man direkt ablesen, dass \mathbf{A}^{-1} invertierbar ist und dass die inverse Matrix $(\mathbf{A}^{-1})^{-1}$ von \mathbf{A}^{-1} gerade \mathbf{A} ist, also $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.

(2) Wir überprüfen, dass $\mathbf{X} := \mathbf{B}^{-1} \cdot \mathbf{A}^{-1}$ die Eigenschaft $(\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{X} = \mathbf{E}_n$ der inversen Matrix besitzt. Da die inverse Matrix zu $\mathbf{A} \cdot \mathbf{B}$ eindeutig bestimmt ist, folgt dann $(\mathbf{A} \cdot \mathbf{B})^{-1} = \mathbf{B}^{-1} \cdot \mathbf{A}^{-1}$.

$$\begin{aligned} (\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{X} &= (\mathbf{A} \cdot \mathbf{B}) \cdot (\mathbf{B}^{-1} \cdot \mathbf{A}^{-1}) = \mathbf{A} \cdot \underbrace{\mathbf{B} \cdot \mathbf{B}^{-1}}_{=\mathbf{E}_n} \cdot \mathbf{A}^{-1} \\ &= \mathbf{A} \cdot \underbrace{\mathbf{E}_n}_{=\mathbf{A}^{-1}} \cdot \mathbf{A}^{-1} = \mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{E}_n. \end{aligned}$$

Also ist gilt in der Tat $(\mathbf{A} \cdot \mathbf{B})^{-1} = \mathbf{B}^{-1} \cdot \mathbf{A}^{-1}$.

(3) Wir wissen aus Satz 12.46, dass $\text{Rang}(\mathbf{A}^T) = \text{Rang}(\mathbf{A})$ gilt. Ist \mathbf{A} invertierbar, so gilt nach Satz 13.9 $\text{Rang}(\mathbf{A}) = n$ und somit auch $\text{Rang}(\mathbf{A}^T) = n$.

Nach Satz 13.9 ist somit auch \mathbf{A}^T invertierbar. Wir zeigen nun, dass $\mathbf{X} := (\mathbf{A}^{-1})^T$ die Inverse von \mathbf{A}^T ist, indem wir die Eigenschaft $\mathbf{A}^T \cdot \mathbf{X} = \mathbf{E}_n$ nachprüfen:

$$\mathbf{A}^T \cdot \mathbf{X} = \mathbf{A}^T \cdot (\mathbf{A}^{-1})^T = (\mathbf{A}^{-1} \cdot \mathbf{A})^T = \mathbf{E}_n^T = \mathbf{E}_n,$$

wobei $\mathbf{A}^T \cdot (\mathbf{A}^{-1})^T = (\mathbf{A}^{-1} \cdot \mathbf{A})^T$ nach Hilfssatz 11.18 (3) folgt. Also ist $\mathbf{X} = (\mathbf{A}^{-1})^T$ die inverse Matrix zu \mathbf{A}^T , d.h.. $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$. \square

13.3 Die Determinante

In diesem Teilkapitel lernen wir die **Determinante** für beliebige quadratische Matrizen \mathbf{A} in $\mathbb{K}^{n \times n}$ kennen. Wir werden sehen, dass wir mit Hilfe der Determinante sehr leicht überprüfen können, ob eine Matrix invertierbar ist.

Wir erinnern uns zunächst an die Determinante von 2×2 -Matrizen bzw. 3×3 -Matrizen, die wir bereits in Kapitel 3 der HM A kennengelernt haben:

Für $\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}$ gilt:

$$\det(\mathbf{A}) = \det \left(\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \right) = a_{1,1} a_{2,2} - a_{1,2} a_{2,1}. \quad (13.2)$$

Für $\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}$ gilt (Regel von Sarrus):

$$\begin{aligned} \det(\mathbf{A}) &= \det \left(\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} \right) \\ &= a_{1,1} a_{2,2} a_{3,3} + a_{1,2} a_{2,3} a_{3,1} + a_{1,3} a_{2,1} a_{3,2} \\ &\quad - a_{1,3} a_{2,2} a_{3,1} - a_{1,1} a_{2,3} a_{3,2} - a_{1,2} a_{2,1} a_{3,3}. \end{aligned} \quad (13.3)$$

Die Tatsache, dass wir nun Matrizen aus $\mathbb{K}^{2 \times 2}$ bzw. $\mathbb{K}^{3 \times 3}$ betrachten, macht für die obigen beiden Formeln keinen Unterschied; diese gelten auch für Matrizen mit komplexen Einträgen.

Weiter haben wir in Kapitel 3 der HM A im Kontext der Cramerschen Regel bewiesen, dass für $n = 2$ und $n = 3$ gilt: Das lineare Gleichungssystem $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}$

mit $\mathbf{A} \in \mathbb{R}^{n \times n}$ ist **genau dann eindeutig lösbar**, wenn $\det(\mathbf{A}) \neq 0$ ist. Auch die Cramersche Regel für $n = 2$ und $n = 3$ gilt für Matrizen $\mathbf{A} \in \mathbb{C}^{n \times n}$ mit komplexen Einträgen, wenn $\mathbf{A} \vec{x} = \vec{b}$ eindeutig lösbar ist. Wir wissen aus der HM A, dass sich dann die eindeutige Lösung des linearen Gleichungssystems mittels der Cramerschen Regel mit Hilfe der Determinante berechnen lässt.

Zusammen mit Satz 13.9 und dem Wissen über den Rang einer Matrix aus Teilkapitel 12.7 erhalten wir somit **für $\mathbf{A} \in \mathbb{K}^{n \times n}$ mit $n = 2$ oder $n = 3$ die folgenden äquivalenten Aussagen:**

- (i) $\det(\mathbf{A}) \neq 0$
- (ii) \mathbf{A} ist invertierbar.
- (iii) $\text{Rang}(\mathbf{A}) = n$
- (iv) Die Spaltenvektoren von \mathbf{A} sind linear unabhängig.
- (v) Die Zeilenvektoren von \mathbf{A} sind linear unabhängig.
- (vi) Das lineare Gleichungssystem $\mathbf{A} \vec{x} = \vec{b}$ ist für jedes $\vec{b} \in \mathbb{K}^n$ eindeutig lösbar.

Wir werden jetzt $\det(\mathbf{A})$ auch für $n \geq 4$ definieren, und wir werden sehen, dass die obigen Äquivalenzen dann für Matrizen in $\mathbb{K}^{n \times n}$ mit beliebigem n gelten (siehe Satz 13.26 weiter hinten).

Definition 13.15. (Determinante – Laplacescher Entwicklungssatz)

Sei $\mathbf{A} = [a_{i,k}] \in \mathbb{K}^{n \times n}$ mit $n \geq 2$. Dann definieren wir die Determinante $\det(\mathbf{A})$ „rekursiv“ durch **Entwickeln nach einer beliebigen Zeile**: Durch **Entwicklung nach der i -ten Zeile** erhalten wir

$$\det(\mathbf{A}) := \sum_{k=1}^n (-1)^{i+k} a_{i,k} \det(\mathbf{A}_{\widehat{i,k}}). \quad (13.4)$$

Alternativ können wir die Determinante $\det(\mathbf{A})$ „rekursiv“ durch **Entwickeln nach einer beliebigen Spalte** definieren: Durch **Entwicklung nach der k -ten Spalte** erhalten wir

$$\det(\mathbf{A}) := \sum_{i=1}^n (-1)^{i+k} a_{i,k} \det(\mathbf{A}_{\widehat{i,k}}). \quad (13.5)$$

Sowohl in (13.4) als auch in (13.5) ist $\mathbf{A}_{\widehat{i,k}}$ diejenige Matrix in $\mathbb{K}^{(n-1) \times (n-1)}$,

die aus \mathbf{A} durch Streichen der i -ten Zeile und k -ten Spalte entsteht, also

$$\mathbf{A}_{\widehat{i,k}} = \begin{bmatrix} a_{1,1} & \cdots & a_{1,k-1} & a_{1,k+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,k-1} & a_{i-1,k+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,k-1} & a_{i+1,k+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,k-1} & a_{n,k+1} & \cdots & a_{n,n} \end{bmatrix}.$$

Bemerkung 13.16. (Laplacescher Entwicklungssatz)

- (1) Mit der „rekursiven“ **Berechnung** ist Folgendes gemeint: Indem wir die Formel (13.4) bzw. (13.5) erneut anwenden, können wir dann ebenfalls $\det(\mathbf{A}_{\widehat{i,k}})$ berechnen. Wir wiederholen diesen Prozess solange, bis wir irgendwann nach Streichen von Zeilen und Spalten Matrizen in $\mathbb{K}^{n \times n}$ mit $n = 3$ oder $n = 2$ erhalten, deren Determinante wir bereits mit (13.3) bzw. (13.2) berechnen können.
- (2) Wichtig ist die (nicht offensichtliche) Information, dass die Formeln (13.4) und (13.5) für jedes i bzw. k den **gleichen Wert** für die Determinante $\det(\mathbf{A})$ liefern.
- (3) Mit der Definition $\det(\mathbf{A}) := a_{1,1}$ für die Determinante einer 1×1 -Matrix $\mathbf{A} = [a_{1,1}]$ gilt der Laplacesche Entwicklungssatz auch für $n = 2$.
- (4) Für $n = 2$ bzw. $n = 3$ liefert der Laplacesche Entwicklungssatz jeweils die gleiche Formel wie (13.2) bzw. (13.3) (siehe auch Spezialfall 13.18 bzw. Spezialfall 13.19).
- (5) In der Mathematik wird die Determinante eigentlich anders über Permutationen eingeführt. Die Formeln (13.4) und (13.5) treten dann als Satz, der sogenannte Laplacesche Entwicklungssatz, auf. Für Anwender ist es aber sinnvoll, die Determinante mit der Formel aus dem Laplaceschen Entwicklungssatz zu definieren.

Betrachten wir zunächst vier Beispiele.

Beispiel 13.17. (Determinante)

(a) Wir wollen die Determinante der Matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

mit dem Laplaceschen Entwicklungssatz berechnen. Wir entwickeln nach der ersten Zeile und erhalten

$$\begin{aligned} \det(\mathbf{A}) &= (-1)^{1+1} \cdot 1 \cdot \det \left(\begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix} \right) + (-1)^{1+2} \cdot 2 \cdot \det \left(\begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix} \right) \\ &\quad + (-1)^{1+3} \cdot 3 \cdot \det \left(\begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix} \right) \\ &= (5 \cdot 9 - 6 \cdot 8) - 2 \cdot (4 \cdot 9 - 6 \cdot 7) + 3 \cdot (4 \cdot 8 - 5 \cdot 7) \\ &= (45 - 48) - 2 \cdot (36 - 42) + 3 \cdot (32 - 35) \\ &= -3 - 2 \cdot (-6) + 3 \cdot (-3) \\ &= -3 + 12 - 9 = 0. \end{aligned}$$

(b) Wir wollen die Determinante der Matrix

$$\mathbf{B} = \begin{bmatrix} -1 & 4 & 2 & 1 \\ 2 & 1 & -2 & 1 \\ 1 & 0 & 0 & 3 \\ 3 & -1 & 2 & -1 \end{bmatrix}$$

mit dem Laplaceschen Entwicklungssatz berechnen. Wir dürfen uns dabei aussuchen, nach welcher Spalte oder Zeile wir entwickeln. Da in der dritten Zeile zwei Einträge null sind, ist die Entwicklung nach der dritten Zeile vom Rechenaufwand her besonders günstig!

$$\begin{aligned} \det(\mathbf{B}) &= (-1)^{3+1} \cdot 1 \cdot \det \left(\begin{bmatrix} 4 & 2 & 1 \\ 1 & -2 & 1 \\ -1 & 2 & -1 \end{bmatrix} \right) \\ &\quad + 0 + 0 + (-1)^{3+4} \cdot 3 \cdot \det \left(\begin{bmatrix} -1 & 4 & 2 \\ 2 & 1 & -2 \\ 3 & -1 & 2 \end{bmatrix} \right). \end{aligned}$$

Mit der Regel von Sarrus erhalten wir

$$\begin{aligned} \det \begin{pmatrix} \begin{bmatrix} 4 & 2 & 1 \\ 1 & -2 & 1 \\ -1 & 2 & -1 \end{bmatrix} \end{pmatrix} &= 4 \cdot (-2) \cdot (-1) + 2 \cdot 1 \cdot (-1) + 1 \cdot 1 \cdot 2 \\ &\quad - 1 \cdot (-2) \cdot (-1) - 4 \cdot 1 \cdot 2 - 2 \cdot 1 \cdot (-1) \\ &= 8 - 2 + 2 - 2 - 8 + 2 = 0, \end{aligned}$$

$$\begin{aligned} \det \begin{pmatrix} \begin{bmatrix} -1 & 4 & 2 \\ 2 & 1 & -2 \\ 3 & -1 & 2 \end{bmatrix} \end{pmatrix} &= (-1) \cdot 1 \cdot 2 + 4 \cdot (-2) \cdot 3 + 2 \cdot 2 \cdot (-1) \\ &\quad - 2 \cdot 1 \cdot 3 - (-1) \cdot (-2) \cdot (-1) - 4 \cdot 2 \cdot 2 \\ &= -2 - 24 - 4 - 6 + 2 - 16 = -50. \end{aligned}$$

Also finden wir

$$\det(\mathbf{B}) = (-1)^{3+1} \cdot 1 \cdot 0 + (-1)^{3+4} \cdot 3 \cdot (-50) = 0 + (-3) \cdot (-50) = 150.$$

(c) Wir wollen die Determinante der Matrix

$$\mathbf{C} = \begin{bmatrix} 2 & 1 & 1 & 4 \\ 1 & 0 & -1 & 2 \\ 4 & 3 & 0 & 1 \\ -1 & 2 & 3 & 0 \end{bmatrix}$$

berechnen und entwickeln dazu nach der dritten Spalte:

$$\begin{aligned} &\det(\mathbf{C}) \\ &= (-1)^{1+3} \cdot 1 \cdot \det \begin{pmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 4 & 3 & 1 \\ -1 & 2 & 0 \end{bmatrix} \end{pmatrix} + (-1)^{2+3} \cdot (-1) \cdot \det \begin{pmatrix} \begin{bmatrix} 2 & 1 & 4 \\ 4 & 3 & 1 \\ -1 & 2 & 0 \end{bmatrix} \end{pmatrix} \\ &\quad + (-1)^{3+3} \cdot 0 \cdot \det \begin{pmatrix} \begin{bmatrix} 2 & 1 & 4 \\ 1 & 0 & 2 \\ -1 & 2 & 0 \end{bmatrix} \end{pmatrix} + (-1)^{4+3} \cdot 3 \cdot \det \begin{pmatrix} \begin{bmatrix} 2 & 1 & 4 \\ 1 & 0 & 2 \\ 4 & 3 & 1 \end{bmatrix} \end{pmatrix} \\ &= 1 \cdot (0 + 0 + 16 + 6 - 2 - 0) + 1 \cdot (0 - 1 + 32 + 12 - 4 - 0) \\ &\quad + 0 + (-3) \cdot (0 + 8 + 12 - 0 - 12 - 1) \\ &= 20 + 39 - 21 = 38. \end{aligned}$$

(d) Die Determinante der Matrix

$$\mathbf{D} = \begin{bmatrix} \frac{3}{2} & 0 & \frac{1}{2} \\ 0 & 3 & 0 \\ \frac{1}{2} & 0 & \frac{3}{2} \end{bmatrix}$$

berechnet man am einfachsten durch Entwicklung nach der zweiten Spalte (oder der zweiten Zeile):

$$\det(\mathbf{D}) = (-1)^{2+2} \cdot 3 \cdot \det \left(\begin{bmatrix} \frac{3}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{3}{2} \end{bmatrix} \right) = 3 \cdot \left(\frac{9}{4} - \frac{1}{4} \right) = 3 \cdot 2 = 6.$$

Als weiteres Beispiel und als Anwendung des Laplaceschen Entwicklungssatzes wollen wir die Formeln für die Determinante einer 2×2 -Matrix und einer 3×3 -Matrix mit dem Laplaceschen Entwicklungssatz herleiten. Daran sehen wir, dass (13.2) und (13.3) nur Sonderfälle der allgemeinen Definition 13.15 sind.

Spezialfall 13.18. (Determinante einer 2×2 -Matrix)

Nach dem Laplaceschen Entwicklungssatz finden wir für die Determinante der beliebigen 2×2 -Matrix

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}$$

bei Entwicklung nach der ersten Zeile

$$\begin{aligned} \det(\mathbf{A}) &= \det \left(\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \right) = (-1)^{1+1} a_{1,1} \det(\mathbf{A}_{\widehat{1,1}}) + (-1)^{1+2} a_{1,2} \det(\mathbf{A}_{\widehat{1,2}}) \\ &= a_{1,1} \det(\mathbf{A}_{\widehat{1,1}}) - a_{1,2} \det(\mathbf{A}_{\widehat{1,2}}). \end{aligned} \quad (13.6)$$

Die beiden Matrizen $\mathbf{A}_{\widehat{1,1}}$ und $\mathbf{A}_{\widehat{1,2}}$ erhält man jeweils durch das Streichen der ersten Zeile und der ersten bzw. zweiten Spalte; also gilt

$$\mathbf{A}_{\widehat{1,1}} = [a_{2,2}] \quad \text{und} \quad \mathbf{A}_{\widehat{1,2}} = [a_{2,1}].$$

Definieren wir nun die Determinante einer 1×1 -Matrix, also einer reellen oder komplexen Zahl, als diese Zahl selber (vgl. auch Bemerkung 13.16 (3)), so haben wir

$$\det(\mathbf{A}_{\widehat{1,1}}) = a_{2,2} \quad \text{und} \quad \det(\mathbf{A}_{\widehat{1,2}}) = a_{2,1}.$$

Einsetzen in (13.6) liefert

$$\det \left(\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \right) = a_{1,1} a_{2,2} - a_{1,2} a_{2,1},$$

und dieses ist gerade die Formel (13.2).

Spezialfall 13.19. (Determinante einer 3×3 -Matrix)

Wir wollen nun die Regel von Sarrus (13.3) mit Hilfe des Laplaceschen Entwicklungssatzes herleiten. Für eine beliebige 3×3 -Matrix

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}$$

finden wir durch Entwicklung nach der ersten Zeile

$$\begin{aligned} \det(\mathbf{A}) &= \left(\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} \right) = (-1)^{1+1} a_{1,1} \det \left(\begin{bmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{bmatrix} \right) \\ &\quad + (-1)^{1+2} a_{1,2} \det \left(\begin{bmatrix} a_{2,1} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{bmatrix} \right) + (-1)^{1+3} a_{1,3} \det \left(\begin{bmatrix} a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{bmatrix} \right) \\ &= a_{1,1} (a_{2,2} a_{3,3} - a_{2,3} a_{3,2}) - a_{1,2} (a_{2,1} a_{3,3} - a_{2,3} a_{3,1}) \\ &\quad + a_{1,3} (a_{2,1} a_{3,2} - a_{2,2} a_{3,1}) \\ &= a_{1,1} a_{2,2} a_{3,3} - a_{1,1} a_{2,3} a_{3,2} - a_{1,2} a_{2,1} a_{3,3} \\ &\quad + a_{1,2} a_{2,3} a_{3,1} + a_{1,3} a_{2,1} a_{3,2} - a_{1,3} a_{2,2} a_{3,1} \\ &= a_{1,1} a_{2,2} a_{3,3} + a_{1,2} a_{2,3} a_{3,1} + a_{1,3} a_{2,1} a_{3,2} \\ &\quad - a_{1,3} a_{2,2} a_{3,1} - a_{1,1} a_{2,3} a_{3,2} - a_{1,2} a_{2,1} a_{3,3}, \end{aligned}$$

und wir erhalten in der Tat die Regel von Sarrus (13.3).

Als Letztes untersuchen wir den Fall von unteren bzw. oberen Dreiecksmatrizen.

Hilfssatz 13.20. (Determinante von unteren bzw. oberen Dreiecksmatrizen)

Sei $A = [a_{i,k}] \in \mathbb{K}^{n \times n}$ mit $n \geq 2$ eine *untere Dreiecksmatrix* bzw. eine

obere Dreiecksmatrix, also

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & 0 & \cdots & 0 \\ a_{2,1} & a_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ a_{n,1} & \cdots & a_{n,n-1} & a_{n,n} \end{bmatrix} \quad \text{bzw.} \quad \mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ 0 & a_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ 0 & \cdots & 0 & a_{n,n} \end{bmatrix}.$$

Dann gilt

$$\det(\mathbf{A}) = a_{1,1} \cdot a_{2,2} \cdot \dots \cdot a_{n,n}, \quad (13.7)$$

d.h. die Determinante ist das **Produkt der Einträge auf der Diagonalen** (von links oben nach rechts unten) der unteren bzw. oberen Dreiecksmatrix.

Betrachten wir einige Beispiele für die Anwendung von Hilfssatz 13.20.

Beispiel 13.21. (Determinante von unteren/oberen Dreiecksmatrizen)

Gegeben seien die Matrizen

$$\mathbf{E}_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ -3 & -5 & 0 & 0 \\ 13 & -19 & 4 & 0 \\ -7 & 1017 & 23 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -1 & 3 & -7 & 9 \\ 0 & 100 & 2 & -13 \\ 0 & 0 & 3 & -11 \\ 0 & 0 & 0 & 5 \end{bmatrix}.$$

Die Matrizen \mathbf{A} und \mathbf{E}_n sind untere Dreiecksmatrizen, und die Matrix \mathbf{B} ist eine obere Dreiecksmatrix. Also können wir Hilfssatz 13.20 anwenden und finden

$$\begin{aligned} \det(\mathbf{E}_n) &= \underbrace{1 \cdot 1 \cdot \dots \cdot 1}_{n\text{-mal}} = 1, \\ \det(\mathbf{A}) &= 2 \cdot (-5) \cdot 4 \cdot 1 = -40, \\ \det(\mathbf{B}) &= (-1) \cdot 100 \cdot 3 \cdot 5 = -1500. \end{aligned}$$

Zuletzt beweisen wir Hilfssatz 13.20 mit Hilfe des Laplaceschen Entwicklungssatzes.

Beweis von Hilfssatz 13.20: Wir geben den Beweis mit vollständiger Induktion über n für $n \geq 2$, und wir betrachten nur obere Dreiecksmatrizen. Der Fall unterer Dreiecksmatrizen kann analog behandelt werden.

(IA) $n = 2$: Sei \mathbf{A} eine beliebige obere Dreiecksmatrix in $\mathbb{K}^{2 \times 2}$, also

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} \\ 0 & a_{2,2} \end{bmatrix}.$$

Nach der Formel (13.2) für die Determinante einer 2×2 -Matrix finden wir

$$\det(\mathbf{A}) = a_{1,1} \cdot a_{2,2} - a_{1,2} \cdot 0 = a_{1,1} \cdot a_{2,2}. \quad \checkmark$$

(IS) $n \rightsquigarrow n + 1$: Sei \mathbf{A} eine obere Dreiecksmatrix in $\mathbb{K}^{(n+1) \times (n+1)}$, also

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & \cdots & \cdots & a_{1,n+1} \\ 0 & a_{2,2} & \cdots & a_{2,n+1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n+1,n+1} \end{bmatrix}.$$

Dann entwickeln wir die Determinante von \mathbf{A} nach der ersten Spalte und erhalten

$$\det(\mathbf{A}) = \underbrace{(-1)^{1+1}}_{=1} a_{1,1} \cdot \det \left(\underbrace{\begin{bmatrix} a_{2,2} & \cdots & \cdots & a_{2,n+1} \\ 0 & a_{3,3} & \cdots & a_{3,n+1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n+1,n+1} \end{bmatrix}}_{n \times n\text{-Matrix}} \right). \quad (13.8)$$

Nach der Induktionsvoraussetzung finden wir für die Determinante der verbleibenden $n \times n$ -oberen Dreiecksmatrix

$$\det \left(\begin{bmatrix} a_{2,2} & \cdots & \cdots & a_{2,n+1} \\ 0 & a_{3,3} & \cdots & a_{3,n+1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n+1,n+1} \end{bmatrix} \right) = a_{2,2} \cdot a_{3,3} \cdot \cdots \cdot a_{n+1,n+1}.$$

Einsetzen in (13.8) liefert nun

$$\det(\mathbf{A}) = a_{1,1} \cdot a_{2,2} \cdot a_{3,3} \cdot \cdots \cdot a_{n+1,n+1},$$

und wir haben bewiesen, dass die Formel (13.7) auch für $n + 1$ gilt. \square

13.4 Rechenregeln für Determinanten

In diesem Teilkapitel lernen wir wichtige Rechenregeln für Determinanten kennen.

Satz 13.22. (Rechenregeln für Determinanten – Teil I)

(1) $\det(\mathbf{A} \cdot \mathbf{B}) = \det(\mathbf{A}) \cdot \det(\mathbf{B})$ für alle $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{n \times n}$.

(2) Falls \mathbf{A} invertierbar ist, gilt $\det(\mathbf{A}) \neq 0$, und

$$\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})}.$$

(3) $\det(\mathbf{A}^T) = \det(\mathbf{A})$ für alle $\mathbf{A} \in \mathbb{K}^{n \times n}$.

Beweis von Satz 13.22:

(1) Diese Formel werden wir nicht beweisen.

(2) Ist \mathbf{A} invertierbar, so gilt $\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{E}_n$ und nach Rechenregel (1) folgt

$$1 = \det(\mathbf{E}_n) = \det(\mathbf{A} \cdot \mathbf{A}^{-1}) = \det(\mathbf{A}) \cdot \det(\mathbf{A}^{-1}).$$

Daraus folgt insbesondere, dass $\det(\mathbf{A})$ und $\det(\mathbf{A}^{-1})$ beide ungleich null sein müssen (sonst wäre die rechte Seite gleich null), also $\det(\mathbf{A}) \neq 0$ und $\det(\mathbf{A}^{-1}) \neq 0$. Division durch $\det(\mathbf{A}) \neq 0$ liefert nun die gesuchte Formel.

(3) Diese Formel kann man mit vollständiger Induktion über $n \geq 2$ beweisen, indem man $\det(\mathbf{A})$ nach der ersten Zeile entwickelt und $\det(\mathbf{A}^T)$ nach der ersten Spalte entwickelt. \square

Betrachten wir einige Beispiele.

Beispiel 13.23. (Rechenregeln für Determinanten – Teil I)

(a) In Beispiel 13.17 (a) haben wir berechnet, dass für die Matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

gilt $\det(\mathbf{A}) = 0$. Nach Satz 13.22 (3) folgt nun für die zu \mathbf{A} transponierte Matrix

$$\mathbf{A}^T = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix},$$

dass gilt $\det(\mathbf{A}^T) = \det(\mathbf{A}) = 0$.

(b) Wir wollen die Determinante der Matrix

$$\mathbf{B} = \begin{bmatrix} \frac{3}{4} & 0 & -\frac{1}{4} \\ 0 & \frac{1}{3} & 0 \\ -\frac{1}{4} & 0 & \frac{3}{4} \end{bmatrix}$$

berechnen. Nach Beispiel 13.12 (b) wissen wir, dass \mathbf{B} die inverse Matrix der Matrix

$$\mathbf{D} = \begin{bmatrix} \frac{3}{2} & 0 & \frac{1}{2} \\ 0 & 3 & 0 \\ \frac{1}{2} & 0 & \frac{3}{2} \end{bmatrix}$$

ist. Deren Determinante haben wir bereits in Beispiel 13.17 (d) berechnet und fanden $\det(\mathbf{D}) = 6$. Also gilt nach Satz 13.22 (2) wegen $\mathbf{D} = \mathbf{B}^{-1}$ und $(\mathbf{B}^{-1})^{-1} = \mathbf{B}$

$$\det(\mathbf{B}) = \frac{1}{\det(\mathbf{B}^{-1})} = \frac{1}{\det(\mathbf{D})} = \frac{1}{6}.$$

(c) Die Determinante der Matrix

$$\begin{bmatrix} 0 & \frac{2}{3} & 2 \\ \frac{3}{2} & \frac{5}{3} & \frac{7}{2} \\ 3 & \frac{8}{3} & 5 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \cdot \begin{bmatrix} \frac{3}{4} & 0 & -\frac{1}{4} \\ 0 & \frac{1}{3} & 0 \\ -\frac{1}{4} & 0 & \frac{3}{4} \end{bmatrix} = \mathbf{A} \cdot \mathbf{B}$$

ist nach Satz 13.22 (1) und Beispielen (a) und (b)

$$\det \left(\begin{bmatrix} 0 & \frac{2}{3} & 2 \\ \frac{3}{2} & \frac{5}{3} & \frac{7}{2} \\ 3 & \frac{8}{3} & 5 \end{bmatrix} \right) = \det(\mathbf{A} \cdot \mathbf{B}) = \det(\mathbf{A}) \cdot \det(\mathbf{B}) = 0 \cdot \frac{1}{6} = 0.$$

Zur Motivation des nächsten Satzes überprüfen wir einige wichtige Eigenschaften der Determinante für den Fall $n = 2$. In Folgenden sei also $\mathbf{A} \in \mathbb{K}^{2 \times 2}$ immer

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}.$$

- Vertauscht man zwei Zeilen (oder Spalten) von A , so ändert sich das Vorzeichen der Determinante:

$$\det \left(\begin{bmatrix} a_{2,1} & a_{2,2} \\ a_{1,1} & a_{1,2} \end{bmatrix} \right) = a_{2,1} a_{1,2} - a_{2,2} a_{1,1} = -[a_{1,1} a_{2,2} - a_{1,2} a_{2,1}] = -\det(\mathbf{A}).$$

- Einen gemeinsamen Faktor $\lambda \in \mathbb{K}$ einer Zeile (oder Spalte) darf man aus $\det(\mathbf{A})$ herausziehen:

$$\begin{aligned} \det \left(\begin{bmatrix} \lambda a_{1,1} & \lambda a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \right) &= \lambda a_{1,1} a_{2,2} - \lambda a_{1,2} a_{2,1} \\ &= \lambda (a_{1,1} a_{2,2} - a_{1,2} a_{2,1}) = \lambda \det(\mathbf{A}). \end{aligned}$$

Insbesondere gilt $\det(\lambda \mathbf{A}) = \lambda^2 \det(\mathbf{A})$.

- Wird zu einer Zeile (oder Spalte) von \mathbf{A} ein Zeilenvektor $\vec{\mathbf{b}}^T = [b_1 \ b_2]$ (bzw. ein Spaltenvektor $\vec{\mathbf{c}} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$) addiert, so gilt

$$\begin{aligned} \det \left(\begin{bmatrix} a_{1,1} + b_1 & a_{1,2} + b_2 \\ a_{2,1} & a_{2,2} \end{bmatrix} \right) &= (a_{1,1} + b_1) a_{2,2} - (a_{1,2} + b_2) a_{2,1} \\ &= (a_{1,1} a_{2,2} - a_{1,2} a_{2,1}) + (b_1 a_{2,2} - b_2 a_{2,1}) \\ &= \det \left(\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \right) + \det \left(\begin{bmatrix} b_1 & b_2 \\ a_{2,1} & a_{2,2} \end{bmatrix} \right), \end{aligned}$$

$$\begin{aligned} \det \left(\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} + b_1 & a_{2,2} + b_2 \end{bmatrix} \right) &= a_{1,1} (a_{2,2} + b_2) - a_{1,2} (a_{2,1} + b_1) \\ &= (a_{1,1} a_{2,2} - a_{1,2} a_{2,1}) + (a_{1,1} b_2 - a_{1,2} b_1) \\ &= \det \left(\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \right) + \det \left(\begin{bmatrix} a_{1,1} & a_{1,2} \\ b_1 & b_2 \end{bmatrix} \right), \end{aligned}$$

und analog für den Fall mit Spalten statt Zeilen.

- Addiert man das μ -fache einer Zeile zu einer **anderen** Zeile, so ändert sich die Determinante nicht:

$$\begin{aligned} \det \left(\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} + \mu a_{1,1} & a_{2,2} + \mu a_{1,2} \end{bmatrix} \right) &= a_{1,1} (a_{2,2} + \mu a_{1,2}) - a_{1,2} (a_{2,1} + \mu a_{1,1}) \\ &= a_{1,1} a_{2,2} - a_{1,2} a_{2,1} = \det(\mathbf{A}), \end{aligned}$$

$$\begin{aligned} \det \left(\begin{bmatrix} a_{1,1} + \mu a_{2,1} & a_{1,2} + \mu a_{2,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \right) &= (a_{1,1} + \mu a_{2,1}) a_{2,2} - (a_{1,2} + \mu a_{2,2}) a_{2,1} \\ &= a_{1,1} a_{2,2} - a_{1,2} a_{2,1} = \det(\mathbf{A}). \end{aligned}$$

Analoges gilt für die Spalten: Addiert man das μ -fache einer Spalte zu einer **anderen** Spalte, so ändert sich die Determinante nicht.

- Nach den Vorüberlegungen zu Beginn von Kapitel 13.3 gilt für $\mathbf{A} \in \mathbb{K}^{2 \times 2}$:

$$\det(\mathbf{A}) = 0 \iff \text{Rang}(\mathbf{A}) < 2 \iff \left(\begin{array}{l} \text{Die Spaltenvektoren von} \\ \mathbf{A} \text{ sind linear abhängig.} \end{array} \right)$$

Die Rechenregeln, die wir oben für die Determinante einer 2×2 -Matrix überprüft haben, gelten auch für Matrizen in $\mathbb{K}^{n \times n}$ mit beliebigem $n \geq 2$.

Satz 13.24. (Rechenregeln für Determinanten – Teil II)

Sei $\mathbf{A} \in \mathbb{K}^{n \times n}$. Zur Formulierung der nachfolgenden Rechenregeln für die Determinante schreiben wir \mathbf{A} auch als

$$\mathbf{A} = [\vec{s}_1 \quad \vec{s}_2 \quad \cdots \quad \vec{s}_n] = \begin{bmatrix} \vec{z}_1^T \\ \vec{z}_2^T \\ \vdots \\ \vec{z}_n^T \end{bmatrix}$$

mit den Spaltenvektoren $\vec{s}_1, \vec{s}_2, \dots, \vec{s}_n$ bzw. den Zeilenvektoren $\vec{z}_1^T, \vec{z}_2^T, \dots, \vec{z}_n^T$ von \mathbf{A} . Dann gelten die folgenden Rechenregeln:

(4) Beim Austausch zweier Zeilen bzw. Spalten von \mathbf{A} wechselt $\det(\mathbf{A})$ das Vorzeichen.

(5) Für jedes $\lambda \in \mathbb{K}$ und jedes $i, k \in \{1, 2, \dots, n\}$ gilt

$$\det \left(\begin{bmatrix} \vec{z}_1^T \\ \vdots \\ \vec{z}_{i-1}^T \\ \lambda \cdot \vec{z}_i^T \\ \vec{z}_{i+1}^T \\ \vdots \\ \vec{z}_n^T \end{bmatrix} \right) = \lambda \cdot \det \left(\begin{bmatrix} \vec{z}_1^T \\ \vdots \\ \vec{z}_{i-1}^T \\ \vec{z}_i^T \\ \vec{z}_{i+1}^T \\ \vdots \\ \vec{z}_n^T \end{bmatrix} \right) = \lambda \cdot \det(\mathbf{A}),$$

$$\begin{aligned} & \det([\vec{s}_1 \quad \cdots \quad \vec{s}_{k-1} \quad \lambda \cdot \vec{s}_k \quad \vec{s}_{k+1} \quad \cdots \quad \vec{s}_n]) \\ &= \lambda \cdot \det([\vec{s}_1 \quad \cdots \quad \vec{s}_{k-1} \quad \vec{s}_k \quad \vec{s}_{k+1} \quad \cdots \quad \vec{s}_n]) = \lambda \cdot \det(\mathbf{A}). \end{aligned}$$

In Worten: Einen gemeinsamen Faktor $\lambda \in \mathbb{K}$ einer Zeile (oder Spalte) darf man aus der Determinante herausziehen.

(6) Sei $\vec{\mathbf{b}}^T \in \mathbb{K}^{1 \times n}$ ein Zeilenvektor und $\vec{\mathbf{c}} \in \mathbb{K}^n$ ein Spaltenvektor. Dann gilt für jedes $i = 1, 2, \dots, n$

$$\det \begin{pmatrix} \begin{bmatrix} \vec{\mathbf{z}}_1^T \\ \vdots \\ \vec{\mathbf{z}}_{i-1}^T \\ \vec{\mathbf{z}}_i^T + \vec{\mathbf{b}}^T \\ \vec{\mathbf{z}}_{i+1}^T \\ \vdots \\ \vec{\mathbf{z}}_n^T \end{bmatrix} \end{pmatrix} = \det \begin{pmatrix} \begin{bmatrix} \vec{\mathbf{z}}_1^T \\ \vdots \\ \vec{\mathbf{z}}_{i-1}^T \\ \vec{\mathbf{z}}_i^T \\ \vec{\mathbf{z}}_{i+1}^T \\ \vdots \\ \vec{\mathbf{z}}_n^T \end{bmatrix} \end{pmatrix} + \det \begin{pmatrix} \begin{bmatrix} \vec{\mathbf{z}}_1^T \\ \vdots \\ \vec{\mathbf{z}}_{i-1}^T \\ \vec{\mathbf{b}}^T \\ \vec{\mathbf{z}}_{i+1}^T \\ \vdots \\ \vec{\mathbf{z}}_n^T \end{bmatrix} \end{pmatrix},$$

und für jedes $k = 1, 2, \dots, n$ gilt

$$\begin{aligned} & \det \left([\vec{\mathbf{s}}_1 \ \cdots \ \vec{\mathbf{s}}_{k-1} \ \vec{\mathbf{s}}_k + \vec{\mathbf{c}} \ \vec{\mathbf{s}}_{k+1} \ \cdots \ \vec{\mathbf{s}}_n] \right) \\ &= \det \left([\vec{\mathbf{s}}_1 \ \cdots \ \vec{\mathbf{s}}_{k-1} \ \vec{\mathbf{s}}_k \ \vec{\mathbf{s}}_{k+1} \ \cdots \ \vec{\mathbf{s}}_n] \right) \\ & \quad + \det \left([\vec{\mathbf{s}}_1 \ \cdots \ \vec{\mathbf{s}}_{k-1} \ \vec{\mathbf{c}} \ \vec{\mathbf{s}}_{k+1} \ \cdots \ \vec{\mathbf{s}}_n] \right). \end{aligned}$$

(7) Sei $\mu \in \mathbb{K}$. Die Addition des μ -fachen der i -ten Zeile (oder Spalte) von \mathbf{A} zur k -ten Zeile (oder Spalte) von \mathbf{A} ändert den Wert von $\det(\mathbf{A})$ nicht, wenn $i \neq k$ ist.

(8) Sind zwei Zeilenvektoren bzw. zwei Spaltenvektoren von \mathbf{A} **linear abhängig**, so gilt $\det(\mathbf{A}) = 0$. Insbesondere: Wenn eine Zeile (oder Spalte) von \mathbf{A} der Nullvektor ist, so gilt $\det(\mathbf{A}) = 0$.

Die Regeln (4) bis (8) aus Satz 13.24 kann man benutzen, um die Determinante einer Matrix konkret zu berechnen: Bei (4), (5) und (7) handelt es sich um **elementare Zeilenoperationen** bzw. um **elementare Spaltenoperationen**, mit Hilfe derer man die Matrix, deren Determinante berechnet werden soll, vereinfachen kann. Um elementare Zeilenoperationen oder elementare Spaltenoperationen an der Matrix anzugeben, nutzen wir wieder die in Teilkapitel 11.3 eingeführte Notation. Die Spalten der Matrix bezeichnen wir dabei mit S_1, S_2, \dots, S_n und notieren elementare Spaltenoperationen analog zu den elementaren Zeilenoperationen.

Natürlich kann man die Regeln (4), (5) und (7) aus Satz 13.24 auch zusammen mit dem Entwicklungssatz von Laplace verwenden, indem man die Matrix zunächst geeignet mit den Regeln (4), (5) und (7) vereinfacht und danach nach einer geeigneten Zeile oder Spalte entwickelt.

Wir üben dieses an einigen Beispielen.

Beispiel 13.25. (Berechnung der Determinante)

- (a) Wir verwenden elementare Zeilenoperationen (d.h. die Regeln (4), (5) und (7) aus Satz 13.24), um die Matrix

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 & 4 \\ 1 & 0 & -1 & 2 \\ 4 & 3 & 0 & 1 \\ -1 & 2 & 3 & 0 \end{bmatrix}$$

auf obere Dreiecksgestalt zu bringen und anschließend ihre Determinante leicht mit Hilfssatz 13.20 berechnen zu können.

$$\begin{aligned} \det(\mathbf{A}) &= \det \left(\begin{bmatrix} 2 & 1 & 1 & 4 \\ 1 & 0 & -1 & 2 \\ 4 & 3 & 0 & 1 \\ -1 & 2 & 3 & 0 \end{bmatrix} \right) \stackrel{\substack{Z_1 \rightarrow Z_1 - 2Z_2 \\ Z_3 \rightarrow Z_3 - 4Z_2 \\ Z_4 \rightarrow Z_4 + Z_2}}{\downarrow} \det \left(\begin{bmatrix} 0 & 1 & 3 & 0 \\ 1 & 0 & -1 & 2 \\ 0 & 3 & 4 & -7 \\ 0 & 2 & 2 & 2 \end{bmatrix} \right) \\ &\stackrel{\substack{Z_1 \leftrightarrow Z_2 \\ Z_4 \rightarrow \frac{1}{2}Z_4}}{\downarrow} -2 \cdot \det \left(\begin{bmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & 3 & 0 \\ 0 & 3 & 4 & -7 \\ 0 & 1 & 1 & 1 \end{bmatrix} \right) \stackrel{\substack{Z_2 \rightarrow Z_2 - Z_4 \\ Z_3 \rightarrow Z_3 - 3Z_4}}{\downarrow} -2 \cdot \det \left(\begin{bmatrix} 1 & 0 & -1 & 2 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 1 & -10 \\ 0 & 1 & 1 & 1 \end{bmatrix} \right) \\ &\stackrel{\substack{Z_2 \leftrightarrow Z_4}}{\downarrow} 2 \cdot \det \left(\begin{bmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & -10 \\ 0 & 0 & 2 & -1 \end{bmatrix} \right) \stackrel{Z_4 \rightarrow Z_4 - 2Z_3}{\downarrow} 2 \cdot \det \left(\begin{bmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & -10 \\ 0 & 0 & 0 & 19 \end{bmatrix} \right) \\ &= 2 \cdot 1 \cdot 1 \cdot 1 \cdot 19 = 38, \end{aligned}$$

wobei wir im letzten Schritt Hilfssatz 13.20 benutzt haben.

Wir hätten die Determinante auch wie folgt berechnen können

$$\det(\mathbf{A}) = \det \left(\begin{bmatrix} 2 & 1 & 1 & 4 \\ 1 & 0 & -1 & 2 \\ 4 & 3 & 0 & 1 \\ -1 & 2 & 3 & 0 \end{bmatrix} \right) \stackrel{\substack{Z_1 \rightarrow Z_1 - 2Z_2 \\ Z_3 \rightarrow Z_3 - 4Z_2 \\ Z_4 \rightarrow Z_4 + Z_2}}{\downarrow} \det \left(\begin{bmatrix} 0 & 1 & 3 & 0 \\ 1 & 0 & -1 & 2 \\ 0 & 3 & 4 & -7 \\ 0 & 2 & 2 & 2 \end{bmatrix} \right)$$

$$\begin{array}{l} \text{Entwicklung} \\ \text{nach der} \\ \text{1. Spalte} \\ \downarrow \\ \cong \end{array} 0 + (-1)^{2+1} \cdot 1 \cdot \det \left(\begin{bmatrix} 1 & 3 & 0 \\ 3 & 4 & -7 \\ 2 & 2 & 2 \end{bmatrix} \right) + 0 + 0$$

$$\begin{array}{l} Z_3 \rightarrow \frac{1}{2} Z_3 \\ \downarrow \\ \cong \end{array} -2 \cdot \det \left(\begin{bmatrix} 1 & 3 & 0 \\ 3 & 4 & -7 \\ 1 & 1 & 1 \end{bmatrix} \right) = -2 \cdot (4 - 21 + 0 - 0 + 7 - 9) = 38.$$

- (b) Wir verwenden elementare Spalten- und Zeilenoperationen (d.h. die Regeln (4), (5) und (7) aus Satz 13.24) und gegebenenfalls den Laplaceschen Entwicklungssatz, um die Determinante der Matrix

$$\mathbf{B} = \begin{bmatrix} 4 & 3 & 0 & 1 \\ 9 & 7 & 2 & 3 \\ 4 & 0 & 2 & 1 \\ 3 & -1 & 4 & 0 \end{bmatrix}$$

zu berechnen. Wir gehen wie folgt vor (es gibt viele andere Varianten die zum Ziel führen):

$$\det(\mathbf{B}) = \det \left(\begin{bmatrix} 4 & 3 & 0 & 1 \\ 9 & 7 & 2 & 3 \\ 4 & 0 & 2 & 1 \\ 3 & -1 & 4 & 0 \end{bmatrix} \right) \begin{array}{l} S_3 \rightarrow \frac{1}{2} S_3 \\ \downarrow \\ \cong \end{array} 2 \cdot \det \left(\begin{bmatrix} 4 & 3 & 0 & 1 \\ 9 & 7 & 1 & 3 \\ 4 & 0 & 1 & 1 \\ 3 & -1 & 2 & 0 \end{bmatrix} \right)$$

$$\begin{array}{l} Z_1 \rightarrow Z_1 - Z_3 \\ \downarrow \\ \cong \end{array} 2 \cdot \det \left(\begin{bmatrix} 0 & 3 & -1 & 0 \\ 9 & 7 & 1 & 3 \\ 4 & 0 & 1 & 1 \\ 3 & -1 & 2 & 0 \end{bmatrix} \right)$$

$$\begin{array}{l} \text{Entwicklung} \\ \text{nach der} \\ \text{1. Zeile} \\ \downarrow \\ \cong \end{array} 0 + 2 \cdot (-1)^{1+2} \cdot 3 \cdot \det \left(\begin{bmatrix} 9 & 1 & 3 \\ 4 & 1 & 1 \\ 3 & 2 & 0 \end{bmatrix} \right) \\ + 2 \cdot (-1)^{1+3} \cdot (-1) \cdot \det \left(\begin{bmatrix} 9 & 7 & 3 \\ 4 & 0 & 1 \\ 3 & -1 & 0 \end{bmatrix} \right) + 0$$

$$= -6 \cdot (0 + 3 + 24 - 9 - 18 - 0) - 2 \cdot (0 + 21 - 12 - 0 + 9 - 0)$$

$$= 0 - 2 \cdot 18 = -36.$$

(c) Wir wollen die Determinante der Matrix

$$\mathbf{C} = \begin{bmatrix} 1 & 2 & -1 & 0 \\ -1 & -2 & 3 & 2 \\ 5 & -7 & 1 & 1 \\ -10 & 14 & 0 & 0 \end{bmatrix}$$

berechnen:

$$\det(\mathbf{C}) = \det \left(\begin{bmatrix} 1 & 2 & -1 & 0 \\ -1 & -2 & 3 & 2 \\ 5 & -7 & 1 & 1 \\ -10 & 14 & 0 & 0 \end{bmatrix} \right)$$

$$\stackrel{\substack{Z_2 \rightarrow Z_2 + Z_1 \\ Z_4 \rightarrow Z_4 + 2Z_3 \\ \Downarrow \\ \equiv}}{\det \left(\begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 0 & 2 & 2 \\ 5 & -7 & 1 & 1 \\ 0 & 0 & 2 & 2 \end{bmatrix} \right)} = 0,$$

wobei die Determinante der neuen Matrix nach Regel (8) in Satz 13.24 null ist, da der zweite und vierte Zeilenvektor linear abhängig sind.

13.5 Einige Anwendungen von Determinanten

Als Erstes halten wir einen Satz fest, den wir zu Beginn von Teilkapitel 13.3 bereits für den Fall $n = 2$ und $n = 3$ formulieren konnten.

Satz 13.26. (äquiv. Charakterisierungen invertierbarer Matrizen)

Für $\mathbf{A} \in \mathbb{K}^{n \times n}$ sind die folgenden Aussagen **äquivalent**:

- (i) $\det(\mathbf{A}) \neq 0$
- (ii) \mathbf{A} ist **invertierbar**.
- (iii) $\text{Rang}(\mathbf{A}) = n$
- (iv) Die Spalten von \mathbf{A} sind **linear unabhängig**.
- (v) Die Zeilen von \mathbf{A} sind **linear unabhängig**.
- (vi) Das lineare Gleichungssystem $\mathbf{A} \vec{x} = \vec{b}$ ist für jedes $\vec{b} \in \mathbb{K}^n$ **eindeutig lösbar**.

Durch Bilden der Kontraposition (vgl. Anhang B im Skript der HM A) folgt, dass die Verneinungen der Aussagen aus Satz 13.26 äquivalent sind.

Folgerung 13.27. (aus Satz 13.26)

Für $\mathbf{A} \in \mathbb{K}^{n \times n}$ sind die folgenden Aussagen **äquivalent**:

- (i) $\det(\mathbf{A}) = 0$
- (ii) \mathbf{A} ist **nicht invertierbar**.
- (iii) $\text{Rang}(\mathbf{A}) < n$
- (iv) Die Spalten von \mathbf{A} sind **linear abhängig**.
- (v) Die Zeilen von \mathbf{A} sind **linear abhängig**.
- (vi) Es gibt mindestens ein $\vec{\mathbf{b}} \in \mathbb{K}^n$, für das das lineare Gleichungssystem $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}$ **nicht lösbar** oder **nicht eindeutig lösbar** ist.

Beweis von Satz 13.26:

- Aus Satz 13.9 zusammen mit Bemerkung 12.41 erhalten wir, dass Aussagen (ii) bis (vi) äquivalent sind. Wir zeigen nun noch (i) \iff (iv).
- *Beweis von (iv) \implies (i):* Sind die Spaltenvektoren von \mathbf{A} linear unabhängig, so ist das lineare Gleichungssystem $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{0}}$ eindeutig lösbar und wir können die Matrix mit elementaren Zeilenumformungen in die reduzierte Stufenform bringen. Die Determinante der reduzierten Stufenform ist 1, und nach Satz 13.24 ist $\det(\mathbf{A})$ ein Vielfaches (mit Faktor $\neq 0$) der Determinante der Matrix in reduzierter Stufenform. Somit gilt $\det(\mathbf{A}) \neq 0$.
- *Beweis von (i) \implies (iv):* Statt „ $\det(\mathbf{A}) \neq 0 \implies$ Die Spalten von \mathbf{A} sind linear unabhängig.“ zeigen wir die äquivalente Aussage „Die Spalten von \mathbf{A} sind linear abhängig. $\implies \det(\mathbf{A}) = 0$ “. Seien also die Spaltenvektoren von \mathbf{A} linear abhängig. Dann können wir durch elementare Spaltenumformungen eine Matrix $\tilde{\mathbf{A}}$ erzeugen mit einer Spalte, die der Nullvektor ist. Nach Satz 13.24 (8) gilt dann $\det(\tilde{\mathbf{A}}) = 0$ und nach Satz 13.24 $\det(\mathbf{A}) = \mu \det(\tilde{\mathbf{A}}) = \mu \cdot 0 = 0$ mit einem $\mu \in \mathbb{K} \setminus \{0\}$. \square

Der vorige Satz liefert viele wichtige Anwendungen. Betrachten einige dieser Anwendungen mit zugehörigen Beispielen.

Anwendung 13.28. (Bilden n Vektoren in \mathbb{K}^n eine Basis von \mathbb{K}^n ?)

Aufgabenstellung: Gegeben seien n Vektoren $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ in \mathbb{K}^n . Ist $(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)$ eine Basis des \mathbb{K}^n ?

Lösungsansatz: Wir schreiben $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ als Spaltenvektoren in eine $n \times n$ -Matrix \mathbf{A} und berechnen die Determinante.

- Ist $\det(\mathbf{A}) \neq 0$, so sind $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ nach Satz 13.26 linear unabhängig und bilden somit eine Basis von \mathbb{K}^n .
- Ist $\det(\mathbf{A}) = 0$, so sind $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ nach Folgerung 13.27 linear abhängig und bilden somit keine Basis von \mathbb{K}^n .

Beispiel 13.29. (Sind 3 gegebene Vektoren in \mathbb{K}^3 eine Basis von \mathbb{K}^3 ?)

Ist $\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} \right)$ eine Basis von \mathbb{R}^3 ?

Aus Beispiel 13.23 (a) wissen wir, dass gilt $\det \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} = 0$.

Also sind die Vektoren $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$ nach Folgerung 13.27

linear abhängig, und somit ist $\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} \right)$ keine Basis von \mathbb{R}^3 .

Anwendung 13.30. (Ist eine Matrix invertierbar?)

Aufgabenstellung: Gegeben sei eine Matrix $\mathbf{A} \in \mathbb{K}^{n \times n}$. Ist \mathbf{A} invertierbar?

Lösungsansatz: Wir berechnen $\det(\mathbf{A})$.

- Gilt $\det(\mathbf{A}) \neq 0$, so ist \mathbf{A} nach Satz 13.26 invertierbar.
- Gilt $\det(\mathbf{A}) = 0$, so ist \mathbf{A} nach Folgerung 13.27 nicht invertierbar.

Beispiel 13.31. (Ist eine Matrix invertierbar?)

Ist die Matrix $\mathbf{A} = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$ invertierbar?

Aus Beispiel 13.23 (a) wissen wir, dass gilt $\det \left(\begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix} \right) = 0$.

Somit ist die Matrix \mathbf{A} nach Folgerung 13.27 nicht invertierbar.

Wir lernen noch weitere Anwendungen der Determinante sind.

Satz 13.32. (Formel zur Berechnung der inversen Matrix)

Sei $\mathbf{A} \in \mathbb{K}^{n \times n}$ mit $\det(\mathbf{A}) \neq 0$, d.h. \mathbf{A} ist invertierbar. Sei $\mathbf{C} = [c_{i,k}] \in \mathbb{K}^{n \times n}$ gegeben durch

$$c_{i,k} := (-1)^{i+k} \cdot \det(\mathbf{A}_{\widehat{i,k}}).$$

Dann gilt

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \cdot \mathbf{C}^T.$$

Für den Spezialfall $n = 2$ erhalten wir die folgende Formel:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \cdot \begin{bmatrix} d & -c \\ -b & a \end{bmatrix}^T = \frac{1}{ad - bc} \cdot \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}. \quad (13.9)$$

Beispiel 13.33. (Berechnung der Inversen mit Satz 13.32)

(a) Für die Matrix $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ gilt

$$\det(\mathbf{A}) = \det \left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \right) = 1 \cdot 4 - 2 \cdot 3 = 4 - 6 = -2 \neq 0,$$

und somit ist \mathbf{A} invertierbar. Mit (13.9) finden wir die Inverse

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^{-1} = \frac{1}{-2} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}.$$

(b) Für die Matrix $\mathbf{A} = \begin{bmatrix} 0 & 1 & -4 \\ 1 & 2 & -1 \\ 1 & 1 & 2 \end{bmatrix}$ gilt

$$\det(\mathbf{A}) = \det \left(\begin{bmatrix} 0 & 1 & -4 \\ 1 & 2 & -1 \\ 1 & 1 & 2 \end{bmatrix} \right) = 0 - 1 - 4 + 8 + 0 - 2 = 1 \neq 0,$$

und somit ist \mathbf{A} invertierbar. Mit Satz 13.32 finden wir die Inverse

$$\begin{aligned} \mathbf{A}^{-1} &= \frac{1}{\det(\mathbf{A})} \cdot \begin{bmatrix} \det(\mathbf{A}_{\widehat{1,1}}) & -\det(\mathbf{A}_{\widehat{1,2}}) & \det(\mathbf{A}_{\widehat{1,3}}) \\ -\det(\mathbf{A}_{\widehat{2,1}}) & \det(\mathbf{A}_{\widehat{2,2}}) & -\det(\mathbf{A}_{\widehat{2,3}}) \\ \det(\mathbf{A}_{\widehat{3,1}}) & -\det(\mathbf{A}_{\widehat{3,2}}) & \det(\mathbf{A}_{\widehat{3,3}}) \end{bmatrix}^T \\ &= \frac{1}{1} \cdot \begin{bmatrix} 5 & -3 & -1 \\ -6 & 4 & 1 \\ 7 & -4 & -1 \end{bmatrix}^T = \begin{bmatrix} 5 & -6 & 7 \\ -3 & 4 & -4 \\ -1 & 1 & -1 \end{bmatrix}, \end{aligned}$$

wobei wir die $\det(\mathbf{A}_{\widehat{i,k}})$ mit der folgenden Nebenrechnung bestimmt haben:

$$\det(\mathbf{A}_{\widehat{1,1}}) = \det \left(\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \right) = 4 + 1 = 5,$$

$$\det(\mathbf{A}_{\widehat{1,2}}) = \det \left(\begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix} \right) = 2 + 1 = 3,$$

$$\det(\mathbf{A}_{\widehat{1,3}}) = \det \left(\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \right) = 1 - 2 = -1,$$

$$\det(\mathbf{A}_{\widehat{2,1}}) = \det \left(\begin{bmatrix} 1 & -4 \\ 1 & 2 \end{bmatrix} \right) = 2 + 4 = 6,$$

$$\det(\mathbf{A}_{\widehat{2,2}}) = \det \left(\begin{bmatrix} 0 & -4 \\ 1 & 2 \end{bmatrix} \right) = 0 + 4 = 4,$$

$$\det(\mathbf{A}_{\widehat{2,3}}) = \det \left(\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \right) = 0 - 1 = -1,$$

$$\det(\mathbf{A}_{\widehat{3,1}}) = \det \left(\begin{bmatrix} 1 & -4 \\ 2 & -1 \end{bmatrix} \right) = -1 + 8 = 7,$$

$$\det(\mathbf{A}_{\widehat{3,2}}) = \det \left(\begin{bmatrix} 0 & -4 \\ 1 & -1 \end{bmatrix} \right) = 0 + 4 = 4,$$

$$\det(\mathbf{A}_{\widehat{3,3}}) = \det \left(\begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \right) = 0 - 1 = -1.$$

Als Letztes lernen wir die **Cramersche Regel** (vgl. Kapitel 3 im HM A Skript) für beliebiges $n \geq 2$ kennen.

Satz 13.34. (Cramersche Regel)

Seien $\mathbf{A} \in \mathbb{K}^{n \times n}$ mit $\det(\mathbf{A}) \neq 0$ und $\vec{\mathbf{b}} \in \mathbb{K}^n$. Dann ist das lineare Gleichungssystem $\mathbf{A} \vec{\mathbf{x}} = \vec{\mathbf{b}}$ **eindeutig lösbar**. Die Lösung

$$\vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

lässt sich wie folgt berechnen: Sind $\vec{\mathbf{a}}_1, \vec{\mathbf{a}}_2, \dots, \vec{\mathbf{a}}_n$ die Spaltenvektoren von \mathbf{A} , so gilt

$$\begin{aligned} x_1 &= \frac{\det([\vec{\mathbf{b}} \ \vec{\mathbf{a}}_2 \ \vec{\mathbf{a}}_3 \ \cdots \ \vec{\mathbf{a}}_n])}{\det(\mathbf{A})}, \\ x_2 &= \frac{\det([\vec{\mathbf{a}}_1 \ \vec{\mathbf{b}} \ \vec{\mathbf{a}}_3 \ \cdots \ \vec{\mathbf{a}}_n])}{\det(\mathbf{A})}, \\ &\vdots \\ x_n &= \frac{\det([\vec{\mathbf{a}}_1 \ \vec{\mathbf{a}}_2 \ \cdots \ \vec{\mathbf{a}}_{n-1} \ \vec{\mathbf{b}}])}{\det(\mathbf{A})}. \end{aligned}$$

Die Cramersche Regel für $n = 2$ bzw. $n = 3$ erhalten wir als Sonderfall von Satz 13.34. In Kapitel 3 der HM A haben wir bereits diverse Beispiele zur Anwendung der Cramerschen Regel für $n = 2$ und $n = 3$ kennengelernt.

13.6 Eigenwerte und Eigenvektoren

In diesem Teilkapitel lernen wir die neuen Begriffe des Eigenwertes und Eigenvektors einer quadratischen Matrix kennen.

Definition 13.35. (Eigenwert und Eigenvektor)

Sei $\mathbf{A} = [a_{i,k}] \in \mathbb{C}^{n \times n}$. $\lambda \in \mathbb{C}$ heißt ein **Eigenwert (EW)** von \mathbf{A} , falls $\vec{\mathbf{x}} \in \mathbb{C}^n \setminus \{\vec{\mathbf{0}}\}$ existiert mit

$$\mathbf{A} \vec{\mathbf{x}} = \lambda \vec{\mathbf{x}}. \quad (13.10)$$

Ein $\vec{\mathbf{x}} \in \mathbb{C}^n \setminus \{\vec{\mathbf{0}}\}$, das (13.10) erfüllt, heißt ein **Eigenvektor (EV)** zum Eigenwert λ .

Betrachten wir zunächst ein Beispiel.

Beispiel 13.36. (Eigenwert und Eigenvektor)

Die Matrix $\mathbf{A} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ hat die Eigenwerte $\lambda_1 = 3$ und $\lambda_2 = -4$, denn

$$\mathbf{A} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

$$\mathbf{A} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -4 \\ 0 \end{bmatrix} = -4 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Also ist $\vec{\mathbf{x}}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ ein Eigenvektor zum Eigenwert $\lambda_1 = 3$,

und $\vec{\mathbf{x}}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ ist ein Eigenvektor zum Eigenwert $\lambda_2 = -4$.

Auch $7\vec{\mathbf{x}}_1 = \begin{bmatrix} 7 \\ 0 \\ 0 \end{bmatrix}$ ist ein Eigenvektor zum Eigenwert $\lambda_1 = 3$, denn

$$\mathbf{A} \begin{bmatrix} 7 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 7 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 21 \\ 0 \\ 0 \end{bmatrix} = 3 \begin{bmatrix} 7 \\ 0 \\ 0 \end{bmatrix}.$$

Wir sehen, dass weiter auch $\vec{\mathbf{x}}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ ein Eigenvektor zum Eigenwert $\lambda_1 = 3$ ist,

denn

$$\mathbf{A} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} = 3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Ein weiterer Eigenvektor zu $\lambda_1 = 3$ ist

$$\begin{bmatrix} -2 \\ 0 \\ 5 \end{bmatrix} = -2 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 5 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = -2\vec{\mathbf{x}}_1 + 5\vec{\mathbf{x}}_3,$$

denn

$$\mathbf{A} \begin{bmatrix} -2 \\ 0 \\ 5 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} -2 \\ 0 \\ 5 \end{bmatrix} = \begin{bmatrix} -6 \\ 0 \\ 15 \end{bmatrix} = 3 \begin{bmatrix} -2 \\ 0 \\ 5 \end{bmatrix}.$$

Am vorigen Beispiel sehen wir, dass eine Matrix mehr als einen Eigenwert haben kann, und dass zu einem Eigenwert mehrere Eigenvektoren existieren können. Für $\lambda_1 = 3$ haben wir sogar zwei linear unabhängige Eigenvektoren $\vec{\mathbf{x}}_1$ und $\vec{\mathbf{x}}_3$ gefunden. Aufgrund unseres Beispiels vermuten wir:

- Vielfache $\alpha \vec{\mathbf{x}}$ mit $\alpha \neq 0$ von einem Eigenvektor $\vec{\mathbf{x}}$ zu einem Eigenwert λ sind ebenfalls Eigenvektoren zu demselben Eigenwert λ . In der Tat gilt

$$\mathbf{A}(\alpha \vec{\mathbf{x}}) = \alpha (\mathbf{A} \vec{\mathbf{x}}) = \alpha (\lambda \vec{\mathbf{x}}) = \lambda (\alpha \vec{\mathbf{x}}).$$

- Nicht-triviale Linearkombinationen $\alpha \vec{\mathbf{x}} + \beta \vec{\mathbf{y}} \neq \vec{\mathbf{0}}$ von Eigenvektoren $\vec{\mathbf{x}}$ und $\vec{\mathbf{y}}$ zu demselben Eigenwert λ liefern wieder einen Eigenvektor zu diesem Eigenwert. In der Tat gilt

$$\begin{aligned} \mathbf{A}(\alpha \vec{\mathbf{x}} + \beta \vec{\mathbf{y}}) &= \mathbf{A} \alpha \vec{\mathbf{x}} + \mathbf{A} \beta \vec{\mathbf{y}} = \alpha \mathbf{A} \vec{\mathbf{x}} + \beta \mathbf{A} \vec{\mathbf{y}} \\ &= \alpha \lambda \vec{\mathbf{x}} + \beta \lambda \vec{\mathbf{y}} = \lambda (\alpha \vec{\mathbf{x}} + \beta \vec{\mathbf{y}}). \end{aligned}$$

Unsere Untersuchungen im Rahmen des vorigen Beispiels führen uns zu dem Begriff des Eigenraums eines Eigenwertes.

Definition 13.37. (Eigenraum eines Eigenwertes)

Sei $\mathbf{A} = [a_{i,k}] \in \mathbb{C}^{n \times n}$, und sei $\lambda \in \mathbb{C}$ ein Eigenwert von \mathbf{A} . Wir nennen die Menge

$$E_{\mathbf{A}}(\lambda) := \{ \vec{x} \in \mathbb{C}^n : \mathbf{A} \vec{x} = \lambda \vec{x} \}$$

den **Eigenraum** zum Eigenwert λ . $E_{\mathbf{A}}(\lambda)$ enthält alle Eigenvektoren zu λ und den Nullvektor $\vec{0}$. Wegen

$$\mathbf{A} \vec{x} = \lambda \vec{x} \quad \iff \quad \mathbf{A} \vec{x} - \lambda \mathbf{E}_n \vec{x} = \vec{0} \quad \iff \quad (\mathbf{A} - \lambda \mathbf{E}_n) \vec{x} = \vec{0}$$

ist also $E_{\mathbf{A}}(\lambda) = \mathbb{L}_{[\mathbf{A} - \lambda \mathbf{E}_n | \vec{0}]}$ die Lösungsmenge des homogenen linearen Gleichungssystems $(\mathbf{A} - \lambda \mathbf{E}_n) \vec{x} = \vec{0}$ und somit ein **Untervektorraum** von \mathbb{C}^n .

Wir kennen nun zwar Eigenwerte, Eigenvektoren und Eigenräume aber noch wissen wir nicht, wie man diese berechnet. Dieses wollen wir jetzt untersuchen:

Ist λ ein Eigenwert von $\mathbf{A} \in \mathbb{C}^{n \times n}$, so hat die Gleichung $\mathbf{A} \vec{x} = \lambda \vec{x}$ eine Lösung $\vec{x} \in \mathbb{C}^n \setminus \{ \vec{0} \}$. Wir schreiben diese Gleichung wie folgt um:

$$\mathbf{A} \vec{x} = \lambda \vec{x} \quad \iff \quad \mathbf{A} \vec{x} - \lambda \mathbf{E}_n \vec{x} = \vec{0} \quad \iff \quad (\mathbf{A} - \lambda \mathbf{E}_n) \vec{x} = \vec{0}$$

Die Gleichung $(\mathbf{A} - \lambda \mathbf{E}_n) \vec{x} = \vec{0}$ hat immer die triviale Lösung $\vec{x} = \vec{0}$. Sie hat **genau dann weitere Lösungen $\vec{x} \neq \vec{0}$, wenn $\det(\mathbf{A} - \lambda \mathbf{E}_n) = 0$ ist**. Also gilt:

$$\lambda \in \mathbb{C} \text{ ist Eigenwert von } \mathbf{A} \in \mathbb{C}^{n \times n} \quad \iff \quad \det(\mathbf{A} - \lambda \mathbf{E}_n) = 0.$$

Wir halten diese Erkenntnis als Satz fest.

Satz 13.38. (Bestimmung der Eigenwerte)

Sei $\mathbf{A} \in \mathbb{C}^{n \times n}$. Die Eigenwerte λ von \mathbf{A} sind genau die Lösungen der **charakteristischen Gleichung**

$$\det(\mathbf{A} - \lambda \mathbf{E}_n) = 0. \quad (13.11)$$

$p_{\mathbf{A}}(\lambda) := \det(\mathbf{A} - \lambda \mathbf{E}_n)$ ist eine komplexe Polynomfunktion n -ten Grades, und man nennt $p_{\mathbf{A}}$ das **charakteristische Polynom** von \mathbf{A} . Nach dem Fundamentalsatz der Algebra (siehe Satz 4.21 Kapitel 4 im Skript der HM A) hat $p_{\mathbf{A}}$ **genau n nicht notwendigerweise verschiedene komplexe Nullstellen**, d.h. (13.11) hat **genau n nicht notwendigerweise verschiedene komplexe Lösungen**.

Jetzt haben wir alle Informationen, um die Eigenwerte, zugehörigen Eigenvektoren und Eigenräume einer Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ praktisch zu berechnen.

Methode 13.39. (Bestimmung der Eigenwerte und Eigenräume)

Zur Bestimmung der Eigenwerte und der zugehörigen Eigenräume einer Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ gehen wir wie folgt vor:

- (1) Da alle Eigenwerte Lösungen der charakteristischen Gleichung

$$\det(\mathbf{A} - \lambda \mathbf{E}_n) = 0.$$

sind, bestimmen wir die Lösungen dieser Gleichung. Dieses liefert uns n nicht notwendigerweise verschiedene komplexe Eigenwerte $\lambda_1, \lambda_2, \dots, \lambda_n$.

- (2) Für jeden Eigenwert λ_i bestimmen wir anschließend alle Lösungen des linearen Gleichungssystems

$$(\mathbf{A} - \lambda_i \mathbf{E}_n) \vec{\mathbf{x}} = \vec{\mathbf{0}}, \quad (13.12)$$

denn $(\mathbf{A} - \lambda_i \mathbf{E}_n) \vec{\mathbf{x}} = \vec{\mathbf{0}} \iff \mathbf{A} \vec{\mathbf{x}} = \lambda_i \vec{\mathbf{x}}$. Die Menge der Lösungen von (13.12) bildet den Eigenraum $E_{\mathbf{A}}(\lambda_i)$, und alle nicht-trivialen Lösungen von (13.12) sind Eigenvektoren von \mathbf{A} zum Eigenwert λ_i .

Beispiel 13.40. (Eigenwerte, Eigenvektoren und Eigenräume)

- (a) Berechnen wir die Eigenwerte, Eigenvektoren und Eigenräume der Matrix

$$\mathbf{A} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

aus Beispiel 13.36 nun systematisch.

$$\begin{aligned} \det(\mathbf{A} - \lambda \mathbf{E}_3) &= \det \left(\begin{bmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \\ &= \det \left(\begin{bmatrix} 3 - \lambda & 0 & 0 \\ 0 & -4 - \lambda & 0 \\ 0 & 0 & 3 - \lambda \end{bmatrix} \right) \end{aligned}$$

$$= (3 - \lambda)(-4 - \lambda)(3 - \lambda) = (3 - \lambda)^2(-4 - \lambda),$$

und wir lesen die Eigenwerte $\lambda_1 = 3$ (mit der algebraischen Vielfachheit 2) und $\lambda_2 = -4$ ab.

Für $\lambda_1 = 3$ finden wir

$$\begin{aligned} (\mathbf{A} - 3\mathbf{E}_3) \vec{x} = \vec{0} &\iff \left(\begin{bmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 3 \end{bmatrix} - 3 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \vec{x} = \vec{0} \\ &\iff \begin{bmatrix} 3-3 & 0 & 0 \\ 0 & -4-3 & 0 \\ 0 & 0 & 3-3 \end{bmatrix} \vec{x} = \vec{0} \iff \begin{bmatrix} 0 & 0 & 0 \\ 0 & -7 & 0 \\ 0 & 0 & 0 \end{bmatrix} \vec{x} = \vec{0}, \end{aligned}$$

und wir haben nur eine nicht-triviale Gleichung $-7x_2 = 0$, also $x_2 = 0$. Wir setzen $x_1 = \alpha$ und $x_3 = \beta$ und erhalten somit den Eigenraum von $\lambda_1 = 3$, also

$$\begin{aligned} E_{\mathbf{A}}(3) &= \left\{ \begin{bmatrix} \alpha \\ 0 \\ \beta \end{bmatrix} \in \mathbb{C}^3 : \alpha, \beta \in \mathbb{C} \right\} \\ &= \left\{ \alpha \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \in \mathbb{C}^3 : \alpha, \beta \in \mathbb{C} \right\} = \text{LH} \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right). \end{aligned}$$

Für $\lambda_2 = -4$ finden wir

$$\begin{aligned} (\mathbf{A} - (-4)\mathbf{E}_3) \vec{x} = \vec{0} &\iff \begin{bmatrix} 3 - (-4) & 0 & 0 \\ 0 & -4 - (-4) & 0 \\ 0 & 0 & 3 - (-4) \end{bmatrix} \vec{x} = \vec{0} \\ &\iff \begin{bmatrix} 7 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 7 \end{bmatrix} \vec{x} = \vec{0} \iff \begin{bmatrix} 7 & 0 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \\ 0 & 0 & 7 & | & 0 \end{bmatrix} \xrightarrow{\substack{Z_1 \rightarrow \frac{1}{7} Z_1 \\ Z_3 \rightarrow \frac{1}{7} Z_3}} \iff \begin{bmatrix} 1 & 0 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \\ 0 & 0 & 1 & | & 0 \end{bmatrix}, \end{aligned}$$

und es gilt $x_1 = x_3 = 0$. Somit ist der Eigenraum von $\lambda_2 = -4$

$$E_{\mathbf{A}}(-4) = \left\{ \begin{bmatrix} 0 \\ \gamma \\ 0 \end{bmatrix} \in \mathbb{C}^3 : \gamma \in \mathbb{C} \right\} = \text{LH} \left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right).$$

(b) Gesucht sind die Eigenwerte, Eigenvektoren und Eigenräume der Matrix

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

Wir berechnen zunächst

$$\mathbf{A} - \lambda \mathbf{E}_2 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} - \lambda \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 - \lambda & -1 \\ 1 & 1 - \lambda \end{bmatrix}. \quad (13.13)$$

Also gilt

$$\begin{aligned} \det(\mathbf{A} - \lambda \mathbf{E}_2) &= \det \left(\begin{bmatrix} 1 - \lambda & -1 \\ 1 & 1 - \lambda \end{bmatrix} \right) = (1 - \lambda)^2 + 1 = (1 - \lambda)^2 - j^2 \\ &= (1 - \lambda - j)(1 - \lambda + j) = (1 - j - \lambda)(1 + j - \lambda), \end{aligned}$$

d.h. $\det(\mathbf{A} - \lambda \mathbf{E}_2) = 0$ gilt genau dann wenn $\lambda = 1 + j$ oder $\lambda = 1 - j$ ist. Die Eigenwerte von \mathbf{A} sind also $\lambda_1 := 1 + j$ und $\lambda_2 := 1 - j$.

Nun berechnen wir den Eigenraum zu $\lambda_1 = 1 + j$: Mit (13.13) finden wir

$$\begin{aligned} (\mathbf{A} - (1 + j) \mathbf{E}_2) \vec{x} = \vec{0} &\iff \begin{bmatrix} 1 - (1 + j) & -1 \\ 1 & 1 - (1 + j) \end{bmatrix} \vec{x} = \vec{0} \\ \iff \begin{bmatrix} -j & -1 \\ 1 & -j \end{bmatrix} \vec{x} = \vec{0} &\iff \begin{bmatrix} -j & -1 & | & 0 \\ 1 & -j & | & 0 \end{bmatrix} \\ \xrightarrow{Z_1 \leftrightarrow j \cdot Z_1} \begin{bmatrix} 1 & -j & | & 0 \\ 1 & -j & | & 0 \end{bmatrix} &\xrightarrow{Z_2 \leftrightarrow Z_2 - Z_1} \begin{bmatrix} 1 & -j & | & 0 \\ 0 & 0 & | & 0 \end{bmatrix}. \end{aligned}$$

Wir setzen $x_2 = \mu$ und finden $x_1 = j x_2 = j \mu$. Somit gilt

$$E_{\mathbf{A}}(1 + j) = \left\{ \begin{bmatrix} j \mu \\ \mu \end{bmatrix} \in \mathbb{C}^2 : \mu \in \mathbb{C} \right\} = \text{LH} \left(\begin{bmatrix} j \\ 1 \end{bmatrix} \right).$$

Nun berechnen wir den Eigenraum zu $\lambda_2 = 1 - j$: Mit (13.13) finden wir

$$\begin{aligned} (\mathbf{A} - (1 - j) \mathbf{E}_2) \vec{x} = \vec{0} &\iff \begin{bmatrix} 1 - (1 - j) & -1 \\ 1 & 1 - (1 - j) \end{bmatrix} \vec{x} = \vec{0} \\ \iff \begin{bmatrix} j & -1 \\ 1 & j \end{bmatrix} \vec{x} = \vec{0} &\iff \begin{bmatrix} j & -1 & | & 0 \\ 1 & j & | & 0 \end{bmatrix} \\ \xrightarrow{Z_1 \leftrightarrow -j \cdot Z_1} \begin{bmatrix} 1 & j & | & 0 \\ 1 & j & | & 0 \end{bmatrix} &\xrightarrow{Z_2 \leftrightarrow Z_2 - Z_1} \begin{bmatrix} 1 & j & | & 0 \\ 0 & 0 & | & 0 \end{bmatrix}. \end{aligned}$$

Wir setzen $x_2 = \mu$ und finden $x_1 = -j x_2 = -j \mu$. Somit gilt

$$E_{\mathbf{A}}(1-j) = \left\{ \begin{bmatrix} -j\mu \\ \mu \end{bmatrix} \in \mathbb{C}^2 : \mu \in \mathbb{C} \right\} = \text{LH} \left(\begin{bmatrix} -j \\ 1 \end{bmatrix} \right).$$

(c) Gesucht sind die Eigenwerte, Eigenvektoren und Eigenräume der Matrix

$$\mathbf{A} = \begin{bmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{bmatrix}.$$

Wir berechnen zunächst

$$\begin{aligned} \mathbf{A} - \lambda \mathbf{E}_3 &= \begin{bmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -\lambda & -1 & 1 \\ -3 & -2-\lambda & 3 \\ -2 & -2 & 3-\lambda \end{bmatrix}. \end{aligned} \quad (13.14)$$

Nun berechnen wir das charakteristische Polynom von \mathbf{A} mittels Entwicklung nach der ersten Spalte (bei der Berechnung von $\det(\mathbf{A} - \lambda \mathbf{E}_3)$)

$$\begin{aligned} p_{\mathbf{A}}(\lambda) &= \det(\mathbf{A} - \lambda \mathbf{E}_3) = \det \left(\begin{bmatrix} -\lambda & -1 & 1 \\ -3 & -2-\lambda & 3 \\ -2 & -2 & 3-\lambda \end{bmatrix} \right) \\ &= (-\lambda) \cdot \det \left(\begin{bmatrix} -2-\lambda & 3 \\ -2 & 3-\lambda \end{bmatrix} \right) - (-3) \cdot \det \left(\begin{bmatrix} -1 & 1 \\ -2 & 3-\lambda \end{bmatrix} \right) \\ &\quad + (-2) \cdot \det \left(\begin{bmatrix} -1 & 1 \\ -2-\lambda & 3 \end{bmatrix} \right) \\ &= -\lambda \cdot [(-2-\lambda)(3-\lambda) + 6] + 3 \cdot [-(3-\lambda) + 2] - 2 \cdot [-3 + 2 + \lambda] \\ &= -\lambda(\lambda^2 - \lambda) + 3(\lambda - 1) - 2(\lambda - 1) = -\lambda^2(\lambda - 1) + (\lambda - 1) \\ &= (\lambda - 1)(-\lambda^2 + 1) = -(\lambda - 1)(\lambda^2 - 1) = -(\lambda - 1)^2(\lambda + 1). \end{aligned}$$

Also gilt

$$p_{\mathbf{A}}(\lambda) = 0 \quad \iff \quad (\lambda = 1 \text{ oder } \lambda = -1).$$

Also sind die Eigenwerte $\lambda_1 = 1$ (mit algebraischer Vielfachheit 2) und $\lambda_2 = -1$.

Für $\lambda_1 = 1$ finden wir mit (13.14)

$$\begin{aligned}
 (\mathbf{A} - 1 \cdot \mathbf{E}_3) \vec{x} = \vec{0} &\iff \begin{bmatrix} -1 & -1 & 1 \\ -3 & -3 & 3 \\ -2 & -2 & 2 \end{bmatrix} \vec{x} = \vec{0} \\
 &\iff \begin{bmatrix} -1 & -1 & 1 & | & 0 \\ -3 & -3 & 3 & | & 0 \\ -2 & -2 & 2 & | & 0 \end{bmatrix} \xrightarrow[\downarrow]{\begin{array}{l} Z_2 \rightarrow Z_2 - 3Z_1 \\ Z_3 \rightarrow Z_3 - 2Z_1 \end{array}} \begin{bmatrix} -1 & -1 & 1 & | & 0 \\ 0 & 0 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix}.
 \end{aligned}$$

Also gilt mit $x_2 = \alpha$ und $x_3 = \beta$, dass $x_1 = -x_2 + x_3 = -\alpha + \beta$. Der Eigenraum von $\lambda_1 = 1$ ist damit

$$\begin{aligned}
 E_{\mathbf{A}}(1) &= \left\{ \begin{bmatrix} -\alpha + \beta \\ \alpha \\ \beta \end{bmatrix} \in \mathbb{C}^3 : \alpha, \beta \in \mathbb{C} \right\} \\
 &= \left\{ \alpha \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \in \mathbb{C}^3 : \alpha, \beta \in \mathbb{C} \right\} = \text{LH} \left(\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right).
 \end{aligned}$$

Für $\lambda_2 = -1$ finden wir mit (13.14)

$$\begin{aligned}
 (\mathbf{A} - (-1) \cdot \mathbf{E}_3) \vec{x} = \vec{0} \\
 &\iff \begin{bmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ -2 & -2 & 4 \end{bmatrix} \vec{x} = \vec{0} \iff \begin{bmatrix} 1 & -1 & 1 & | & 0 \\ -3 & -1 & 3 & | & 0 \\ -2 & -2 & 4 & | & 0 \end{bmatrix} \\
 &\xrightarrow[\downarrow]{\begin{array}{l} Z_2 \rightarrow Z_2 + 3Z_1 \\ Z_3 \rightarrow Z_3 + 2Z_1 \end{array}} \begin{bmatrix} 1 & -1 & 1 & | & 0 \\ 0 & -4 & 6 & | & 0 \\ 0 & -4 & 6 & | & 0 \end{bmatrix} \xrightarrow[\downarrow]{Z_3 \rightarrow Z_3 - Z_2} \begin{bmatrix} 1 & -1 & 1 & | & 0 \\ 0 & -4 & 6 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} \\
 &\xrightarrow[\downarrow]{Z_2 \rightarrow -\frac{1}{4}Z_2} \begin{bmatrix} 1 & -1 & 1 & | & 0 \\ 0 & 1 & -\frac{3}{2} & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} \xrightarrow[\downarrow]{Z_1 \rightarrow Z_1 + Z_2} \begin{bmatrix} 1 & 0 & -\frac{1}{2} & | & 0 \\ 0 & 1 & -\frac{3}{2} & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix}.
 \end{aligned}$$

Also gilt mit $x_3 = \gamma$, dass $x_1 = \frac{1}{2}\gamma$ und $x_2 = \frac{3}{2}\gamma$. Der Eigenraum von

$\lambda_2 = -1$ ist damit

$$E_{\mathbf{A}}(-1) = \left\{ \begin{bmatrix} \frac{1}{2}\gamma \\ \frac{3}{2}\gamma \\ \gamma \end{bmatrix} \in \mathbb{C}^3 : \gamma \in \mathbb{C} \right\} = \text{LH} \left(\begin{bmatrix} \frac{1}{2} \\ \frac{3}{2} \\ 1 \end{bmatrix} \right) = \text{LH} \left(\begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \right).$$

Wir führen noch eine weitere Bezeichnung ein.

Definition 13.41. (algebraische/geometrische Vielfachheit eines Eigenwerts)

Sei $\mathbf{A} \in \mathbb{C}^{n \times n}$, und sei λ ein Eigenwert von \mathbf{A} .

- (1) Ist λ eine k -fache Nullstelle der charakteristischen Gleichung $p_{\mathbf{A}}(\lambda) = 0$, so heißt k die **algebraische Vielfachheit** von λ .
- (2) Die Dimension $\dim(E_{\mathbf{A}}(\lambda))$ des Eigenraums $E_{\mathbf{A}}(\lambda)$ heißt die **geometrische Vielfachheit** von λ .

Wir bemerken, dass immer für jeden Eigenwert λ von \mathbf{A} gilt:

$$(\text{geometrische Vielfachheit von } \lambda) \leq (\text{algebraische Vielfachheit von } \lambda). \quad (13.15)$$

Betrachten wir auch hierzu einige Beispiele

Beispiel 13.42. (algebraische/geometrische Vielfachheit eines EWs)

(a) Für die Matrix aus Beispiel 13.40 (a),

$$\mathbf{A} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 3 \end{bmatrix},$$

fanden wir, dass der Eigenwert $\lambda_1 = 3$ die algebraische und geometrische Vielfachheit 2 hat. Der Eigenwert $\lambda_2 = -4$ hat die algebraische und geometrische Vielfachheit 1.

(b) Für die Matrix aus Beispiel 13.40 (b),

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix},$$

fanden wir die beiden Eigenwerte $\lambda_1 = 1 + j$, $\lambda_2 = 1 - j$. Für beide Eigenwerte gilt:

$$\text{geometrische Vielfachheit} = \text{algebraische Vielfachheit} = 1.$$

(c) Für die Matrix aus Beispiel 13.40 (c),

$$\mathbf{A} = \begin{bmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{bmatrix}$$

fanden wir die beiden Eigenwerte $\lambda_1 = 1$, $\lambda_2 = -1$.

Für $\lambda_1 = 1$ gilt: geometrische Vielfachheit = algebraische Vielfachheit = 2.

Für $\lambda_2 = -1$ gilt: geometrische Vielfachheit = algebraische Vielfachheit = 1.

(d) Für die Matrix $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$ berechnen wir das charakteristische Polynom:

$$p_{\mathbf{A}}(\lambda) = \det \left(\begin{bmatrix} 2 - \lambda & 1 \\ 0 & 2 - \lambda \end{bmatrix} \right) = (2 - \lambda)^2 - 0 = (2 - \lambda)^2$$

$\implies \mathbf{A}$ hat den Eigenwert $\lambda = 2$ mit der algebraischen Vielfachheit = 2.

Wir berechnen nun den Eigenraum zu $\lambda = 2$:

$$(\mathbf{A} - 2\mathbf{E}_2) \vec{x} = \vec{0} \iff \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \vec{x} = \vec{0} \iff \left[\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \end{array} \right],$$

und wir lesen ab, dass $x_2 = 0$ und $x_1 \in \mathbb{C}$ beliebig ist. Also ist

$$E_{\mathbf{A}}(2) = \text{LH} \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) \implies \text{geometrische Vielfachheit} = 1$$

Hier gilt also für den Eigenwert $\lambda = 2$

$$\text{geometrische Vielfachheit} = 1 < 2 = \text{algebraische Vielfachheit}.$$

Als Letztes halten wir noch einen nützlichen Satz fest, den wir in Kapitel 15 benötigen werden.

Satz 13.43. (EVEN zu verschiedenen EWEen sind linear unabhängig)

$\mathbf{A} \in \mathbb{C}^{n \times n}$ habe genau $m \leq n$ **verschiedene** Eigenwerte. Seien $\lambda_1, \lambda_2, \dots, \lambda_m \in \mathbb{C}$ diese m verschiedenen Eigenwerte von \mathbf{A} , und seien $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m \in \mathbb{C}^n \setminus \{\vec{0}\}$ zugehörige Eigenvektoren, also $\mathbf{A} \vec{x}_i = \lambda_i \vec{x}_i$, $i = 1, 2, \dots, m$. Dann sind die Vektoren $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m$ **linear unabhängig**.

Beweis von Satz 13.43: Wir zeigen nur, dass zwei Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig sind. Man kann die Idee verallgemeinern und dann mit vollständiger Induktion über die Anzahl der betrachteten Eigenvektoren zu verschiedenen Eigenwerten zeigen, dass m Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig sind.

Seien also $\lambda_1 \neq \lambda_2$ zwei verschiedene Eigenwerte von \mathbf{A} , und seien \vec{x}_1 bzw. \vec{x}_2 jeweils ein Eigenvektor zu λ_1 bzw. λ_2 . Wir betrachten nun

$$c_1 \vec{x}_1 + c_2 \vec{x}_2 = \vec{0}. \quad (13.16)$$

Multiplikation mit der Matrix \mathbf{A} von links liefert:

$$c_1 \underbrace{\mathbf{A} \vec{x}_1}_{=\lambda_1 \vec{x}_1} + c_2 \underbrace{\mathbf{A} \vec{x}_2}_{=\lambda_2 \vec{x}_2} = \mathbf{A} \vec{0} \quad \iff \quad c_1 \lambda_1 \vec{x}_1 + c_2 \lambda_2 \vec{x}_2 = \vec{0}. \quad (13.17)$$

Da $\lambda_1 \neq \lambda_2$ ist, muss mindestens einer dieser beiden Eigenwerte von null verschieden sein. Wir nehmen an, dass $\lambda_1 \neq 0$ gilt. Wir multiplizieren nun (13.16) mit λ_1 und erhalten

$$c_1 \lambda_1 \vec{x}_1 + c_2 \lambda_1 \vec{x}_2 = \vec{0}. \quad (13.18)$$

Subtrahieren der Gleichung (13.18) von der Gleichung (13.17) liefert

$$c_2 \lambda_2 \vec{x}_2 - c_2 \lambda_1 \vec{x}_2 = \vec{0} \quad \iff \quad c_2 \underbrace{(\lambda_2 - \lambda_1)}_{\neq 0} \underbrace{\vec{x}_2}_{\neq \vec{0}} = \vec{0}. \quad \implies \quad c_2 = 0,$$

und durch Einsetzen von $c_2 = 0$ in (13.16) folgt

$$c_1 \underbrace{\vec{x}_1}_{\neq \vec{0}} = \vec{0} \quad \implies \quad c_1 = 0.$$

Also folgt aus (13.16), dass $c_1 = c_2 = 0$ gelten muss, und wir haben gezeigt, dass die beiden Eigenvektoren \vec{x}_1, \vec{x}_2 linear unabhängig sind. \square

Teil V

Lineare Differentialgleichungen

Lineare Differentialgleichungen höherer Ordnung

In Kapitel 9 der HM A haben wir bereits gelernt, wie man lineare Differentialgleichungen erster Ordnung löst. Dieses wird in diesem Kapitel verallgemeinert, und wir lernen, wie man lineare Differentialgleichungen zweiter und höherer Ordnung mit konstanten Koeffizienten löst. Dabei verallgemeinern wir die Ideen, die bereits in Kapitel 9 der HM A vorkamen, und nutzen unser neu erworbenes Wissen aus der Linearen Algebra.

In den Vorlesungen wird dieses Kapitel in Absprache mit dem Dozenten der GET B nach Teilkapitel 12.5 eingeschoben.

14.1 Einführung zu linearen Differentialgleichungen höherer Ordnung

In diesem Teilkapitel lernen wir zunächst die neue Notation und Terminologie kennen.

Definition 14.1. (lineare Differentialgleichung n -ter Ordnung)

Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, sei $n \in \mathbb{N}$, und seien $a_0, a_1, \dots, a_{n-1}, b : I \rightarrow \mathbb{R}$ stetige Funktionen. Eine Gleichung der Form

$$y^{(n)} + a_{n-1}(t) y^{(n-1)} + \dots + a_1(t) y' + a_0(t) y = b(t)$$

heißt eine **lineare Differentialgleichung (DGL) n -ter Ordnung**.

Betrachten wir einige Beispiele.

Beispiel 14.2. (lineare Differentialgleichung n -ter Ordnung)

- (a) $y'' + \sin(t)y' + \cos(t)y = e^t$ ist eine lineare Differentialgleichung zweiter Ordnung.
- (b) $y' - 5y = 7$ ist eine lineare Differentialgleichung erster Ordnung.
- (c) $y^{(4)} + \frac{1}{t}y'' = 0$ ist eine lineare Differentialgleichung vierter Ordnung.

Bemerkung 14.3. (lineare DGL n -ter Ordnung in Anwendungen)

Für die Anwendungen sind besonders die folgenden Fälle interessant:

(1) $n = 1$: $y' + a_0(t)y = b(t) \iff y' = -a_0(t)y + b(t)$

Dieser Fall wurde bereits in Kapitel 9 der HM A behandelt.

(2) $n = 2$: $y'' + a_1(t)y' + a_0(t)y = b(t)$

(Für Anwendungsbeispiele siehe Skript der GET B, Kapitel 7)

Wir lernen nun, was man unter einer Lösung einer linearen Differentialgleichung n -ter Ordnung versteht.

Definition 14.4. (Lösung einer linearen DGL n -ter Ordnung)

Es seien die Voraussetzungen und Notation wie in Definition 14.1. Eine Funktion $y : I \rightarrow \mathbb{R}$ heißt eine **Lösung der Differentialgleichung**

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \dots + a_1(t)y' + a_0(t)y = b(t)$$

falls

(i) y auf I n -mal stetig differenzierbar ist und

(ii) $y^{(n)}(t) + a_{n-1}(t)y^{(n-1)}(t) + \dots + a_1(t)y'(t) + a_0(t)y(t) = b(t)$ für alle $t \in I$ erfüllt ist.

Beispiel 14.5. (Lösung einer linearen DGL n -ter Ordnung)

Die lineare DGL 2ter Ordnung $y'' + y = 2e^t$ hat beispielsweise die Lösung

$$y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) := \sin(t) + \cos(t) + e^t,$$

denn mit $y'(t) = \cos(t) - \sin(t) + e^t$, $y''(t) = -\sin(t) - \cos(t) + e^t$ gilt

$$y'' + y = (-\sin(t) - \cos(t) + e^t) + (\sin(t) + \cos(t) + e^t) = 2e^t \quad \text{für alle } t \in \mathbb{R}.$$

Als Letztes lernen wir analog zu den linearen Differentialgleichungen erster Ordnung (siehe Kapitel 9 im HM A Skript) auch noch die Begriffe homogen bzw. inhomogen kennen.

Definition 14.6. (homogene bzw. inhomogene lineare DGL n -ter Ordnung)

Es seien die Voraussetzungen und Notation wie in Definition 14.1. Die Differentialgleichung

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \dots + a_1(t)y' + a_0(t)y = b(t) \quad (14.1)$$

*heißt **homogen**, falls $b(t) = 0$ für alle $t \in I$. Andernfalls heißt sie **inhomogen**.*

Falls (14.1) eine inhomogene lineare Differentialgleichung n -ter Ordnung ist (also wenn $b(t)$ auf I ungleich der Nullfunktion ist), so nennen wir

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \dots + a_1(t)y' + a_0(t)y = 0$$

*die zu (14.1) gehörende **homogene lineare Differentialgleichung**.*

Betrachten wir auch hier zwei Beispiele.

Beispiel 14.7. (homogene bzw. inhomogene lin. DGL n -ter Ordn.)

(a) $y^{(3)} + 4ty'' + 17y = 0$ ist eine homogene lineare DGL dritter Ordnung.

(b) $y'' + \sin(t)y' + \cos(t)y = e^t$ ist eine inhomogene lineare DGL zweiter Ordnung.

Die zugehörige homogene lineare DGL ist $y'' + \sin(t)y' + \cos(t)y = 0$.

14.2 Lösungstheorie für die homogene lineare Differentialgleichung

In diesem Teilkapitel lernen wir die Lösungstheorie für homogene lineare Differentialgleichungen n -ter Ordnung kennen.

Dabei gelten in diesem Teilkapitel **immer die folgenden Voraussetzungen**:

Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, sei $n \in \mathbb{N}$, und seien $a_0, a_1, \dots, a_{n-1} : I \rightarrow \mathbb{R}$ stetige Funktionen. Wir betrachten die **homogene lineare Differentialgleichung n -ter Ordnung**

$$y^{(n)} + a_{n-1}(t) y^{(n-1)} + \dots + a_1(t) y' + a_0(t) y = 0. \quad (\text{H}_n)$$

Satz 14.8. (Lösungsmenge von (H_n))

Die Lösungsmenge $\mathbb{L}_{(\text{H}_n)}$ der homogenen linearen Differentialgleichung (H_n) ist ein **Untervektorraum** des \mathbb{R} -Vektorraums $\mathcal{C}^n(I)$ der n -mal stetig differenzierbaren Funktionen auf I .

Wir beweisen den Satz für den Sonderfall $n = 2$.

Beweis von Satz 14.8 (für $n = 2$): Wir betrachten die Differentialgleichung

$$y'' + a_1(t) y' + a_0(t) y = 0. \quad (\text{H}_2)$$

Sei $\mathbb{L}_{(\text{H}_2)} \subseteq \mathcal{C}^2(I)$ die Lösungsmenge von (H_2) . Wir verwenden Satz 12.7:

(i) Die Nullfunktion $y(t) = 0$ für alle $t \in I$ ist in $\mathbb{L}_{(\text{H}_2)}$, denn

$$y''(t) + a_1(t) y'(t) + a_0(t) y(t) = 0 + a_1(t) \cdot 0 + a_0(t) \cdot 0 = 0 \quad \text{für alle } t \in I.$$

(ii) Seien $y, z \in \mathbb{L}_{(\text{H}_2)}$. $\implies y + z \in \mathbb{L}_{(\text{H}_2)}$, denn:

$$\begin{aligned} & (y + z)''(t) + a_1(t) (y + z)'(t) + a_0(t) (y + z)(t) \\ &= (y''(t) + z''(t)) + a_1(t) (y'(t) + z'(t)) + a_0(t) (y(t) + z(t)) \\ &= y''(t) + z''(t) + a_1(t) y'(t) + a_1(t) z'(t) + a_0(t) y(t) + a_0(t) z(t) \\ &= \underbrace{y''(t) + a_1(t) y'(t) + a_0(t) y(t)}_{=0} + \underbrace{z''(t) + a_1(t) z'(t) + a_0(t) z(t)}_{=0} \\ &= 0 + 0 = 0 \quad (\text{weil } y \text{ und } z \text{ jeweils Lösungen von } (\text{H}_2) \text{ sind}) \end{aligned}$$

(iii) Seien $y \in \mathbb{L}_{(H_2)}$, $\lambda \in \mathbb{R}$. $\implies \lambda y \in \mathbb{L}_{(H_2)}$, denn:

$$\begin{aligned} & (\lambda y)''(t) + a_1(t) (\lambda y)'(t) + a_0(t) (\lambda y)(t) \\ &= \lambda y''(t) + a_1(t) \lambda y'(t) + a_0(t) \lambda y(t) \\ &= \lambda \underbrace{[y''(t) + a_1(t) y'(t) + a_0(t) y(t)]}_{=0} \\ &= \lambda \cdot 0 = 0 \quad (\text{weil } y \text{ Lösung von } (H_2) \text{ ist}) \end{aligned}$$

Damit haben wir die drei Eigenschaften eines Untervektorraums nachgewiesen und somit gezeigt, dass $\mathbb{L}_{(H_2)}$ ein Untervektorraum von $\mathcal{C}^2(I)$ ist. \square

Als Untervektorraum des \mathbb{R} -Vektorraums $\mathcal{C}^n(I)$ ist die Lösungsmenge $\mathbb{L}_{(H_n)}$ der homogenen linearen Differentialgleichung (H_n) selber ein \mathbb{R} -Vektorraum. Wir sprechen daher im Folgenden auch vom **Lösungsraum** statt von der Lösungsmenge. **Welche Dimension hat dieser \mathbb{R} -Vektorraum?** Die Antwort gibt folgender Satz, den wir hier nicht beweisen.

Satz 14.9. (Lösungsraum und Fundamentalsystem)

Der **Lösungsraum** $\mathbb{L}_{(H_n)}$ von (H_n) hat die **Dimension** n . Eine **Basis** des Lösungsraums bezeichnet man als **Fundamentalsystem (FS)** für (H_n) . Es besteht aus n **linear unabhängigen Lösungen** von (H_n) .

Spezialfall $n = 2$: Der Lösungsraum $\mathbb{L}_{(H_2)}$ von (H_2) hat die Dimension 2. Jedes Fundamentalsystem für (H_2) besteht aus zwei linear unabhängigen Lösungen von (H_2) . Es reicht also, zwei linear unabhängige reelle Lösungen y_1 und y_2 von (H_2) zu finden. Dann ist die reelle allgemeine Lösung von (H_2) durch

$$y : I \rightarrow \mathbb{R}, \quad y(t) = c_1 y_1(t) + c_2 y_2(t), \quad c_1, c_2 \in \mathbb{R},$$

gegeben, also

$$\mathbb{L}_{(H_2)} = \text{LH}(y_1, y_2) = \{y : I \rightarrow \mathbb{R}, y(t) = c_1 y_1(t) + c_2 y_2(t) : c_1, c_2 \in \mathbb{R}\}.$$

Beispiel 14.10. (Lösungsraum und Fundamentalsystem)

Gesucht ist die allgemeine Lösung von

$$y'' + y = 0. \tag{14.2}$$

Die homogene lineare DGL zweiter Ordnung (14.2) hat die beiden Lösungen

$y_1(t) = \sin(t)$ und $y_2(t) = \cos(t)$, denn

$$\begin{aligned}\sin''(t) + \sin(t) &= \cos'(t) + \sin(t) = -\sin(t) + \sin(t) = 0 && \text{für alle } t \in \mathbb{R}, \\ \cos''(t) + \cos(t) &= -\sin'(t) + \cos(t) = -\cos(t) + \cos(t) = 0 && \text{für alle } t \in \mathbb{R}.\end{aligned}$$

Nach Beispiel 12.24 (b) sind $y_1(t) = \sin(t)$ und $y_2(t) = \cos(t)$ linear unabhängig. Also ist $(\sin(t), \cos(t))$ ein Fundamentalsystem für (14.2). Die allgemeine Lösung von (14.2) ist somit

$$y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) = c_1 \sin(t) + c_2 \cos(t), \quad c_1, c_2 \in \mathbb{R},$$

und der Lösungsraum ist

$$\begin{aligned}\mathbb{L}_{(\mathbb{H}_2)} &= \text{LH}(\sin(t), \cos(t)) \\ &= \{y : \mathbb{R} \rightarrow \mathbb{R}, y(t) := c_1 \sin(t) + c_2 \cos(t) : c_1, c_2 \in \mathbb{R}\}.\end{aligned}$$

Bei der Differentialgleichung (14.2) handelt es sich um eine homogene lineare Differentialgleichung zweiter Ordnung **mit konstanten Koeffizienten**, da die Koeffizientenfunktionen $a_1(t) = 0$ und $a_0(t) = 1$ hier konstant sind. Wie man im Fall konstanter Koeffizienten(funktionen) $a_0, a_1, \dots, a_{n-1} \in \mathbb{R}$ in (\mathbb{H}_n) ein Fundamentalsystem bestimmen kann, lernen wir in Teilkapiteln 14.4 und 14.5.

14.3 Lösungstheorie für die inhomogene lineare Differentialgleichung

In diesem Teilkapitel lernen wir die Lösungstheorie für inhomogene lineare Differentialgleichungen n -ter Ordnung kennen.

Dabei gelten in diesem Teilkapitel **immer die folgenden Voraussetzungen**:

Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, sei $n \in \mathbb{N}$, und seien $a_0, a_1, \dots, a_{n-1}, b : I \rightarrow \mathbb{R}$ stetige Funktionen. Wir betrachten die **inhomogene lineare Differentialgleichung n -ter Ordnung**

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \dots + a_1(t)y' + a_0(t)y = b(t) \quad (\text{IH}_n)$$

und die **zugehörige homogene lineare Differentialgleichung**

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \dots + a_1(t)y' + a_0(t)y = 0. \quad (\text{H}_n)$$

Der nachfolgende Satz stellt den Zusammenhang zwischen den Lösungen von (IH_n) und (H_n) her.

Satz 14.11. (Zusammenhang zw. den Lösungen von (IH_n) und (H_n))
Sind w und z Lösungen von (IH_n) , so ist $w - z$ eine Lösung von (H_n) .

Beweis von Satz 14.11 (für $n = 2$): Seien w und z Lösungen von (IH_n) mit $n = 2$. Dann gilt

$$\begin{aligned} & (w - z)''(t) + a_1(t) (w - z)'(t) + a_0(t) (w - z)(t) \\ &= w''(t) - z''(t) + a_1(t) w'(t) - a_1(t) z'(t) + a_0(t) w(t) - a_0(t) z(t) \\ &= \underbrace{[w''(t) + a_1(t) w'(t) + a_0(t) w(t)]}_{=b(t)} - \underbrace{[z''(t) + a_1(t) z'(t) + a_0(t) z(t)]}_{=b(t)} \\ &= b(t) - b(t) = 0. \quad (\text{weil } w \text{ und } z \text{ Lösungen von } (IH_n) \text{ mit } n = 2 \text{ sind}) \end{aligned}$$

Also ist $w - z$ eine Lösung von (H_n) mit $n = 2$. □

Mit Hilfe von Satz 14.11 lässt sich nun leicht der folgende Satz herleiten.

Satz 14.12. (Darstellung der Lösungsmenge von (IH_n))

*Kennt man eine **spezielle Lösung** y_S von (IH_n) und ein **Fundamentalsystem** (y_1, y_2, \dots, y_n) von (H_n) , so kennt man die **allgemeine Lösung** von (IH_n) :*

$$y = y_S + c_1 y_1 + c_2 y_2 + \dots + c_n y_n \quad \text{mit } c_1, c_2, \dots, c_n \in \mathbb{R}.$$

*Die **Lösungsmenge** von (IH_n) ist also*

$$\begin{aligned} \mathbb{L}_{(IH_n)} &= \{y = y_S + c_1 y_1 + c_2 y_2 + \dots + c_n y_n : c_1, c_2, \dots, c_n \in \mathbb{R}\} \\ &= \{y = y_S + y_H : y_H \in \mathbb{L}_{(H_n)}\} \end{aligned} \quad (14.3)$$

Es ist zu beachten, dass die Lösungsmenge (14.3) der inhomogenen Differentialgleichung **kein** Untervektorraum von $\mathcal{C}^n(I)$ ist.

Den Beweis von Satz 14.12 führen wir in einer Übungsaufgabe durch.

Betrachten wir ein Beispiel.

Beispiel 14.13. (allgemeine Lösung einer inhom. lin. DGL 2. Ordn.)

Gesucht ist die Lösung von

$$y'' + y = 5.$$

Mit Inspizieren, sieht man leicht, dass $y_S : \mathbb{R} \rightarrow \mathbb{R}$, $y_S(t) := 5$, eine spezielle Lösung ist. (In der Tat gilt $y_S''(t) + y_S(t) = 0 + 5 = 5$.)

Aus Beispiel 14.10 wissen wir bereits, dass

$$(\sin(t), \cos(t))$$

ein Fundamentalsystem für die zugehörige homogene Gleichung

$$y'' + y = 0$$

ist. Nach Satz 14.12 ist daher

$$y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) = 5 + c_1 \sin(t) + c_2 \cos(t), \quad c_1, c_2 \in \mathbb{R},$$

die allgemeine Lösung von $y'' + y = 5$. Die Lösungsmenge ist somit

$$\mathbb{L}_{(\mathbb{H}_2)} = \{y : \mathbb{R} \rightarrow \mathbb{R}, y(t) := 5 + c_1 \sin(t) + c_2 \cos(t) : c_1, c_2 \in \mathbb{R}\}.$$

14.4 Homogene lineare Differentialgleichung mit konstanten Koeffizienten: Der Fall $n = 2$

In diesem Teilkapitel betrachten wir **homogene lineare Differentialgleichungen zweiter Ordnung mit konstanten Koeffizienten** (d.h. die Koeffizientenfunktionen a_0 und a_1 sind jeweils konstant, also $a_0(t) = a_0$ und $a_1(t) = a_1$ für alle $t \in I$):

$$y'' + a_1 y' + a_0 y = 0 \quad (\mathbb{H}_2)$$

mit $a_0, a_1 \in \mathbb{R}$. Um diese Differentialgleichung zu lösen, macht man den **komplexen Exponentialansatz**

$$y(t) = e^{st}, \quad s \in \mathbb{C}.$$

Ableiten liefert

$$y'(t) = s e^{st} \quad \text{und} \quad y''(t) = s^2 e^{st}.$$

Einsetzen in (\mathbb{H}_2) ergibt nun:

$$\begin{aligned} s^2 e^{st} + a_1 s e^{st} + a_0 e^{st} = 0 &\iff (s^2 + a_1 s + a_0) e^{st} = 0 \\ &\iff s^2 + a_1 s + a_0 = 0 \end{aligned} \quad (14.4)$$

$$\begin{aligned} &\iff s^2 + 2 \frac{a_1}{2} s + \left(\frac{a_1}{2}\right)^2 - \left(\frac{a_1}{2}\right)^2 + a_0 = 0 \\ &\iff \left(s + \frac{a_1}{2}\right)^2 - \frac{1}{4}(a_1^2 - 4a_0) = 0 \end{aligned}$$

Die Gleichung (14.4) heißt die **charakteristische Gleichung von** (H_2) . Für jede Lösung s von (14.4) ist e^{st} eine Lösung von (H_2) . Wir müssen also (14.4) lösen. Dabei können **drei Fälle** auftreten:

- *Fall 1:* (14.4) hat zwei verschiedene reelle Lösungen s_1, s_2 (falls gilt $a_1^2 - 4a_0 > 0$).
- *Fall 2:* (14.4) hat eine zweifache reelle Lösung s (falls gilt $a_1^2 - 4a_0 = 0$). Dann gilt $s = -a_1/2$.
- *Fall 3:* (14.4) hat keine reellen Lösungen (falls gilt $a_1^2 - 4a_0 < 0$). Dann hat (14.4) zwei konjugiert komplexe, nicht-reelle Lösungen der Form

$$s_1 = \alpha + \beta j \quad \text{und} \quad s_2 = \alpha - \beta j \quad \text{mit } \alpha, \beta \in \mathbb{R}, \text{ wobei } \beta \neq 0.$$

(Dabei sind $\alpha, \beta \in \mathbb{R}$ passend zu bestimmen.)

Anmerkung: Andere Fälle können nicht auftreten, d.h. wir können keine „doppelte“ komplexe, nicht-reelle Lösung finden und auch keine zwei verschiedenen komplexen, nicht-reellen Lösungen, die nicht konjugiert komplex zueinander sind. Das sieht man wie folgt: Sie $s_1 \in \mathbb{C} \setminus \mathbb{R}$ eine komplexe, nicht-reelle Lösung von (14.4). Dann gilt (indem wir die Gleichung komplex konjugieren)

$$\begin{aligned} s_1^2 + a_1 s_1 + a_0 = 0 &\iff \overline{s_1^2 + a_1 s_1 + a_0} = 0 \\ &\iff \overline{s_1}^2 + \overline{a_1} \overline{s_1} + \overline{a_0} = 0 \\ &\iff \overline{s_1}^2 + a_1 \overline{s_1} + a_0 = 0 \quad (\text{weil } a_0, a_1 \in \mathbb{R} \text{ sind}). \end{aligned}$$

Also ist $s_2 := \overline{s_1}$ ebenfalls eine Lösung von (14.4). Mehr als zwei verschiedene Lösungen kann es bei einer quadratischen Gleichung nicht geben.

Fall 1 (d.h. $a_1^2 > 4a_0$): Die charakteristische Gleichung (14.4) von (H_2) habe **zwei verschiedene reelle** Lösungen s_1, s_2 . Dann sind

$$y_1(t) = e^{s_1 t} \quad \text{und} \quad y_2(t) = e^{s_2 t}$$

Lösungen von (H_2) . Da $s_1 \neq s_2$ ist, sind diese Lösungen linear unabhängig. Um die lineare Unabhängigkeit von y_1 und y_2 nachzuweisen, betrachten wir für alle $t \in \mathbb{R}$:

$$\lambda_1 e^{s_1 t} + \lambda_2 e^{s_2 t} = 0 \iff \lambda_1 + \lambda_2 e^{s_2 t} e^{-s_1 t} = 0 \iff \lambda_1 = -\lambda_2 e^{(s_2 - s_1)t}.$$

Wir haben nun auf der linken Seite die Konstante λ_1 und auf der rechten Seite $-\lambda_2$ multipliziert mit der (wegen $s_2 \neq s_1$) nicht-konstanten Funktion $e^{(s_2-s_1)t}$. Diese Gleichung kann für alle $t \in \mathbb{R}$ nur gelten, wenn $\lambda_1 = \lambda_2 = 0$ ist. Damit gilt der nachfolgende Satz.

Satz 14.14. (zwei verschiedene reelle Lösungen der char. Gleichung)

*Hat die charakteristische Gleichung (14.4) von (H_2) zwei verschiedene reelle Lösungen $s_1, s_2 \in \mathbb{R}$, so ist $(e^{s_1 t}, e^{s_2 t})$ ein **Fundamentalsystem** für (H_2) . Die reelle **allgemeine Lösung** von (H_2) ist dann*

$$y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) := c_1 e^{s_1 t} + c_2 e^{s_2 t}, \quad c_1, c_2 \in \mathbb{R}.$$

Beispiel 14.15. (zwei verschiedene reelle Lösungen der char. Gleichung)

Wir betrachten die homogene lineare Differentialgleichung 2-ter Ordnung

$$y'' - 4y' + 3y = 0$$

Der Exponentialansatz $y(t) = e^{st}$ (und damit $y'(t) = s e^{st}$ und $y''(t) = s^2 e^{st}$) liefert

$$s^2 e^{st} - 4s e^{st} + 3e^{st} = 0 \iff (s^2 - 4s + 3)e^{st} = 0 \iff s^2 - 4s + 3 = 0.$$

Die charakteristische Gleichung

$$0 = s^2 - 4s + 3 = (s - 1)(s - 3)$$

hat die Lösungen $s_1 = 1$ und $s_2 = 3$. Damit ist (e^t, e^{3t}) ein Fundamentalsystem, und die reelle allgemeine Lösung ist

$$y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) := c_1 e^t + c_2 e^{3t}, \quad c_1, c_2 \in \mathbb{R}.$$

Fall 2 (d.h. $a_1^2 = 4a_0$): Die charakteristische Gleichung (14.4) von (H_2) habe die (**einzige**) reelle Lösung $s = -a_1/2$. In diesem Fall liefert der Exponentialansatz nur eine Lösung von (H_2) , nämlich

$$y_1(t) = e^{st} = e^{-\frac{a_1}{2}t}.$$

Wir zeigen nun, dass

$$y_2(t) = t e^{st}$$

ebenfalls eine Lösung von (H_2) ist: Mit

$$y_2'(t) = (t e^{st})' = e^{st} + t s e^{st} = (1 + st) e^{st},$$

$$y_2''(t) = ((1 + st) e^{st})' = s e^{st} + (1 + st) s e^{st} = (2s + s^2 t) e^{st}$$

folgt

$$\begin{aligned} y_2''(t) + a_1 y_2'(t) + a_0 y_2(t) &= (2s + s^2 t) e^{st} + a_1 (1 + st) e^{st} + a_0 t e^{st} \\ &= \left[\underbrace{(s^2 + a_1 s + a_0)}_{=0} t + \underbrace{(2s + a_1)}_{=0} \right] e^{st} = 0, \end{aligned}$$

da $s = -a_1/2$ eine Lösung von $s^2 + a_1 s + a_0 = 0$ und $2s + a_1 = 0$ ist. Die Lösungen

$$y_1(t) = e^{st} \quad \text{und} \quad y_2(t) = t e^{st}$$

sind linear unabhängig, denn betrachten wir

$$0 = \lambda_1 e^{st} + \lambda_2 t e^{st} = (\lambda_1 + \lambda_2 t) e^{st} \quad \text{für alle } t \in \mathbb{R},$$

so folgt $\lambda_1 + \lambda_2 t = 0$ für alle $t \in \mathbb{R}$ (weil $e^{st} \neq 0$ für alle $t \in \mathbb{R}$) und somit $\lambda_1 = \lambda_2 = 0$. Damit haben wir den nachfolgenden Satz bewiesen.

Satz 14.16. (genau eine zweifache reelle Lösung der char. Gleichung)

*Hat die charakteristische Gleichung von (H_2) die **einzigste (zweifache) reelle Lösung** $s = -a_1/2$, so ist $(e^{st}, t e^{st})$ ein **Fundamentalsystem** für (H_2) . Die **reelle allgemeine Lösung** von (H_2) ist dann*

$$y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) := c_1 e^{st} + c_2 t e^{st} = (c_1 + c_2 t) e^{st}, \quad c_1, c_2 \in \mathbb{R}.$$

Beispiel 14.17. (genau eine zweifache reelle Lösung der char. Gleichung)

Wir betrachten die homogene lineare Differentialgleichung 2-ter Ordnung

$$y'' + y' + \frac{1}{4} y = 0.$$

Der Exponentialansatz $y(t) = e^{st}$ (und damit $y'(t) = s e^{st}$ und $y''(t) = s^2 e^{st}$) liefert

$$s^2 e^{st} + s e^{st} + \frac{1}{4} e^{st} = 0 \iff \left(s^2 + s + \frac{1}{4} \right) e^{st} = 0 \iff s^2 + s + \frac{1}{4} = 0.$$

Die charakteristische Gleichung

$$0 = s^2 + s + \frac{1}{4} = \left(s + \frac{1}{2} \right)^2$$

hat genau eine reelle Lösung $s = -1/2$. Also ist $(e^{-\frac{1}{2}t}, t e^{-\frac{1}{2}t})$ ein Fundamentalsystem. Die reelle allgemeine Lösung ist damit

$$y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) := (c_1 + c_2 t) e^{-\frac{1}{2}t}, \quad c_1, c_2 \in \mathbb{R}.$$

Fall 3 (d.h. $a_1^2 < 4a_0$): Die charakteristische Gleichung von (H_2) habe die **zwei konjugiert komplexen Lösungen** $s_1 = \alpha + \beta j$ und $s_2 = \alpha - \beta j$ (mit $\alpha, \beta \in \mathbb{R}$, wobei $\beta \neq 0$). Damit erhalten wir die folgenden zwei linear unabhängigen komplexen Lösungen von (H_2) :

$$\begin{aligned} z_1(t) &= e^{s_1 t} = e^{(\alpha + j\beta)t} = e^{\alpha t} e^{j\beta t} \\ &= e^{\alpha t} [\cos(\beta t) + j \sin(\beta t)] = e^{\alpha t} \cos(\beta t) + j e^{\alpha t} \sin(\beta t), \\ z_2(t) &= e^{s_2 t} = e^{(\alpha - j\beta)t} = e^{\alpha t} e^{-j\beta t} \\ &= e^{\alpha t} [\cos(-\beta t) + j \sin(-\beta t)] = e^{\alpha t} \cos(\beta t) - j e^{\alpha t} \sin(\beta t), \end{aligned}$$

wobei wir die Euler-Formel

$$e^{j\phi} = \cos(\phi) + j \sin(\phi)$$

(vgl. Kapitel 4 der HM A) und $\sin(-x) = -\sin(x)$ und $\cos(-x) = \cos(x)$ benutzt haben. (In einer Übungsaufgabe zeigen wir, dass die beiden Lösungen auch tatsächlich linear unabhängig sind.) Wir bemerken, dass gilt $\overline{z_1(t)} = z_2(t)$ und $\overline{z_2(t)} = z_1(t)$.

Wir sind aber eigentlich an den reellen Lösungen (und nicht an komplexwertigen Lösungen) von (H_2) interessiert. Wir beobachten, dass die reellen Funktionen

$$\begin{aligned} y_1(t) &:= \frac{1}{2} [z_1(t) + z_2(t)] = e^{\alpha t} \cos(\beta t), \\ y_2(t) &:= \frac{1}{2j} [z_1(t) - z_2(t)] = e^{\alpha t} \sin(\beta t) \end{aligned}$$

als Linearkombinationen von z_1 und z_2 nach Satz 14.8 ebenfalls Lösungen von (H_2) sind. (*Anmerkung:* Es gilt $y_1(t) = \operatorname{Re}(z_1(t)) = \operatorname{Re}(z_2(t))$ und $y_2(t) = \operatorname{Im}(z_1(t)) = -\operatorname{Im}(z_2(t))$.) Da $\beta \neq 0$ ist, kann man zeigen, dass die beiden Lösungen linear unabhängig sind und somit ein reelles Fundamentalsystem von (H_2) bilden. Wir weisen die lineare Unabhängigkeit von y_1, y_2 in einer Übungsaufgabe nach. Damit gilt der nachfolgende Satz.

Satz 14.18. (zwei zueinander konjugiert komplexe, nicht reelle Lösungen der charakteristischen Gleichung)

Hat die charakteristische Gleichung von (H_2) die **beiden konjugiert komplexen Lösungen** $s_1 = \alpha + \beta j$ und $s_2 = \alpha - \beta j$ mit $\beta \neq 0$, so gelten:

(1) $(e^{s_1 t}, e^{s_2 t})$ ist ein **komplexes Fundamentalsystem** für (H_2) . Die **allgemeine komplexe Lösung** von (H_2) ist

$$\begin{aligned} z : \mathbb{R} \rightarrow \mathbb{C}, \quad z(t) &:= k_1 e^{s_1 t} + k_2 e^{s_2 t} = k_1 e^{(\alpha + \beta j)t} + k_2 e^{(\alpha - \beta j)t} \\ &= e^{\alpha t} [k_1 e^{j\beta t} + k_2 e^{-j\beta t}], \quad k_1, k_2 \in \mathbb{C}. \end{aligned}$$

(2) $(e^{\alpha t} \cos(\beta t), e^{\alpha t} \sin(\beta t))$ ist ein **reelles Fundamentalsystem** für (H_2) . Die **allgemeine reelle Lösung** von (H_2) ist

$$\begin{aligned} y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) &:= c_1 e^{\alpha t} \cos(\beta t) + c_2 e^{\alpha t} \sin(\beta t) \\ &= e^{\alpha t} [c_1 \cos(\beta t) + c_2 \sin(\beta t)], \quad c_1, c_2 \in \mathbb{R}. \end{aligned}$$

Betrachten wir auch hierzu ein Beispiel.

Beispiel 14.19. (zwei zueinander konjugiert komplexe, nicht reelle Lösungen der charakteristischen Gleichung)

Wir betrachten die homogene lineare Differentialgleichung 2-ter Ordnung

$$y'' + 2y' + 2y = 0.$$

Der Exponentialansatz $y(t) = e^{st}$ (und damit $y'(t) = s e^{st}$ und $y''(t) = s^2 e^{st}$) liefert

$$s^2 e^{st} + 2s e^{st} + 2e^{st} = 0 \iff (s^2 + 2s + 2)e^{st} = 0 \iff s^2 + 2s + 2 = 0.$$

Die charakteristische Gleichung

$$0 = s^2 + 2s + 2 = (s + 1)^2 + 1 = (s + 1)^2 - j^2 = (s + 1 - j)(s + 1 + j)$$

hat die zwei zueinander konjugiert komplexen Lösungen

$$s_1 = -1 + j \quad \text{und} \quad s_2 = -1 - j.$$

Wir finden als ein komplexes Fundamentalsystem

$$(e^{(-1+j)t}, e^{(-1-j)t}) = (e^{-t} e^{jt}, e^{-t} e^{-jt}),$$

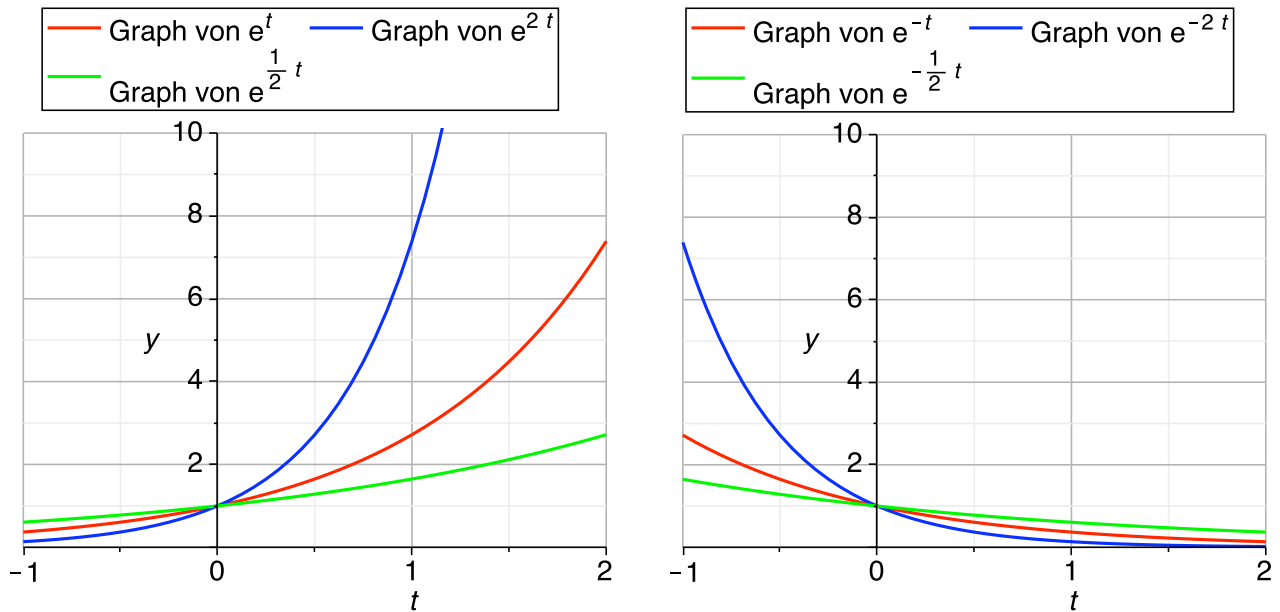


Abbildung 14.1: Funktionen der Form $e^{\alpha t}$ mit $\alpha > 0$ wachsen exponentiell. Funktionen der Form $e^{\alpha t}$ mit $\alpha < 0$ streben exponentiell (wie e^{-t}) gegen null.

und die allgemeine komplexe Lösung ist

$$\begin{aligned} z : \mathbb{R} \rightarrow \mathbb{C}, \quad z(t) &:= k_1 e^{(-1+j)t} + k_2 e^{(-1-j)t} = k_1 e^{-t} e^{jt} + k_2 e^{-t} e^{-jt} \\ &= e^{-t} [k_1 e^{jt} + k_2 e^{-jt}], \quad k_1, k_2 \in \mathbb{C}. \end{aligned}$$

Wir finden als ein reelles Fundamentalsystem

$$(e^{-t} \cos(t), e^{-t} \sin(t)),$$

und die reelle allgemeine reelle Lösung ist

$$\begin{aligned} y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) &:= c_1 e^{-t} \cos(t) + c_2 e^{-t} \sin(t) \\ &= e^{-t} [c_1 \cos(t) + c_2 \sin(t)], \quad c_1, c_2 \in \mathbb{R}. \end{aligned}$$

Bemerkung 14.20. (Zshg. zwischen reeller und komplexer Lösung)

Es gilt für eine geeignete Wahl der Konstanten $k_1, k_2 \in \mathbb{C}$ bzw. $c_1, c_2 \in \mathbb{R}$:

$$\begin{aligned} e^{\alpha t} [k_1 e^{j\beta t} + k_2 e^{-j\beta t}] &= e^{\alpha t} [c_1 \cos(\beta t) + c_2 \sin(\beta t)] \\ \iff k_1 e^{j\beta t} + k_2 e^{-j\beta t} &= c_1 \cos(\beta t) + c_2 \sin(\beta t) \\ \iff k_1 [\cos(\beta t) + j \sin(\beta t)] + k_2 [\cos(\beta t) - j \sin(\beta t)] \\ &= c_1 \cos(\beta t) + c_2 \sin(\beta t) \\ \iff (k_1 + k_2) \cos(\beta t) + j (k_1 - k_2) \sin(\beta t) &= c_1 \cos(\beta t) + c_2 \sin(\beta t) \end{aligned}$$

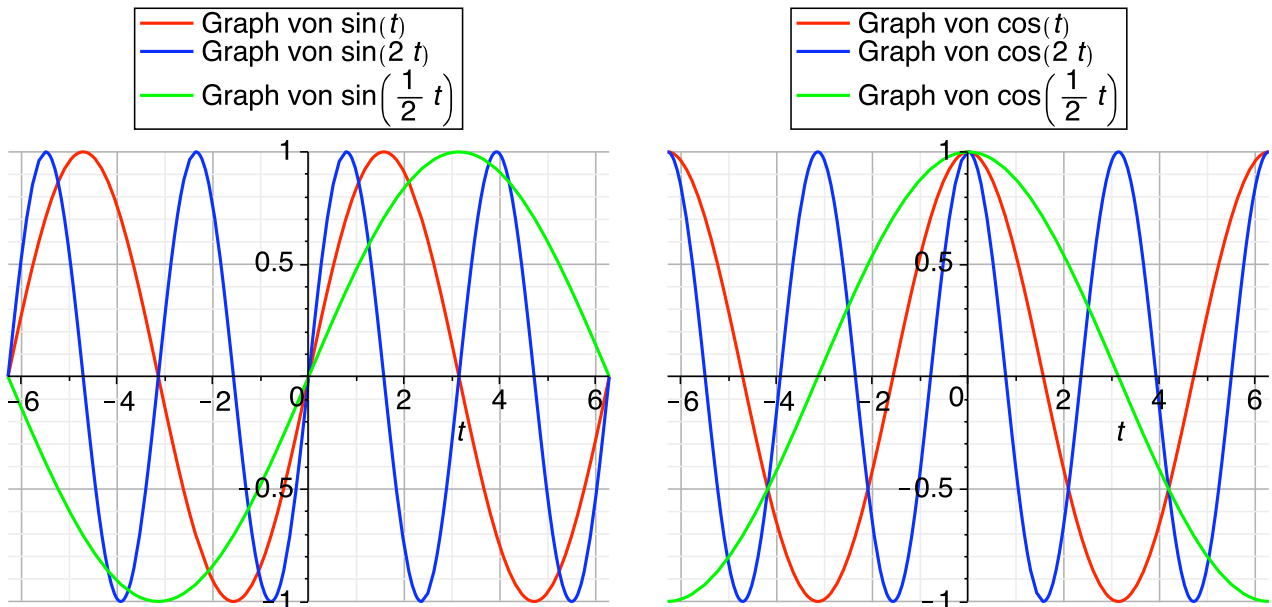


Abbildung 14.2: Funktionen der Form $\sin(\beta t)$ und $\cos(\beta t)$ sind klassische Wellen oder Schwingungen. Der Wert β steuert die Frequenz der Schwingung.

$$\begin{aligned}
 &\Leftrightarrow [(k_1 + k_2) - c_1] \cos(\beta t) + [j(k_1 - k_2) - c_2] \sin(\beta t) = 0 \\
 &\Leftrightarrow (k_1 + k_2) - c_1 = 0 \quad \text{und} \quad j(k_1 - k_2) - c_2 = 0 \\
 &\Leftrightarrow c_1 = k_1 + k_2 \quad \text{und} \quad (c_2 = j(k_1 - k_2) \Leftrightarrow j c_2 = k_2 - k_1) \\
 &\Leftrightarrow k_1 = \frac{1}{2}(c_1 - j c_2) \quad \text{und} \quad k_2 = \frac{1}{2}(c_1 + j c_2), \tag{14.5}
 \end{aligned}$$

wobei wir von der viertletzten Zeile in die drittletzte Zeile die lineare Unabhängigkeit von $\sin(\beta t)$ und $\cos(\beta t)$ benutzt haben. Es folgt aus dieser Rechnung: Die komplexe Lösung

$$y(t) = e^{\alpha t} [k_1 e^{j\beta t} + k_2 e^{-j\beta t}]$$

ist genau dann reellwertig, wenn (14.5) mit $c_1, c_2 \in \mathbb{R}$ gilt, also wenn $k_2 = \bar{k}_1$ gilt.

Abschließend schauen wir uns in Abbildung 14.1 bis Abbildung 14.3 Bilder der verschiedenen Typen von Lösungen an, um ein Gefühl für die durch solche Funktionen beschriebenen Phänomene zu bekommen.

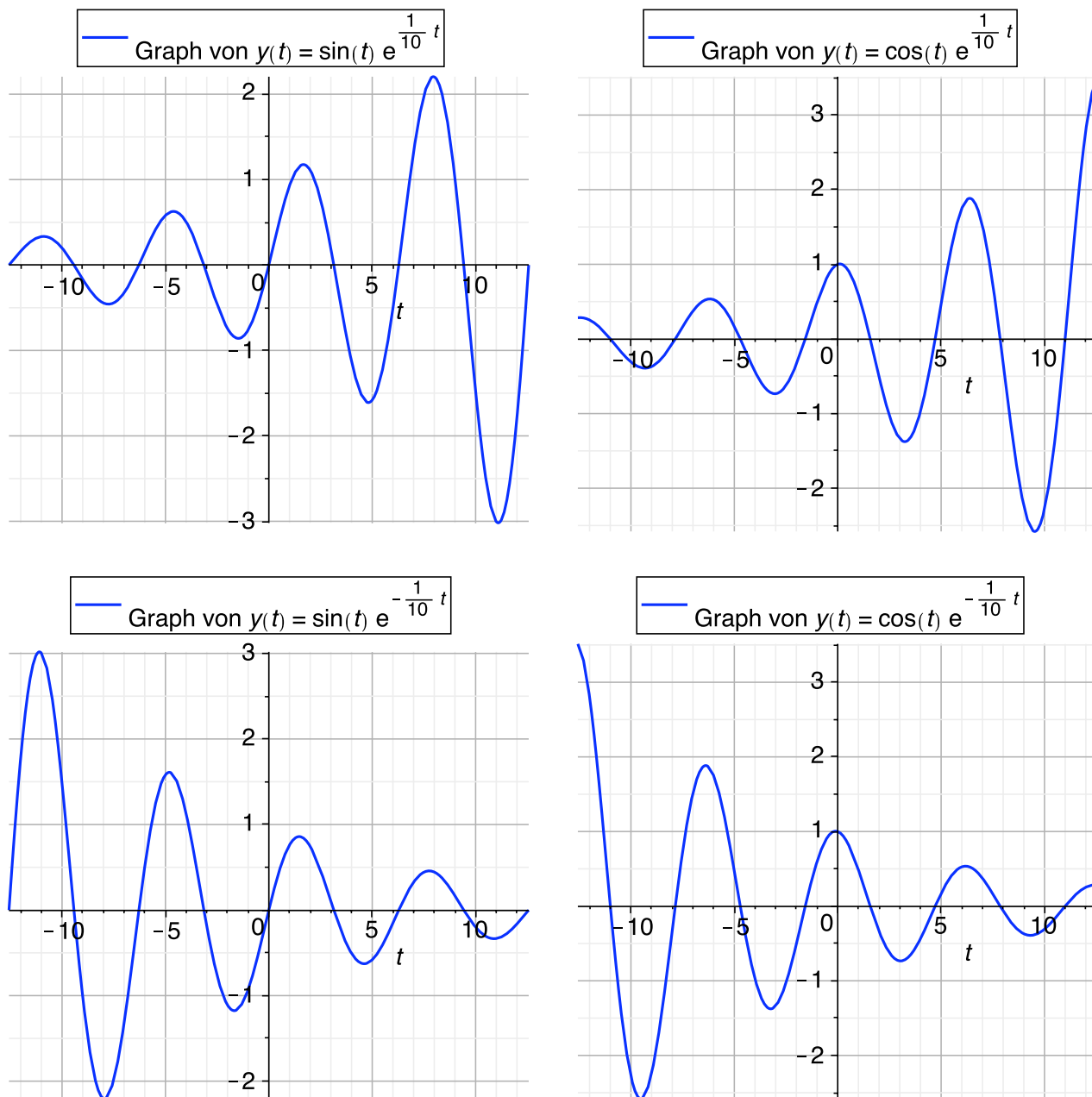


Abbildung 14.3: Funktionen der Form $\sin(\beta t) e^{\alpha t}$ und $\cos(\beta t) e^{\alpha t}$ sind Schwingungen, deren Amplitude exponentiell wächst, wenn $\alpha > 0$ ist, bzw. deren Amplitude exponentiell (wie e^{-t}) gedämpft wird, wenn $\alpha < 0$ ist.

14.5 Homogene lineare Differentialgleichung mit konstanten Koeffizienten: Der Fall $n \in \mathbb{N}$ beliebig

In diesem Abschnitt betrachten wir **homogene lineare Differentialgleichungen n -ter Ordnung mit konstanten Koeffizienten**,

$$y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y' + a_0 y = 0, \quad (\text{H}_n)$$

mit $a_0, a_1, \dots, a_{n-1} \in \mathbb{R}$. Wie im Fall $n = 2$ machen wir den **komplexen Exponentialansatz**

$$y(t) = e^{st}, \quad s \in \mathbb{C}.$$

Ableiten liefert nun

$$y'(t) = s e^{st}, \quad y''(t) = s^2 e^{st}, \quad \dots, \quad y^{(n-1)}(t) = s^{n-1} e^{st}, \quad y^{(n)}(t) = s^n e^{st},$$

und Einsetzen in (H_n) liefert

$$\begin{aligned} & s^n e^{st} + a_{n-1} s^{n-1} e^{st} + \dots + a_1 s e^{st} + a_0 e^{st} = 0 \\ \iff & (s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0) e^{st} = 0 \\ \iff & \underbrace{s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0}_{=: p_n(s)} = 0. \end{aligned} \quad (14.6)$$

Die Gleichung (14.6) heißt die **charakteristische Gleichung** von (H_n) . Der Fundamentalsatz der Algebra (siehe Satz 4.21 in Kapitel 4 der HM A) besagt, dass es n (nicht notwendigerweise verschiedene) $s_1, s_2, \dots, s_n \in \mathbb{C}$ gibt mit

$$p_n(s) = (s - s_1)(s - s_2) \cdot \dots \cdot (s - s_n). \quad (14.7)$$

Diese komplexen Zahlen s_1, s_2, \dots, s_n sind also **Lösungen der charakteristischen Gleichung (14.6)**, d.h. die Funktionen

$$e^{s_1 t}, \quad e^{s_2 t}, \quad \dots, \quad e^{s_n t}$$

sind **Lösungen der Differentialgleichung (H_n)** .

Definition 14.21. (*k -fache Lösungen der charakteristischen Glg.*)

Ist $z \in \mathbb{C}$ eine Lösung der charakteristischen Gleichung (14.6), so kommt z

unter dem $s_1, s_2, \dots, s_n \in \mathbb{C}$ in (14.7) vor. Wir zählen, wie oft der Linearfaktor $(s - z)$ in der Faktorisierung

$$p_n(s) = (s - s_1)(s - s_2) \cdot \dots \cdot (s - s_n)$$

vorkommt: Kommt er genau k -mal vor (wobei $k \in \mathbb{N}$ und $k \leq n$), so sagen wir, z ist eine k -fache **Lösung der charakteristischen Gleichung (14.6)**.

Betrachten wir zunächst zwei Beispiele.

Beispiel 14.22. (mehrfache Lösungen der charakteristischen Gleichung bei homogenen linearen DGLen höherer Ordnung)

(a) Wir betrachten die homogene lineare Differentialgleichung zweiter Ordnung

$$y'' + 2y' + y = 0.$$

Der komplexe Exponentialansatz $y(t) = e^{st}$ (und damit $y'(t) = s e^{st}$ und $y''(t) = s^2 e^{st}$) liefert

$$\begin{aligned} s^2 e^{st} + 2s e^{st} + e^{st} = 0 & \iff (s^2 + 2s + 1) e^{st} = 0 \\ & \iff s^2 + 2s + 1 = 0. \end{aligned}$$

Die charakteristische Gleichung

$$0 = s^2 + 2s + 1 = (s + 1)^2$$

hat also eine 2-fache (oder doppelte) Lösung $s_1 = s_2 = -1$. Mit den Kenntnissen aus dem vorigen Teilkapitel wissen wir bereits, dass die reelle allgemeine Lösung der Differentialgleichung durch

$$y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) := c_1 e^{-t} + c_2 t e^{-t}, \quad c_1, c_2 \in \mathbb{R},$$

gegeben ist.

(b) Wir betrachten die homogene lineare Differentialgleichung dritter Ordnung

$$y''' - 4y'' + 4y' = 0.$$

Der komplexe Exponentialansatz $y(t) = e^{st}$ (und damit $y'(t) = s e^{st}$, $y''(t) = s^2 e^{st}$ und $y'''(t) = s^3 e^{st}$) liefert

$$s^3 e^{st} - 4s^2 e^{st} + 4s e^{st} = 0 \iff (s^3 - 4s^2 + 4s) e^{st} = 0$$

$$\iff s^3 - 4s^2 + 4s = 0.$$

An der charakteristischen Gleichung

$$0 = s^3 - 4s^2 + 4s = s(s^2 - 4s + 4) = s(s - 2)^2$$

lesen wir ab:

- $s_1 = 0$ ist eine einfache Lösung von $s^3 - 4s^2 + 4s = 0$.
- $s_2 = s_3 = 2$ ist eine 2-fache (oder doppelte) Lösung der charakteristischen Gleichung $s^3 - 4s^2 + 4s = 0$.

Was ist hier die reelle allgemeine Lösung? Wir vermuten in Analogie zu dem Fall zweiter Ordnung, dass die reelle allgemeine Lösung

$$\begin{aligned} y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) &:= c_1 \underbrace{e^{0t}}_{=1} + c_2 e^{2t} + c_3 t e^{2t} \\ &= c_1 + c_2 e^{2t} + c_3 t e^{2t}, \quad c_1, c_2, c_3 \in \mathbb{R}, \end{aligned}$$

ist. Dieses ist auch korrekt, wie man durch Einsetzen von $y_3(t) := t e^{2t}$ in die Differentialgleichung und Überprüfen der Funktionen $y_1(t) := 1$, $y_2(t) := e^{2t}$ und $y_3(t) = t e^{2t}$ auf lineare Unabhängigkeit zeigen kann.

Nachdem wir nun das Konzept der Vielfachheit einer Lösung der charakteristischen Gleichung verstanden haben, kommen wir auf die allgemeine Lösung einer homogenen linearen Differentialgleichung mit konstanten Koeffizienten zurück.

Satz 14.23. (allgemeine Lösung einer homogenen linearen DGL n -ter Ordnung mit konstanten Koeffizienten)

Wir betrachten die homogene lineare Differentialgleichung n -ter Ordnung

$$y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y' + a_0 y = 0 \quad (\text{H}_n)$$

mit konstanten Koeffizienten $a_0, a_1, \dots, a_{n-1} \in \mathbb{R}$ und deren charakteristische Gleichung

$$s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0 = (s - s_1)(s - s_2) \cdot \dots \cdot (s - s_n) = 0 \quad (14.8)$$

mit den n (nicht notwendigerweise verschiedenen) komplexen Lösungen $s_1, s_2, \dots, s_n \in \mathbb{C}$.

(1) *Ist s eine k -fache Lösung der charakteristischen Gleichung (14.8), so sind*

$$e^{st}, \quad t e^{st}, \quad \dots, \quad t^{k-1} e^{st}$$

k **linear unabhängige** Lösungen von (H_n) . Damit haben wir k (im Allgemeinen komplexe) linear unabhängige Lösungen von (H_n) gefunden.

(2) Betrachtet man alle Lösungen von (14.8), so erhält man mit (1) **insgesamt n (komplexe) Lösungen von (H_n)** . Man kann zeigen (vgl. Kapitel 15 dieses Skripts), dass diese n Lösungen **linear unabhängig** sind. Also bilden sie ein (komplexes) Fundamentalsystem für (H_n) . Sind s_1, s_2, \dots, s_n alle reell, so haben wir damit bereits ein reelles Fundamentalsystem gefunden.

(3) Ist $z = \alpha + \beta j$ (mit $\alpha, \beta \in \mathbb{R}$ und $\beta \neq 0$) eine **k -fache komplexe** Lösung der charakteristischen Gleichung (14.8), so ist ebenfalls $\bar{z} = \alpha - \beta j$ eine **k -fache komplexe** Lösung der charakteristischen Gleichung. In diesem Fall sind die $2k$ reellen Funktionen

$$\begin{array}{ccccccc} e^{\alpha t} \cos(\beta t), & t e^{\alpha t} \cos(\beta t), & \dots, & t^{k-1} e^{\alpha t} \cos(\beta t) \\ e^{\alpha t} \sin(\beta t), & t e^{\alpha t} \sin(\beta t), & \dots, & t^{k-1} e^{\alpha t} \sin(\beta t) \end{array}$$

$2k$ **linear unabhängige Lösungen von (H_n)** . Sie entstehen durch Bildung von Real- bzw. Imaginärteil aus den komplexen Lösungen (siehe (1)) zu z bzw. \bar{z} . Mit dieser Beobachtung können wir ein komplexes Fundamentalsystem leicht in ein reelles umwandeln.

Betrachten wir einige Beispiele.

Beispiel 14.24. (allgemeine Lösung einer homogenen linearen DGL n -ter Ordnung mit konstanten Koeffizienten)

(a) Wir betrachten die homogene lineare Differentialgleichung dritter Ordnung

$$y''' - 2y'' + y' = 0.$$

Der komplexe Exponentialansatz $y(t) = e^{st}$ (und damit $y'(t) = s e^{st}$, $y''(t) = s^2 e^{st}$ und $y'''(t) = s^3 e^{st}$) liefert

$$s^3 e^{st} - 2s^2 e^{st} + s e^{st} = 0 \quad \iff \quad (s^3 - 2s^2 + s) e^{st} = 0,$$

d.h. wir finden die charakteristische Gleichung

$$0 = s^3 - 2s^2 + s = s(s^2 - 2s + 1) = s(s-1)^2.$$

Deren Lösungen sind $s_1 = 0$ (1-fach), $s_2 = s_3 = 1$ (2-fach). Somit finden wir ein reelles Fundamentalsystem

$$(e^{0 \cdot t}, e^{1 \cdot t}, t e^{1 \cdot t}) = (1, e^t, t e^t)$$

und die reelle allgemeine Lösung

$$y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) := c_1 \cdot 1 + c_2 e^t + c_3 t e^t = c_1 + (c_2 + c_3 t) e^t, \quad c_1, c_2, c_3 \in \mathbb{R}.$$

(b) Wir betrachten die homogene lineare Differentialgleichung vierter Ordnung

$$y^{(4)} - y = 0.$$

Der komplexe Exponentialansatz $y(t) = e^{st}$ (und damit $y'(t) = s e^{st}$, $y''(t) = s^2 e^{st}$, $y'''(t) = s^3 e^{st}$ und $y^{(4)}(t) = s^4 e^{st}$) liefert

$$s^4 e^{st} - e^{st} = 0 \quad \iff \quad (s^4 - 1) e^{st} = 0,$$

d.h. wir finden die charakteristische Gleichung

$$\begin{aligned} 0 &= s^4 - 1 = (s^2 - 1)(s^2 + 1) = (s^2 - 1)(s^2 - j^2) \\ &= (s - 1)(s + 1)(s - j)(s + j). \end{aligned}$$

Deren Lösungen sind $s_1 = 1$, $s_2 = -1$, $s_3 = j$, $s_4 = -j$ (jeweils 1-fach). Damit finden wir ein komplexes Fundamentalsystem

$$(e^t, e^{-t}, e^{jt}, e^{-jt}),$$

und das zugehörige reelle Fundamentalsystem ist

$$(e^t, e^{-t}, \cos(t), \sin(t)).$$

(*Erklärung:* Es gilt mit der Euler-Formel $e^{jt} = \cos(t) + j \sin(t)$ und $e^{-jt} = \cos(t) - j \sin(t)$, und der Realteil bzw. der Imaginärteil dieser Funktionen sind $\cos(t)$ bzw. $\pm \sin(t)$.)

Die reelle allgemeine Lösung ist also

$$y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) := c_1 e^t + c_2 e^{-t} + c_3 \cos(t) + c_4 \sin(t), \quad c_1, c_2, c_3, c_4 \in \mathbb{R}.$$

(c) Wir betrachten die homogene lineare Differentialgleichung vierter Ordnung

$$y^{(4)} + 2y'' + y = 0.$$

Der komplexe Exponentialansatz $y(t) = e^{st}$ (und damit $y'(t) = s e^{st}$, $y''(t) = s^2 e^{st}$, $y'''(t) = s^3 e^{st}$ und $y^{(4)}(t) = s^4 e^{st}$) liefert

$$s^4 e^{st} + 2s^2 e^{st} + e^{st} = 0 \quad \iff \quad (s^4 + 2s^2 + 1) e^{st} = 0,$$

d.h. wir finden die charakteristische Gleichung

$$0 = s^4 + 2s^2 + 1 = (s^2 + 1)^2 = (s^2 - j^2)^2$$

$$= [(s-j)(s+j)]^2 = (s-j)^2 (s+j)^2.$$

Deren Lösungen sind $s_1 = s_2 = j$ (2-fach) und $s_3 = s_4 = -j$ (2-fach). Wir finden also ein komplexes Fundamentalsystem

$$(e^{jt}, t e^{jt}, e^{-jt}, t e^{-jt}).$$

Das zugehörige reelle Fundamentalsystem ist

$$(\cos(t), t \cos(t), \sin(t), t \sin(t)).$$

(*Erklärung:* Es gilt mit der Euler-Formel $e^{jt} = \cos(t) + j \sin(t)$ und $e^{-jt} = \cos(t) - j \sin(t)$, und der Realteil bzw. der Imaginärteil dieser Funktionen sind $\cos(t)$ bzw. $\pm \sin(t)$. Analog gilt mit der Euler-Formel $t e^{jt} = t \cos(t) + j t \sin(t)$ und $t e^{-jt} = t \cos(t) - j t \sin(t)$, und der Realteil bzw. der Imaginärteil dieser Funktionen sind $t \cos(t)$ bzw. $\pm t \sin(t)$.)

Die reelle allgemeine Lösung ist somit

$$\begin{aligned} y: \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) &:= c_1 \cos(t) + c_2 t \cos(t) + c_3 \sin(t) + c_4 t \sin(t) \\ &= (c_1 + c_2 t) \cos(t) + (c_3 + c_4 t) \sin(t), \quad c_1, c_2, c_3, c_4 \in \mathbb{R}. \end{aligned}$$

(d) Wir betrachten die homogene lineare Differentialgleichung vierter Ordnung

$$y^{(4)} - 4y''' + 13y'' = 0.$$

Der komplexe Exponentialansatz $y(t) = e^{st}$ (und damit $y'(t) = s e^{st}$, $y''(t) = s^2 e^{st}$, $y'''(t) = s^3 e^{st}$ und $y^{(4)}(t) = s^4 e^{st}$) liefert

$$s^4 e^{st} - 4s^3 e^{st} + 13s^2 e^{st} = 0 \quad \iff \quad (s^4 - 4s^3 + 13s^2) e^{st} = 0,$$

d.h. wir finden die charakteristische Gleichung

$$\begin{aligned} 0 &= s^4 - 4s^3 + 13s^2 = s^2 (s^2 - 4s + 13) = s^2 [(s-2)^2 + 9] \\ &= s^2 [(s-2)^2 - (3j)^2] = s^2 (s-2-3j)(s-2+3j). \end{aligned}$$

Deren Lösungen sind $s_1 = s_2 = 0$ (2-fach) und $s_3 = 2 + 3j$, $s_4 = 2 - 3j$. Also finden wir ein komplexes Fundamentalsystem

$$\left(\underbrace{e^{0 \cdot t}}_{=1}, \underbrace{t e^{0 \cdot t}}_{=t \cdot 1}, e^{(2+3j)t}, e^{(2-3j)t} \right) = (1, t, e^{2t} e^{j3t}, e^{2t} e^{-j3t}).$$

Das zugehörige reelle Fundamentalsystem ist

$$(1, t, e^{2t} \cos(3t), e^{2t} \sin(3t)).$$

(*Erklärung:* Es gilt mit der Euler-Formel

$$\begin{aligned} e^{(2+3j)t} &= e^{2t} e^{j3t} = e^{2t} [\cos(3t) + j \sin(3t)] \\ &= e^{2t} \cos(3t) + j e^{2t} \sin(3t), \\ e^{(2-3j)t} &= e^{2t} e^{-j3t} = e^{2t} [\cos(-3t) + j \sin(-3t)] \\ &= e^{2t} \cos(3t) - j e^{2t} \sin(3t), \end{aligned}$$

und der Realteil bzw. der Imaginärteil dieser Funktionen sind $e^{2t} \cos(3t)$ bzw. $\pm e^{2t} \sin(3t)$.)

Die reelle allgemeine Lösung ist somit

$$\begin{aligned} y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) &:= c_1 + c_2 t + c_3 e^{2t} \cos(3t) + c_4 e^{2t} \sin(3t) \\ &= c_1 + c_2 t + (c_3 \cos(3t) + c_4 \sin(3t)) e^{2t}, \quad c_1, c_2, c_3, c_4 \in \mathbb{R}. \end{aligned}$$

(e) Wir betrachten die homogene lineare Differentialgleichung erster Ordnung

$$y' + a y = 0 \quad \text{mit } a \in \mathbb{R}.$$

Der komplexe Exponentialansatz $y(t) = e^{st}$ (und damit $y'(t) = s e^{st}$) liefert

$$s e^{st} + a e^{st} = 0 \quad \iff \quad (s + a) e^{st} = 0,$$

d.h. wir finden die charakteristische Gleichung

$$s + a = 0 \quad \iff \quad s = -a.$$

Damit finden wir das reelle Fundamentalsystem (e^{-at}) und die reelle allgemeine Lösung

$$y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) := c e^{-at}, \quad c \in \mathbb{R}.$$

Die homogene lineare Differentialgleichung $y' + a y = 0$ können wir bereits mit den Kenntnissen aus Kapitel 9 der HM A lösen, und wir erhalten die bereits aus der HM A bekannte Lösung als Sonderfall für $n = 1$ der allgemeinen Theorie für homogene lineare Differentialgleichungen n -ter Ordnung mit konstanten Koeffizienten.

14.6 Inhomogene lineare Differentialgleichung mit konstanten Koeffizienten

Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, sei $n \in \mathbb{N}$, und sei $b : I \rightarrow \mathbb{R}$ stetig. Wir betrachten die **inhomogene lineare Differentialgleichung n -ter Ordnung**

$$y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y' + a_0 y = b(t) \quad (\text{IH}_n)$$

mit konstanten Koeffizienten $a_0, a_1, \dots, a_{n-1} \in \mathbb{R}$ und die **zugehörige homogene lineare Differentialgleichung**

$$y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y' + a_0 y = 0. \quad (\text{H}_n)$$

Im Folgenden seien $s_1, s_2, \dots, s_n \in \mathbb{C}$ die Lösungen der charakteristischen Gleichung

$$s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0 = 0 \quad (14.9)$$

zu (H_n) und (y_1, y_2, \dots, y_n) das daraus resultierende reelle Fundamentalsystem (wie in Teilkapitel 14.5 erarbeitet). Nach Satz 14.12 reicht es, eine spezielle (reelle) Lösung y_S von (IH_n) zu finden. Dann ist die allgemeine (reelle) Lösung von (IH_n) nach Satz 14.12 gegeben durch

$$y(t) = y_S(t) + c_1 y_1(t) + c_2 y_2(t) + \dots + c_n y_n(t), \quad c_1, c_2, \dots, c_n \in \mathbb{R}.$$

Wie kann man eine spezielle Lösung y_S von (IH_n) finden?

Methode 1: Raten. Bitte hier **immer** nachweisen, dass es sich tatsächlich um eine Lösung von (IH_n) handelt, indem Sie Ihre geratene Lösung in die Differentialgleichung (IH_n) einsetzen!

Methode 2: „Methode der unbestimmten Koeffizienten“ mit einem Ansatz nach Tabelle 14.1. Zuerst ist zu beachten, dass diese Methode nur angewendet werden kann, wenn die rechte Seite $b(t)$ der Differentialgleichung von einer Form ist, die in der Tabelle aufgeführt ist. Ist dieses der Fall, so wählt man gemäß der Tabelle einen Ansatz y_S für die spezielle Lösung und setzt diesen in die inhomogene lineare Differentialgleichung (IH_n) ein, um die Koeffizienten (d.h. die Konstanten) im Ansatz zu bestimmen.

Bemerkung 14.25. (zu Tabelle 14.1)

- (1) Merken muss man sich nur die letzte Tabellenzeile. Alle anderen Zeilen sind darin als Spezialfälle enthalten.
- (2) Ist $b(t)$ eine Linearkombination der in der linken Spalte genannten Funktionen, so wählt man als Ansatz für y_S die Summe aus den entsprechenden Ansätzen in der rechten Spalte. Dabei müssen die Konstanten (also die „Koeffizienten“) alle unterschiedlich benannt sein.

Betrachten wir einige Beispiele zur Anwendung der „Methode der unbestimmten Koeffizienten“.

rechte Seite $b(t)$	Ansatz für $y_S(t)$
$b_0 + b_1 t + \dots + b_m t^m$	falls 0 keine Lösung der charakteristischen Gleichung ist: $A_0 + A_1 t + \dots + A_m t^m$ falls 0 eine k -fache Lösung der charakteristischen Gleichung ist: $t^k (A_0 + A_1 t + \dots + A_m t^m)$
$(b_0 + b_1 t + \dots + b_m t^m) e^{\alpha t}$	falls α keine Lösung der charakteristischen Gleichung ist: $(A_0 + A_1 t + \dots + A_m t^m) e^{\alpha t}$ falls α eine k -fache Lösung der charakteristischen Gleichung ist: $t^k (A_0 + A_1 t + \dots + A_m t^m) e^{\alpha t}$
$(b_0 + b_1 t + \dots + b_m t^m) \cos(\beta t)$ oder $(b_0 + b_1 t + \dots + b_m t^m) \sin(\beta t)$ oder eine Summe dieser Funktionentypen	falls βj keine Lösung der charakteristischen Gleichung ist: $(A_0 + A_1 t + \dots + A_m t^m) \cos(\beta t) + (B_0 + B_1 t + \dots + B_m t^m) \sin(\beta t)$ falls βj eine k -fache Lösung der charakteristischen Gleichung ist: $t^k [(A_0 + A_1 t + \dots + A_m t^m) \cos(\beta t) + (B_0 + B_1 t + \dots + B_m t^m) \sin(\beta t)]$
$(b_0 + b_1 t + \dots + b_m t^m) e^{\alpha t} \cos(\beta t)$ oder $(b_0 + b_1 t + \dots + b_m t^m) e^{\alpha t} \sin(\beta t)$ oder eine Summe dieser Funktionentypen	falls $\alpha + \beta j$ keine Lösung der charakteristischen Gleichung ist: $[(A_0 + A_1 t + \dots + A_m t^m) \cos(\beta t) + (B_0 + B_1 t + \dots + B_m t^m) \sin(\beta t)] e^{\alpha t}$ falls $\alpha + \beta j$ eine k -fache Lösung der charakteristischen Gleichung ist: $t^k [(A_0 + A_1 t + \dots + A_m t^m) \cos(\beta t) + (B_0 + B_1 t + \dots + B_m t^m) \sin(\beta t)] e^{\alpha t}$

Tabelle 14.1: Damit die „Methode der unbestimmten Koeffizienten“ angewendet werden kann, muss die rechte Seite $b(t)$ der inhomogenen linearen Differentialgleichung (IH_n) in der linken Spalte der Tabelle vertreten sein. Ist dieses der Fall, so wählt man einen Ansatz y_S für die spezielle Lösung gemäß der rechten Spalte der Tabelle.

Beispiel 14.26. (Bestimmung einer speziellen Lösung der inhomogenen linearen DGL mit der „Methode der unbestimmten Koeffizienten“)

- (a) Wir betrachten die inhomogene lineare Differentialgleichung zweiter Ordnung

$$y'' - 2y' + y = e^{3t}.$$

Schritt 1: Bestimmung des Fundamentalsystems der zugeh. hom. lin. DGL. Mit dem komplexen Exponentialansatz $y(t) = e^{st}$ finden wir die charakteristische Gleichung der zugehörigen homogenen linearen Differentialgleichung

$$y'' - 2y' + y = 0.$$

Die charakteristische Gleichung ist

$$0 = s^2 - 2s + 1 = (s - 1)^2, \quad \text{d.h. } s_1 = s_2 = 1 \quad (\text{Vielfachheit: } 2).$$

Damit finden wir ein reelles Fundamentalsystem $(e^t, t e^t)$.

Schritt 2: Finden einer speziellen Lösung. Es gilt $b(t) = e^{3t}$, und $s = 1$ ist die einzige (und eine zweifache Lösung) der charakteristischen Gleichung. Damit ist $s = 3$ also keine Lösung der charakteristischen Gleichung. Gemäß Tabelle 14.1 wählen wir daher den folgenden Ansatz:

$$y_S(t) = \gamma e^{3t}, \quad y'_S(t) = 3\gamma e^{3t}, \quad y''_S(t) = 9\gamma e^{3t}.$$

Einsetzen in die inhomogene lineare Differentialgleichung liefert:

$$9\gamma e^{3t} - 2(3\gamma e^{3t}) + \gamma e^{3t} = e^{3t} \iff 4\gamma e^{3t} = e^{3t} \iff \gamma = \frac{1}{4}.$$

Also ist

$$y_S : \mathbb{R} \rightarrow \mathbb{R}, \quad y_S(t) := \frac{1}{4} e^{3t},$$

eine spezielle Lösung.

Ergebnis: Die reelle allgemeine Lösung der inhomogenen linearen Differentialgleichung $y'' - 2y' + y = e^{3t}$ ist also

$$y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) := \frac{1}{4} e^{3t} + c_1 e^t + c_2 t e^t = \frac{1}{4} e^{3t} + (c_1 + c_2 t) e^t, \quad c_1, c_2 \in \mathbb{R}.$$

- (b) Wir betrachten die inhomogene lineare Differentialgleichung zweiter Ordnung

$$y'' - 2y' + y = e^t.$$

Schritt 1: Bestimmung des Fundamentalsystems der zugeh. hom. lin. DGL. Mit dem komplexen Exponentialansatz $y(t) = e^{st}$ finden wir die charakteristische Gleichung der zugehörigen homogenen linearen Differentialgleichung

$$y'' - 2y' + y = 0.$$

Die charakteristische Gleichung ist

$$0 = s^2 - 2s + 1 = (s - 1)^2, \quad \text{d.h. } s_1 = s_2 = 1 \quad (\text{Vielfachheit: } 2).$$

Damit finden wir ein reelles Fundamentalsystem $(e^t, t e^t)$.

Schritt 2: Finden einer speziellen Lösung. Es gilt $b(t) = e^t = e^{1t}$, und $s = 1$ ist eine zweifache Lösung der charakteristischen Gleichung. Gemäß Tabelle 14.1 wählen wir daher den folgenden Ansatz:

$$y_S(t) = t^2 \gamma e^t, \quad y'_S(t) = (2t + t^2) \gamma e^t, \quad y''_S(t) = (2 + 4t + t^2) \gamma e^t.$$

Einsetzen in die inhomogene lineare Differentialgleichung liefert:

$$\begin{aligned} & (2 + 4t + t^2) \gamma e^t - 2(2t + t^2) \gamma e^t + t^2 \gamma e^t = e^t \\ \iff & (2 + 4t + t^2 - 4t - 2t^2 + t^2) \gamma e^t = e^t \\ \iff & 2\gamma e^t = e^t \quad \iff \quad 2\gamma = 1 \quad \iff \quad \gamma = \frac{1}{2}. \end{aligned}$$

Also ist eine spezielle Lösung

$$y_S : \mathbb{R} \rightarrow \mathbb{R}, \quad y_S(t) := \frac{1}{2} t^2 e^t.$$

Ergebnis: Die reelle allgemeine Lösung der inhomogenen linearen Differentialgleichung $y'' - 2y' + y = e^t$ ist also

$$\begin{aligned} y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) &:= \frac{1}{2} t^2 e^t + (c_1 + c_2 t) e^t \\ &= \left(c_1 + c_2 t + \frac{1}{2} t^2 \right) e^t, \quad c_1, c_2 \in \mathbb{R}. \end{aligned}$$

- (c) Wir betrachten die inhomogene lineare Differentialgleichung zweiter Ordnung

$$y'' - 2y' + 2y = e^t \cos(t).$$

Schritt 1: Bestimmung des Fundamentalsystems der zugeh. hom. lin. DGL. Mit dem komplexen Exponentialansatz $y(t) = e^{st}$ finden wir die charakteristische Gleichung der zugehörigen homogenen linearen Differentialgleichung

$$y'' - 2y' + 2y = 0.$$

Die charakteristische Gleichung ist

$$\begin{aligned} 0 &= s^2 - 2s + 2 = (s - 1)^2 + 1 = (s - 1)^2 - j^2 = (s - 1 - j)(s - 1 + j) \\ \implies & \quad s_1 = 1 + j, \quad s_2 = 1 - j, \end{aligned}$$

d.h. $\alpha = 1$ und $\beta = 1$. Damit erhalten wir ein reelles Fundamentalsystem $(e^t \cos(t), e^t \sin(t))$.

Schritt 2: Finden einer speziellen Lösung. Es gilt $b(t) = e^t \cos(t)$, und $s = 1 + j$ ist eine einfache Lösung der charakteristischen Gleichung. Gemäß Tabelle 14.1 wählen wir daher den folgenden Ansatz:

$$y_S(t) = t \underbrace{(A \cos(t) + B \sin(t))}_{=y_H(t)} e^t = t y_H(t),$$

$$y'_S(t) = y_H(t) + t y'_H(t),$$

$$y''_S(t) = y'_H(t) + y'_H(t) + t y''_H(t) = 2 y'_H(t) + t y''_H(t),$$

wobei wir bereits wissen, dass die mit y_H abgekürzte Funktion eine Lösung der zugehörigen homogenen linearen Differentialgleichung ist. Einsetzen in die inhomogene lineare Differentialgleichung und Ausnutzen, dass y_H eine Lösung der zugehörigen homogenen Gleichung ist, liefert

$$\begin{aligned} & y'_S(t)' - 2 y'_S(t) + 2 y_S(t) = e^t \cos(t) \\ \iff & 2 y'_H(t) + t y''_H(t) - 2 (y_H(t) + t y'_H(t)) + 2 t y_H(t) = e^t \cos(t) \\ \iff & 2 y'_H(t) - 2 y_H(t) + t \underbrace{(y''_H(t) - 2 y'_H(t) + 2 y_H(t))}_{=0} = e^t \cos(t) \\ \iff & 2 y'_H(t) - 2 y_H(t) = e^t \cos(t). \end{aligned}$$

Mit

$$y_H(t) = (A \cos(t) + B \sin(t)) e^t,$$

$$y'_H(t) = (-A \sin(t) + B \cos(t)) e^t + (A \cos(t) + B \sin(t)) e^t$$

finden wir

$$\begin{aligned} & 2 y'_H(t) - 2 y_H(t) = e^t \cos(t) \\ \iff & 2 (-A \sin(t) + B \cos(t)) e^t + 2 (A \cos(t) + B \sin(t)) e^t \\ & \quad - 2 (A \cos(t) + B \sin(t)) e^t = e^t \cos(t) \\ \iff & 2 (-A \sin(t) + B \cos(t)) e^t = e^t \cos(t) \\ \iff & 2 (-A \sin(t) + B \cos(t)) = \cos(t) \\ \iff & -2A \sin(t) + (2B - 1) \cos(t) = 0 \\ \iff & -2A = 0 \quad \text{und} \quad 2B - 1 = 0 \\ \iff & A = 0 \quad \text{und} \quad B = \frac{1}{2}. \end{aligned}$$

Also ist

$$y_S : \mathbb{R} \rightarrow \mathbb{R}, \quad y_S(t) := \frac{1}{2} t e^t \sin(t),$$

eine spezielle Lösung der inhomogenen linearen Differentialgleichung.

Ergebnis: Die reelle allgemeine Lösung der inhomogenen linearen Differentialgleichung $y'' - 2y' + 2y = e^t \cos(t)$ ist also

$$y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) := \frac{1}{2} t e^t \sin(t) + (c_1 \cos(t) + c_2 \sin(t)) e^t, \quad c_1, c_2 \in \mathbb{R}.$$

In der Literatur finden Sie weitere Methoden zur Bestimmung einer speziellen Lösung der inhomogenen linearen Differentialgleichung, beispielsweise:

- Variation der Konstanten,
- den Potenzreihenansatz.

Diese werden aber hier nicht weiter behandelt.

14.7 Anfangswertaufgaben

In diesem letzten Teilkapitel betrachten wir nun noch Anfangswertprobleme.

Definition 14.27. (Anfangswertproblem für eine lineare DGL n -ter Ordnung)

Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, sei $n \in \mathbb{N}$ und seien $a_0, a_1, \dots, a_{n-1}, b : I \rightarrow \mathbb{R}$ stetige Funktionen. Seien $y_0, y_1, \dots, y_{n-1} \in \mathbb{R}$ und $t_0 \in I$. Die lineare Differentialgleichung n -ter Ordnung

$$y^{(n)} + a_{n-1}(t) y^{(n-1)} + \dots + a_1(t) y' + a_0(t) y = b(t)$$

zusammen mit den **Anfangsbedingungen**

$$y(t_0) = y_0, \quad y'(t_0) = y_1, \quad \dots, \quad y^{(n-1)}(t_0) = y_{n-1}$$

heißt ein **Anfangswertproblem (AWP)** oder eine **Anfangswertaufgabe**.

Durch die Vorgabe der Anfangsbedingungen werden die n Konstanten in der allgemeinen Lösung festgelegt. Betrachten wir zwei Beispiele für Anfangswertprobleme.

Beispiel 14.28. (Anfangswertprobleme)

(a) Betrachten wir das Anfangswertproblem

$$y'' + 2y' + 2y = 0, \quad y(0) = 1, \quad y'(0) = 2.$$

Schritt 1: Bestimmung des Fundamentalsystems der homogenen lin. DGL.
Mit dem komplexen Exponentialansatz $y(t) = e^{st}$ finden wir

$$s^2 e^{st} + 2s e^{st} + 2e^{st} = 0 \quad \iff \quad s^2 + 2s + 2 = 0,$$

und die charakteristische Gleichung kann wie folgt faktorisiert werden:

$$\begin{aligned} 0 &= s^2 + 2s + 2 = (s + 1)^2 + 1 = (s + 1)^2 - j^2 = (s + 1 - j)(s + 1 + j) \\ \implies & \quad s_1 = -1 + j, \quad s_2 = -1 - j, \end{aligned}$$

d.h. $\alpha = -1$ und $\beta = \pm 1$. Damit finden wir ein reelles Fundamentalsystem

$$(e^{-t} \cos(t), e^{-t} \sin(t)),$$

und die reelle allgemeine Lösung ist

$$\begin{aligned} y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) &:= c_1 e^{-t} \cos(t) + c_2 e^{-t} \sin(t) \\ &= e^{-t} [c_1 \cos(t) + c_2 \sin(t)], \quad c_1, c_2 \in \mathbb{R}. \end{aligned}$$

Schritt 2: Anfangsbedingung einarbeiten. Mit

$$\begin{aligned} y(t) &= e^{-t} [c_1 \cos(t) + c_2 \sin(t)], \\ y'(t) &= e^{-t} [-c_1 \cos(t) - c_2 \sin(t) - c_1 \sin(t) + c_2 \cos(t)], \end{aligned}$$

finden wir für die gegebenen Anfangsbedingungen in $t = 0$:

$$\begin{aligned} 1 &= y(0) = e^0 [c_1 \cos(0) + c_2 \sin(0)] = c_1, \\ 2 &= y'(0) = e^0 [-c_1 \cos(0) - c_2 \sin(0) - c_1 \sin(0) + c_2 \cos(0)] \\ &= -c_1 + c_2. \end{aligned}$$

Also gilt $c_1 = 1$ und $c_2 = 2 + c_1 = 3$.

Ergebnis: Die Lösung des Anfangswertproblems ist

$$y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) := e^{-t} [\cos(t) + 3 \sin(t)].$$

(b) Betrachten wir das Anfangswertproblem

$$y'' - 6y' + 9y = 2e^t + 9t - 15, \quad y(0) = y'(0) = 0.$$

Schritt 1: Bestimmung des Fundamentalsystems der zugeh. hom. lin. DGL. Wir finden die charakteristische Gleichung der zugehörigen homogenen linearen Differentialgleichung

$$y'' - 6y' + 9y = 0$$

mit dem komplexen Exponentialansatz $y(t) = e^{st}$:

$$s^2 e^{st} - 6s e^{st} + 9 e^{st} = 0 \quad \Longleftrightarrow \quad s^2 - 6s + 9 = 0$$

Die charakteristische Gleichung ist

$$0 = s^2 - 6s + 9 = (s - 3)^2 \quad \Longrightarrow \quad s_1 = s_2 = 3 \quad (\text{Vielfachheit: } 2).$$

Somit finden wir ein reelles Fundamentalsystem $(e^{3t}, t e^{3t})$.

Schritt 2: Finden einer speziellen Lösung. Gemäß Tabelle 14.1 wählen wir den folgenden Ansatz:

$$y_S(t) = A e^t + B_0 + B_1 t, \quad y'_S(t) = A e^t + B_1, \quad y''_S(t) = A e^t.$$

Einsetzen in die inhomogene lineare Differentialgleichung liefert:

$$\begin{aligned} y''_S(t) - 6y'_S(t) + 9y_S(t) &= 2e^t + 9t - 15, \\ \Longleftrightarrow A e^t - 6(A e^t + B_1) + 9(A e^t + B_0 + B_1 t) &= 2e^t + 9t - 15 \\ \Longleftrightarrow (4A - 2)e^t + (9B_0 - 6B_1 + 15) + (9B_1 - 9)t &= 0 \\ \Longleftrightarrow 4A - 2 = 0 \quad \text{und} \quad 9B_0 - 6B_1 + 15 = 0 \quad \text{und} \quad 9B_1 - 9 = 0 \\ \Longleftrightarrow A = \frac{1}{2}, \quad B_1 = 1, \quad B_0 = \frac{1}{9}(6B_1 - 15) = \frac{1}{9}(6 - 15) &= -1. \end{aligned}$$

Also ist

$$y_S : \mathbb{R} \rightarrow \mathbb{R}, \quad y_S(t) := \frac{1}{2} e^t + t - 1$$

eine spezielle Lösung.

Zwischenergebnis: Die reelle allgemeine Lösung der inhomogenen linearen Differentialgleichung ist

$$y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) := \frac{1}{2} e^t + t - 1 + (c_1 + c_2 t) e^{3t}, \quad c_1, c_2 \in \mathbb{R}.$$

Schritt 3: Anfangsbedingungen einarbeiten. Mit

$$\begin{aligned} y(t) &= \frac{1}{2} e^t + t - 1 + (c_1 + c_2 t) e^{3t}, \\ y'(t) &= \frac{1}{2} e^t + 1 + (c_2 + 3c_1 + 3c_2 t) e^{3t} \end{aligned}$$

finden wir für die gegebenen Anfangsbedingungen in $t = 0$:

$$0 = y(0) = \frac{1}{2} e^0 + 0 - 1 + (c_1 + c_2 \cdot 0) e^0 = \frac{1}{2} - 1 + c_1 = -\frac{1}{2} + c_1,$$

$$\begin{aligned} 0 = y'(0) &= \frac{1}{2} e^0 + 1 + (c_2 + 3c_1 + 3c_2 \cdot 0) e^0 \\ &= \frac{1}{2} + 1 + c_2 + 3c_1 = \frac{3}{2} + 3c_1 + c_2, \end{aligned}$$

also $c_1 = \frac{1}{2}$ und $c_2 = -\frac{3}{2} - 3c_1 = -\frac{3}{2} - \frac{3}{2} = -3$.

Ergebnis: Die Lösung des Anfangswertproblems ist

$$y : \mathbb{R} \rightarrow \mathbb{R}, \quad y(t) := \frac{1}{2} e^t + t - 1 + \left(\frac{1}{2} - 3t \right) e^{3t}.$$

Systeme linearer Differentialgleichungen

In diesem Kapitel betrachten wir Systeme linearer Differentialgleichungen erster Ordnung. Die Lösungstheorie linearer Differentialgleichungen erster Ordnung mit konstanten Koeffizienten führt auf quadratische Matrizen und nutzt die Eigenwerte, Eigenvektoren und Eigenräume quadratischer Matrizen, die wir in Teilkapitel 13.6 kennengelernt haben. Auch hier spielt ein komplexer Exponentialansatz eine Rolle. Am Ende lernen wir, dass man lineare Differentialgleichungen n -ter Ordnung leicht in ein System linearer Differentialgleichungen erster Ordnung überführen kann.

15.1 Wege in \mathbb{R}^n

Wir starten mit einem Teilkapitel über Wege im \mathbb{R}^n , weil man die Lösung eines linearen Differentialgleichungssystems als einen solchen Weg interpretieren und veranschaulichen kann.

Definition 15.1. (Weg in \mathbb{R}^n)

Seien $I \subseteq \mathbb{R}$ ein Intervall und $n \in \mathbb{N}$. Eine Funktion

$$\vec{x} : I \rightarrow \mathbb{R}^n, \quad \vec{x}(t) := \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix},$$

heißt ein **Weg** in \mathbb{R}^n .

Ein Weg ist also eine Funktion einer Variable mit Werten in \mathbb{R}^n . Sie sollten sich einen Weg $\vec{x}(t)$ als eine Kurve im \mathbb{R}^n vorstellen, die mit der Zeit durchlaufen wird. Zum Zeitpunkt t befindet man sich am Ort $\vec{x}(t)$. Natürlich können wir auch nach der Geschwindigkeit dieser Bewegung fragen – darauf kommen wir weiter unten zurück.

Betrachten wir einige Beispiele.

Beispiel 15.2. (Wege im \mathbb{R}^n)

(a) Seien $\vec{p}, \vec{v} \in \mathbb{R}^n$ mit $\vec{v} \neq \vec{0}$. Der Weg

$$\vec{x} : \mathbb{R} \rightarrow \mathbb{R}^n, \quad \vec{x}(t) := \vec{p} + t \vec{v},$$

durchläuft eine Gerade in \mathbb{R}^n .

(b) Der Weg

$$\vec{x} : [0, 2\pi] \rightarrow \mathbb{R}^2, \quad \vec{x}(t) := \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix},$$

durchläuft die Kreislinie in der Ebene um $(0, 0)$ mit Radius 1 genau einmal im Gegenuhrzeigersinn.

(c) Der Weg

$$\vec{x} : \mathbb{R} \rightarrow \mathbb{R}^3, \quad \vec{x}(t) := \begin{bmatrix} \cos(t) \\ \sin(t) \\ t \end{bmatrix},$$

beschreibt eine Schraubenlinie in \mathbb{R}^3 .

Da Wege als Komponenten (oder Einträge) reellwertige Funktionen haben können wie die üblichen Begriffe der Analysis, also Stetigkeit, Differenzierbarkeit, etc. auf Wege übertragen, indem wir verlangen, dass die entsprechende Eigenschaft für alle Komponenten(funktionen) gelten.

Definition 15.3. (stetiger und differenzierbarer Weg, C^1 -Weg)

Seien I ein Intervall, $\vec{x} : I \rightarrow \mathbb{R}^n$ ein Weg und $t_0 \in I$.

- (1) \vec{x} heißt **stetig in** t_0 , falls für jedes $k \in \{1, 2, \dots, n\}$ die Funktion $x_k : I \rightarrow \mathbb{R}$ stetig in t_0 ist.

- (2) \vec{x} heißt **stetig in I** , falls \vec{x} in jedem $t_0 \in I$ stetig ist.
- (3) \vec{x} heißt **differenzierbar in t_0** , falls für jedes $k \in \{1, 2, \dots, n\}$ die Funktion $x_k : I \rightarrow \mathbb{R}$ differenzierbar in t_0 ist. In diesem Fall heißt

$$\vec{x}'(t_0) = \begin{bmatrix} x'_1(t_0) \\ x'_2(t_0) \\ \vdots \\ x'_n(t_0) \end{bmatrix}$$

die **Ableitung von \vec{x} an der Stelle t_0** .

- (4) \vec{x} heißt **differenzierbar in I** , falls \vec{x} in jedem $t_0 \in I$ differenzierbar ist. $\vec{x}' : I \rightarrow \mathbb{R}^n$ heißt dann die **Ableitung von \vec{x}** .
- (5) Ist \vec{x} differenzierbar in I und ist die Ableitung \vec{x}' stetig in I , dann heißt \vec{x} ein **stetig differenzierbarer Weg** oder kurz ein **\mathcal{C}^1 -Weg**.

Wir halten zunächst einige Beobachtungen zu differenzierbaren Wegen fest.

Bemerkung 15.4. (differenzierbare Wege)

- (1) Ist ein Weg $\vec{x} : I \rightarrow \mathbb{R}^n$ **differenzierbar in t_0** , so ist \vec{x} auch **stetig in t_0** .
- (2) Wir haben Differenzierbarkeit in der HM A nur auf offenen Intervallen betrachtet. Falls I ein **abgeschlossenes Intervall** (oder ein **halboffenes Intervall**) ist, so verlangen wir bei einem **in I differenzierbaren Weg** (bzw. bei einem **\mathcal{C}^1 -Weg auf I**) $\vec{x} : I \rightarrow \mathbb{R}^n$, dass dieser auf einem offenen Intervall $]a, b[$ mit $I \subseteq]a, b[$ definiert werden kann und dort in allen Punkten $t \in I$ differenzierbar (bzw. stetig differenzierbar) ist. ($\vec{x} :]a, b[\rightarrow \mathbb{R}^n$ ist dann eine Fortsetzung des Weges \vec{x} von I auf das offene Intervall $]a, b[$.)
- (3) Analog ist $\vec{x} : I \rightarrow \mathbb{R}^n$ **in $t_0 \in I$ differenzierbar**, wenn \vec{x} auf einem offenen Intervall $]t_0 - \varepsilon, t_0 + \varepsilon[$ (mit einem beliebig kleinen festen $\varepsilon > 0$, das in der Regel in Abhängigkeit von t_0 gewählt ist) definiert ist und in t_0 differenzierbar ist.

Beispiel 15.5. (stetig differenzierbare Wege)

Die Wege \vec{x} aus Beispiel 15.2 sind alle stetig differenzierbar. Wir bestimmen jeweils die Ableitung \vec{x}' .

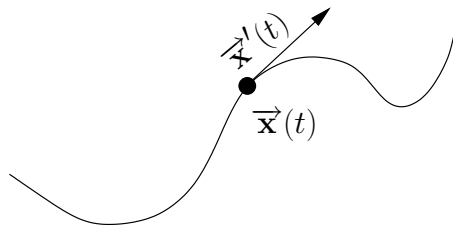


Abbildung 15.1: Veranschaulichung der Ableitung eines \mathcal{C}^1 -Wegs als Tangentialvektor oder Geschwindigkeitsvektor.

$$\vec{x} : \mathbb{R} \rightarrow \mathbb{R}^n, \vec{x}(t) := \vec{p} + t \vec{v}, \text{ mit } \vec{p}, \vec{v} \in \mathbb{R}^n, \vec{v} \neq \vec{0} \implies \vec{x}'(t) = \vec{v}$$

$$(b) \vec{x} : [0, 2\pi] \rightarrow \mathbb{R}^2, \vec{x}(t) := \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix} \implies \vec{x}'(t) = \begin{bmatrix} -\sin(t) \\ \cos(t) \end{bmatrix}$$

(Bei diesem auf dem abgeschlossenen Intervall $[0, 2\pi]$ definierten Weg, betrachten wir für das Differenzieren z.B. einfach seine Fortsetzung auf das offene Intervall \mathbb{R} .)

$$(c) \vec{x} : \mathbb{R} \rightarrow \mathbb{R}^3, \vec{x}(t) := \begin{bmatrix} \cos(t) \\ \sin(t) \\ t \end{bmatrix} \implies \vec{x}'(t) = \begin{bmatrix} -\sin(t) \\ \cos(t) \\ 1 \end{bmatrix}$$

In der nächsten Bemerkung halten wir die geometrische und die physikalische Interpretation der Ableitung eines \mathcal{C}^1 -Wegs fest.

Bemerkung 15.6. (geometrische und physikalische Interpretation der Ableitung eines \mathcal{C}^1 -Wegs)

- (1) **Geometrische Interpretation der Ableitung:** Sei \vec{x} ein \mathcal{C}^1 -Weg. $\vec{x}'(t)$ kann man als **Tangentialvektor im Punkt $\vec{x}(t)$** interpretieren. Man denkt sich $\vec{x}'(t)$ oft als Pfeil, der **tangential** an den Punkt $\vec{x}(t)$ angeheftet ist (vgl. Abbildung 15.1).
- (2) **Physikalische Interpretation der Ableitung:** Ist \vec{x} ein \mathcal{C}^1 -Weg, so kann man $\vec{x}'(t)$ als **Geschwindigkeitsvektor** interpretieren (vgl. Abbildung 15.1). Da die Ableitung, also der Geschwindigkeitsvektor, stetig ist ändert sich die Geschwindigkeit kontinuierlich und nicht abrupt.

Als Letztes halten wir Rechenregeln für die Ableitung von Wegen fest. Diese Rechenregeln folgen direkt aus den Rechenregeln für die Ableitungen reellwertiger

Funktionen, die wir in der HM A (siehe Kapitel 7 des Skripts der HM A) kennengelernt haben.

Satz 15.7. (Rechenregeln für die Ableitung differenzierbarer Wege)

Seien $\vec{x} : I \rightarrow \mathbb{R}^n$, $\vec{y} : I \rightarrow \mathbb{R}^n$ differenzierbare Wege und $\lambda \in \mathbb{R}$.

(1) **Linearität:** $\vec{x} + \vec{y}$ und $\lambda \cdot \vec{x}$ sind differenzierbar in I und

$$\begin{aligned}(\vec{x} + \vec{y})'(t) &= \vec{x}'(t) + \vec{y}'(t), \\ (\lambda \vec{x})'(t) &= \lambda \vec{x}'(t).\end{aligned}$$

(2) **Produktregeln:**

- Ist $\varphi : I \rightarrow \mathbb{R}$ eine differenzierbare Funktion, so ist $\varphi \vec{x}$ ein differenzierbarer Weg und

$$(\varphi \vec{x})'(t) = \varphi'(t) \vec{x}(t) + \varphi(t) \vec{x}'(t)$$

- Das Standardskalarprodukt $\langle \vec{x} | \vec{y} \rangle : I \rightarrow \mathbb{R}$ der Wege \vec{x} und \vec{y} ist eine differenzierbare reellwertige Funktion und die Ableitung ist

$$\frac{d}{dt} \langle \vec{x}(t) | \vec{y}(t) \rangle = \langle \vec{x}'(t) | \vec{y}(t) \rangle + \langle \vec{x}(t) | \vec{y}'(t) \rangle.$$

- Ist $n = 3$, dann ist das Vektorprodukt $\vec{x} \times \vec{y} : I \rightarrow \mathbb{R}^3$ differenzierbar und

$$(\vec{x} \times \vec{y})'(t) = \vec{x}'(t) \times \vec{y}(t) + \vec{x}(t) \times \vec{y}'(t).$$

(3) **Kettenregel:** Ist $J \subseteq \mathbb{R}$ ebenfalls ein Intervall und $\alpha : J \rightarrow I$ differenzierbar, so ist der Weg $\vec{x} \circ \alpha : J \rightarrow \mathbb{R}^n$ differenzierbar und

$$(\vec{x} \circ \alpha)'(t) = \alpha'(t) \vec{x}'(\alpha(t)).$$

Bemerkung 15.8. ($\mathcal{C}^1(I, \mathbb{R}^n)$ ist ein \mathbb{R} -Vektorraum)

Die Menge $\mathcal{C}^1(I, \mathbb{R}^n)$ aller \mathcal{C}^1 -Wege $\vec{x} : I \rightarrow \mathbb{R}^n$ mit der üblichen Addition und der üblichen Multiplikation mit Skalaren ist ein \mathbb{R} -Vektorraum. Man überprüft dieses, indem man die Vektorraumeigenschaften nachrechnet.

Die Tatsache, dass die Menge aller \mathcal{C}^1 -Wege ein \mathbb{R} -Vektorraum ist, werden wir im

nächsten Teilkapitel benutzen.

15.2 Allgemeine Systeme linearer Differentialgleichungen

Als Vorbereitung erinnern wir uns zunächst an den **Lösungsweg für inhomogene lineare Differentialgleichungen erster Ordnung** (vgl. Kapitel 9 im Skript der HM A)

$$y' = a(t)y + b(t).$$

Schritt 1 des Lösungswegs: Lösen der zugehörigen homogenen linearen Differentialgleichung $y' = a(t)y$.

Schritt 2 des Lösungswegs: Variation der Konstanten, also wir wählen den Ansatz $y_S(t) = u(t)y_H(t)$, wobei y_H eine nichttriviale Lösung der zugehörigen homogenen linearen Differentialgleichung $y' = a(t)y$ ist.

Unser Ziel in diesem Teilkapitel ist es, diesen Ansatz auf Systeme linearer Differentialgleichungen zu verallgemeinern.

Notation 15.9. (Voraussetzungen für dieses Teilkapitel)

In diesem Teilkapitel seien immer $I \subseteq \mathbb{R}$ ein Intervall und für $t \in I$

$$\mathbf{A}(t) := \begin{bmatrix} a_{1,1}(t) & a_{1,2}(t) & \cdots & a_{1,n}(t) \\ a_{2,1}(t) & a_{2,2}(t) & \cdots & a_{2,n}(t) \\ \vdots & & \ddots & \vdots \\ a_{n,1}(t) & a_{n,2}(t) & \cdots & a_{n,n}(t) \end{bmatrix} \in \mathbb{R}^{n \times n} \quad \text{und} \quad \vec{\mathbf{b}}(t) := \begin{bmatrix} b_1(t) \\ b_2(t) \\ \vdots \\ b_n(t) \end{bmatrix} \in \mathbb{R}^n,$$

wobei die $a_{i,k} : I \rightarrow \mathbb{R}$ und $b_i : I \rightarrow \mathbb{R}$, $i = 1, 2, \dots, n$; $k = 1, 2, \dots, n$, stetige Funktionen seien.

In anderen Worten: $\vec{\mathbf{b}} : I \rightarrow \mathbb{R}^n$ ist ein stetiger Weg und $\mathbf{A} : I \rightarrow \mathbb{R}^{n \times n}$ ist eine stetige matrixwertige Funktion.

Zunächst definieren wir Systeme linearer Differentialgleichungen und die entsprechenden Anfangswertprobleme

Ordnung mit einem Anfangswert, also für ein Anfangswertproblem.

Beispiel 15.11. (Anfangswertproblem)

Sei $I =]0, \infty[$, $t_0 = 2$, und wir betrachten das Anfangswertproblem

$$\vec{x}' = \begin{bmatrix} 0 & 1 \\ -\frac{2}{t^2} & \frac{2}{t} \end{bmatrix} \vec{x} + \begin{bmatrix} t^4 \\ t^3 \end{bmatrix}, \quad \vec{x}(2) = \begin{bmatrix} 1 \\ 4 \end{bmatrix}.$$

Ausgeschrieben lautet dieses

$$\begin{aligned} x_1' &= x_2 + t^4, \\ x_2' &= -\frac{2}{t^2} x_1 + \frac{2}{t} x_2 + t^3 \end{aligned}$$

mit dem Anfangswerten $x_1(2) = 1$ und $x_2(2) = 4$.

Bemerkung 15.12. (Lösbarkeit eines Anfangswertproblems)

Man kann zeigen, dass jedes Anfangswertproblem für (15.1) **genau eine** Lösung hat.

Satz 15.13. (Lösungsmenge des homogenen Systems linearer DGLen, Fundamentalsystem, Fundamentalmatrix, allgemeine Lösung)

- (1) Die **Lösungsmenge** $\mathbb{L}_{(H_n)}$ des homogenen (zu (15.1) gehörenden) Systems linearer Differentialgleichungen

$$\vec{x}' = \mathbf{A}(t) \vec{x} \quad (H_n)$$

ist ein n -dimensionaler \mathbb{R} -Vektorraum. Eine Basis $(\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n)$ von $\mathbb{L}_{(H_n)}$ heißt ein **Fundamentalsystem (FS)** für (H_n) . Die daraus gebildete Matrix

$$\Phi(t) := [\vec{\varphi}_1(t) \quad \vec{\varphi}_2(t) \quad \cdots \quad \vec{\varphi}_n(t)], \quad t \in I,$$

heißt eine **Fundamentalmatrix (FM)** für (H_n) .

- (2) Ist $(\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n)$ ein (reelles) Fundamentalsystem für (H_n) und $\Phi(t)$ die zugehörige Fundamentalmatrix, so ist

$$\vec{x}_H : I \rightarrow \mathbb{R}^n, \quad \vec{x}_H(t) := c_1 \vec{\varphi}_1(t) + c_2 \vec{\varphi}_2(t) + \cdots + c_n \vec{\varphi}_n(t)$$

$$= \Phi(t) \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}, \quad c_1, c_2, \dots, c_n \in \mathbb{R},$$

die **allgemeine Lösung** von (H_n) .

(3) Ist $\vec{x}_S : I \rightarrow \mathbb{R}^n$ irgendeine („spezielle“) **Lösung von (15.1)** und ist \vec{x}_H die **allgemeine Lösung** von (H_n) aus (2), so ist $\vec{x} = \vec{x}_S + \vec{x}_H$ die **allgemeine Lösung** von (15.1).

Beweis von Satz 15.13:

(1) Zunächst zeigt man, dass die Lösungsmenge $\mathbb{L}_{(H_n)}$ von (H_n) ein Untervektorraum des \mathbb{R} -Vektorraums aller \mathcal{C}^1 -Wege auf I ist. Wir zeigen dieses in einer Übungsaufgabe.

Wir zeigen nun, dass $\dim(\mathbb{L}_{(H_n)}) = n$ gilt.

Sei $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ die Standardbasis von \mathbb{R}^n und sei $t_0 \in I$. Nach Bemerkung 15.12 hat jedes der n Anfangswertprobleme

$$\vec{x}' = \mathbf{A}(t) \vec{x}, \quad \vec{x}(t_0) = \vec{e}_k, \quad k = 1, 2, \dots, n,$$

genau eine Lösung $\vec{\varphi}_k$.

Wir zeigen nun, dass $(\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n)$ eine Basis von $\mathbb{L}_{(H_n)}$ ist:

(i) $\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n$ sind linear unabhängig, denn:

Seien $c_1, c_2, \dots, c_n \in \mathbb{R}$ mit

$$c_1 \vec{\varphi}_1 + c_2 \vec{\varphi}_2 + \dots + c_n \vec{\varphi}_n = \vec{0}.$$

Dann gilt

$$c_1 \vec{\varphi}_1(t) + c_2 \vec{\varphi}_2(t) + \dots + c_n \vec{\varphi}_n(t) = \vec{0} \quad \text{für alle } t \in I,$$

und insbesondere folgt für $t = t_0$

$$\begin{aligned} c_1 \underbrace{\vec{\varphi}_1(t_0)}_{=\vec{e}_1} + c_2 \underbrace{\vec{\varphi}_2(t_0)}_{=\vec{e}_2} + \dots + c_n \underbrace{\vec{\varphi}_n(t_0)}_{=\vec{e}_n} &= \vec{0} \\ \iff c_1 \vec{e}_1 + c_2 \vec{e}_2 + \dots + c_n \vec{e}_n &= \vec{0}, \end{aligned}$$

und wegen der linearen Unabhängigkeit von $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ folgt, dass $c_1 = c_2 = \dots = c_n = 0$ ist. Also sind $\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n$ linear unabhängig.

- (ii) Jedes $\vec{x} \in \mathbb{L}_{(H_n)}$ ist eine Linearkombination von $\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n$, denn:
Sei $\vec{x} \in \mathbb{L}_{(H_n)}$ und sei

$$\vec{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} := \vec{x}(t_0) \in \mathbb{R}^n.$$

Dann ist

$$\vec{w} := c_1 \vec{\varphi}_1 + c_2 \vec{\varphi}_2 + \dots + c_n \vec{\varphi}_n$$

eine Lösung von (H_n) , da $\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n$ Lösungen von (H_n) sind und da $\mathbb{L}_{(H_n)}$ ein \mathbb{R} -Vektorraum ist. Außerdem gilt

$$\vec{w}(t_0) = c_1 \vec{\varphi}_1(t_0) + c_2 \vec{\varphi}_2(t_0) + \dots + c_n \vec{\varphi}_n(t_0) = c_1 \vec{e}_1 + c_2 \vec{e}_2 + \dots + c_n \vec{e}_n = \vec{c}.$$

Also sind sowohl \vec{x} als auch \vec{w} Lösungen des Anfangswertproblems

$$\vec{x}' = \mathbf{A}(t) \vec{x}, \quad \vec{x}(t_0) = \vec{c}.$$

Nach Bemerkung 15.12 gibt es aber genau eine Lösung dieses Anfangswertproblems, d.h. es folgt

$$\vec{x} = \vec{w} = c_1 \vec{\varphi}_1 + c_2 \vec{\varphi}_2 + \dots + c_n \vec{\varphi}_n.$$

Also gilt $\vec{x} \in \text{LH}(\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n)$.

Damit haben wir gezeigt, dass $(\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n)$ ein Basis für $\mathbb{L}_{(H_n)}$ ist und dass somit insbesondere $\mathbb{L}_{(H_n)}$ ein n -dimensionaler \mathbb{R} -Vektorraum ist.

- (2) Satz 15.13 (2) folgt direkt aus den Eigenschaften einer Basis, weil jedes Fundamentalsystem eine Basis von $\mathbb{L}_{(H)}$ ist.
- (3) Um Satz 15.13 (3) zu zeigen, betrachten wir eine beliebige Lösung \vec{x} des inhomogenen Systems linearer Differentialgleichungen (15.1). Dann gilt für $\vec{y} := \vec{x} - \vec{x}_S$

$$\begin{aligned} \mathbf{A}(t) \vec{y} &= \mathbf{A}(t) (\vec{x} - \vec{x}_S) = \mathbf{A}(t) \vec{x} - \mathbf{A}(t) \vec{x}_S \\ &= \underbrace{(\mathbf{A}(t) \vec{x} + \vec{b}(t))}_{=\vec{x}'(t)} - \underbrace{(\mathbf{A}(t) \vec{x}_S + \vec{b}(t))}_{=\vec{x}_S'(t)} \\ &= \vec{x}'(t) - \vec{x}_S'(t) = \vec{y}'(t), \end{aligned}$$

d.h. \vec{y} ist eine Lösung von (H_n) . Also gilt $\vec{x} = \vec{x}_S + \vec{y}$ mit $\vec{y} \in \mathbb{L}_{(H_n)}$.

Umgekehrt ist jeder \mathcal{C}^1 -Weg der Form $\vec{x} := \vec{x}_S + \vec{x}_H$ mit $\vec{x}_H \in \mathbb{L}_{(H_n)}$ eine Lösung von (15.1), denn:

$$\begin{aligned} \mathbf{A}(t) \vec{x} + \vec{b}(t) &= \mathbf{A}(t) (\vec{x}_S + \vec{x}_H) + \vec{b}(t) \\ &= \underbrace{\mathbf{A}(t) \vec{x}_S + \vec{b}(t)}_{=\vec{x}_S'(t)} + \underbrace{\mathbf{A}(t) \vec{x}_H}_{=\vec{x}_H'(t)} \\ &= \vec{x}_S'(t) + \vec{x}_H'(t) = \vec{x}'(t). \end{aligned}$$

Daraus folgt dann, dass $\vec{x} := \vec{x}_S + \vec{x}_H$ mit $\vec{x}_H \in \mathbb{L}_{(H_n)}$ die allgemeine Lösung von (15.1) ist. \square

Bemerkung 15.14. (Wronski-Determinante)

Seien $\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n$ Lösungen von (H_n) . Um herauszufinden, ob $(\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n)$ ein Fundamentalsystem für (H_n) ist, bildet man die **Wronski-Determinante**

$$W(t) := \det \left(\begin{bmatrix} \vec{\varphi}_1(t) & \vec{\varphi}_2(t) & \cdots & \vec{\varphi}_n(t) \end{bmatrix} \right), \quad t \in I.$$

Falls $W(t) \neq 0$ für irgendein $t = t_0 \in I$ gilt, so sind $\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n$ linear unabhängig.

(*Erklärung:* Aus der Gleichung $c_1 \vec{\varphi}_1 + c_2 \vec{\varphi}_2 + \dots + c_n \vec{\varphi}_n = \vec{0}$ folgt das insbesondere

$$\begin{aligned} c_1 \vec{\varphi}_1(t) + c_2 \vec{\varphi}_2(t) + \dots + c_n \vec{\varphi}_n(t) &= \vec{0} \\ \iff \begin{bmatrix} \vec{\varphi}_1(t) & \vec{\varphi}_2(t) & \cdots & \vec{\varphi}_n(t) \end{bmatrix} \vec{c} &= \vec{0} \end{aligned}$$

für jedes $t \in I$. Aus Kapitel 13 wissen wir, dass dieses lineare Gleichungssystem genau dann nur die Lösung $\vec{c} = \vec{0}$ hat, wenn

$$\underbrace{\det \left(\begin{bmatrix} \vec{\varphi}_1(t) & \vec{\varphi}_2(t) & \cdots & \vec{\varphi}_n(t) \end{bmatrix} \right)}_{=W(t)} \neq 0 \quad (15.2)$$

ist. Finden wir ein $t = t_0 \in I$ mit (15.2), so folgt $c_1 = c_2 = \dots = c_n = 0$. \square

Beispiel 15.15. (homogenes System linearer Differentialgleichungen, Fundamentalsystem, Fundamentalmatrix, allgemeine Lösung)

Sei $I =]0, \infty[$, und wir betrachten das homogene System linearer Differentialglei-

chungen erster Ordnung

$$\vec{x}' = \mathbf{A}(t) \vec{x} \quad \text{mit} \quad \mathbf{A}(t) := \begin{bmatrix} 0 & 1 \\ -\frac{2}{t^2} & \frac{2}{t} \end{bmatrix},$$

also ausgeschrieben

$$\begin{aligned} x_1' &= x_2, \\ x_2' &= -\frac{2}{t^2} x_1 + \frac{2}{t} x_2. \end{aligned}$$

Hier sind

$$\vec{\varphi}_1 : \mathbb{R} \rightarrow \mathbb{R}^2, \quad \vec{\varphi}_1(t) := \begin{bmatrix} t \\ 1 \end{bmatrix}, \quad \text{und} \quad \vec{\varphi}_2 : \mathbb{R} \rightarrow \mathbb{R}^2, \quad \vec{\varphi}_2(t) := \begin{bmatrix} t^2 \\ 2t \end{bmatrix},$$

zwei linear unabhängige Lösungen, denn:

(i) Es gilt

$$\begin{aligned} \mathbf{A}(t) \vec{\varphi}_1(t) &= \begin{bmatrix} 0 & 1 \\ -\frac{2}{t^2} & \frac{2}{t} \end{bmatrix} \begin{bmatrix} t \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -\frac{2}{t} + \frac{2}{t} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \vec{\varphi}_1'(t), \\ \mathbf{A}(t) \vec{\varphi}_2(t) &= \begin{bmatrix} 0 & 1 \\ -\frac{2}{t^2} & \frac{2}{t} \end{bmatrix} \begin{bmatrix} t^2 \\ 2t \end{bmatrix} = \begin{bmatrix} 2t \\ -2 + 4 \end{bmatrix} = \begin{bmatrix} 2t \\ 2 \end{bmatrix} = \vec{\varphi}_2'(t). \end{aligned}$$

Also sind $\vec{\varphi}_1, \vec{\varphi}_2$ Lösungen von $\vec{x}' = \mathbf{A}(t) \vec{x}$.

(ii) Weiter gilt

$$W(t) = \det([\vec{\varphi}_1(t) \quad \vec{\varphi}_2(t)]) = \det\left(\begin{bmatrix} t & t^2 \\ 1 & 2t \end{bmatrix}\right) = 2t^2 - t^2 = t^2 \neq 0$$

für alle $t \in I$, d.h. $\vec{\varphi}_1, \vec{\varphi}_2$ sind linear unabhängig.

Also ist

$$(\vec{\varphi}_1, \vec{\varphi}_2) = \left(\begin{bmatrix} t \\ 1 \end{bmatrix}, \begin{bmatrix} t^2 \\ 2t \end{bmatrix} \right)$$

ein Fundamentalsystem für $\vec{x}' = \mathbf{A}(t) \vec{x}$. Die Fundamentalmatrix ist

$$\Phi(t) := \begin{bmatrix} t & t^2 \\ 1 & 2t \end{bmatrix},$$

und die allgemeine Lösung ist

$$\vec{x} : \mathbb{R} \rightarrow \mathbb{R}^2, \quad \vec{x}(t) := c_1 \begin{bmatrix} t \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} t^2 \\ 2t \end{bmatrix} = \begin{bmatrix} t & t^2 \\ 1 & 2t \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}, \quad c_1, c_2 \in \mathbb{R}.$$

Bemerkung 15.16. (DGL der Fundamentalmatrix)

Ist $(\vec{\varphi}_1, \vec{\varphi}_2, \dots, \vec{\varphi}_n)$ ein Fundamentalsystem für das homogene System linearer Differentialgleichungen (H_n) und ist

$$\Phi(t) = [\vec{\varphi}_1(t) \quad \vec{\varphi}_2(t) \quad \cdots \quad \vec{\varphi}_n(t)], \quad t \in I,$$

die zugehörige Fundamentalmatrix für (H_n) , so erfüllt diese die Differentialgleichung

$$\Phi'(t) = \mathbf{A}(t) \cdot \Phi(t) \quad \text{für alle } t \in I. \quad (15.3)$$

Dabei ist die Fundamentalmatrix komponentenweise zu differenzieren, also

$$\Phi'(t) = [\vec{\varphi}_1'(t) \quad \vec{\varphi}_2'(t) \quad \cdots \quad \vec{\varphi}_n'(t)].$$

Aus $\vec{\varphi}_k'(t) = \mathbf{A}(t) \vec{\varphi}_k(t)$, $k = 1, 2, \dots, n$, folgt (15.3) dann direkt.

Der nächste Satz erklärt, wie man die allgemeine Lösung des inhomogenen Systems linearer Differentialgleichungen findet, wenn man bereits ein Fundamentalsystem des zugehörigen homogenen Systems linearer Differentialgleichungen (H_n) kennt. Allerdings müssen wir später noch lernen, wie man ein solches Fundamentalsystem findet. – Die Methode „Variation der Konstanten“, die wir im nachfolgenden Satz kennenlernen, ist eine Verallgemeinerung der Methode gleichen Namens für inhomogene lineare Differentialgleichungen erster Ordnung (vgl. Kapitel 9 im Skript der HM A).

Satz 15.17. (Variation der Konstanten)

Sei $\Phi = [\vec{\varphi}_1 \quad \vec{\varphi}_2 \quad \cdots \quad \vec{\varphi}_n]$ eine Fundamentalmatrix für $\vec{x}' = \mathbf{A}(t)\vec{x}$. Dann ist die **allgemeine Lösung** von

$$\vec{x}' = \mathbf{A}(t) \vec{x} + \vec{b}(t)$$

gegeben durch

$$\vec{x}(t) := \Phi(t) \int (\Phi(t))^{-1} \vec{b}(t) dt. \quad (15.4)$$

Dabei ist das Integral so zu lesen, dass wir jede Komponente der vektorwertigen Funktion $(\Phi(t))^{-1} \vec{b}(t)$ einzeln (unbestimmt) über t integrieren, also für jede Komponente der vektorwertigen Funktion $(\Phi(t))^{-1} \vec{b}(t)$ eine Stammfunktion plus eine Integrationskonstante angeben.

Teilbeweis von Satz 15.17: Nach Bemerkung 15.16 gilt

$$\Phi'(t) = \mathbf{A}(t) \Phi(t) \quad \text{für alle } t \in I. \quad (15.5)$$

Sei nun \vec{x} durch (15.4) gegeben. Wir leiten \vec{x} ab und zeigen mit Hilfe von (15.5), dass \vec{x} eine Lösung von $\vec{x}' = \mathbf{A}(t) \vec{x} + \vec{b}$ ist: Mit der Produktregel finden wir

$$\begin{aligned} \vec{x}'(t) &= \frac{d}{dt} \left(\Phi(t) \int (\Phi(t))^{-1} \vec{b}(t) dt \right) \\ &= \left(\frac{d}{dt} \Phi(t) \right) \int (\Phi(t))^{-1} \vec{b}(t) dt + \Phi(t) \left(\frac{d}{dt} \int (\Phi(t))^{-1} \vec{b}(t) dt \right) \\ &= \Phi'(t) \int (\Phi(t))^{-1} \vec{b}(t) dt + \underbrace{\Phi(t) (\Phi(t))^{-1}}_{= \mathbf{E}_n} \vec{b}(t) \\ &\stackrel{(15.5)}{=} \underbrace{\mathbf{A}(t) \Phi(t) \int (\Phi(t))^{-1} \vec{b}(t) dt}_{= \vec{x}(t)} + \vec{b}(t) \\ &= \mathbf{A}(t) \vec{x}(t) + \vec{b}(t). \end{aligned}$$

Damit haben wir gezeigt, dass \vec{x} , gegeben durch (15.4), eine (spezielle) Lösung von $\vec{x}' = \mathbf{A}(t) \vec{x} + \vec{b}$ ist. In dieser Lösung kommen n Integrationskonstanten von den n Integralen über die Komponentenfunktionen in (15.4) vor, und die Lösung (15.4) ist von der Form

$$\vec{x} = \vec{x}_S + c_1 \vec{\varphi}_1 + c_2 \vec{\varphi}_2 + \dots + c_n \vec{\varphi}_n, \quad c_1, c_2, \dots, c_n \in \mathbb{R}. \quad (15.6)$$

Letzteres werden wir allerdings hier nicht beweisen. Nach Satz 15.13 ist (15.6) die allgemeine Lösung von $\vec{x}' = \mathbf{A}(t) \vec{x} + \vec{b}$. \square

Betrachten wir ein Beispiel.

Beispiel 15.18. (Variation der Konstanten)

Sei $I =]0, \infty[$, und wir betrachten

$$\vec{x}' = \begin{bmatrix} 0 & 1 \\ -\frac{2}{t^2} & \frac{2}{t} \end{bmatrix} \vec{x} + \begin{bmatrix} t^4 \\ t^3 \end{bmatrix}, \quad \vec{x}(2) = \begin{bmatrix} 1 \\ 4 \end{bmatrix}.$$

Schritt 1: Finden der allgemeinen Lösung des zugehörigen homogenen Systems linearer Differentialgleichungen.

Wie man diesen Schritt durchführt lernen wir in den nächsten Teilkapiteln für den Fall konstanter Koeffizienten. Für das gegebene Beispiel wissen wir allerdings bereits aus Beispiel 15.15, dass eine Fundamentalmatrix des zugehörigen homogenen

Systems linearer Differentialgleichungen durch

$$\Phi(t) := \begin{bmatrix} t & t^2 \\ 1 & 2t \end{bmatrix}$$

gegeben ist.

Schritt 2: Variation der Konstanten.

Mit der Wronski-Determinante

$$W(t) = \det(\Phi(t)) = \det \begin{pmatrix} t & t^2 \\ 1 & 2t \end{pmatrix} = 2t^2 - t^2 = t^2$$

gilt

$$(\Phi(t))^{-1} = \frac{1}{t^2} \begin{bmatrix} 2t & -t^2 \\ -1 & t \end{bmatrix} = \begin{bmatrix} \frac{2}{t} & -1 \\ -\frac{1}{t^2} & \frac{1}{t} \end{bmatrix} \quad (15.7)$$

und

$$\int \begin{bmatrix} \frac{2}{t} & -1 \\ -\frac{1}{t^2} & \frac{1}{t} \end{bmatrix} \begin{bmatrix} t^4 \\ t^3 \end{bmatrix} dt = \int \begin{bmatrix} t^3 \\ 0 \end{bmatrix} dt = \begin{bmatrix} \int t^3 dt \\ \int 0 dt \end{bmatrix} = \begin{bmatrix} \frac{1}{4}t^4 + c_1 \\ 0 + c_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{4}t^4 \\ 0 \end{bmatrix} + \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

mit $c_1, c_2 \in \mathbb{R}$. Nach Satz 15.17 ist

$$\begin{aligned} \vec{x}(t) &= \begin{bmatrix} t & t^2 \\ 1 & 2t \end{bmatrix} \left(\int \begin{bmatrix} \frac{2}{t} & -1 \\ -\frac{1}{t^2} & \frac{1}{t} \end{bmatrix} \begin{bmatrix} t^4 \\ t^3 \end{bmatrix} dt \right) = \begin{bmatrix} t & t^2 \\ 1 & 2t \end{bmatrix} \left(\begin{bmatrix} \frac{1}{4}t^4 \\ 0 \end{bmatrix} + \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \right) \\ &= \begin{bmatrix} \frac{1}{4}t^5 \\ \frac{1}{4}t^4 \end{bmatrix} + c_1 \begin{bmatrix} t \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} t^2 \\ 2t \end{bmatrix} = \frac{1}{4} \begin{bmatrix} t^5 \\ t^4 \end{bmatrix} + \Phi(t) \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \end{aligned} \quad (15.8)$$

die allgemeine Lösung des inhomogenen Systems linearer Differentialgleichungen.

Schritt 3: Anfangsbedingung einarbeiten.

Wir setzen $t = 2$ in die allgemeine Lösung (15.8) ein und verlangen, dass die Anfangsbedingung gilt:

$$\begin{aligned} \begin{bmatrix} 1 \\ 4 \end{bmatrix} &= \vec{x}(2) = \frac{1}{4} \begin{bmatrix} 2^5 \\ 2^4 \end{bmatrix} + \Phi(2) \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix} + \Phi(2) \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \\ \iff & \begin{bmatrix} -7 \\ 0 \end{bmatrix} = \Phi(2) \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \\ \iff & \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = (\Phi(2))^{-1} \begin{bmatrix} -7 \\ 0 \end{bmatrix} \stackrel{(15.7)}{=} \begin{bmatrix} 1 & -1 \\ -\frac{1}{4} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} -7 \\ 0 \end{bmatrix} = \begin{bmatrix} -7 \\ \frac{7}{4} \end{bmatrix}, \end{aligned}$$

d.h. $c_1 = -7$ und $c_2 = 7/4$

Ergebnis: Die Lösung des Anfangswertproblems ist

$$\vec{x} : \mathbb{R} \rightarrow \mathbb{R}^2, \quad \vec{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \frac{1}{4} \begin{bmatrix} t^5 \\ t^4 \end{bmatrix} - 7 \begin{bmatrix} t \\ 1 \end{bmatrix} + \frac{7}{4} \begin{bmatrix} t^2 \\ 2t \end{bmatrix} = \begin{bmatrix} \frac{1}{4}t^5 + \frac{7}{4}t^2 - 7t \\ \frac{1}{4}t^4 + \frac{7}{2}t - 7 \end{bmatrix}.$$

15.3 Homogene Systeme mit konstanten Koeffizienten: Der Fall $n = 2$

In diesem Teilkapitel lernen wir für den Fall $n = 2$, also für ein **homogenes** System **zweier** linearer Differentialgleichungen erster Ordnung mit **konstanten Koeffizienten**, wie man dieses löst. Dabei werden wir unser Wissen über die Eigenwerte und Eigenvektoren quadratischer Matrizen aus Teilkapitel 13.6 anwenden.

Im Folgenden sei $\mathbf{A} \in \mathbb{R}^{2 \times 2}$ eine Matrix mit konstanten Einträgen. Wir betrachten in diesem Teilkapitel:

$$\begin{aligned} \vec{x}' = \mathbf{A} \vec{x} &\iff \begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} & \text{(H}_2\text{)} \\ &\iff \begin{cases} x_1'(t) = a_{1,1} x_1(t) + a_{1,2} x_2(t) \\ x_2'(t) = a_{2,1} x_1(t) + a_{2,2} x_2(t) \end{cases} \end{aligned}$$

In Verallgemeinerung des komplexen Exponentialansatzes aus Kapitel 14 wählen wir den folgenden (**vektorwertigen**) **komplexen Exponentialansatz**:

$$\vec{x}(t) = e^{\lambda t} \vec{u} = \begin{bmatrix} e^{\lambda t} u_1 \\ e^{\lambda t} u_2 \end{bmatrix} \quad \text{mit} \quad \lambda \in \mathbb{C}, \quad \vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \in \mathbb{C}^2 \setminus \{\vec{0}\}. \quad (15.9)$$

Die Ableitung dieses Ansatzes ist

$$\vec{x}'(t) = \lambda e^{\lambda t} \vec{u} = \begin{bmatrix} \lambda e^{\lambda t} u_1 \\ \lambda e^{\lambda t} u_2 \end{bmatrix}. \quad (15.10)$$

Einsetzen des Ansatzes (15.9) und seiner Ableitung (15.10) in (H₂) ergibt:

$$\lambda e^{\lambda t} \vec{u} = \mathbf{A} (e^{\lambda t} \vec{u}) = e^{\lambda t} \mathbf{A} \vec{u} \quad \Big| : e^{\lambda t} \iff \lambda \vec{u} = \mathbf{A} \vec{u}. \quad (15.11)$$

An (15.11) lesen wir ab: $\vec{x}(t) = e^{\lambda t} \vec{u}$ ist **genau dann eine nicht-triviale Lösung von (H₂)**, wenn die folgenden Bedingungen beide gelten:

- (i) λ ist ein Eigenwert von $\mathbf{A} \in \mathbb{R}^{2 \times 2}$, und
- (ii) $\vec{\mathbf{u}}$ ist ein zu λ gehöriger Eigenvektor.

Um (H_2) zu lösen, bestimmt man also die Eigenwerte und Eigenvektoren von \mathbf{A} . Da das charakteristische Polynom $p_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda \mathbf{E}_2)$ eine reelle Polynomfunktion zweiten Grades ist, können die folgenden drei Fälle auftreten:

Fall 1: \mathbf{A} hat zwei verschiedene reelle Eigenwerte λ_1, λ_2 , oder

Fall 2: \mathbf{A} hat einen reellen Eigenwert λ (mit algebraischer Vielfachheit 2), oder

Fall 3: \mathbf{A} hat zwei zueinander konjugiert komplexe, nicht-reelle Eigenwerte $\lambda_1 = \alpha + j\beta$ und $\lambda_2 = \alpha - j\beta$.

Wir betrachten diese drei Fälle nun jeweils separat. Dabei ist vieles ähnlich zu dem Fall einer homogenen linearen Differentialgleichung zweiter Ordnung mit konstanten Koeffizienten (vgl. Teilkapitel 14.4).

Satz 15.19. (Fall 1: zwei verschiedene reelle Eigenwerte)

$\mathbf{A} \in \mathbb{R}^{2 \times 2}$ habe zwei verschiedene reelle Eigenwerte $\lambda_1, \lambda_2 \in \mathbb{R}$ mit $\lambda_1 \neq \lambda_2$. Seien $\vec{\mathbf{u}}_1, \vec{\mathbf{u}}_2 \in \mathbb{R}^2$ zwei zugehörige Eigenvektoren. Dann sind

$$\vec{\varphi}_1: \mathbb{R} \rightarrow \mathbb{R}^2, \quad \vec{\varphi}_1(t) := e^{\lambda_1 t} \vec{\mathbf{u}}_1, \quad \text{und} \quad \vec{\varphi}_2: \mathbb{R} \rightarrow \mathbb{R}^2, \quad \vec{\varphi}_2(t) := e^{\lambda_2 t} \vec{\mathbf{u}}_2,$$

zwei linear unabhängige reelle Lösungen von (H_2) . $(\vec{\varphi}_1, \vec{\varphi}_2)$ ist ein reelles Fundamentalsystem für (H_2) , und die reelle allgemeine Lösung von (H_2) ist

$$\vec{\mathbf{x}}: \mathbb{R} \rightarrow \mathbb{R}^2, \quad \vec{\mathbf{x}}(t) := c_1 \vec{\varphi}_1(t) + c_2 \vec{\varphi}_2(t) = c_1 e^{\lambda_1 t} \vec{\mathbf{u}}_1 + c_2 e^{\lambda_2 t} \vec{\mathbf{u}}_2, \quad c_1, c_2 \in \mathbb{R}.$$

Betrachten wir ein Beispiel.

Beispiel 15.20. (Fall 1: zwei verschiedene reelle Eigenwerte)

Wir betrachten

$$\begin{bmatrix} x_1' = 2x_1 + 2x_2 \\ x_2' = 2x_1 - x_2 \end{bmatrix} \iff \begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

$$\text{also} \quad \mathbf{A} = \begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix}. \quad (15.12)$$

Schritt 1: Bestimmung der Eigenwerte von \mathbf{A}

$$\begin{aligned} \det(\mathbf{A} - \lambda \mathbf{E}_2) &= \det \left(\begin{bmatrix} 2 - \lambda & 2 \\ 2 & -1 - \lambda \end{bmatrix} \right) = (2 - \lambda)(-1 - \lambda) - 4 \\ &= \lambda^2 - \lambda - 6 = (\lambda + 2)(\lambda - 3), \end{aligned}$$

d.h. die Eigenwerte sind $\lambda_1 = -2$ und $\lambda_2 = 3$.

Schritt 2: Bestimmung der Eigenräume von \mathbf{A}

$$\lambda_1 = -2 : \quad \begin{bmatrix} 4 & 2 & | & 0 \\ 2 & 1 & | & 0 \end{bmatrix} \xrightarrow[\begin{smallmatrix} Z_1 \rightarrow \frac{1}{2} Z_1 \\ \downarrow \\ Z_2 \rightarrow Z_2 - \frac{1}{2} Z_1 \end{smallmatrix}]{\iff} \begin{bmatrix} 2 & 1 & | & 0 \\ 0 & 0 & | & 0 \end{bmatrix} \implies 2u_1 = -u_2$$

$$\implies E_{\mathbf{A}}(-2) = \text{LH} \left(\begin{bmatrix} 1 \\ -2 \end{bmatrix} \right), \quad \text{d.h.} \quad \vec{\varphi}_1(t) = e^{-2t} \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$$\lambda_2 = 3 : \quad \begin{bmatrix} -1 & 2 & | & 0 \\ 2 & -4 & | & 0 \end{bmatrix} \xrightarrow[\begin{smallmatrix} Z_1 \rightarrow -Z_1 \\ \downarrow \\ Z_2 \rightarrow Z_2 + 2Z_1 \end{smallmatrix}]{\iff} \begin{bmatrix} 1 & -2 & | & 0 \\ 0 & 0 & | & 0 \end{bmatrix} \implies u_1 = 2u_2$$

$$\implies E_{\mathbf{A}}(3) = \text{LH} \left(\begin{bmatrix} 2 \\ 1 \end{bmatrix} \right), \quad \text{d.h.} \quad \vec{\varphi}_2(t) = e^{3t} \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Ergebnis: Ein Fundamentalsystem von (15.12) ist

$$\left(e^{-2t} \begin{bmatrix} 1 \\ -2 \end{bmatrix}, e^{3t} \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right),$$

und die zugehörige Fundamentalmatrix ist

$$\Phi(t) := \begin{bmatrix} e^{-2t} & 2e^{3t} \\ -2e^{-2t} & e^{3t} \end{bmatrix}.$$

Also ist die allgemeine Lösung von (15.12):

$$\begin{aligned} \vec{\mathbf{x}} : \mathbb{R} &\rightarrow \mathbb{R}^2, \quad \vec{\mathbf{x}}(t) := c_1 e^{-2t} \begin{bmatrix} 1 \\ -2 \end{bmatrix} + c_2 e^{3t} \begin{bmatrix} 2 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} e^{-2t} & 2e^{3t} \\ -2e^{-2t} & e^{3t} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \Phi(t) \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}, \quad c_1, c_2 \in \mathbb{R}. \end{aligned}$$

Zur Übung überprüfen wir für das Beispiel noch, dass es sich tatsächlich um ein Fundamentalsystem handelt (also dass die beiden Lösungen $\vec{\varphi}_1, \vec{\varphi}_2$ linear unabhängig sind), indem wir die Wronski-Determinante berechnen:

$$\begin{aligned} W(t) &= \det(\Phi(t)) = \det\left(\begin{bmatrix} e^{-2t} & 2e^{3t} \\ -2e^{-2t} & e^{3t} \end{bmatrix}\right) \\ &= e^{-2t} e^{3t} - 2e^{3t} (-2e^{-2t}) = e^t + 4e^t = 5e^t \neq 0 \quad \text{für alle } t \in \mathbb{R}. \end{aligned}$$

Bevor wir Fall 2 betrachten, wollen wir Satz 15.19 noch beweisen.

Beweis von Satz 15.19: Aus unseren Vorüberlegungen auf Seiten 190 bis 191 wissen wir bereits, dass

$$\vec{x}(t) := c_1 \vec{\varphi}_1(t) + c_2 \vec{\varphi}_2(t) = c_1 e^{\lambda_1 t} \vec{u}_1 + c_2 e^{\lambda_2 t} \vec{u}_2.$$

als Linearkombination der beiden Lösungen $\vec{\varphi}_1, \vec{\varphi}_2$ von (H_2) ebenfalls eine Lösung von (H_2) ist. Zu zeigen ist also nun noch, dass die beiden Lösungen linear unabhängig sind. Um dieses nachzuweisen, berechnen wir die Wronski-Determinante der beiden Lösungen. Dabei bezeichnen wir die Einträge von \vec{u}_1 bzw. \vec{u}_2 jeweils mit

$$\vec{u}_1 = \begin{bmatrix} u_{1,1} \\ u_{2,1} \end{bmatrix} \quad \text{bzw.} \quad \vec{u}_2 = \begin{bmatrix} u_{1,2} \\ u_{2,2} \end{bmatrix}.$$

Wir finden

$$\begin{aligned} W(t) &= \det([\vec{\varphi}_1(t) \quad \vec{\varphi}_2(t)]) = \det\left(\begin{bmatrix} e^{\lambda_1 t} u_{1,1} & e^{\lambda_2 t} u_{1,2} \\ e^{\lambda_1 t} u_{2,1} & e^{\lambda_2 t} u_{2,2} \end{bmatrix}\right) \\ &= e^{\lambda_1 t} u_{1,1} e^{\lambda_2 t} u_{2,2} - e^{\lambda_2 t} u_{1,2} e^{\lambda_1 t} u_{2,1} = e^{(\lambda_1 + \lambda_2)t} (u_{1,1} u_{2,2} - u_{1,2} u_{2,1}) \\ &= \underbrace{e^{(\lambda_1 + \lambda_2)t}}_{\neq 0} \det\left(\begin{bmatrix} u_{1,1} & u_{1,2} \\ u_{2,1} & u_{2,2} \end{bmatrix}\right) = \underbrace{e^{(\lambda_1 + \lambda_2)t}}_{\neq 0} \underbrace{\det([\vec{u}_1 \quad \vec{u}_2])}_{\neq 0} \neq 0 \end{aligned}$$

für alle $t \in \mathbb{R}$, wobei wir genutzt haben, dass nach Satz 13.43 die Eigenvektoren \vec{u}_1, \vec{u}_2 linear unabhängig sind, da sie zu verschiedenen Eigenwerten gehören. \square

Satz 15.21. (Fall 2: ein reeller Eigenwert)

$\mathbf{A} \in \mathbb{R}^{2 \times 2}$ habe **nur einen reellen Eigenwert** $\lambda \in \mathbb{R}$. Dieser Eigenwert hat dann die algebraische Vielfachheit 2. Nach (13.15) in Teilkapitel 13.6 kann die geometrische Vielfachheit von λ (d.h. die Dimension des Eigenraums $E_{\mathbf{A}}(\lambda)$) entweder 1 oder 2 sein.

(1) Ist $\dim(E_{\mathbf{A}}(\lambda)) = 2$, so gilt $A = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$.

In diesem Fall sind

$$\vec{\varphi}_1: \mathbb{R} \rightarrow \mathbb{R}^2, \quad \vec{\varphi}_1(t) := e^{\lambda t} \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \text{und} \quad \vec{\varphi}_2: \mathbb{R} \rightarrow \mathbb{R}^2, \quad \vec{\varphi}_2(t) := e^{\lambda t} \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

zwei linear unabhängige reelle Lösungen, $(\vec{\varphi}_1, \vec{\varphi}_2)$ ist ein **reelles Fundamentalsystem für (H_2)** , und die **reelle allgemeine Lösung von (H_2)** ist

$$\begin{aligned} \vec{x}: \mathbb{R} \rightarrow \mathbb{R}^2, \quad \vec{x}(t) &:= c_1 \vec{\varphi}_1(t) + c_2 \vec{\varphi}_2(t) \\ &= c_1 e^{\lambda t} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 e^{\lambda t} \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad c_1, c_2 \in \mathbb{R}. \end{aligned}$$

(2) Ist $\dim(E_{\mathbf{A}}(\lambda)) = 1$, so wählt man einen Eigenvektor $\vec{u} \in \mathbb{R}^2$ zum Eigenwert λ . Danach bestimmt man $\vec{v} \in \mathbb{R}^2$ als Lösung des linearen Gleichungssystems

$$(\mathbf{A} - \lambda \mathbf{E}_2) \vec{v} = \vec{u}.$$

Dann sind

$$\begin{aligned} \vec{\varphi}_1: \mathbb{R} \rightarrow \mathbb{R}^2, \quad \vec{\varphi}_1(t) &:= e^{\lambda t} \vec{u}, \quad \text{und} \\ \vec{\varphi}_2: \mathbb{R} \rightarrow \mathbb{R}^2, \quad \vec{\varphi}_2(t) &:= e^{\lambda t} (t \vec{u} + \vec{v}), \end{aligned}$$

zwei linear unabhängige reelle Lösungen von (H_2) . $(\vec{\varphi}_1, \vec{\varphi}_2)$ ist ein **reelles Fundamentalsystem für (H_2)** , und die **reelle allgemeine Lösung von (H_2)** ist

$$\begin{aligned} \vec{x}: \mathbb{R} \rightarrow \mathbb{R}^2, \quad \vec{x}(t) &:= c_1 \vec{\varphi}_1(t) + c_2 \vec{\varphi}_2(t) \\ &= c_1 e^{\lambda t} \vec{u} + c_2 e^{\lambda t} (t \vec{u} + \vec{v}), \quad c_1, c_2 \in \mathbb{R}. \end{aligned}$$

Beispiel 15.22. (Fall 2: ein reeller Eigenwert)

Wir betrachten

$$\begin{bmatrix} x'_1 = 4x_1 - 3x_2 \\ x'_2 = 3x_1 - 2x_2 \end{bmatrix} \iff \begin{bmatrix} x'_1 \\ x'_2 \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

$$\text{also } \mathbf{A} = \begin{bmatrix} 4 & -3 \\ 3 & -2 \end{bmatrix}. \quad (15.13)$$

Schritt 1: Bestimmung der Eigenwerte von A

$$\det(\mathbf{A} - \lambda \mathbf{E}_2) = \det \left(\begin{bmatrix} 4 - \lambda & -3 \\ 3 & -2 - \lambda \end{bmatrix} \right) = (4 - \lambda)(-2 - \lambda) + 9 \\ = \lambda^2 - 2\lambda + 1 = (\lambda - 1)^2,$$

d.h. $\lambda = 1$ ist der einzige reelle Eigenwert.

Schritt 2a: Bestimmung des Eigenraums zu $\lambda = 1$

$$\lambda = 1 : \quad \begin{bmatrix} 3 & -3 & | & 0 \\ 3 & -3 & | & 0 \end{bmatrix} \xrightarrow[\begin{smallmatrix} Z_1 \rightarrow \frac{1}{3} Z_1 \\ \leftarrow \end{smallmatrix}]{\begin{smallmatrix} Z_2 \rightarrow Z_2 - Z_1 \\ \leftarrow \end{smallmatrix}} \begin{bmatrix} 1 & -1 & | & 0 \\ 0 & 0 & | & 0 \end{bmatrix} \implies u_1 = u_2 \\ \implies E_{\mathbf{A}}(1) = \text{LH} \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \right), \quad \text{d.h.} \quad \vec{\varphi}_1 : \mathbb{R} \rightarrow \mathbb{R}^2, \quad \vec{\varphi}_1(t) := e^t \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Schritt 2b: Finden der zweiten von $\vec{\varphi}_1$ linear unabhängigen Lösung

Finde **eine** Lösung von $(\mathbf{A} - 1 \cdot \mathbf{E}_2) \vec{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

$$(\mathbf{A} - 1 \cdot \mathbf{E}_2) \vec{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \iff \begin{bmatrix} 3 & -3 & | & 1 \\ 3 & -3 & | & 1 \end{bmatrix} \xrightarrow[\begin{smallmatrix} Z_1 \rightarrow \frac{1}{3} Z_1 \\ \leftarrow \end{smallmatrix}]{\begin{smallmatrix} Z_2 \rightarrow Z_2 - Z_1 \\ \leftarrow \end{smallmatrix}} \begin{bmatrix} 1 & -1 & | & \frac{1}{3} \\ 0 & 0 & | & 0 \end{bmatrix},$$

und z.B. ist $\vec{v} = \begin{bmatrix} \frac{1}{3} \\ 0 \end{bmatrix}$ eine Lösung.

Also ist eine zweite linear unabhängige Lösung

$$\vec{\varphi}_2 : \mathbb{R} \rightarrow \mathbb{R}^2, \quad \vec{\varphi}_2(t) := e^t (t \vec{u} + \vec{v}) = e^t \left(t \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} \frac{1}{3} \\ 0 \end{bmatrix} \right) = e^t \begin{bmatrix} t + \frac{1}{3} \\ t \end{bmatrix}.$$

Ergebnis: Ein reelles Fundamentalsystem von (15.13) ist

$$\left(e^t \begin{bmatrix} 1 \\ 1 \end{bmatrix}, e^t \begin{bmatrix} t + \frac{1}{3} \\ t \end{bmatrix} \right),$$

und eine reelle Fundamentalmatrix ist

$$\Phi(t) := \begin{bmatrix} e^t & e^t (t + \frac{1}{3}) \\ e^t & e^t t \end{bmatrix}.$$

Also ist die reelle allgemeine Lösung von (15.13)

$$\begin{aligned}\vec{x} : \mathbb{R} &\rightarrow \mathbb{R}^2, & \vec{x}(t) &:= c_1 e^t \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 e^t \begin{bmatrix} t + \frac{1}{3} \\ t \end{bmatrix} \\ & & &= e^t \left(c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} t + \frac{1}{3} \\ t \end{bmatrix} \right), & c_1, c_2 &\in \mathbb{R}.\end{aligned}$$

Zur Übung berechnen wir noch die Wronski-Determinante, um uns zu überzeugen, dass $\vec{\varphi}_1, \vec{\varphi}_2$ in der Tat linear unabhängig sind.

$$\begin{aligned}W(t) &= \det(\Phi(t)) = e^t e^t t - e^t \left(t + \frac{1}{3}\right) e^t \\ &= t e^{2t} - \left(t + \frac{1}{3}\right) e^{2t} = -\frac{1}{3} e^{2t} \neq 0 \quad \text{für alle } t \in \mathbb{R}.\end{aligned}$$

Beweis von Satz 15.21:

- (1) Es gelte $\dim(E_{\mathbf{A}}(\lambda)) = 2$. Da der einzige zwei-dimensionale Untervektorraum von \mathbb{C}^2 der Vektorraum \mathbb{C}^2 selber ist, folgt $E_{\mathbf{A}}(\lambda) = \mathbb{C}^2$. Also sind alle Vektoren von \mathbb{C}^2 Eigenvektoren von \mathbf{A} , d.h. $\mathbf{A} \vec{x} = \lambda \vec{x}$ für alle $\vec{x} \in \mathbb{C}^2$. Die einzige Matrix in $\mathbb{C}^{2 \times 2}$ mit dieser Eigenschaft ist $\mathbf{A} = \lambda \mathbf{E}_2$. Insbesondere sind die Vektoren \vec{e}_1, \vec{e}_2 der Standardbasis Eigenvektoren von \mathbf{A} . Nach unseren Vorüberlegungen auf Seiten 190 bis 191 sind somit

$$\vec{\varphi}_1(t) = e^{\lambda t} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{und} \quad \vec{\varphi}_2(t) = e^{\lambda t} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

beides Lösungen von (H_2) . Wir berechnen die Wronski-Determinante, um zu zeigen, dass diese beiden Lösungen in der Tat linear unabhängig sind und somit ein Fundamentalsystem bilden:

$$W(t) = \det(\Phi(t)) = \det \left(\begin{bmatrix} e^{\lambda t} & 0 \\ 0 & e^{\lambda t} \end{bmatrix} \right) = e^{\lambda t} e^{\lambda t} - 0 = e^{2\lambda t} \neq 0$$

für alle $t \in \mathbb{R}$.

- (2) Falls $\dim E_{\mathbf{A}}(\lambda) = 1$, wählen wir einen zu λ gehörenden Eigenvektor $\vec{u} \in \mathbb{R}^2$. Nach unseren Vorüberlegungen auf Seiten 190 bis 191 ist dann $\vec{\varphi}_1(t) = e^{\lambda t} \vec{u}$ eine Lösung von (H_2) . Um eine zweite linear unabhängige Lösung $\vec{\varphi}_2$ zu finden zu finden, machen wir den Ansatz:

$$\vec{x}(t) = e^{\lambda t} (t \vec{u} + \vec{v}) \quad \text{mit} \quad \vec{v} \in \mathbb{R}^n,$$

$$\vec{x}'(t) = e^{\lambda t} (\lambda t \vec{u} + \lambda \vec{v} + \vec{u}).$$

Einsetzen in (H_2) liefert

$$\begin{aligned} e^{\lambda t} (\lambda t \vec{u} + \lambda \vec{v} + \vec{u}) &= \mathbf{A} [e^{\lambda t} (t \vec{u} + \vec{v})] \\ \iff e^{\lambda t} (\lambda t \vec{u} + \lambda \vec{v} + \vec{u}) &= e^{\lambda t} [t \mathbf{A} \vec{u} + \mathbf{A} \vec{v}] \quad \Big| : e^{\lambda t} \\ \iff t \lambda \vec{u} + \lambda \vec{v} + \vec{u} &= t \mathbf{A} \vec{u} + \mathbf{A} \vec{v}. \end{aligned}$$

Da \vec{u} Eigenvektor zum Eigenwert λ ist, gilt $\mathbf{A} \vec{u} = \lambda \vec{u}$ und somit

$$\begin{aligned} t \lambda \vec{u} + \lambda \vec{v} + \vec{u} = t \lambda \vec{u} + \mathbf{A} \vec{v} &\iff \lambda \vec{v} + \vec{u} = \mathbf{A} \vec{v} \\ &\iff \vec{u} = (\mathbf{A} - \lambda \mathbf{E}_2) \vec{v}. \end{aligned}$$

Ist also \vec{v} eine Lösung des linearen Gleichungssystems $(\mathbf{A} - \lambda \mathbf{E}_2) \vec{v} = \vec{u}$, so ist $\vec{\varphi}_2(t) = e^{\lambda t} (t \vec{u} + \vec{v})$ eine Lösung von (H_2) .

Wir müssen noch zeigen, dass die beiden Lösungen $\vec{\varphi}_1$ und $\vec{\varphi}_2$ von (H_2) linear unabhängig sind und somit ein Fundamentalsystem bilden. Dazu berechnen wir die Wronski-Determinante. Vorab überlegen wir uns, dass \vec{u} und \vec{v} linear unabhängig sind: Wir wissen, dass $\vec{u} \neq \vec{0}$ ist, und \vec{v} muss ebenfalls ungleich dem Nullvektor sein, denn ansonsten wäre $(\mathbf{A} - \lambda \mathbf{E}_2) \vec{v} = \vec{0} \neq \vec{u}$. \vec{u} und \vec{v} sind genau dann linear abhängig, wenn ein Vektor ein Vielfaches des anderen ist. Da beide Vektoren vom Nullvektor verschieden sind, folgt, dass diese genau dann linear abhängig sind, wenn es eine Konstante $\mu \in \mathbb{R}$ gibt mit $\vec{v} = \mu \vec{u}$. Dann gilt aber

$$(\mathbf{A} - \lambda \mathbf{E}_2) \vec{v} = (\mathbf{A} - \lambda \mathbf{E}_2) (\mu \vec{u}) = \mu \underbrace{(\mathbf{A} - \lambda \mathbf{E}_2) \vec{u}}_{= \vec{0} \text{ weil } \vec{u} \text{ Eigenvektor zu } \lambda} = \vec{0},$$

und diese Gleichung ist ein Widerspruch zu $(\mathbf{A} - \lambda \mathbf{E}_2) \vec{v} = \vec{u}$. Also sind \vec{u} und \vec{v} linear unabhängig.

Mit den Rechenregeln für Determinanten finden wir nun für die Wronski-Determinante:

$$\begin{aligned} W(t) &= \det([\vec{\varphi}_1(t) \quad \vec{\varphi}_2(t)]) = \det([e^{\lambda t} \vec{u} \quad e^{\lambda t} (t \vec{u} + \vec{v})]) \\ &= e^{2\lambda t} \det([\vec{u} \quad t \vec{u} + \vec{v}]) \\ &= e^{2\lambda t} \underbrace{\det([\vec{u} \quad t \vec{u}])}_{=0} + e^{2\lambda t} \det([\vec{u} \quad \vec{v}]) \\ &= e^{2\lambda t} \det([\vec{u} \quad \vec{v}]) \neq 0 \end{aligned}$$

für alle $t \in \mathbb{R}$, da die beiden Vektoren \vec{u} und \vec{v} linear unabhängig sind. \square

Satz 15.23. (Fall 3: zwei konjugiert komplexe Eigenwerte)

$\mathbf{A} \in \mathbb{R}^{2 \times 2}$ habe zwei zueinander konjugiert komplexe, nicht-reelle Eigenwerte $\lambda_1 = \lambda = \alpha + \beta j$ und $\lambda_2 = \bar{\lambda} = \alpha - \beta j$, wobei $\alpha, \beta \in \mathbb{R}$ mit $\beta \neq 0$.

Ist $\vec{\mathbf{w}} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \in \mathbb{C}^2$ ein Eigenvektor zu λ , so ist $\overline{\vec{\mathbf{w}}} = \begin{bmatrix} \overline{w_1} \\ \overline{w_2} \end{bmatrix}$ ein Eigenvektor

zu $\bar{\lambda}$. Dann sind

$$\begin{aligned} \vec{\psi}_1 : \mathbb{R} &\rightarrow \mathbb{C}^2, & \vec{\psi}_1(t) &:= e^{\lambda t} \vec{\mathbf{w}} = e^{\alpha t} e^{j\beta t} \vec{\mathbf{w}}, & \text{und} \\ \vec{\psi}_2 : \mathbb{R} &\rightarrow \mathbb{C}^2, & \vec{\psi}_2(t) &:= e^{\bar{\lambda} t} \overline{\vec{\mathbf{w}}} = e^{\alpha t} e^{-j\beta t} \overline{\vec{\mathbf{w}}} = \overline{\vec{\psi}_1(t)}, \end{aligned}$$

zwei linear unabhängige komplexe Lösungen von (H_2) . Zwei reelle linear unabhängige Lösungen von (H_2) erhält man, indem man den Realteil und den Imaginärteil dieser beiden Lösungen bildet, also

$$\begin{aligned} \vec{\varphi}_1 : \mathbb{R} &\rightarrow \mathbb{R}^2, & \vec{\varphi}_1(t) &:= \operatorname{Re}(\vec{\psi}_1(t)) = \operatorname{Re}(e^{\alpha t} e^{j\beta t} \vec{\mathbf{w}}), \\ \vec{\varphi}_2 : \mathbb{R} &\rightarrow \mathbb{R}^2, & \vec{\varphi}_2(t) &:= \operatorname{Im}(\vec{\psi}_1(t)) = \operatorname{Im}(e^{\alpha t} e^{j\beta t} \vec{\mathbf{w}}). \end{aligned}$$

Dann ist $(\vec{\varphi}_1, \vec{\varphi}_2)$ ein reelles Fundamentalsystem für (H_2) , und die reelle allgemeine Lösung ist

$$\begin{aligned} \vec{\mathbf{x}} : \mathbb{R} &\rightarrow \mathbb{R}^2, & \vec{\mathbf{x}}(t) &:= c_1 \vec{\varphi}_1(t) + c_2 \vec{\varphi}_2(t) \\ & & &= c_1 \operatorname{Re}(e^{\alpha t} e^{j\beta t} \vec{\mathbf{w}}) + c_2 \operatorname{Im}(e^{\alpha t} e^{j\beta t} \vec{\mathbf{w}}), & c_1, c_2 \in \mathbb{R}. \end{aligned}$$

Betrachten wir ein Beispiel.

Beispiel 15.24. (Fall 3: zwei konjugiert komplexe Eigenwerte)

Wir betrachten

$$\begin{bmatrix} x'_1 = x_1 - x_2 \\ x'_2 = x_1 + x_2 \end{bmatrix} \iff \begin{bmatrix} x'_1 \\ x'_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \text{also} \quad \mathbf{A} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}. \quad (15.14)$$

Schritt 1: Bestimmung der Eigenwerte von \mathbf{A}

$$\det(\mathbf{A} - \lambda \mathbf{E}_2) = \det \left(\begin{bmatrix} 1 - \lambda & -1 \\ 1 & 1 - \lambda \end{bmatrix} \right) = (1 - \lambda)^2 + 1$$

$$= (1 - \lambda)^2 - j^2 = (1 - \lambda - j)(1 - \lambda + j),$$

d.h. wir haben zwei zueinander konjugiert komplexe Eigenwerte $\lambda_1 = 1 + j$ und $\lambda_2 = 1 - j = \overline{\lambda_1}$.

Schritt 2a: Bestimmung des Eigenraums zu $\lambda_1 = 1 + j$

$$\left[\begin{array}{cc|c} -j & -1 & 0 \\ 1 & -j & 0 \end{array} \right] \xrightarrow[Z_2 \rightarrow Z_2 - j Z_1]{\downarrow} \left[\begin{array}{cc|c} -j & -1 & 0 \\ 0 & 0 & 0 \end{array} \right] \implies w_2 = -j w_1,$$

d.h. es gilt

$$E_{\mathbf{A}}(1 + j) = \left\{ \begin{bmatrix} \mu \\ -j \mu \end{bmatrix} : \mu \in \mathbb{C} \right\} \stackrel{\mu=j}{=} \text{LH} \left(\begin{bmatrix} j \\ 1 \end{bmatrix} \right).$$

Damit erhalten wir

$$\vec{\psi}_1 : \mathbb{R} \rightarrow \mathbb{C}^2, \quad \vec{\psi}_1(t) := e^{(1+j)t} \begin{bmatrix} j \\ 1 \end{bmatrix} = e^t e^{jt} \begin{bmatrix} j \\ 1 \end{bmatrix}.$$

(Weiter erhalten wir nach dem Satz 15.23

$$E_{\mathbf{A}}(1 - j) = \text{LH} \left(\begin{bmatrix} \overline{j} \\ 1 \end{bmatrix} \right) = \text{LH} \left(\begin{bmatrix} -j \\ 1 \end{bmatrix} \right) \quad \text{und}$$

$$\vec{\psi}_2 : \mathbb{R} \rightarrow \mathbb{C}^2, \quad \vec{\psi}_2(t) := \overline{\vec{\psi}_1(t)} = e^t e^{jt} \begin{bmatrix} j \\ 1 \end{bmatrix} = e^t e^{-jt} \begin{bmatrix} -j \\ 1 \end{bmatrix}.$$

Diese Information brauchen wir allerdings nicht zu der weiteren Lösung des Problems.)

Schritt 2b: Bestimmung zweier reeller linear unabhängiger Lösungen

Wir berechnen den Realteil und Imaginärteil von $\vec{\psi}_1(t)$: Mit der Euler-Formel gilt

$$\begin{aligned} \vec{\psi}_1(t) &= e^t e^{jt} \begin{bmatrix} j \\ 1 \end{bmatrix} = e^t (\cos(t) + j \sin(t)) \begin{bmatrix} j \\ 1 \end{bmatrix} \\ &= e^t \begin{bmatrix} -\sin(t) + j \cos(t) \\ \cos(t) + j \sin(t) \end{bmatrix} = \underbrace{e^t \begin{bmatrix} -\sin(t) \\ \cos(t) \end{bmatrix}}_{= \text{Re}(\vec{\psi}_1(t)) = \vec{\varphi}_1(t)} + j \underbrace{e^t \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}}_{= \text{Im}(\vec{\psi}_1(t)) = \vec{\varphi}_2(t)}. \end{aligned}$$

Zwei reelle linear unabhängige Lösungen von (15.14) sind also

$$\vec{\varphi}_1 : \mathbb{R} \rightarrow \mathbb{R}^2, \quad \vec{\varphi}_1(t) := e^t \begin{bmatrix} -\sin(t) \\ \cos(t) \end{bmatrix}, \quad \vec{\varphi}_2 : \mathbb{R} \rightarrow \mathbb{R}^2, \quad \vec{\varphi}_2(t) := e^t \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}.$$

Ergebnis: Ein reelles Fundamentalsystem von (15.14) ist somit

$$\left(e^t \begin{bmatrix} -\sin(t) \\ \cos(t) \end{bmatrix}, e^t \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix} \right),$$

und die zugehörige reelle Fundamentalmatrix ist

$$\Phi(t) := \begin{bmatrix} -e^t \sin(t) & e^t \cos(t) \\ e^t \cos(t) & e^t \sin(t) \end{bmatrix}.$$

Die reelle allgemeine Lösung von (15.14) ist also

$$\begin{aligned} \vec{x} : \mathbb{R} \rightarrow \mathbb{R}^2, \quad \vec{x}(t) &:= c_1 e^t \begin{bmatrix} -\sin(t) \\ \cos(t) \end{bmatrix} + c_2 e^t \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix} \\ &= e^t \left(c_1 \begin{bmatrix} -\sin(t) \\ \cos(t) \end{bmatrix} + c_2 \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix} \right), \quad c_1, c_2 \in \mathbb{R}. \end{aligned}$$

Zur Übung berechnen wir noch die Wronski-Determinante, um uns zu überzeugen, dass die beiden reellen Lösungen $\vec{\varphi}_1, \vec{\varphi}_2$ wirklich linear unabhängig sind:

$$\begin{aligned} W(t) &= \det(\Phi(t)) = \det \left(\begin{bmatrix} -e^t \sin(t) & e^t \cos(t) \\ e^t \cos(t) & e^t \sin(t) \end{bmatrix} \right) \\ &= -e^{2t} \sin^2(t) - e^{2t} \cos^2(t) \\ &= -e^{2t} [\sin^2(t) + \cos^2(t)] = -e^{2t} \neq 0 \quad \text{für alle } t \in \mathbb{R}. \end{aligned}$$

Beweis von Satz 15.23: Wir zeigen zunächst, dass der Vektor $\overline{\vec{w}}$, wie behauptet, ein Eigenvektor zu $\lambda_2 = \overline{\lambda_1}$ ist: Durch Bilden des konjugiert Komplexen auf beiden Seiten von $\mathbf{A} \vec{w} = \lambda_1 \vec{w}$ folgt, dass

$$\overline{\mathbf{A} \vec{w}} = \overline{\lambda_1 \vec{w}} \quad \iff \quad \overline{\mathbf{A} \vec{w}} = \overline{\lambda_1} \overline{\vec{w}} \quad \iff \quad \mathbf{A} \overline{\vec{w}} = \lambda_2 \overline{\vec{w}},$$

weil \mathbf{A} eine reelle Matrix ist und weil $\lambda_2 = \overline{\lambda_1}$ ist. Also ist $\overline{\vec{w}}$ in der Tat ein Eigenvektor zu $\lambda_2 = \overline{\lambda_1}$.

Mit den Berechnungen aus dem Beweis von Satz 15.19 folgt direkt, dass $\vec{\psi}_1, \vec{\psi}_2$ zwei linear unabhängige Lösungen von (H_2) sind. (Eine Inspektion des Arguments in Satz 15.19 zeigt nämlich, dass die Argumentation auch für komplexes $\lambda_1 \neq \lambda_2$ funktioniert.) Wir müssen nur noch zeigen, dass $\vec{\varphi}_1, \vec{\varphi}_2$ ebenfalls zwei linear unabhängige Lösungen von (H_2) sind. Wir zeigen zunächst, dass $\overline{\vec{\psi}_1(t)} = \vec{\psi}_2(t)$ ist. In der Tat gilt

$$\overline{\vec{\psi}_1(t)} = \overline{e^{\alpha t} e^{j\beta t} \vec{w}} = e^{\alpha t} \overline{e^{j\beta t} \vec{w}} = e^{\alpha t} e^{-j\beta t} \vec{w} = \vec{\psi}_2(t).$$

Wegen $\operatorname{Re}(z) = \frac{1}{2}(z + \bar{z})$ und $\operatorname{Im}(z) = \frac{1}{2j}(z - \bar{z})$ für alle $z \in \mathbb{C}$ folgt

$$\begin{aligned}\vec{\varphi}_1(t) &= \operatorname{Re}(\vec{\psi}_1(t)) = \frac{1}{2} \left(\vec{\psi}_1(t) + \overline{\vec{\psi}_1(t)} \right) \\ &= \frac{1}{2} \left(\vec{\psi}_1(t) + \vec{\psi}_2(t) \right) = \frac{1}{2} \vec{\psi}_1(t) + \frac{1}{2} \vec{\psi}_2(t), \\ \vec{\varphi}_2(t) &= \operatorname{Im}(\vec{\psi}_1(t)) = \frac{1}{2j} \left(\vec{\psi}_1(t) - \overline{\vec{\psi}_1(t)} \right) \\ &= \frac{1}{2j} \left(\vec{\psi}_1(t) - \vec{\psi}_2(t) \right) = \frac{1}{2j} \vec{\psi}_1(t) - \frac{1}{2j} \vec{\psi}_2(t),\end{aligned}$$

d.h. $\vec{\varphi}_1, \vec{\varphi}_2$ sind als Linearkombinationen von $\vec{\psi}_1, \vec{\psi}_2$ Lösungen von (H_2) . Aus

$$\begin{aligned}\vec{0} &= c_1 \vec{\varphi}_1 + c_2 \vec{\varphi}_2 = c_1 \left[\frac{1}{2} \vec{\psi}_1(t) + \frac{1}{2} \vec{\psi}_2(t) \right] + c_2 \left[\frac{1}{2j} \vec{\psi}_1(t) - \frac{1}{2j} \vec{\psi}_2(t) \right] \\ &= \left[\frac{1}{2} c_1 + \frac{1}{2j} c_2 \right] \vec{\psi}_1(t) + \left[\frac{1}{2} c_1 - \frac{1}{2j} c_2 \right] \vec{\psi}_2(t) \quad \text{für alle } t \in \mathbb{R}\end{aligned}$$

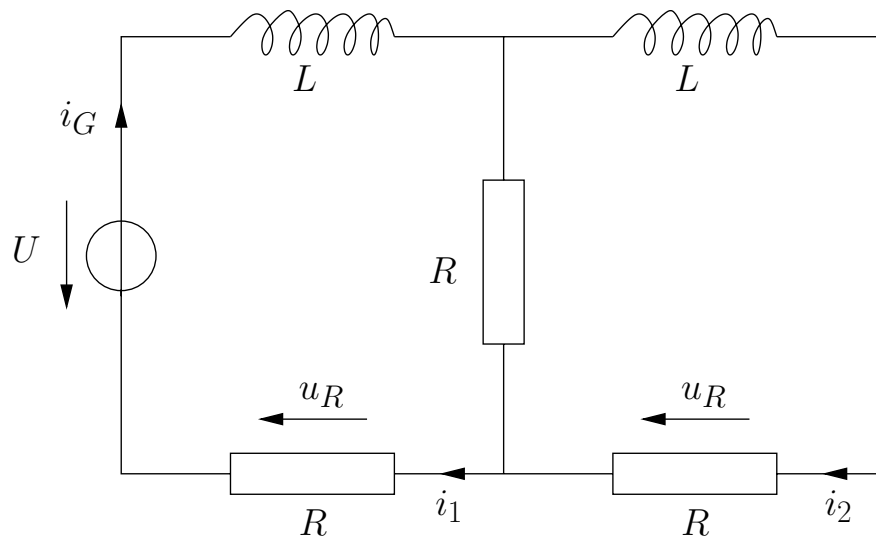
folgt wegen der linearen Unabhängigkeit von $\vec{\psi}_1, \vec{\psi}_2$, dass

$$\left[\begin{array}{l} \frac{1}{2} c_1 + \frac{1}{2j} c_2 = 0 \quad \text{(I)} \\ \frac{1}{2} c_1 - \frac{1}{2j} c_2 = 0 \quad \text{(II)} \end{array} \right] \implies \begin{cases} \text{aus (I) + (II):} & c_1 = 0, \\ \text{einsetzen in (I):} & \frac{1}{2j} c_2 \implies c_2 = 0 \end{cases}$$

gelten muss, d.h. $\vec{\varphi}_1, \vec{\varphi}_2$ sind linear unabhängig. □

15.4 Ein Anwendungsbeispiel

In diesem Teilkapitel betrachten wir als Anwendung aus der Elektrotechnik einen Schaltkreis, der in der nachfolgenden Zeichnung skizziert ist.



Weiter soll die folgende Anfangsbedingung

$$i_1(0) = i_2(0) = 0$$

an die beiden Ströme i_1 und i_2 gelten.

Das Aufstellen der Gleichungen liefert (für die beiden Maschen):

$$\begin{bmatrix} L \frac{di_1}{dt} + R(i_1 - i_2) + R i_1 = U \\ L \frac{di_2}{dt} + R(i_2 - i_1) + R i_2 = 0 \end{bmatrix} \iff \begin{bmatrix} i_1' = -2 \frac{R}{L} i_1 + \frac{R}{L} i_2 + \frac{U}{L} \\ i_2' = \frac{R}{L} i_1 - 2 \frac{R}{L} i_2 \end{bmatrix}$$

Mit $k := R/L$ ergibt dieses das folgende inhomogene lineare Differentialgleichungssystem erster Ordnung mit konstanten Koeffizienten

$$\begin{bmatrix} i_1' \\ i_2' \end{bmatrix} = \underbrace{\begin{bmatrix} -2k & k \\ k & -2k \end{bmatrix}}_{=\mathbf{A}} \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} + \begin{bmatrix} \frac{U}{L} \\ 0 \end{bmatrix}. \quad (15.15)$$

Schritt 1: Bestimmen der Eigenwerte von \mathbf{A}

$$\begin{aligned} \det \left(\begin{bmatrix} -2k - \lambda & k \\ k & -2k - \lambda \end{bmatrix} \right) &= (-2k - \lambda)^2 - k^2 = (\lambda + 2k)^2 - k^2 \\ &= (\lambda + 2k - k)(\lambda + 2k + k) = (\lambda + k)(\lambda + 3k) \end{aligned}$$

Also sind die Eigenwerte von \mathbf{A} :

$$\lambda_1 = -k = -\frac{R}{L}, \quad \lambda_2 = -3k = -3\frac{R}{L}.$$

Schritt 2: Bestimmen der Eigenräume von λ_1 bzw. λ_2

$$\lambda_1 = -k : \quad \left[\begin{array}{cc|c} -k & k & 0 \\ k & -k & 0 \end{array} \right] \xrightarrow[\downarrow]{\begin{array}{l} Z_2 \rightarrow Z_2 + Z_1 \\ Z_1 \rightarrow -\frac{1}{k} Z_1 \end{array}} \left[\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array} \right]$$

$$\implies i_1 = i_2 \implies E_{\mathbf{A}}(-k) = \text{LH} \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)$$

$$\lambda_2 = -3k : \quad \left[\begin{array}{cc|c} k & k & 0 \\ k & k & 0 \end{array} \right] \xrightarrow[\downarrow]{\begin{array}{l} Z_2 \rightarrow Z_2 - Z_1 \\ Z_1 \rightarrow \frac{1}{k} Z_1 \end{array}} \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array} \right]$$

$$\implies i_1 = -i_2 \implies E_{\mathbf{A}}(-3k) = \text{LH} \left(\begin{bmatrix} -1 \\ 1 \end{bmatrix} \right)$$

Zwischenergebnis: Die allgemeine Lösung des zugehörigen homogenen Systems linearer Differentialgleichungen

$$\begin{bmatrix} i_1' \\ i_2' \end{bmatrix} = \begin{bmatrix} -2k & k \\ k & -2k \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \end{bmatrix}$$

ist also

$$\vec{\mathbf{i}}_H : \mathbb{R} \rightarrow \mathbb{R}^2, \quad \vec{\mathbf{i}}_H(t) := c_1 e^{-kt} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 e^{-3kt} \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad c_1, c_2 \in \mathbb{R}.$$

Ein Fundamentalsystem ist also

$$\left(e^{-kt} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, e^{-3kt} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right),$$

und die zugehörige Fundamentalmatrix ist

$$\Phi(t) := \begin{bmatrix} e^{-kt} & -e^{-3kt} \\ e^{-kt} & e^{-3kt} \end{bmatrix}.$$

Schritt 3: Bestimmen einer speziellen Lösung von (15.15)

Da die Inhomogenität $\begin{bmatrix} U \\ L \\ 0 \end{bmatrix}$ nicht von t abhängt, wählen wir als Ansatz für eine

spezielle Lösung $\vec{\mathbf{i}}_S$ eine konstante Funktion:

$$\vec{\mathbf{i}}_S(t) = \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \implies \vec{\mathbf{i}}_S'(t) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Einsetzen in (15.15) mit $k = R/L$ liefert

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -2k & k \\ k & -2k \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} + \begin{bmatrix} U/L \\ 0 \end{bmatrix} \stackrel{k=R/L}{\iff} \begin{bmatrix} -2\frac{R}{L} & \frac{R}{L} \\ \frac{R}{L} & -2\frac{R}{L} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} -\frac{U}{L} \\ 0 \end{bmatrix},$$

und wir berechnen wie gewohnt die Lösung durch elementare Zeilenumformungen:

$$\begin{array}{c} \begin{bmatrix} -2\frac{R}{L} & \frac{R}{L} & | & -\frac{U}{L} \\ \frac{R}{L} & -2\frac{R}{L} & | & 0 \end{bmatrix} \xrightarrow{\substack{Z_1 \rightarrow \frac{1}{R} Z_1 \\ Z_2 \rightarrow \frac{1}{R} Z_2}} \begin{bmatrix} -2 & 1 & | & -\frac{U}{R} \\ 1 & -2 & | & 0 \end{bmatrix} \xrightarrow{Z_1 \leftrightarrow Z_2} \begin{bmatrix} 1 & -2 & | & 0 \\ -2 & 1 & | & -\frac{U}{R} \end{bmatrix} \\ \xrightarrow{Z_2 \rightarrow Z_2 + 2Z_1} \begin{bmatrix} 1 & -2 & | & 0 \\ 0 & -3 & | & -\frac{U}{R} \end{bmatrix} \xrightarrow{Z_2 \rightarrow -\frac{1}{3} Z_2} \begin{bmatrix} 1 & -2 & | & 0 \\ 0 & 1 & | & \frac{U}{3R} \end{bmatrix} \xrightarrow{Z_1 \rightarrow Z_1 + 2Z_2} \begin{bmatrix} 1 & 0 & | & \frac{2U}{3R} \\ 0 & 1 & | & \frac{U}{3R} \end{bmatrix}, \end{array}$$

d.h. $I_1 = \frac{2U}{3R}$ und $I_2 = \frac{U}{3R}$. Also ist

$$\vec{\mathbf{i}}_S : \mathbb{R} \rightarrow \mathbb{R}^2, \quad \vec{\mathbf{i}}_S(t) := \frac{U}{3R} \begin{bmatrix} 2 \\ 1 \end{bmatrix},$$

eine spezielle Lösung von (15.15).

Alternative in Schritt 3: Bestimmen einer speziellen Lösung von (15.15)

Wir nutzen die Methode Variation der Konstanten (vgl. Satz 15.17) zur Bestimmung der allgemeinen Lösung der inhomogenen Gleichung. Wir bestimmen zuerst die Inverse $(\Phi(t))^{-1}$ der Fundamentalmatrix $\Phi(t)$:

$$\begin{aligned} W(t) = \det(\Phi(t)) &= \det \left(\begin{bmatrix} e^{-kt} & -e^{-3kt} \\ e^{-kt} & e^{-3kt} \end{bmatrix} \right) \\ &= e^{-kt} e^{-3kt} + e^{-kt} e^{-3kt} = 2e^{-4kt} \neq 0 \quad \text{für alle } t \in \mathbb{R}. \end{aligned}$$

Somit gilt

$$(\Phi(t))^{-1} = \frac{1}{2e^{-4kt}} \begin{bmatrix} e^{-3kt} & e^{-3kt} \\ -e^{-kt} & e^{-kt} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} e^{kt} & e^{kt} \\ -e^{3kt} & e^{3kt} \end{bmatrix}.$$

Nach Satz 15.17 ist die allgemeine Lösung von (15.15) durch

$$\vec{\mathbf{i}}(t) := \Phi(t) \int (\Phi(t))^{-1} \begin{bmatrix} \frac{U}{L} \\ 0 \end{bmatrix} dt$$

gegeben. Wir berechnen vorab den Integranden

$$(\Phi(t))^{-1} \begin{bmatrix} \frac{U}{L} \\ 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} e^{kt} & e^{kt} \\ -e^{3kt} & e^{3kt} \end{bmatrix} \begin{bmatrix} \frac{U}{L} \\ 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} e^{kt} \frac{U}{L} \\ -e^{3kt} \frac{U}{L} \end{bmatrix} = \frac{1}{2} \frac{U}{L} \begin{bmatrix} e^{kt} \\ -e^{3kt} \end{bmatrix}.$$

Damit erhalten wir

$$\begin{aligned} \vec{\mathbf{i}}(t) &= \begin{bmatrix} e^{-kt} & -e^{-3kt} \\ e^{-kt} & e^{-3kt} \end{bmatrix} \int \frac{1}{2} \frac{U}{L} \begin{bmatrix} e^{kt} \\ -e^{3kt} \end{bmatrix} dt \\ &= \frac{1}{2} \frac{U}{L} \begin{bmatrix} e^{-kt} & -e^{-3kt} \\ e^{-kt} & e^{-3kt} \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{k} e^{kt} + c_1 \\ -\frac{1}{3k} e^{3kt} + c_2 \end{bmatrix} \\ &= \frac{1}{2} \frac{U}{L} \begin{bmatrix} e^{-kt} & -e^{-3kt} \\ e^{-kt} & e^{-3kt} \end{bmatrix} \cdot \left(\begin{bmatrix} \frac{1}{k} e^{kt} \\ -\frac{1}{3k} e^{3kt} \end{bmatrix} + \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \right) \\ &= \frac{1}{2} \frac{U}{L} \begin{bmatrix} \frac{1}{k} + \frac{1}{3k} \\ \frac{1}{k} - \frac{1}{3k} \end{bmatrix} + \frac{1}{2} \frac{U}{L} \begin{bmatrix} e^{-kt} & -e^{-3kt} \\ e^{-kt} & e^{-3kt} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \\ &= \frac{1}{2} \frac{U}{L} \begin{bmatrix} \frac{4}{3k} \\ \frac{2}{3k} \end{bmatrix} + \frac{1}{2} \frac{U}{L} \begin{bmatrix} e^{-kt} & -e^{-3kt} \\ e^{-kt} & e^{-3kt} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \\ &= \frac{1}{3k} \frac{U}{L} \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \underbrace{\frac{1}{2} \frac{U}{L} c_1}_{=\tilde{c}_1} e^{-kt} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \underbrace{\frac{1}{2} \frac{U}{L} c_2}_{=\tilde{c}_2} e^{-3kt} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ &\stackrel{k=R/L}{=} \frac{U}{3R} \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \tilde{c}_1 e^{-\frac{R}{L}t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \tilde{c}_2 e^{-\frac{3R}{L}t} \begin{bmatrix} -1 \\ 1 \end{bmatrix}. \end{aligned}$$

Vorläufiges Ergebnis: Die allgemeine Lösung von (15.15) ist nun durch

$$\vec{\mathbf{i}} : \mathbb{R} \rightarrow \mathbb{R}, \quad \vec{\mathbf{i}}(t) = \begin{bmatrix} i_1(t) \\ i_2(t) \end{bmatrix} := \frac{U}{3R} \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_1 e^{-\frac{R}{L}t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 e^{-\frac{3R}{L}t} \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad c_1, c_2 \in \mathbb{R}, \quad (15.16)$$

gegeben.

Schritt 4: Einarbeiten der Anfangsbedingung $i_1(0) = i_2(0) = 0$

Wir setzen $t = 0$ in (15.16) ein und fordern, dass die Anfangsbedingung $i_1(0) = i_2(0) = 0$ erfüllt ist:

$$\begin{aligned} \begin{bmatrix} 0 \\ 0 \end{bmatrix} &= \begin{bmatrix} i_1(0) \\ i_2(0) \end{bmatrix} = \frac{U}{3R} \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ \Leftrightarrow & -\frac{U}{3R} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ \Leftrightarrow & -\frac{U}{3R} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \\ \Leftrightarrow & \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = -\frac{U}{3R} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = -\frac{U}{3R} \cdot \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \frac{U}{6R} \begin{bmatrix} -3 \\ 1 \end{bmatrix}. \end{aligned}$$

Ergebnis: Die Lösung des Anfangswertproblems (15.15) mit $i_1(0) = i_2(0) = 0$ ist also:

$$\begin{aligned} \vec{i} : \mathbb{R} \rightarrow \mathbb{R}^2, \quad \vec{i}(t) &= \begin{bmatrix} i_1(t) \\ i_2(t) \end{bmatrix} := \frac{U}{3R} \begin{bmatrix} 2 \\ 1 \end{bmatrix} - \frac{3U}{6R} e^{-\frac{R}{L}t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{U}{6R} e^{-\frac{3R}{L}t} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ \Leftrightarrow & \begin{bmatrix} i_1 : \mathbb{R} \rightarrow \mathbb{R}, \quad i_1(t) := \frac{U}{6R} \left(4 - 3e^{-\frac{R}{L}t} - e^{-\frac{3R}{L}t} \right) \\ i_2 : \mathbb{R} \rightarrow \mathbb{R}, \quad i_2(t) := \frac{U}{6R} \left(2 - 3e^{-\frac{R}{L}t} + e^{-\frac{3R}{L}t} \right) \end{bmatrix}. \end{aligned}$$

15.5 Homogene Systeme mit konstanten Koeffizienten: Der allgemeine Fall

Sei $n \in \mathbb{N}$ und $\mathbf{A} \in \mathbb{R}^{n \times n}$. In diesem Teilkapitel lernen wir, wie man für beliebiges n vorgeht, um die **reelle allgemeine Lösung des homogenen Systems linearer Differentialgleichungen erster Ordnung**

$$\vec{x}' = \mathbf{A} \vec{x} \quad (\text{H}_n)$$

zu finden.

Satz 15.25. (allgemeine reelle Lösung von (H_n))

Sei λ ein Eigenwert der Matrix \mathbf{A} in (H_n) mit dem Eigenraum $E_{\mathbf{A}}(\lambda)$, der algebraischen Vielfachheit m und der geometrischen Vielfachheit $\ell = \dim(E_{\mathbf{A}}(\lambda))$. Es gilt dann

$$1 \leq \ell \leq m \leq n.$$

- (1) Analog zu Teilkapitel 15.3 ist $\vec{x}(t) = e^{\lambda t} \vec{u}$ für jedes $\vec{u} \in E_{\mathbf{A}}(\lambda)$ eine (komplexe) Lösung von (H_n) . Auf diese Weise erhalten wir ℓ linear unabhängige Lösungen von (H_n) .
- (2) Ist $\ell < m$, so kann man mit Hilfe von Hauptvektoren $m - \ell$ weitere (komplexe) linear unabhängige Lösungen von (H_n) berechnen (siehe Satz 15.27 weiter unten für die Bestimmung der Hauptvektoren). Zusammen mit den Lösungen aus (1) hat man dann genau m linear unabhängige Lösungen.
- (3) Ist $\lambda \notin \mathbb{R}$, so ist auch $\bar{\lambda}$ ein Eigenwert von \mathbf{A} mit der gleichen algebraischen und geometrischen Vielfachheit wie λ . Wie in Satz 15.23 kann man komplexe Lösungen in reelle Lösungen umwandeln.

Betrachtet man auf diese Weise alle Eigenwerte von \mathbf{A} , so erhält man **insgesamt n linear unabhängige reelle Lösungen von (H_n)** , also ein reelles Fundamentalsystem von (H_n) .

Diesen Satz werden wir nicht beweisen. Wir betrachten zunächst ein Beispiel. Danach lernen wir die bereits erwähnten Hauptvektoren kennen.

Beispiel 15.26. (allgemeine reelle Lösung von (H_n))

Wir betrachten das folgende homogene System linearer Differentialgleichungen erster Ordnung mit konstanten Koeffizienten:

$$\begin{bmatrix} x_1' = & & -x_2 + & x_3 \\ x_2' = -3x_1 - & 2x_2 + & 3x_3 \\ x_3' = -2x_1 - & 2x_2 + & 3x_3 \end{bmatrix} \iff \begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix} = \begin{bmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (15.17)$$

Es gilt also

$$\mathbf{A} = \begin{bmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{bmatrix}.$$

Schritt 1: Berechnung der Eigenwerte von A

Aus Beispiel 13.40 (c) und Beispiel 13.42 (c) wissen wir, dass \mathbf{A} die Eigenwerte

$$\lambda_1 = 1 \quad \text{mit der algebraischen und geometrischen Vielfachheit } 2,$$

$$\lambda_2 = -1 \quad \text{mit der algebraischen und geometrischen Vielfachheit } 1$$

hat.

Schritt 2: Bestimmung der zugehörigen Eigenräume

Aus Beispiel 13.40 (c) wissen wir, dass \mathbf{A} die zugehörigen Eigenräume

$$E_{\mathbf{A}}(1) = \text{LH} \left(\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right) \quad \text{bzw.} \quad E_{\mathbf{A}}(-1) = \text{LH} \left(\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \right)$$

sind. $E_{\mathbf{A}}(1)$ liefert nach Satz 15.25 die folgenden zwei linear unabhängigen Lösungen

$$\vec{\varphi}_1 : \mathbb{R} \rightarrow \mathbb{R}^3, \quad \vec{\varphi}_1(t) := e^t \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad \text{und} \quad \vec{\varphi}_2 : \mathbb{R} \rightarrow \mathbb{R}^3, \quad \vec{\varphi}_2(t) := e^t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},$$

von (15.17), und $E_{\mathbf{A}}(-1)$ liefert nach Satz 15.25 die folgende Lösung

$$\vec{\varphi}_3 : \mathbb{R} \rightarrow \mathbb{R}^3, \quad \vec{\varphi}_3(t) := e^{-t} \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix},$$

von (15.17). Nach Satz 15.25 sind die so gefunden drei Lösungen linear unabhängig.

Ergebnis:

$$\left(e^t \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, e^t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, e^{-t} \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \right)$$

ist ein Fundamentalsystem von (15.17), und die reelle allgemeine Lösung von (15.17) ist durch

$$\vec{\mathbf{x}} : \mathbb{R} \rightarrow \mathbb{R}^3, \quad \vec{\mathbf{x}}(t) := e^t \left(c_1 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right) + c_3 e^{-t} \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \quad c_1, c_2, c_3 \in \mathbb{R},$$

$$\iff \left[\begin{array}{l} x_1 : \mathbb{R} \rightarrow \mathbb{R}, \quad x_1(t) := (-c_1 + c_2) e^t + c_3 e^{-t} \\ x_2 : \mathbb{R} \rightarrow \mathbb{R}, \quad x_2(t) := c_1 e^t + 3 c_3 e^{-t} \\ x_3 : \mathbb{R} \rightarrow \mathbb{R}, \quad x_3(t) := c_2 e^t + 2 c_3 e^{-t} \end{array} \right] \quad \text{mit } c_1, c_2, c_3 \in \mathbb{R}$$

gegeben.

Im nächsten Satz lernen wir die bereits erwähnten Hauptvektoren kennen.

Satz 15.27. (Hauptvektoren)

Sei λ ein Eigenwert der Matrix \mathbf{A} in (\mathbb{H}_n)

(1) $\vec{\mathbf{u}} \in \mathbb{C}^n$ heißt ein **Hauptvektor von \mathbf{A} zu λ** , falls $k \in \mathbb{N}$ existiert mit

$$(\mathbf{A} - \lambda \mathbf{E}_n)^k \vec{\mathbf{u}} = \vec{\mathbf{0}}.$$

(2) Jeder Eigenvektor ist Hauptvektor, aber nicht umgekehrt.

(3) Ist m die algebraische Vielfachheit von λ , so gibt es m **linear unabhängige Hauptvektoren von \mathbf{A} zu λ** .

(4) **Spezialfall:** $m \geq 2$ und die geometrische Vielfachheit ist $\ell = 1$:
Hier kann man Hauptvektoren berechnen, indem man nacheinander die folgenden linearen Gleichungssysteme löst:

$$(\mathbf{A} - \lambda \mathbf{E}_n) \vec{\mathbf{u}} = \vec{\mathbf{0}} \quad \implies \quad \text{Die Lösung sei } \vec{\mathbf{u}}_1 \text{ (Eigenvektor).}$$

$$(\mathbf{A} - \lambda \mathbf{E}_n) \vec{\mathbf{u}} = \vec{\mathbf{u}}_1 \quad \implies \quad \text{Die Lösung sei } \vec{\mathbf{u}}_2.$$

$$(\mathbf{A} - \lambda \mathbf{E}_n) \vec{\mathbf{u}} = \vec{\mathbf{u}}_2 \quad \implies \quad \text{Die Lösung sei } \vec{\mathbf{u}}_3.$$

$$\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots$$

$$(\mathbf{A} - \lambda \mathbf{E}_n) \vec{\mathbf{u}} = \vec{\mathbf{u}}_{m-1} \quad \implies \quad \text{Die Lösung sei } \vec{\mathbf{u}}_m.$$

Die Vektoren $\vec{\mathbf{u}}_1, \vec{\mathbf{u}}_2, \dots, \vec{\mathbf{u}}_m$ sind dann **linear unabhängige Hauptvektoren von \mathbf{A} zu λ** , und

$$\vec{\varphi}_1(t) := e^{\lambda t} \vec{\mathbf{u}}_1,$$

$$\vec{\varphi}_2(t) := e^{\lambda t} (t \vec{\mathbf{u}}_1 + \vec{\mathbf{u}}_2),$$

$$\vec{\varphi}_3(t) := e^{\lambda t} \left(\frac{t^2}{2} \vec{\mathbf{u}}_1 + t \vec{\mathbf{u}}_2 + \vec{\mathbf{u}}_3 \right),$$

$$\vdots$$

$$\vec{\varphi}_m(t) := e^{\lambda t} \left(\frac{t^{m-1}}{(m-1)!} \vec{\mathbf{u}}_1 + \frac{t^{m-2}}{(m-2)!} \vec{\mathbf{u}}_2 + \dots + t \vec{\mathbf{u}}_{m-1} + \vec{\mathbf{u}}_m \right)$$

sind m **linear unabhängige Lösungen** von (H_n) .

(5) Für den allgemeinen Fall siehe die Fachliteratur zu Systemen von linearen Differentialgleichungen.

Betrachten wir ein Beispiel für den Spezialfall (Satz 15.27 (4)), dass die geometrische Vielfachheit jedes Eigenwertes genau 1 ist.

Beispiel 15.28. (Hauptvektoren)

Wir betrachten das folgende homogene System linearer Differentialgleichungen erster Ordnung mit konstanten Koeffizienten:

$$\begin{bmatrix} x_1' = x_1 - 2x_2 + x_3 \\ x_2' = -x_2 - x_3 \\ x_3' = 4x_2 + 3x_3 \end{bmatrix} \iff \begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix} = \begin{bmatrix} 1 & -2 & 1 \\ 0 & -1 & -1 \\ 0 & 4 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (15.18)$$

Es gilt also

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 1 \\ 0 & -1 & -1 \\ 0 & 4 & 3 \end{bmatrix}.$$

Schritt 1: Bestimmung der Eigenwerte von \mathbf{A}

Mittels Entwickeln nach der ersten Spalte finden wir

$$\begin{aligned} p_{\mathbf{A}}(\lambda) &= \det(\mathbf{A} - \lambda \mathbf{E}_n) = \det \left(\begin{bmatrix} 1 - \lambda & -2 & 1 \\ 0 & -1 - \lambda & -1 \\ 0 & 4 & 3 - \lambda \end{bmatrix} \right) \\ &= (-1)^{1+1} (1 - \lambda) \det \left(\begin{bmatrix} -1 - \lambda & -1 \\ 4 & 3 - \lambda \end{bmatrix} \right) \\ &= (1 - \lambda) [(-1 - \lambda)(3 - \lambda) + 4] \\ &= (1 - \lambda) [\lambda^2 - 2\lambda + 1] \\ &= (1 - \lambda) (\lambda - 1)^2 \\ &= (1 - \lambda) (1 - \lambda)^2 = (1 - \lambda)^3. \end{aligned}$$

Also ist $\lambda = 1$ der einzige Eigenwert von \mathbf{A} , und er hat die algebraische Vielfachheit 3.

Schritt 2: Bestimmung des Eigenraums zu $\lambda = 1$

$$\begin{aligned}
 (\mathbf{A} - 1 \mathbf{E}_3) \vec{\mathbf{u}} = \vec{\mathbf{0}} &\iff \left[\begin{array}{ccc|c} 1 & -1 & -2 & 1 \\ 0 & -1 & -1 & -1 \\ 0 & 4 & 3 & -1 \end{array} \right] \iff \\
 \left[\begin{array}{ccc|c} 0 & -2 & 1 & 0 \\ 0 & -2 & -1 & 0 \\ 0 & 4 & 2 & 0 \end{array} \right] &\xrightarrow[\downarrow]{\begin{array}{l} Z_3 \rightarrow Z_3 + 2Z_2 \\ Z_1 \rightarrow Z_1 - Z_2 \end{array}} \left[\begin{array}{ccc|c} 0 & 0 & 2 & 0 \\ 0 & -2 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] \xrightarrow[\downarrow]{\begin{array}{l} Z_1 \rightarrow \frac{1}{2} Z_1 \\ Z_2 \rightarrow -\frac{1}{2} Z_2 \end{array}} \\
 \left[\begin{array}{ccc|c} 0 & 0 & 1 & 0 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] &\xrightarrow[\downarrow]{Z_2 \rightarrow Z_2 - \frac{1}{2} Z_1} \left[\begin{array}{ccc|c} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] \xrightarrow[\downarrow]{Z_1 \leftrightarrow Z_2} \left[\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right],
 \end{aligned}$$

und es folgt $x_2 = x_3 = 0$, und x_1 kann beliebig in \mathbb{C} gewählt werden. Also ist der Eigenraum zu $\lambda = 1$

$$E_{\mathbf{A}}(1) = \left\{ \begin{bmatrix} \alpha \\ 0 \\ 0 \end{bmatrix} \in \mathbb{C}^3 : \alpha \in \mathbb{C} \right\} = \text{LH} \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right),$$

und die geometrische Vielfachheit von $\lambda = 1$ ist $\ell = \dim(E_{\mathbf{A}}(1)) = 1 < 3$.

Schritt 3: Bestimmung der Hauptvektoren

- Aus Schritt 2 kennen wir bereits den ersten Hauptvektor von $\lambda = 1$, nämlich den Eigenvektor

$$\vec{\mathbf{u}}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

und dieser liefert und die folgende Lösung von (15.18) (vgl. Satz 15.27 (4))

$$\vec{\varphi}_1 : \mathbb{R} \rightarrow \mathbb{R}^3, \quad \vec{\varphi}_1(t) := e^t \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

- Nach Satz 15.27 (4) lösen wir nun das lineare Gleichungssystem

$$(\mathbf{A} - 1 \mathbf{E}_3) \vec{\mathbf{u}} = \vec{\mathbf{u}}_1,$$

um einen zweiten Hauptvektor zu finden:

$$\begin{aligned} \left[\begin{array}{ccc|c} 1 & -1 & -2 & 1 \\ 0 & -1 & -1 & 0 \\ 0 & 4 & 3 & -1 \end{array} \right] & \iff \left[\begin{array}{ccc|c} 0 & -2 & 1 & 1 \\ 0 & -2 & -1 & 0 \\ 0 & 4 & 2 & 0 \end{array} \right] \\ \begin{array}{l} Z_3 \rightarrow Z_3 + 2Z_2 \\ Z_1 \rightarrow Z_1 - Z_2 \\ \iff \\ \downarrow \end{array} & \left[\begin{array}{ccc|c} 0 & 0 & 2 & 1 \\ 0 & -2 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] & \begin{array}{l} Z_1 \rightarrow \frac{1}{2}Z_1 \\ Z_2 \rightarrow -\frac{1}{2}Z_2 \\ \iff \\ \downarrow \end{array} & \left[\begin{array}{ccc|c} 0 & 0 & 1 & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] \\ \begin{array}{l} Z_2 \rightarrow Z_2 - \frac{1}{2}Z_1 \\ \iff \\ \downarrow \end{array} & \left[\begin{array}{ccc|c} 0 & 0 & 1 & \frac{1}{2} \\ 0 & 1 & 0 & -\frac{1}{4} \\ 0 & 0 & 0 & 0 \end{array} \right] & \begin{array}{l} Z_1 \leftrightarrow Z_2 \\ \iff \\ \downarrow \end{array} & \left[\begin{array}{ccc|c} 0 & 1 & 0 & -\frac{1}{4} \\ 0 & 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \end{array} \right], \end{aligned}$$

und wir finden $x_2 = -\frac{1}{4}$, $x_3 = \frac{1}{2}$ und $x_1 \in \mathbb{C}$ beliebig. Somit ist beispielsweise

$$\vec{\mathbf{u}}_2 = \begin{bmatrix} 0 \\ -\frac{1}{4} \\ \frac{1}{2} \end{bmatrix}$$

ein zweiter Hauptvektor, und dieser liefert (nach Satz 15.27 (4)) eine weitere Lösung von (15.18)

$$\vec{\varphi}_2 : \mathbb{R} \rightarrow \mathbb{R}^3, \quad \vec{\varphi}_2(t) := e^t \left(t \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -\frac{1}{4} \\ \frac{1}{2} \end{bmatrix} \right) = e^t \begin{bmatrix} t \\ -\frac{1}{4} \\ \frac{1}{2} \end{bmatrix}.$$

- Nach Satz 15.27 (4) lösen wir nun das lineare Gleichungssystem

$$(\mathbf{A} - 1\mathbf{E}_3) \vec{\mathbf{u}} = \vec{\mathbf{u}}_2,$$

um einen dritten Hauptvektor zu finden:

$$\begin{aligned} \left[\begin{array}{ccc|c} 1 & -1 & -2 & 0 \\ 0 & -1 & -1 & -\frac{1}{4} \\ 0 & 4 & 3 & \frac{1}{2} \end{array} \right] & \iff \left[\begin{array}{ccc|c} 0 & -2 & 1 & 0 \\ 0 & -2 & -1 & -\frac{1}{4} \\ 0 & 4 & 2 & \frac{1}{2} \end{array} \right] \\ \begin{array}{l} Z_3 \rightarrow Z_3 + 2Z_2 \\ Z_1 \rightarrow Z_1 - Z_2 \\ \iff \\ \downarrow \end{array} & \left[\begin{array}{ccc|c} 0 & 0 & 2 & \frac{1}{4} \\ 0 & -2 & -1 & -\frac{1}{4} \\ 0 & 0 & 0 & 0 \end{array} \right] & \begin{array}{l} Z_1 \rightarrow \frac{1}{2}Z_1 \\ Z_2 \rightarrow -\frac{1}{2}Z_2 \\ \iff \\ \downarrow \end{array} & \left[\begin{array}{ccc|c} 0 & 0 & 1 & \frac{1}{8} \\ 0 & 1 & \frac{1}{2} & \frac{1}{8} \\ 0 & 0 & 0 & 0 \end{array} \right] \end{aligned}$$

$$\begin{array}{ccc} Z_2 \rightarrow Z_2 - \frac{1}{2} Z_1 & & Z_1 \leftrightarrow Z_2 \\ \downarrow & & \downarrow \\ \longleftrightarrow & & \longleftrightarrow \end{array} \left[\begin{array}{ccc|c} 0 & 0 & 1 & \frac{1}{8} \\ 0 & 1 & 0 & \frac{1}{16} \\ 0 & 0 & 0 & 0 \end{array} \right] \left[\begin{array}{ccc|c} 0 & 1 & 0 & \frac{1}{16} \\ 0 & 0 & 1 & \frac{1}{8} \\ 0 & 0 & 0 & 0 \end{array} \right],$$

und wir finden $x_2 = \frac{1}{16}$, $x_3 = \frac{1}{8}$ und $x_1 \in \mathbb{C}$ beliebig. Somit ist beispielsweise

$$\vec{\mathbf{u}}_3 = \begin{bmatrix} 0 \\ \frac{1}{16} \\ \frac{1}{8} \end{bmatrix}$$

ein dritter Hauptvektor, und dieser liefert (nach Satz 15.27 (4)) eine dritte Lösung von (15.18)

$$\vec{\varphi}_3 : \mathbb{R} \rightarrow \mathbb{R}^3, \quad \vec{\varphi}_3(t) := e^t \left(\frac{t^2}{2} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ -\frac{1}{4} \\ \frac{1}{2} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{16} \\ \frac{1}{8} \end{bmatrix} \right) = e^t \begin{bmatrix} \frac{t^2}{2} \\ -\frac{1}{4}t + \frac{1}{16} \\ \frac{1}{2}t + \frac{1}{8} \end{bmatrix}.$$

Ergebnis: Nach Satz 15.27 (4) wissen wir, dass die gefundenen Lösungen $\vec{\varphi}_1, \vec{\varphi}_2, \vec{\varphi}_3$ linear unabhängig sind und somit ein reelles Fundamentalsystem von (15.18) bilden. Also ist ein reelles Fundamentalsystem von (15.18) durch

$$\left(e^t \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, e^t \begin{bmatrix} t \\ -\frac{1}{4} \\ \frac{1}{2} \end{bmatrix}, e^t \begin{bmatrix} \frac{t^2}{2} \\ -\frac{1}{4}t + \frac{1}{16} \\ \frac{1}{2}t + \frac{1}{8} \end{bmatrix} \right)$$

gegeben. Multiplizieren wir die Vektoren in \mathbb{R}^3 mit geeigneten reellen Zahlen, um die Brüche loszuwerden, so erhalten wir ein „einfacheres“ reelles Fundamentalsystem

$$\left(e^t \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, e^t \begin{bmatrix} 4t \\ -1 \\ 2 \end{bmatrix}, e^t \begin{bmatrix} 8t^2 \\ 1 - 4t \\ 2 + 8t \end{bmatrix} \right).$$

Die reelle allgemeine Lösung von (15.18) ist somit

$$\vec{\mathbf{x}} : \mathbb{R} \rightarrow \mathbb{R}^3, \quad \vec{\mathbf{x}}(t) := e^t \left(c_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 4t \\ -1 \\ 2 \end{bmatrix} + c_3 \begin{bmatrix} 8t^2 \\ 1 - 4t \\ 2 + 8t \end{bmatrix} \right), \quad c_1, c_2, c_3 \in \mathbb{R}.$$

15.6 Zusammenhang mit linearen Differentialgleichungen höherer Ordnung

In diesem letzten Teilkapitel lernen wir, wie man eine lineare Differentialgleichung n -ter Ordnung in ein System linearer Differentialgleichungen erster Ordnung umwandeln kann.

Seien $n \in \mathbb{N}$, $I \subseteq \mathbb{R}$ ein Intervall und $a_0, a_1, \dots, a_{n-1}, b : I \rightarrow \mathbb{R}$ stetige Funktionen. Die **lineare Differentialgleichung n -ter Ordnung**

$$y^{(n)} + a_{n-1}(t) y^{(n-1)} + \dots + a_2(t) y'' + a_1(t) y' + a_0(t) y = b(t) \quad (15.19)$$

lässt sich **in ein System linearer Differentialgleichungen erster Ordnung umwandeln**, indem man setzt

$$x_1 := y, \quad x_2 := y', \quad x_3 := y'', \quad \dots, \quad x_n := y^{(n-1)}. \quad (15.20)$$

Damit erhält man das folgende System linearer Differentialgleichungen ersten Ordnung für x_1, x_2, \dots, x_n :

$$\begin{bmatrix} x_1' = x_2 \\ x_2' = x_3 \\ \vdots \\ x_{n-1}' = x_n \\ x_n' = -a_0(t)x_1 - a_1(t)x_2 - \dots - a_{n-1}(t)x_n + b(t) \end{bmatrix}$$

Mit

$$\mathbf{A}(t) = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & 1 \\ -a_0(t) & -a_1(t) & \dots & \dots & -a_{n-1}(t) \end{bmatrix}, \quad \vec{\mathbf{d}}(t) = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ b(t) \end{bmatrix}, \quad \vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix}$$

ergibt sich

$$\vec{\mathbf{x}}' = \mathbf{A}(t) \vec{\mathbf{x}} + \vec{\mathbf{d}}(t). \quad (15.21)$$

Es gelten dann:

$$(1) \quad y \text{ löst (15.19).} \quad \implies \quad \vec{\mathbf{x}} := \begin{bmatrix} y \\ y' \\ y'' \\ \vdots \\ y^{(n-1)} \end{bmatrix} \text{ löst (15.21).}$$

$$(2) \quad \vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \text{ löst (15.21).} \quad \implies \quad y := x_1 \text{ löst (15.19).}$$

Also sind (15.19) und (15.21) äquivalente Probleme.

Abschließend betrachten wir noch die **zu (15.19) gehörige homogene Gleichung** für den **Spezialfall konstanter Koeffizienten**, also:

$$y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_2 y'' + a_1 y' + a_0 y = 0. \quad (15.22)$$

Mit (15.20) liefert diese das homogene System linearer Differentialgleichungen erster Ordnung mit konstanten Koeffizienten

$$\vec{x}' = \mathbf{A} \vec{x} \quad (15.23)$$

mit der Matrix

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ -a_0 & -a_1 & \cdots & \cdots & -a_{n-1} \end{bmatrix}.$$

- (1) Mit dem komplexen Exponentialansatz $y(t) = e^{st}$ folgt, dass die **charakteristische Gleichung von (15.22)** durch

$$s^n + a_{n-1} s^{n-1} + \dots + a_2 s^2 + a_1 s + a_0 = 0$$

gegeben ist.

- (2) Das **charakteristische Polynom für die Matrix \mathbf{A}** in (15.23) ist

$$p_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda \mathbf{E}_n) = (-1)^n (\lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_2 \lambda^2 + a_1 \lambda + a_0).$$

Dieses zeigt man mit vollständiger Induktion, indem man

$$\det(\mathbf{A} - \lambda \mathbf{E}_n) = \det \left(\begin{bmatrix} -\lambda & 1 & 0 & \cdots & 0 \\ 0 & -\lambda & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & -\lambda & 1 \\ -a_0 & -a_1 & \cdots & \cdots & -a_{n-1} - \lambda \end{bmatrix} \right)$$

nach der ersten Spalte entwickelt.

Fazit: Die Eigenwerte von \mathbf{A} in (15.23) sind gerade die Lösungen der charakteristischen Gleichung von (15.22).

Betrachten wir hierzu ein Beispiel.

Beispiel 15.29. (Umwandeln einer homogenen linearen DGL 4-ter Ordnung in ein System linearer DGLen 1-ter Ordnung)

Wir betrachten die homogene lineare Differentialgleichung vierter Ordnung

$$y^{(4)} - y = 0 \quad \Longleftrightarrow \quad y^{(4)} = y \quad (15.24)$$

aus Beispiel 14.24 (b), deren reelle allgemeine Lösung

$$y(t) = \gamma_1 e^t + \gamma_2 e^{-t} + \gamma_3 \cos(t) + \gamma_4 \sin(t), \quad \gamma_1, \gamma_2, \gamma_3, \gamma_4 \in \mathbb{R}. \quad (15.25)$$

wir bereits aus Beispiel 14.24 (b) kennen.

Wir wollen diese nun in ein System von homogenen linearen Differentialgleichungen erster Ordnung umwandeln und dieses dann wie gewohnt lösen: Wir setzen

$$x_1 := y, \quad x_2 := y', \quad x_3 := y'', \quad x_4 := y'''$$

und erhalten somit:

$$\begin{aligned} \begin{bmatrix} x_1' = x_2 \\ x_2' = x_3 \\ x_3' = x_4 \\ x_4' = x_1 \end{bmatrix} &\Longleftrightarrow \begin{bmatrix} x_1' \\ x_2' \\ x_3' \\ x_4' \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}}_{=: \mathbf{A}} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \\ &\Longleftrightarrow \underbrace{\vec{x}' = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}}_{= \mathbf{A}} \vec{x} \end{aligned} \quad (15.26)$$

Schritt 1: Berechnung der Eigenwerte von \mathbf{A}

Wir entwickeln nach der ersten Spalte

$$\det(\mathbf{A} - \lambda \mathbf{E}_4) = \det \left(\begin{bmatrix} -\lambda & 1 & 0 & 0 \\ 0 & -\lambda & 1 & 0 \\ 0 & 0 & -\lambda & 1 \\ 1 & 0 & 0 & -\lambda \end{bmatrix} \right)$$

$$\begin{aligned}
&= (-1)^{1+1} (-\lambda) \det \underbrace{\begin{pmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 0 & 0 & -\lambda \end{pmatrix}}_{=(-\lambda)^3} + (-1)^{4+1} \cdot 1 \cdot \det \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \end{pmatrix}}_{=1^3=1} \\
&= (-\lambda)^4 - 1 = \lambda^4 - 1 = (\lambda^2 - 1)(\lambda^2 + 1) = (\lambda - 1)(\lambda + 1)(\lambda - j)(\lambda + j)
\end{aligned}$$

und finden die vier verschiedenen Eigenwerte $\lambda_1 = 1$, $\lambda_2 = -1$, $\lambda_3 = j$ und $\lambda_4 = -j$. (Die Determinanten wurden mit der Formel für die Determinante einer oberen bzw. einer unteren Dreiecksmatrix (vgl. Hilfssatz 13.20) berechnet. Jeder dieser vier Eigenwerte hat die algebraische und geometrische Vielfachheit 1.

Schritt 2: Berechnung der Eigenräume

$$\begin{aligned}
\lambda_1 = 1 : & \begin{bmatrix} -1 & 1 & 0 & 0 & | & 0 \\ 0 & -1 & 1 & 0 & | & 0 \\ 0 & 0 & -1 & 1 & | & 0 \\ 1 & 0 & 0 & -1 & | & 0 \end{bmatrix} \xrightarrow[Z_4 \rightarrow Z_4 + Z_1]{\Leftrightarrow} \begin{bmatrix} -1 & 1 & 0 & 0 & | & 0 \\ 0 & -1 & 1 & 0 & | & 0 \\ 0 & 0 & -1 & 1 & | & 0 \\ 0 & 1 & 0 & -1 & | & 0 \end{bmatrix} \\
& \xrightarrow[Z_4 \rightarrow Z_4 + Z_2]{\Leftrightarrow} \begin{bmatrix} -1 & 1 & 0 & 0 & | & 0 \\ 0 & -1 & 1 & 0 & | & 0 \\ 0 & 0 & -1 & 1 & | & 0 \\ 0 & 0 & 1 & -1 & | & 0 \end{bmatrix} \xrightarrow[Z_4 \rightarrow Z_4 + Z_3]{\Leftrightarrow} \begin{bmatrix} -1 & 1 & 0 & 0 & | & 0 \\ 0 & -1 & 1 & 0 & | & 0 \\ 0 & 0 & -1 & 1 & | & 0 \\ 0 & 0 & 0 & 0 & | & 0 \end{bmatrix} \\
& \xrightarrow[Z_2 \rightarrow Z_2 + Z_3]{\Leftrightarrow} \begin{bmatrix} -1 & 1 & 0 & 0 & | & 0 \\ 0 & -1 & 0 & 1 & | & 0 \\ 0 & 0 & -1 & 1 & | & 0 \\ 0 & 0 & 0 & 0 & | & 0 \end{bmatrix} \xrightarrow[Z_1 \rightarrow Z_1 + Z_2]{\Leftrightarrow} \begin{bmatrix} -1 & 0 & 0 & 1 & | & 0 \\ 0 & -1 & 0 & 1 & | & 0 \\ 0 & 0 & -1 & 1 & | & 0 \\ 0 & 0 & 0 & 0 & | & 0 \end{bmatrix} \\
& \xrightarrow[Z_1 \rightarrow -Z_1]{\Leftrightarrow} \begin{bmatrix} 1 & 0 & 0 & -1 & | & 0 \\ 0 & 1 & 0 & -1 & | & 0 \\ 0 & 0 & 1 & -1 & | & 0 \\ 0 & 0 & 0 & 0 & | & 0 \end{bmatrix} \xrightarrow[Z_2 \rightarrow -Z_2]{\Leftrightarrow} \begin{bmatrix} 1 & 0 & 0 & -1 & | & 0 \\ 0 & 1 & 0 & -1 & | & 0 \\ 0 & 0 & 1 & -1 & | & 0 \\ 0 & 0 & 0 & 0 & | & 0 \end{bmatrix} \xrightarrow[Z_3 \rightarrow -Z_3]{\Leftrightarrow} \begin{bmatrix} u_1 = u_4 \\ u_2 = u_4 \\ u_3 = u_4 \\ u_4 \text{ bel.} \end{bmatrix} \implies E_{\mathbf{A}}(1) = \text{LH} \left(\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right)
\end{aligned}$$

Damit finden wir die folgende Lösung von (15.26):

$$\vec{\varphi}_1 : \mathbb{R} \rightarrow \mathbb{R}^4, \quad \vec{\varphi}_1(t) := e^t \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} e^t \\ e^t \\ e^t \\ e^t \end{bmatrix}.$$

$$\begin{array}{l}
\lambda_2 = -1 : \\
\begin{array}{ccc}
\begin{array}{c} \downarrow \\ \Leftrightarrow \end{array} & \begin{bmatrix} 1 & 1 & 0 & 0 & | & 0 \\ 0 & 1 & 1 & 0 & | & 0 \\ 0 & 0 & 1 & 1 & | & 0 \\ 1 & 0 & 0 & 1 & | & 0 \end{bmatrix} & \begin{array}{c} Z_4 \rightarrow Z_4 - Z_1 \\ \downarrow \\ \Leftrightarrow \end{array} & \begin{bmatrix} 1 & 1 & 0 & 0 & | & 0 \\ 0 & 1 & 1 & 0 & | & 0 \\ 0 & 0 & 1 & 1 & | & 0 \\ 0 & -1 & 0 & 1 & | & 0 \end{bmatrix} \\
\begin{array}{c} \downarrow \\ \Leftrightarrow \end{array} & \begin{bmatrix} 1 & 1 & 0 & 0 & | & 0 \\ 0 & 1 & 1 & 0 & | & 0 \\ 0 & 0 & 1 & 1 & | & 0 \\ 0 & 0 & 1 & 1 & | & 0 \end{bmatrix} & \begin{array}{c} Z_4 \rightarrow Z_4 - Z_3 \\ \downarrow \\ \Leftrightarrow \end{array} & \begin{bmatrix} 1 & 1 & 0 & 0 & | & 0 \\ 0 & 1 & 1 & 0 & | & 0 \\ 0 & 0 & 1 & 1 & | & 0 \\ 0 & 0 & 0 & 0 & | & 0 \end{bmatrix} \\
\begin{array}{c} \downarrow \\ \Leftrightarrow \end{array} & \begin{bmatrix} 1 & 1 & 0 & 0 & | & 0 \\ 0 & 1 & 0 & -1 & | & 0 \\ 0 & 0 & 1 & 1 & | & 0 \\ 0 & 0 & 0 & 0 & | & 0 \end{bmatrix} & \begin{array}{c} Z_1 \rightarrow Z_1 - Z_2 \\ \downarrow \\ \Leftrightarrow \end{array} & \begin{bmatrix} 1 & 0 & 0 & 1 & | & 0 \\ 0 & 1 & 0 & -1 & | & 0 \\ 0 & 0 & 1 & 1 & | & 0 \\ 0 & 0 & 0 & 0 & | & 0 \end{bmatrix} \\
\Rightarrow & \begin{bmatrix} u_1 = -u_4 \\ u_2 = -u_4 \\ u_3 = -u_4 \\ u_4 \text{ bel.} \end{bmatrix} & \Rightarrow & E_{\mathbf{A}}(-1) = \text{LH} \left(\begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \right)
\end{array}
\end{array}$$

Damit finden wir die folgende Lösung von (15.26):

$$\vec{\varphi}_2 : \mathbb{R} \rightarrow \mathbb{R}^4, \quad \vec{\varphi}_2(t) := e^{-t} \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -e^{-t} \\ e^{-t} \\ -e^{-t} \\ e^{-t} \end{bmatrix}.$$

Für $\lambda_3 = j$ finden wir schließlich:

$$\begin{array}{l}
\lambda_3 = j : \\
\begin{array}{ccc}
\begin{array}{c} \downarrow \\ \Leftrightarrow \end{array} & \begin{bmatrix} -j & 1 & 0 & 0 & | & 0 \\ 0 & -j & 1 & 0 & | & 0 \\ 0 & 0 & -j & 1 & | & 0 \\ 1 & 0 & 0 & -j & | & 0 \end{bmatrix} & \begin{array}{c} Z_1 \rightarrow j Z_1 \\ \downarrow \\ \Leftrightarrow \end{array} & \begin{bmatrix} 1 & j & 0 & 0 & | & 0 \\ 0 & -j & 1 & 0 & | & 0 \\ 0 & 0 & -j & 1 & | & 0 \\ 1 & 0 & 0 & -j & | & 0 \end{bmatrix} \\
\begin{array}{c} \downarrow \\ \Leftrightarrow \end{array} & \begin{bmatrix} 1 & j & 0 & 0 & | & 0 \\ 0 & -j & 1 & 0 & | & 0 \\ 0 & 0 & -j & 1 & | & 0 \\ 0 & -j & 0 & -j & | & 0 \end{bmatrix} & \begin{array}{c} Z_4 \rightarrow Z_4 - Z_2 \\ \downarrow \\ \Leftrightarrow \end{array} & \begin{bmatrix} 1 & j & 0 & 0 & | & 0 \\ 0 & -j & 1 & 0 & | & 0 \\ 0 & 0 & -j & 1 & | & 0 \\ 0 & 0 & -1 & -j & | & 0 \end{bmatrix}
\end{array}
\end{array}$$

$$\begin{array}{ccc}
\begin{array}{c} Z_4 \rightarrow Z_4 + j Z_3 \\ \Downarrow \\ \Longleftrightarrow \end{array} & \begin{bmatrix} 1 & j & 0 & 0 & | & 0 \\ 0 & -j & 1 & 0 & | & 0 \\ 0 & 0 & -j & 1 & | & 0 \\ 0 & 0 & 0 & 0 & | & 0 \end{bmatrix} & \begin{array}{c} Z_1 \rightarrow Z_1 + Z_2 \\ Z_3 \rightarrow j Z_3 \\ \Downarrow \\ \Longleftrightarrow \end{array} & \begin{bmatrix} 1 & 0 & 1 & 0 & | & 0 \\ 0 & -j & 1 & 0 & | & 0 \\ 0 & 0 & 1 & j & | & 0 \\ 0 & 0 & 0 & 0 & | & 0 \end{bmatrix} \\
\\
\begin{array}{c} Z_1 \rightarrow Z_1 - Z_3 \\ Z_2 \rightarrow Z_2 - Z_3 \\ \Downarrow \\ \Longleftrightarrow \end{array} & \begin{bmatrix} 1 & 0 & 0 & -j & | & 0 \\ 0 & -j & 0 & -j & | & 0 \\ 0 & 0 & 1 & j & | & 0 \\ 0 & 0 & 0 & 0 & | & 0 \end{bmatrix} & \begin{array}{c} Z_2 \rightarrow j Z_2 \\ \Downarrow \\ \Longleftrightarrow \end{array} & \begin{bmatrix} 1 & 0 & 0 & -j & | & 0 \\ 0 & 1 & 0 & 1 & | & 0 \\ 0 & 0 & 1 & j & | & 0 \\ 0 & 0 & 0 & 0 & | & 0 \end{bmatrix} \\
\\
\Rightarrow & \begin{bmatrix} u_1 = j u_4 \\ u_2 = -u_4 \\ u_3 = -j u_4 \\ u_4 \text{ bel.} \end{bmatrix} & \Rightarrow & E_{\mathbf{A}}(j) = \text{LH} \left(\begin{bmatrix} j \\ -1 \\ -j \\ 1 \end{bmatrix} \right)
\end{array}$$

Damit finden wir die folgende komplexe Lösung von (15.26):

$$\begin{aligned}
\vec{\psi}_3 : \mathbb{R} \rightarrow \mathbb{C}^4, \quad \vec{\psi}_3(t) &:= e^{jt} \begin{bmatrix} j \\ -1 \\ -j \\ 1 \end{bmatrix} = \begin{bmatrix} j e^{jt} \\ -e^{jt} \\ -j e^{jt} \\ e^{jt} \end{bmatrix} = \begin{bmatrix} e^{j\frac{\pi}{2}} e^{jt} \\ -e^{jt} \\ e^{-j\frac{\pi}{2}} e^{jt} \\ e^{jt} \end{bmatrix} = \begin{bmatrix} e^{j(t+\frac{\pi}{2})} \\ -e^{jt} \\ e^{j(t-\frac{\pi}{2})} \\ e^{jt} \end{bmatrix} \\
&= \begin{bmatrix} \cos(t + \frac{\pi}{2}) + j \sin(t + \frac{\pi}{2}) \\ -\cos(t) - j \sin(t) \\ \cos(t - \frac{\pi}{2}) + j \sin(t - \frac{\pi}{2}) \\ \cos(t) + j \sin(t) \end{bmatrix} = \begin{bmatrix} -\sin(t) + j \cos(t) \\ -\cos(t) - j \sin(t) \\ \sin(t) - j \cos(t) \\ \cos(t) + j \sin(t) \end{bmatrix},
\end{aligned}$$

wobei wir die Euler-Formel und die folgenden Additionstheoreme genutzt haben:

$$\begin{aligned}
\cos\left(t + \frac{\pi}{2}\right) &= \cos(t) \cos\left(\frac{\pi}{2}\right) - \sin(t) \sin\left(\frac{\pi}{2}\right) = -\sin(t), \\
\cos\left(t - \frac{\pi}{2}\right) &= \cos(t) \cos\left(-\frac{\pi}{2}\right) - \sin(t) \sin\left(-\frac{\pi}{2}\right) = \sin(t), \\
\sin\left(t + \frac{\pi}{2}\right) &= \sin(t) \cos\left(\frac{\pi}{2}\right) + \cos(t) \sin\left(\frac{\pi}{2}\right) = \cos(t), \\
\sin\left(t - \frac{\pi}{2}\right) &= \sin(t) \cos\left(-\frac{\pi}{2}\right) + \cos(t) \sin\left(-\frac{\pi}{2}\right) = -\cos(t).
\end{aligned}$$

Damit finden wir die folgenden zwei linear unabhängigen reellen Lösungen von

(15.26):

$$\vec{\varphi}_3 : \mathbb{R} \rightarrow \mathbb{R}^4, \quad \vec{\varphi}_3(t) := \operatorname{Re}(\vec{\psi}_3(t)) = \begin{bmatrix} -\sin(t) \\ -\cos(t) \\ \sin(t) \\ \cos(t) \end{bmatrix},$$

$$\vec{\varphi}_4 : \mathbb{R} \rightarrow \mathbb{R}^4, \quad \vec{\varphi}_4(t) := \operatorname{Im}(\vec{\psi}_3(t)) = \begin{bmatrix} \cos(t) \\ -\sin(t) \\ -\cos(t) \\ \sin(t) \end{bmatrix}.$$

Ergebnis: Ein reelles Fundamentalsystem ist durch

$$\left(\begin{bmatrix} e^t \\ e^t \\ e^t \\ e^t \end{bmatrix}, \begin{bmatrix} -e^{-t} \\ e^{-t} \\ -e^{-t} \\ e^{-t} \end{bmatrix}, \begin{bmatrix} -\sin(t) \\ -\cos(t) \\ \sin(t) \\ \cos(t) \end{bmatrix}, \begin{bmatrix} \cos(t) \\ -\sin(t) \\ -\cos(t) \\ \sin(t) \end{bmatrix} \right)$$

gegeben, und die reelle allgemeine Lösung von (15.26) ist $\vec{x} : \mathbb{R} \rightarrow \mathbb{R}^4$,

$$\vec{x}(t) := c_1 \begin{bmatrix} e^t \\ e^t \\ e^t \\ e^t \end{bmatrix} + c_2 \begin{bmatrix} -e^{-t} \\ e^{-t} \\ -e^{-t} \\ e^{-t} \end{bmatrix} + c_3 \begin{bmatrix} -\sin(t) \\ -\cos(t) \\ \sin(t) \\ \cos(t) \end{bmatrix} + c_4 \begin{bmatrix} \cos(t) \\ -\sin(t) \\ -\cos(t) \\ \sin(t) \end{bmatrix}, \quad c_1, c_2, c_3, c_4 \in \mathbb{R}.$$

Wir haben insbesondere

$$x_1 : \mathbb{R} \rightarrow \mathbb{R}, \quad x_1(t) = c_1 e^t - c_2 e^{-t} - c_3 \sin(t) + c_4 \cos(t), \quad c_1, c_2, c_3, c_4 \in \mathbb{R}.$$

Mit der Umbenennung $\gamma_1 := c_1$, $\gamma_2 := -c_2$, $\gamma_3 := c_4$ und $\gamma_4 := -c_3$ ist dieses in der Tat die Lösung (15.25) von (15.24).

Teil VI

Differentialrechnung in mehreren Variablen

KAPITEL 16

Konvergenz, Stetigkeit und Differenzierbarkeit

In diesem Kapitel betrachten wir Funktionen mehrerer reeller Variablen (also z.B. eine Funktion der drei Ortsvariablen x, y, z und der Zeit t) und mit Werten in \mathbb{R} oder auch in \mathbb{R}^m . Wir werden die bereits aus der HM A bekannten Begriffe Grenzwert, Stetigkeit und Differenzierbarkeit (vgl. Kapitel 5 bis 7 der HM A) auf solche Funktionen ausdehnen. Der Begriff des Integrals wird erst in der HM C für Funktionen mehrerer Variablen besprochen. In dem nachfolgenden Kapitel nutzen wir dann die Ableitungen solcher Funktionen, um Extremwertprobleme ohne und mit Nebenbedingungen zu lösen.

16.1 Offene Teilmengen und Konvergenz in \mathbb{R}^n

Wir starten dieses Kapitel mit ein bisschen elementarer Topologie (d.h. der Begriff einer offenen Menge und weitere Begriffe aus diesem Umfeld), die wir in den nachfolgenden Teilkapiteln benötigen. Weiter definieren wir den Begriff der Konvergenz von Folgen in \mathbb{R}^n .

Notation 16.1. (Standardnorm und offene Kugel)

(1) Mit $\|\cdot\|$ bezeichnen wir die **Standardnorm** (oder **euklidische Norm**)

in \mathbb{R}^n , also für $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ ist

$$\|\vec{x}\| := \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$$

die **Standardnorm** in \mathbb{R}^n .

$\|\vec{x} - \vec{y}\|$ ist dann der **Abstand** von \vec{x} und \vec{y} in \mathbb{R}^n .

(2) Für $\vec{a} \in \mathbb{R}^n$ und $r > 0$ ist

$$U_r(\vec{a}) := \{ \vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{a}\| < r \}$$

die **offene Kugel** um \vec{a} mit **Radius** r .

Mit dieser Notation können wir nun den ersten neuen Begriff einführen.

Definition 16.2. (offene Menge)

$\Omega \subseteq \mathbb{R}^n$ heißt **offen**, wenn für jedes $\vec{x} \in \Omega$ ein $\varepsilon > 0$ existiert mit $U_\varepsilon(\vec{x}) \subseteq \Omega$.

In Worten: Um jeden Punkt $\vec{x} \in \Omega$ kann man eine kleine offene Kugel $U_\varepsilon(\vec{x})$ legen, die ganz in Ω enthalten ist (vgl. das linke Bild in Abbildung 16.1). Der Radius $\varepsilon > 0$ dieser Kugel wird dabei in der Regel mit dem Punkt $\vec{x} \in \Omega$ variieren.

Betrachten wir einige Beispiele.

Beispiel 16.3. (offene Mengen)

Die Idee der Begründung in den nachfolgenden Beispielen ist jeweils aus der zugehörigen Skizze für den Fall von \mathbb{R}^2 ersichtlich.

(a) Die leere Menge \emptyset und \mathbb{R}^n sind offen.

Begründungen:

- Die leere Menge \emptyset enthält keine Punkte; also ist für jedes \vec{x} aus \emptyset die Bedingung an die Punkte einer offenen Menge erfüllt. Somit ist die leere Menge offen.
- Für jedes $\vec{x} \in \mathbb{R}^n$ gilt für jeden Radius $\varepsilon > 0$, dass $U_\varepsilon(\vec{x}) \subseteq \mathbb{R}^n$. Also ist \mathbb{R}^n offen.

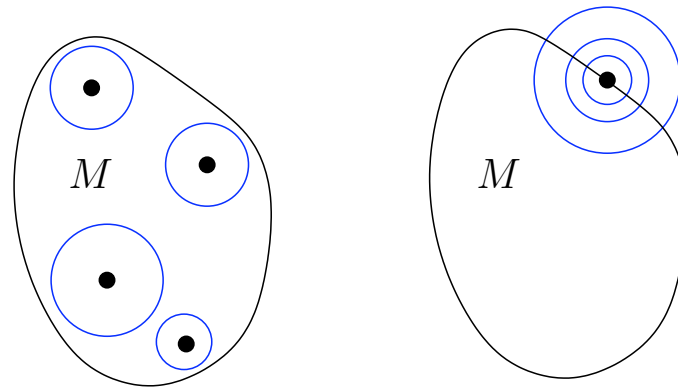


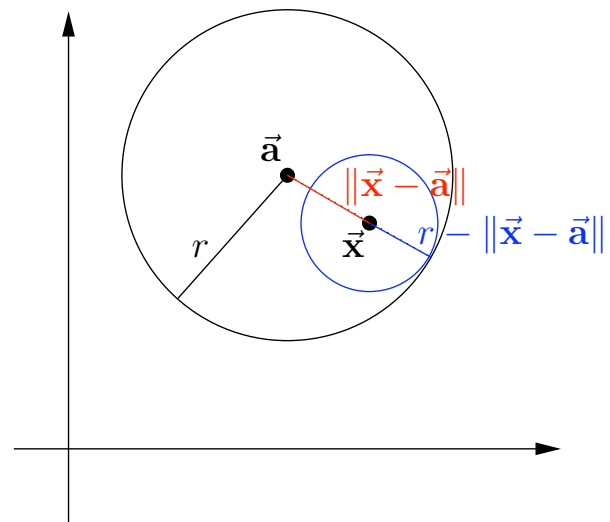
Abbildung 16.1: In beiden Bildern sehen wir eine Menge $M \subseteq \mathbb{R}^2$, die innerhalb der schwarzen geschlossenen Kurve liegt. Ist diese Menge M offen, so gehört der „Rand“, also die Randkurve, nicht mit zu M , und man kann um jedem Punkt in M eine kleine (offene) Kugel legen, die ganz in der Menge M liegt (linkes Bild). Im rechten Bild ist eine typische Situation bei einem Randpunkt zu sehen. Ein Teil jeder (offenen) Kugel um den Randpunkt gehört zur Menge M und ein Teil gehört zu $\mathbb{R}^2 \setminus M$.

(b) Für jedes $\vec{a} \in \mathbb{R}^n$ und jedes $r > 0$ ist die offene Kugel $U_r(\vec{a})$ offen.

Begründung: Sei $\vec{x} \in U_r(\vec{a})$ beliebig. Wir zeigen nun, dass die offene Kugel $U_\varepsilon(\vec{x})$ mit $\varepsilon := r - \|\vec{x} - \vec{a}\|$ ganz in $U_r(\vec{a})$ liegt (siehe Skizze).

Sei dazu $\vec{y} \in U_\varepsilon(\vec{x})$ beliebig. Dann gilt $\|\vec{y} - \vec{x}\| < \varepsilon$, und mit der Dreiecksungleichung folgt

$$\begin{aligned} & \|\vec{y} - \vec{a}\| \\ &= \|(\vec{y} - \vec{x}) + (\vec{x} - \vec{a})\| \\ &\leq \underbrace{\|\vec{y} - \vec{x}\|}_{< \varepsilon} + \|\vec{x} - \vec{a}\| \\ &< \varepsilon + \|\vec{x} - \vec{a}\| \\ &= (r - \|\vec{x} - \vec{a}\|) + \|\vec{x} - \vec{a}\| \\ &= r, \end{aligned}$$



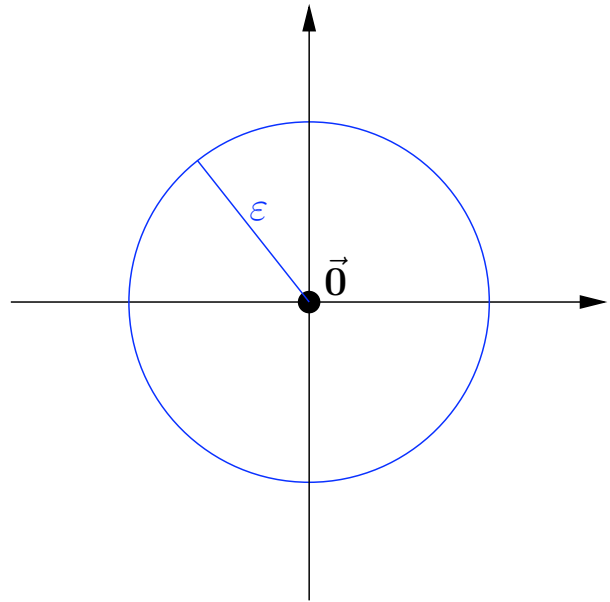
also $\|\vec{y} - \vec{a}\| < r$, d.h. $\vec{y} \in U_r(\vec{a})$. Da $\vec{y} \in U_\varepsilon(\vec{x})$ beliebig war, folgt, dass $U_\varepsilon(\vec{x}) \subseteq U_r(\vec{a})$. Also ist $U_r(\vec{a})$ offen.

(c) $\{\vec{0}\} \subseteq \mathbb{R}^n$ ist nicht offen.

Begründung (siehe Skizze):

Für jeden Radius $\varepsilon > 0$ ist die offene Kugel $U_\varepsilon(\vec{0})$ nicht in $\{\vec{0}\}$ enthalten.

Also gibt es keine offene Kugel $U_\varepsilon(\vec{0})$ mit $U_\varepsilon(\vec{0}) \subseteq \{\vec{0}\}$, und $\{\vec{0}\}$ ist nicht offen.



(d) $\mathbb{R}^n \setminus \{\vec{0}\}$ ist offen.

Begründung: Sei $\vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\}$ beliebig. Dann ist $\vec{x} \neq \vec{0}$ und somit $\|\vec{x}\| > 0$. Wir zeigen nun, dass die offene Kugel $U_\varepsilon(\vec{x})$ mit $\varepsilon := \|\vec{x}\|$ ganz in $\mathbb{R}^n \setminus \{\vec{0}\}$ liegt; damit folgt dann, dass $\mathbb{R}^n \setminus \{\vec{0}\}$ offen ist (siehe Skizze).

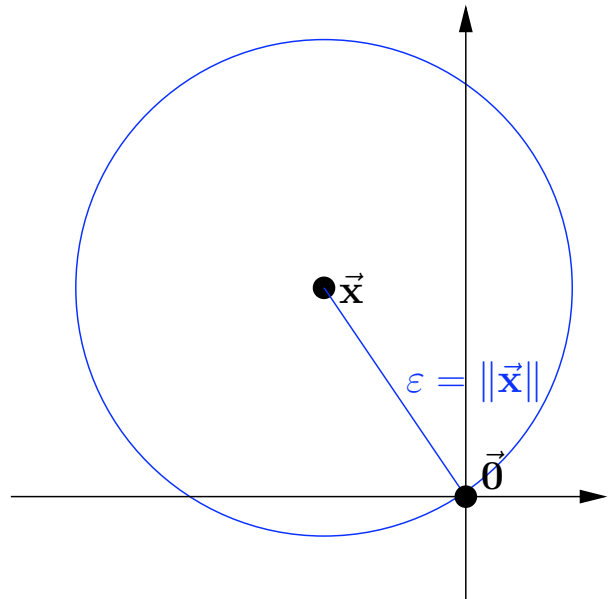
Sei also $\vec{y} \in U_\varepsilon(\vec{x})$ beliebig.

Dann gilt

$$\|\vec{y} - \vec{x}\| < \varepsilon = \|\vec{x}\|.$$

Mit der unteren Dreiecksungleichung folgt

$$\begin{aligned} \|\vec{y}\| &= \|\vec{x} + (\vec{y} - \vec{x})\| \\ &\geq \|\vec{x}\| - \|\vec{y} - \vec{x}\| \\ &\geq \|\vec{x}\| - \underbrace{\|\vec{y} - \vec{x}\|}_{< \varepsilon = \|\vec{x}\|} \\ &> \|\vec{x}\| - \|\vec{x}\| = 0, \end{aligned}$$



also $\|\vec{y}\| > 0$, d.h. $\vec{y} \in \mathbb{R}^n \setminus \{\vec{0}\}$. Da $\vec{y} \in U_\varepsilon(\vec{x})$ beliebig war, haben wir gezeigt, dass die offene Kugel $U_\varepsilon(\vec{x})$ mit $\varepsilon = \|\vec{x}\|$ ganz in $\mathbb{R}^n \setminus \{\vec{0}\}$ liegt. Also ist $\mathbb{R}^n \setminus \{\vec{0}\}$ offen.

(e) $\{\vec{x} \in \mathbb{R}^n : \|\vec{x}\| = 1\}$ und $\{\vec{x} \in \mathbb{R}^n : \|\vec{x}\| \leq 1\}$ sind beide nicht offen.

Begründung: Seien $M_1 := \{\vec{x} \in \mathbb{R}^n : \|\vec{x}\| = 1\}$ und $M_2 := \{\vec{x} \in \mathbb{R}^n : \|\vec{x}\| \leq 1\}$. Wir betrachten ein $\vec{x} \in \mathbb{R}^n$ mit $\|\vec{x}\| = 1$. Dieses gehört zu M_1 und M_2 . Wir zeigen nun, dass es für ein solches \vec{x} keine offene Kugel $U_\varepsilon(\vec{x})$

gibt mit $U_\varepsilon(\vec{x}) \subseteq M_1$ bzw. $U_\varepsilon(\vec{x}) \subseteq M_2$.

Sei dazu $\vec{y} := \left(1 + \frac{\varepsilon}{2}\right) \vec{x}$. Dann ist einerseits

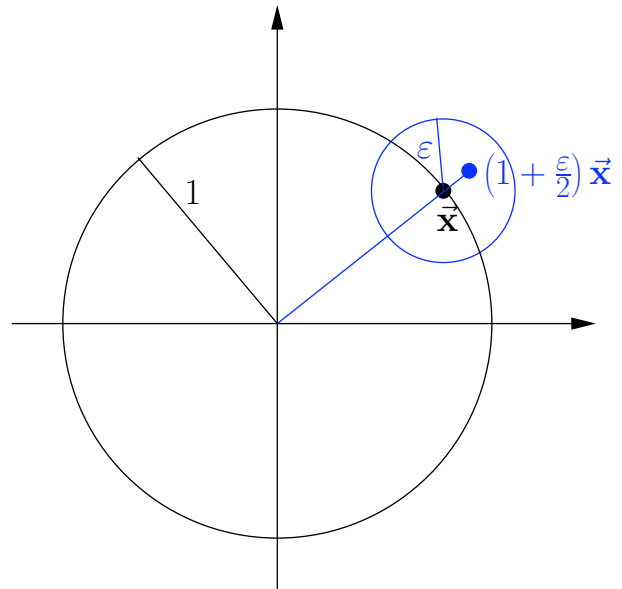
$$\begin{aligned} \|\vec{y}\| &= \left\| \left(1 + \frac{\varepsilon}{2}\right) \vec{x} \right\| \\ &= \left(1 + \frac{\varepsilon}{2}\right) \underbrace{\|\vec{x}\|}_{=1} \\ &= 1 + \frac{\varepsilon}{2} > 1, \end{aligned}$$

d.h. \vec{y} gehört weder zu M_1 noch zu M_2 . Andererseits gilt aber

$$\|\vec{y} - \vec{x}\| = \left\| \left(1 + \frac{\varepsilon}{2}\right) \vec{x} - \vec{x} \right\| = \left\| \frac{\varepsilon}{2} \vec{x} \right\| = \frac{\varepsilon}{2} \underbrace{\|\vec{x}\|}_{=1} = \frac{\varepsilon}{2} < \varepsilon,$$

d.h. \vec{y} gehört zu $U_\varepsilon(\vec{x})$.

Also ist $U_\varepsilon(\vec{x})$ weder eine Teilmenge von M_1 noch eine Teilmenge von M_2 . Da $\varepsilon > 0$ beliebig war, haben wir gezeigt, dass es für kein \vec{x} mit $\|\vec{x}\| = 1$ (aus M_1 und M_2) eine offene Kugel $U_\varepsilon(\vec{x})$ gibt mit $U_\varepsilon(\vec{x}) \subseteq M_1$ oder $U_\varepsilon(\vec{x}) \subseteq M_2$. Also ist weder M_1 noch M_2 offen.



Definition 16.4. (konvergente Folge in \mathbb{R}^n)

(1) Eine Folge $(\vec{x}_{(k)})$ in \mathbb{R}^n heißt **konvergent gegen den Grenzwert** $\vec{x} \in \mathbb{R}^n$, falls gilt $\|\vec{x}_{(k)} - \vec{x}\| \xrightarrow{k \rightarrow \infty} 0$.

Schreibweise: $\vec{x}_{(k)} \xrightarrow{k \rightarrow \infty} \vec{x}$ oder $\lim_{k \rightarrow \infty} \vec{x}_{(k)} = \vec{x}$.

(2) Ist eine Folge $(\vec{x}_{(k)})$ in \mathbb{R}^n **nicht konvergent gegen (irgendein) $\vec{x} \in \mathbb{R}^n$** , so nennt man die Folge **divergent**.

Der nächste Hilfssatz führt die Konvergenz von Folgen im \mathbb{R}^n auf die Konvergenz von Folgen in \mathbb{R} zurück.

Hilfssatz 16.5. (Konvergenz von Folgen in \mathbb{R}^n)

Seien $\vec{\mathbf{x}}_{(k)} = \begin{bmatrix} x_{(k),1} \\ \vdots \\ x_{(k),n} \end{bmatrix}$, $k \in \mathbb{N}$, und $\vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ alle in \mathbb{R}^n .

- (1) Dann gilt $\vec{\mathbf{x}}_{(k)} \xrightarrow{k \rightarrow \infty} \vec{\mathbf{x}}$ genau dann, wenn $x_{(k),i} \xrightarrow{k \rightarrow \infty} x_i$ für alle $i = 1, 2, \dots, n$.
- (2) $(\vec{\mathbf{x}}_{(k)})$ ist **divergent genau dann**, wenn mindestens eine Komponentenfolge $(x_{(k),i})$ divergent ist.

Beweis von Hilfssatz 16.5: Nach der Definition 16.4 gilt:

$$\begin{aligned} \vec{\mathbf{x}}_{(k)} \xrightarrow{k \rightarrow \infty} \vec{\mathbf{x}} &\iff \sqrt{(x_{(k),1} - x_1)^2 + \dots + (x_{(k),n} - x_n)^2} \xrightarrow{k \rightarrow \infty} 0 \\ &\iff (x_{(k),1} - x_1)^2 + \dots + (x_{(k),n} - x_n)^2 \xrightarrow{k \rightarrow \infty} 0 \\ &\iff x_{(k),1} \xrightarrow{k \rightarrow \infty} x_1, \quad \dots, \quad x_{(k),n} \xrightarrow{k \rightarrow \infty} x_n \quad \square \end{aligned}$$

Mit Hilfssatz 16.5 können wir über die Komponentenfolgen $(x_{(k),i})$, $i = 1, 2, \dots, n$, einer Folge $(\vec{\mathbf{x}}_{(k)})$ in \mathbb{R}^n die Resultate über Folgen in \mathbb{R} (vgl. Kapitel 5 der HM A) nutzen, um analoge Resultate für Folgen in \mathbb{R}^n zu bekommen. Dieses ist eine ganz wichtige Erkenntnis und wird in der nachfolgenden Bemerkung im Detail erläutert.

Bemerkung 16.6. (Resultate über konvergente Folgen in \mathbb{R}^n)

Hilfssatz 16.5 ermöglicht es, die **Resultate über konvergente Folgen in \mathbb{R} direkt auf konvergente Folgen in \mathbb{R}^n zu übertragen**. Einige dieser Resultate sind hier angegeben:

- (1) Der Grenzwert einer konvergenten Folge $(\vec{\mathbf{x}}_{(k)})$ in \mathbb{R}^n ist **eindeutig bestimmt**.

Erklärung: Ist $\vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ ein Grenzwert der konvergenten Folge $(\vec{\mathbf{x}}_{(k)}) =$

$\left(\begin{bmatrix} x_{(k),1} \\ \vdots \\ x_{(k),n} \end{bmatrix} \right)$ in \mathbb{R}^n , so wissen wir nach Hilfssatz 16.5, dass für jedes

$i = 1, 2, \dots, n$ die i -te Komponente x_i von \vec{x} der Grenzwert der konvergenten i -ten Komponentenfolge $(x_{(k),i})$ von $(\vec{x}_{(k)})$ ist. Dieser Grenzwert x_i ist für jedes $i = 1, 2, \dots, n$ eindeutig bestimmt, und somit ist \vec{x} eindeutig bestimmt.

- (2) **Summe/Differenz konvergenter Folgen ist eine konvergente Folge:** Sind $(\vec{x}_{(k)})$ und $(\vec{y}_{(k)})$ zwei konvergente Folgen in \mathbb{R}^n mit den Grenzwerten \vec{x} bzw. \vec{y} , so sind auch $(\vec{x}_{(k)} + \vec{y}_{(k)})$ und $(\vec{x}_{(k)} - \vec{y}_{(k)})$ konvergente Folgen in \mathbb{R}^n mit den Grenzwerten $\vec{x} + \vec{y}$ bzw. $\vec{x} - \vec{y}$.

Erklärung: Da $(\vec{x}_{(k)})$ und $(\vec{y}_{(k)})$ konvergent sind mit den Grenzwerten \vec{x} bzw. \vec{y} , sind nach Hilfssatz 16.5 für jedes $i = 1, 2, \dots, n$ die Komponentenfolgen $(x_{(k),i})$ und $(y_{(k),i})$ konvergent mit den Grenzwerten x_i bzw. y_i . Die Folgen $(\vec{x}_{(k)} + \vec{y}_{(k)})$ bzw. $(\vec{x}_{(k)} - \vec{y}_{(k)})$ haben die Komponentenfolgen $(x_{(k),i} + y_{(k),i})$, $i = 1, 2, \dots, n$, bzw. $(x_{(k),i} - y_{(k),i})$, $i = 1, 2, \dots, n$, und diese sind als Summe bzw. Differenz der konvergenten Komponentenfolgen $(x_{(k),i})$ und $(y_{(k),i})$, $i = 1, 2, \dots, n$, konvergent mit dem Grenzwerten $x_i + y_i$ bzw. $x_i - y_i$. Nach Hilfssatz 16.5 sind daher $(\vec{x}_{(k)} + \vec{y}_{(k)})$ und $(\vec{x}_{(k)} - \vec{y}_{(k)})$ konvergente Folgen in \mathbb{R}^n mit den Grenzwerten $\vec{x} + \vec{y}$ bzw. $\vec{x} - \vec{y}$.

- (3) Sei $(\vec{x}_{(k)})$ eine konvergente Folge in \mathbb{R}^n mit dem Grenzwert \vec{x} , und sei (a_k) eine konvergente Folge in \mathbb{R} mit dem Grenzwert a . Dann ist die Folge $(a_k \vec{x}_{(k)})$ eine konvergente Folge in \mathbb{R}^n mit dem Grenzwert $a \vec{x}$. Sind die a_k und der Grenzwert a alle ungleich Null, so ist die Folge $(\vec{x}_{(k)}/a_k)$ eine konvergente Folge in \mathbb{R}^n mit dem Grenzwert \vec{x}/a .

Erklärung: Nach Hilfssatz 16.5 ist jede Komponentenfolge $(x_{(k),i})$, $i = 1, 2, \dots, n$, von $(\vec{x}_{(k)})$ konvergent mit dem Grenzwert x_i , $i = 1, 2, \dots, n$. Somit sind für jedes $i = 1, 2, \dots, n$ auch $(a_k x_{(k),i})$ und, unter den zusätzlichen Voraussetzungen an (a_k) , auch $(x_{(k),i}/a_k)$ konvergent mit den Grenzwerten $a x_i$ bzw. x_i/a . Nach Hilfssatz 16.5 folgt, dass $(a_k \vec{x}_{(k)})$ eine konvergente Folge in \mathbb{R}^n ist mit dem Grenzwert $a \vec{x}$ und, unter den zusätzlichen Voraussetzungen, dass $(\vec{x}_{(k)}/a_k)$ eine konvergente Folge in \mathbb{R}^n ist mit dem Grenzwert \vec{x}/a .

- (4) Ist eine Folge $(\vec{x}_{(k)})$ in \mathbb{R}^n **konvergent**, so ist sie **beschränkt**, d.h. es gibt ein $R > 0$, so dass $\|\vec{x}_{(k)}\| \leq R$ für alle $k \in \mathbb{N}$.

Umgekehrt: Ist eine Folge $(\vec{x}_{(k)})$ **unbeschränkt** (d.h. wenn die $\|\vec{x}_{(k)}\|$ beliebig groß werden können), dann ist die Folge $(\vec{x}_{(k)})$ **divergent**.

Erklärung: Ist $(\vec{x}_{(k)})$ eine konvergente Folge in \mathbb{R}^n mit Grenzwert \vec{x} , so ist nach Hilfssatz 16.5 jede Komponentenfolge $(x_{(k),i})$, $i = 1, 2, \dots, n$, konvergent und somit insbesondere eine beschränkte Folge. Also gibt es

für jedes $i = 1, 2, \dots, n$ eine Konstante $R_i > 0$ mit

$$|(x_{(k),i})| \leq R_i \quad \text{für alle } k \in \mathbb{N}.$$

Daraus folgt für alle $k \in \mathbb{N}$

$$\|\overrightarrow{\mathbf{x}_{(k)}}\| = \sqrt{x_{(k),1}^2 + x_{(k),2}^2 + \dots + x_{(k),n}^2} \leq \underbrace{\sqrt{R_1^2 + R_2^2 + \dots + R_n^2}}_{=: R},$$

d.h. $(\overrightarrow{\mathbf{x}_{(k)}})$ ist beschränkt. – Die zweite Aussage folgt direkt aus der ersten, denn gäbe es konvergente unbeschränkte Folgen, so wäre dieses ein Widerspruch zu der ersten Aussage.

- (5) Eine Folge $(\overrightarrow{\mathbf{x}_{(k)}})$ in \mathbb{R}^n ist genau dann eine **Nullfolge** (also hat den Grenzwert $\vec{\mathbf{0}}$), wenn jede der Komponentenfolgen $(x_{(k),i})$, $i = 1, 2, \dots, n$, eine Nullfolge in \mathbb{R} ist.

Erklärung: Dieses folgt direkt aus Hilfssatz 16.5 für den Sonderfall, dass der Grenzwert $\vec{\mathbf{x}} = \vec{\mathbf{0}}$ ist.

Betrachten wir einige Beispiele konvergenter Folgen, die wir mit den Überlegungen aus Bemerkung 16.6 nun leicht analysieren können.

Beispiel 16.7. (konvergente Folgen)

In den nachfolgenden Beispielen nutzen wir jeweils Hilfssatz 16.5, um die Konvergenz bzw. die Divergenz der Folge zu bestimmen.

(a) Die Folge $(\overrightarrow{\mathbf{x}_{(k)}})_{k \geq 1} = \left(\begin{bmatrix} \frac{1}{k^2} \\ -3 + \frac{1}{k} \\ 2^{-k} + 1 \end{bmatrix} \right)_{k \geq 1}$ ist konvergent gegen $\begin{bmatrix} 0 \\ -3 \\ 1 \end{bmatrix}$, also

$$\overrightarrow{\mathbf{x}_{(k)}} = \begin{bmatrix} \frac{1}{k^2} \\ -3 + \frac{1}{k} \\ 2^{-k} + 1 \end{bmatrix} \xrightarrow{k \rightarrow \infty} \begin{bmatrix} 0 \\ -3 \\ 1 \end{bmatrix}.$$

Begründung: Die Komponentenfolgen sind alle konvergent mit den Grenzwerten

$$\lim_{k \rightarrow \infty} x_{(k),1} = \lim_{k \rightarrow \infty} \frac{1}{k^2} = 0,$$

$$\lim_{k \rightarrow \infty} x_{(k),2} = \lim_{k \rightarrow \infty} \left(-3 + \frac{1}{k} \right) = -3 + 0 = -3,$$

$$\lim_{k \rightarrow \infty} x_{(k),3} = \lim_{k \rightarrow \infty} (2^{-k} + 1) = \lim_{k \rightarrow \infty} \left(\frac{1}{2^k} + 1 \right) = 0 + 1 = 1.$$

Mit Hilfssatz 16.5 folgt somit die Konvergenz von

$$(\overrightarrow{\mathbf{x}}_{(k)})_{k \geq 1} = \left(\begin{bmatrix} \frac{1}{k^2} \\ -3 + \frac{1}{k} \\ 2^{-k} + 1 \end{bmatrix} \right)_{k \geq 1} \quad \text{gegen} \quad \begin{bmatrix} 0 \\ -3 \\ 1 \end{bmatrix}.$$

(b) Die Folge $(\overrightarrow{\mathbf{x}}_{(k)})_{k \geq 1} = \left(\begin{bmatrix} \sin(k\pi) \\ \frac{1}{k} \\ e^{-k} \end{bmatrix} \right)_{k \geq 1}$ ist konvergent gegen $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, also

$$\lim_{k \rightarrow \infty} \overrightarrow{\mathbf{x}}_{(k)} = \lim_{k \rightarrow \infty} \begin{bmatrix} \sin(k\pi) \\ \frac{1}{k} \\ e^{-k} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Begründung: Die Komponentenfolgen sind alle konvergent mit den Grenzwerten

$$\lim_{k \rightarrow \infty} x_{(k),1} = \lim_{k \rightarrow \infty} \sin(k\pi) = \lim_{k \rightarrow \infty} 0 = 0,$$

$$\lim_{k \rightarrow \infty} x_{(k),2} = \lim_{k \rightarrow \infty} \frac{1}{k} = 0,$$

$$\lim_{k \rightarrow \infty} x_{(k),3} = \lim_{k \rightarrow \infty} e^{-k} = 0.$$

Mit Hilfssatz 16.5 folgt somit die Konvergenz von

$$(\overrightarrow{\mathbf{x}}_{(k)})_{k \geq 1} = \left(\begin{bmatrix} \sin(k\pi) \\ \frac{1}{k} \\ e^{-k} \end{bmatrix} \right)_{k \geq 1} \quad \text{gegen} \quad \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

(c) Die Folge $(\overrightarrow{\mathbf{x}}_{(k)})_{k \geq 1} = \left(\begin{bmatrix} k \\ 17 \\ \frac{1}{k} \end{bmatrix} \right)_{k \geq 1}$ ist divergent, denn die Komponentenfolge

$(x_{(k),1})_{k \geq 1} = (k)_{k \geq 1}$ ist unbeschränkt und somit divergent.

Definition 16.8. (Randpunkte und Rand einer Menge im \mathbb{R}^n)

Sei $M \subseteq \mathbb{R}^n$.

(1) $\vec{x} \in \mathbb{R}^n$ heißt ein **Randpunkt von M** , wenn für jedes $\varepsilon > 0$ gilt:

(i) $U_\varepsilon(\vec{x}) \cap M \neq \emptyset$, und (ii) $U_\varepsilon(\vec{x}) \cap (\mathbb{R}^n \setminus M) \neq \emptyset$.

In Worten: Jede offene Kugel $U_\varepsilon(\vec{x})$ um einen Randpunkt \vec{x} enthält sowohl mindestens einen Punkt aus M also auch mindestens einen Punkt aus $\mathbb{R}^n \setminus M$.

(2) Die Menge der Randpunkte von M heißt der **Rand von M** und wird mit ∂M bezeichnet.

(3) $\overline{M} := M \cup \partial M$ heißt der **Abschluss von M** .

Die typische Situation eines Randpunktes einer Menge M ist im rechten Bild in Abbildung 16.1 gezeichnet. Zu beachten ist aber, dass dieses Bild nicht immer (sondern nur im Normalfall) das Richtige ist. Wir werden dieses noch an den Beispielen unten sehen.

Betrachten wir einige Beispiele.

Beispiel 16.9. (Randpunkte und Rand einer Menge im \mathbb{R}^n)

(a) Die offene Kugel $U_\varepsilon(\vec{a})$ um \vec{a} mit Radius ε hat den Rand

$$\partial U_\varepsilon(\vec{a}) = \{ \vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{a}\| = \varepsilon \}$$

und den Abschluss

$$\overline{U_\varepsilon(\vec{a})} = U_\varepsilon(\vec{a}) \cup \partial U_\varepsilon(\vec{a}) = \{ \vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{a}\| \leq \varepsilon \}.$$

Begründung: Dieses zeigen wir weiter unten in Beispiel 16.11.

(b) \mathbb{R}^n hat den Rand $\partial \mathbb{R}^n = \emptyset$ und den Abschluss $\overline{\mathbb{R}^n} = \mathbb{R}^n$.

Begründung: Für jeden Punkt $\vec{x} \in \mathbb{R}^n$ ist die offene Kugel $U_1(\vec{x})$ ganz in \mathbb{R}^n . Also kann kein Punkt \vec{x} in \mathbb{R}^n ein Randpunkt sein, d.h. $\partial \mathbb{R}^n = \emptyset$. Damit folgt direkt $\overline{\mathbb{R}^n} = \mathbb{R}^n \cup \partial \mathbb{R}^n = \mathbb{R}^n \cup \emptyset = \mathbb{R}^n$.

(c) $\Omega := \mathbb{R}^n \setminus \{\vec{0}\}$ hat den Rand $\partial \Omega = \{\vec{0}\}$ und den Abschluss $\overline{\Omega} = \mathbb{R}^n$.

Begründung: Wir wissen aus Beispiel 16.3 (d) bereits, dass $\mathbb{R}^n \setminus \{\vec{0}\}$ offen ist. Also gibt es für jedes $\vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\}$ eine offene Kugel $U_\varepsilon(\vec{x})$ mit $U_\varepsilon(\vec{x}) \subseteq \mathbb{R}^n \setminus \{\vec{0}\}$. Damit kann $\vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\}$ kein Randpunkt von $\mathbb{R}^n \setminus \{\vec{0}\}$ sein. – Betrachten wir nun noch den Punkt $\vec{0}$. Für jede offene Kugel $U_\varepsilon(\vec{0})$ gilt

$$\{\vec{0}\} = U_\varepsilon(\vec{0}) \cap \{\vec{0}\} = U_\varepsilon(\vec{0}) \cap (\mathbb{R}^n \setminus (\mathbb{R}^n \setminus \{\vec{0}\})),$$

$$\{\vec{x} \in \mathbb{R}^n : 0 < \|\vec{x}\| < \varepsilon\} = U_\varepsilon(\vec{0}) \cap (\mathbb{R}^n \setminus \{\vec{0}\}),$$

und somit ist $\vec{0}$ ein Randpunkt von $\mathbb{R}^n \setminus \{\vec{0}\}$. Also gilt $\partial(\mathbb{R}^n \setminus \{\vec{0}\}) = \{\vec{0}\}$ und

$$\overline{\mathbb{R}^n \setminus \{\vec{0}\}} = (\mathbb{R}^n \setminus \{\vec{0}\}) \cup \partial(\mathbb{R}^n \setminus \{\vec{0}\}) = (\mathbb{R}^n \setminus \{\vec{0}\}) \cup \{\vec{0}\} = \mathbb{R}^n.$$

(d) Das offene Quadrat

$$\begin{aligned} Q &:= \{\vec{x} \in \mathbb{R}^2 : \max\{|x_1|, |x_2|\} < 1\} \\ &= \{\vec{x} \in \mathbb{R}^2 : -1 < x_i < 1 \text{ für } i = 1, 2\} \end{aligned}$$

mit Ecken in $(-1, 1)$, $(-1, -1)$, $(1, -1)$, $(1, 1)$ und Kantenlänge 2 hat den Rand

$$\partial Q = \{\vec{x} \in \mathbb{R}^2 : \max\{|x_1|, |x_2|\} = 1\}$$

und dem Abschluss

$$\overline{Q} = Q \cup \partial Q = \{\vec{x} \in \mathbb{R}^2 : \max\{|x_1|, |x_2|\} \leq 1\}.$$

Begründung: Dieses Beispiel untersuchen in einer Übungsaufgabe.

Hilfssatz 16.10. $(\vec{x} \in \overline{M} \iff \text{es gibt } (\overline{\mathbf{x}}_{(k)}) \text{ in } M \text{ mit } \overline{\mathbf{x}}_{(k)} \xrightarrow{k \rightarrow \infty} \vec{x})$

Sei $M \subseteq \mathbb{R}^n$. Es gilt $\vec{x} \in \overline{M}$ genau dann, wenn eine Folge in M existiert, die gegen \vec{x} konvergiert.

Wir zeigen diesen Beweis, weil die Beweisidee auch das Konzept eines Randpunktes klarer macht.

Beweis von Hilfssatz 16.10:

Wir zeigen zunächst die Hinrichtung: „ $\vec{x} \in \overline{M} \implies$ Es existiert eine Folge in M , die gegen \vec{x} konvergiert.“

Beweis der Hinrichtung: Ist $\vec{x} \in \overline{M} = M \cup \partial M$, so können zwei Fälle auftreten: $\vec{x} \in M$ oder $\vec{x} \in \overline{M} \setminus M = (M \cup \partial M) \setminus M = \partial M \setminus M$. Wir betrachten diese beiden Fälle separat.

Fall 1: Der Punkt \vec{x} liegt in M .

Wähle $\overline{\mathbf{x}}_{(k)} := \vec{x}$, $k \in \mathbb{N}$. Dann ist die Folge $(\overline{\mathbf{x}}_{(k)}) = (\vec{x})$ konstant und für alle k gleich dem Punkt \vec{x} . Weil $\vec{x} \in M$ gilt, liegt die konstante Folge $(\overline{\mathbf{x}}_{(k)}) = (\vec{x})$ in M , und sie konvergiert offensichtlich gegen \vec{x} .

Fall 2: Der Punkt \vec{x} liegt nicht in M , d.h. $\vec{x} \in \partial M \setminus M$.

Wähle dann $\vec{x}_{(k)} \in U_{\frac{1}{k}}(\vec{x}) \cap M$. Wir wissen, dass diese Mengen jeweils nicht leer sind, weil \vec{x} ein Randpunkt von M ist. Dann gilt für jedes $k \in \mathbb{N}$, dass $\vec{x}_{(k)} \in M$ und

$$\|\vec{x} - \vec{x}_{(k)}\| < \frac{1}{k}, \quad \text{weil } \vec{x}_{(k)} \in U_{\frac{1}{k}}(\vec{x}),$$

und es folgt $\vec{x}_{(k)} \xrightarrow{k \rightarrow \infty} \vec{x}$.

Damit haben wir die Hinrichtung gezeigt.

Die Rückrichtung „Es existiert eine Folge in M , die gegen \vec{x} konvergiert. $\implies \vec{x} \in \overline{M}$ “ ist äquivalent zu ihrer Kontraposition „ $\vec{x} \notin \overline{M} \implies$ Es existiert keine Folge in M , die gegen \vec{x} konvergiert.“ Wir zeigen diese äquivalente Aussage:.

Beweis der Kontraposition der Rückrichtung: Sei $\vec{x} \notin \overline{M}$, d.h. $\vec{x} \in \mathbb{R}^n \setminus \overline{M}$.

Dann gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(\vec{x}) \cap \overline{M} = \emptyset$, denn: $U_\varepsilon(\vec{x}) \cap (\mathbb{R}^n \setminus \overline{M})$ enthält immer \vec{x} , und würde für jedes $\varepsilon > 0$ gelten, dass $U_\varepsilon(\vec{x}) \cap \overline{M} \neq \emptyset$, so wäre $\vec{x} \in \partial M$. Letzteres ist aber ein Widerspruch zu $\vec{x} \notin \overline{M}$.

Aus $U_\varepsilon(\vec{x}) \cap \overline{M} = \emptyset$ folgt, dass für alle $\vec{y} \in \overline{M}$ gilt $\vec{y} \in \mathbb{R}^n \setminus U_\varepsilon(\vec{x})$, d.h. $\|\vec{y} - \vec{x}\| \geq \varepsilon$. Also kann es keine Folge in \overline{M} geben, die gegen \vec{x} konvergiert. \square

Wir kommen nun auf Beispiel 16.9 (a) zurück.

Beispiel 16.11. (Abschluss der offenen Kugel $U_\varepsilon(\vec{a})$)

Die offene Kugel $U_\varepsilon(\vec{a})$ um \vec{a} mit Radius ε hat den Rand

$$\partial U_\varepsilon(\vec{a}) = \{ \vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{a}\| = \varepsilon \}$$

und den Abschluss

$$\overline{U_\varepsilon(\vec{a})} = U_\varepsilon(\vec{a}) \cup \partial U_\varepsilon(\vec{a}) = \{ \vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{a}\| \leq \varepsilon \}.$$

Nachweis: Wir bestimmen zunächst den Abschluss $\overline{U_\varepsilon(\vec{a})}$ mit Hilfe von Hilfssatz 16.10.

Sei $\vec{x} \in \{ \vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{a}\| \leq \varepsilon \}$.

Gilt $\|\vec{x} - \vec{a}\| < \varepsilon$, so ist $\vec{x} \in U_\varepsilon(\vec{a})$ und die konstante Folge (\vec{x}) konvergiert gegen \vec{x} . Also ist $\vec{x} \in \overline{U_\varepsilon(\vec{a})}$ nach Hilfssatz 16.10.

Gilt $\|\vec{x} - \vec{a}\| = \varepsilon$, so wählen wir die Folge $(\vec{a} + (1 - \frac{1}{k}) \cdot (\vec{x} - \vec{a}))$. Wegen

$$\lim_{k \rightarrow \infty} \left(\vec{a} + \left(1 - \frac{1}{k}\right) \cdot (\vec{x} - \vec{a}) \right) = \vec{a} + 1 \cdot (\vec{x} - \vec{a}) = \vec{x}$$

konvergiert diese Folge gegen \vec{x} , und wegen

$$\begin{aligned} \left\| \left(\vec{a} + \left(1 - \frac{1}{k} \right) \cdot (\vec{x} - \vec{a}) \right) - \vec{a} \right\| &= \left\| \left(1 - \frac{1}{k} \right) \cdot (\vec{x} - \vec{a}) \right\| \\ &= \left(1 - \frac{1}{k} \right) \cdot \|\vec{x} - \vec{a}\| = \left(1 - \frac{1}{k} \right) \cdot \varepsilon < \varepsilon \end{aligned}$$

ist diese Folge eine Folge in $U_\varepsilon(\vec{a})$. Also ist $\vec{x} \in \overline{U_\varepsilon(\vec{a})}$ nach Hilfssatz 16.10.

Damit haben wir gezeigt, dass

$$\{\vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{a}\| \leq \varepsilon\} \subseteq \overline{U_\varepsilon(\vec{a})}$$

gilt. Wenn wir zeigen, dass alle $\vec{x} \in \mathbb{R}^n \setminus \{\vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{a}\| \leq \varepsilon\}$ nicht in $\overline{U_\varepsilon(\vec{a})}$ liegen, dann folgt

$$\{\vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{a}\| \leq \varepsilon\} = \overline{U_\varepsilon(\vec{a})}.$$

Sei also $\vec{x} \in \mathbb{R}^n \setminus \{\vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{a}\| \leq \varepsilon\}$, d.h. $\|\vec{x} - \vec{a}\| > \varepsilon$. Dann gilt für $U_\delta(\vec{x})$ mit $\delta := \|\vec{x} - \vec{a}\| - \varepsilon$, dass $U_\delta(\vec{x}) \cap U_\varepsilon(\vec{a}) = \emptyset$, denn:

Sei $\vec{y} \in U_\delta(\vec{x})$ beliebig. Dann gilt $\|\vec{y} - \vec{x}\| < \delta = \|\vec{x} - \vec{a}\| - \varepsilon$, und mit der unteren Dreiecksungleichung folgt

$$\begin{aligned} \|\vec{y} - \vec{a}\| &= \|(\vec{x} - \vec{a}) + (\vec{y} - \vec{x})\| \geq \|\vec{x} - \vec{a}\| - \|\vec{y} - \vec{x}\| \\ &\geq \|\vec{x} - \vec{a}\| - \underbrace{\|\vec{y} - \vec{x}\|}_{< \delta = \|\vec{x} - \vec{a}\| - \varepsilon} > \|\vec{x} - \vec{a}\| - (\|\vec{x} - \vec{a}\| - \varepsilon) = \varepsilon, \end{aligned}$$

also $\|\vec{y} - \vec{a}\| > \varepsilon$, d.h. $\vec{y} \notin U_\varepsilon(\vec{a})$.

Aus $U_\delta(\vec{x}) \cap U_\varepsilon(\vec{a}) = \emptyset$ folgt aber, dass \vec{x} kein Randpunkt von $U_\varepsilon(\vec{a})$ sein kann, und \vec{x} ist auch nicht in $U_\varepsilon(\vec{a})$. Also folgt, dass $\vec{x} \notin \overline{U_\varepsilon(\vec{a})}$ ist.

Wir wissen nun, dass

$$\overline{U_\varepsilon(\vec{a})} = \{\vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{a}\| \leq \varepsilon\}.$$

Weiter gilt

$$U_\varepsilon(\vec{a}) \cup \partial U_\varepsilon(\vec{a}) = \overline{U_\varepsilon(\vec{a})}$$

und

$$\underbrace{\{\vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{a}\| < \varepsilon\} \cup \{\vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{a}\| = \varepsilon\}}_{= U_\varepsilon(\vec{a})}$$

$$= \underbrace{\{\vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{a}\| \leq \varepsilon\}}_{= U_\varepsilon(\vec{a})}.$$

Da $U_\varepsilon(\vec{a})$ offen ist, kann $U_\varepsilon(\vec{a})$ keine Randpunkte enthalten. Also schließen wir, dass

$$\partial U_\varepsilon(\vec{a}) = \{\vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{a}\| = \varepsilon\}$$

gilt. □

Als Letztes wollen wir noch eine sehr nützliche Notation einführen.

Definition 16.12. (kartesisches Produkt)

Seien A, B, C Mengen.

(1) Das **kartesische Produkt** $A \times B$ von A und B ist die Menge

$$A \times B := \{(a, b) : a \in A, b \in B\}.$$

(2) Das **kartesische Produkt** $A \times B \times C$ von A, B und C ist die Menge

$$A \times B \times C := \{(a, b, c) : a \in A, b \in B, c \in C\}.$$

(3) Analog kann man für n Mengen A_1, A_2, \dots, A_n das **kartesische Produkt** $A_1 \times A_2 \times \dots \times A_n$ durch

$$A_1 \times A_2 \times \dots \times A_n := \{(a_1, a_2, \dots, a_n) : a_i \in A_i \text{ für } i = 1, 2, \dots, n\}$$

definieren.

(Achtung: Beim kartesischen Produkt spielt die **Reihenfolge der Mengen** eine Rolle. In der Regel gilt $A \times B \neq B \times A$.)

Betrachten wir einige Beispiele.

Beispiel 16.13. (kartesisches Produkt)

(a) Seien $A := \{1, 2, 3\}$ und $B := \{7, 8\}$. Dann sind

$$A \times B = \{(1, 7), (1, 8), (2, 7), (2, 8), (3, 7), (3, 8)\},$$

$$B \times A = \{(7, 1), (7, 2), (7, 3), (8, 1), (8, 2), (8, 3)\}.$$

In dem Beispiel gilt $A \times B \neq B \times A$, d.h. das kartesische Produkt ist **nicht** kommutativ.

(b) $\mathbb{R}^2 := \mathbb{R} \times \mathbb{R}$ ist die Menge aller Punkte in der Ebene, denn

$$\mathbb{R} \times \mathbb{R} = \{(x, y) : x \in \mathbb{R}, y \in \mathbb{R}\}.$$

(c) $\mathbb{R}^3 := \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ ist der dreidimensionale Anschauungsraum, denn

$$\mathbb{R} \times \mathbb{R} \times \mathbb{R} = \{(x, y, z) : x, y, z \in \mathbb{R}\}.$$

(d) $[0, 1] \times [0, 1] = \{(x, y) : x \in [0, 1], y \in [0, 1]\}$

ist das Quadrat in der Ebene mit Seitenlänge 1 und dem Eckpunkten $(0, 0)$, $(1, 0)$, $(1, 1)$ und $(0, 1)$.

(e) $\mathbb{R} \times [0, \infty[$ ist die obere Halbebene von \mathbb{R}^2 mit Rand, denn

$$\mathbb{R} \times [0, \infty[= \{(x, y) : x, y \in \mathbb{R} \text{ mit } y \geq 0\}.$$

(f) $W := [-1, 1] \times [-1, 1] \times [-1, 1]$

ist der Würfel mit Seitenlänge 2 und „Zentrum“ in $(0, 0, 0)$.

16.2 Konvergenz und Stetigkeit für Funktionen

Wir starten, indem wir Funktionen mit mehreren Variablen und mit Werten in \mathbb{R}^m (also mit mehreren Komponentenfunktionen) einführen. Ein physikalisches Beispiel für solche Funktionen sind orts- und zeitabhängige Vektorfelder, die z.B. die magnetische Feldstärke (in den Punkten des dreidimensionalen Raums als Funktion der Zeit) beschreiben.

Notation 16.14. (vektorwertige Funktionen mehrerer Variablen)

Seien $n, m \in \mathbb{N}$, und sei $D \subseteq \mathbb{R}^n$.

(1) Wir betrachten Funktionen $\mathbf{f} : D \rightarrow \mathbb{R}^m$, genauer

$$\vec{\mathbf{f}}(\vec{\mathbf{x}}) = \vec{\mathbf{f}}(x_1, x_2, \dots, x_n) = \begin{bmatrix} f_1(x_1, x_2, \dots, x_n) \\ f_2(x_1, x_2, \dots, x_n) \\ \vdots \\ f_m(x_1, x_2, \dots, x_n) \end{bmatrix} = \begin{bmatrix} f_1(\vec{\mathbf{x}}) \\ f_2(\vec{\mathbf{x}}) \\ \vdots \\ f_m(\vec{\mathbf{x}}) \end{bmatrix}.$$

Dabei bezeichnen wir die reellwertigen Funktionen $f_i(\vec{\mathbf{x}}) = f_i(x_1, x_2, \dots, x_n)$ als die **Komponentenfunktionen von $\vec{\mathbf{f}}$** .

Hier identifizieren wir in $\vec{f}(\vec{x}) = \vec{f}(x_1, x_2, \dots, x_n)$ den Spaltenvektor

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \text{ und den Punkt } (x_1, \dots, x_n) \text{ in } \mathbb{R}^n, \text{ weil es keine Rolle spielt,}$$

ob wir uns die n Variablen x_1, \dots, x_n der Funktion $\vec{f} : D \rightarrow \mathbb{R}^m$ als Spaltenvektor in \mathbb{R}^n oder als Punkt im \mathbb{R}^n angeordnet vorstellen. $\vec{f}(\vec{x})$ stellen wir dagegen immer als Spaltenvektor dar.

- (2) Falls $n = 2$ ist, schreiben wir oft $f(x, y)$ statt $f(x_1, x_2)$, bzw. $\vec{f}(x, y)$ statt $\vec{f}(x_1, x_2)$.
- (3) Falls $n = 3$ ist, schreiben wir oft $f(x, y, z)$ statt $f(x_1, x_2, x_3)$, bzw. $\vec{f}(x, y, z)$ statt $\vec{f}(x_1, x_2, x_3)$.

Betrachten wir zwei Beispiele.

Beispiel 16.15. (vektorierte Funktionen mehrerer Variablen)

(a) $D = \mathbb{R}^2 \setminus \{(0, 0)\}$, $f : D \rightarrow \mathbb{R}$, $f(x, y) := \frac{2xy}{x^2 + y^2}$

(b) $D = \mathbb{R} \times [0, \infty[$, $\vec{f} : D \rightarrow \mathbb{R}^3$, $\vec{f}(x, y) := \begin{bmatrix} x + \sqrt{y} \\ e^x - y \\ \sin(x\sqrt{y}) \end{bmatrix}$

Wir verallgemeinern nun den Begriff der Konvergenz einer Funktion $f(x)$ für $x \rightarrow a$ (vgl. Kapitel 6 im Skript der HM A) auf vektorwertige Funktionen mehrerer Variablen.

Definition 16.16. (Konvergenz von \vec{f} für \vec{x} gegen \vec{a})

Sei $\Omega \subseteq \mathbb{R}^n$ offen und $D \subseteq \mathbb{R}^n$ so, dass $\Omega \subseteq D \subseteq \bar{\Omega}$, und sei $\vec{f} : D \rightarrow \mathbb{R}^m$.

Sei $\vec{a} \in \bar{\Omega}$. \vec{f} konvergiert für $\vec{x} \rightarrow \vec{a}$ gegen $\vec{b} \in \mathbb{R}^m$, falls für jede Folge $(\vec{x}_{(k)})$ in $D \setminus \{\vec{a}\}$ mit $\vec{x}_{(k)} \xrightarrow{k \rightarrow \infty} \vec{a}$ gilt:

$$\vec{f}(\vec{x}_{(k)}) \xrightarrow{k \rightarrow \infty} \vec{b}.$$

Schreibweise: $\lim_{\vec{x} \rightarrow \vec{a}} \vec{f}(\vec{x}) = \vec{b}$

(Anmerkung: Die Voraussetzung $\Omega \subseteq D \subseteq \bar{\Omega}$ an $D \subseteq \mathbb{R}^n$ und die offene Menge $\Omega \subseteq \mathbb{R}^n$ bedeuten, dass D gerade Ω selber oder Ω vereinigt mit einigen oder vereinigt mit allen Randpunkten von Ω ist. Diese Voraussetzung garantiert, dass jedes $\vec{a} \in \bar{\Omega}$ der Grenzwert einer Folge in $(\vec{x}_{(k)})$ in $D \setminus \{\vec{a}\}$ ist.)

Mit unserem Verständnis von Folgen in \mathbb{R}^m erhalten wir direkt den folgenden Hilfssatz.

Hilfssatz 16.17. (Konvergenz von \vec{f} für \vec{x} gegen \vec{a} mittels Komponentenfunktionen)

Sei $\Omega \subseteq \mathbb{R}^n$ offen und $D \subseteq \mathbb{R}^n$ so, dass $\Omega \subseteq D \subseteq \bar{\Omega}$, und sei $\vec{f} : D \rightarrow \mathbb{R}^m$. Sei $\vec{a} \in \bar{\Omega}$. \vec{f} konvergiert für $\vec{x} \rightarrow \vec{a}$ gegen $\vec{b} \in \mathbb{R}^m$ genau dann, wenn die Komponentenfunktionen f_1, f_2, \dots, f_m von \vec{f} für $\vec{x} \rightarrow \vec{a}$ gegen jeweils gegen die Komponenten b_1, b_2, \dots , bzw. b_m von \vec{b} konvergieren.

Betrachten wir drei Beispiele.

Beispiel 16.18. (Konvergenz von \vec{f} für \vec{x} gegen \vec{a})

(a) $D = \Omega = \mathbb{R}^2 \setminus \{(0, 0)\}$, $f : D \rightarrow \mathbb{R}$, $f(x, y) := \frac{2xy}{x^2 + y^2}$, $(a_1, a_2) = (0, 0)$.

Für $\left(\frac{1}{k}, 0\right) \xrightarrow{k \rightarrow \infty} (0, 0)$: $f\left(\frac{1}{k}, 0\right) = \frac{2 \cdot \frac{1}{k} \cdot 0}{\left(\frac{1}{k}\right)^2 + 0^2} = 0 \xrightarrow{k \rightarrow \infty} 0$,

aber für $\left(\frac{1}{k}, \frac{1}{k}\right) \xrightarrow{k \rightarrow \infty} (0, 0)$: $f\left(\frac{1}{k}, \frac{1}{k}\right) = \frac{2 \cdot \frac{1}{k} \cdot \frac{1}{k}}{\left(\frac{1}{k}\right)^2 + \left(\frac{1}{k}\right)^2} = \frac{\frac{2}{k^2}}{\frac{2}{k^2}} = 1 \xrightarrow{k \rightarrow \infty} 1$.

Also existiert $\lim_{(x,y) \rightarrow (0,0)} f(x, y)$ nicht!

(b) $D = \Omega = \mathbb{R}^2 \setminus \{(0, 0)\}$, $f : D \rightarrow \mathbb{R}$, $f(x, y) := \frac{2xy}{\sqrt{x^2 + y^2}}$, $(a_1, a_2) = (0, 0)$

Es gilt $\lim_{(x,y) \rightarrow (0,0)} f(x, y) = 0$, denn:

Sei $((x_k, y_k))$ mit $(x_k, y_k) \xrightarrow{k \rightarrow \infty} (0, 0)$, d.h. $x_k \xrightarrow{k \rightarrow \infty} 0, y_k \xrightarrow{k \rightarrow \infty} 0$. Dann folgt

$$|f(x_k, y_k)| = \left| \frac{2x_k y_k}{\sqrt{x_k^2 + y_k^2}} \right| = \frac{2|x_k||y_k|}{\|(x_k, y_k)\|} = 2|x_k| \cdot \underbrace{\frac{|y_k|}{\|(x_k, y_k)\|}}_{\leq 1} \leq 2|x_k| \xrightarrow{k \rightarrow \infty} 0,$$

wobei wir ausgenutzt haben, dass gilt

$$|y_k| = \sqrt{y_k^2} \leq \sqrt{x_k^2 + y_k^2} = \|(x_k, y_k)\| \quad \Longrightarrow \quad \frac{|y_k|}{\|(x_k, y_k)\|} \leq 1.$$

$$(c) \quad D = \mathbb{R} \times]0, \infty[, \quad \vec{f} : D \rightarrow \mathbb{R}^3, \quad \vec{f}(x, y) := \begin{bmatrix} x + \sqrt{y} \\ e^x - y \\ \sin(x\sqrt{y}) \end{bmatrix}$$

Hier ist D die obere Halbebene in \mathbb{R}^2 mit Rand und Ω ist die obere Halbebene ohne Rand, also $\Omega = \mathbb{R} \times]0, \infty[$. Es gilt $\bar{\Omega} = D = \mathbb{R} \times [0, \infty[$. Für jedes $(a_1, a_2) \in \mathbb{R}^2 \times [0, \infty[$ gilt nun

$$\lim_{(x,y) \rightarrow (a_1, a_2)} \vec{f}(x, y) = \lim_{(x,y) \rightarrow (a_1, a_2)} \begin{bmatrix} x + \sqrt{y} \\ e^x - y \\ \sin(x\sqrt{y}) \end{bmatrix} = \begin{bmatrix} a_1 + \sqrt{a_2} \\ e^{a_1} - a_2 \\ \sin(a_1\sqrt{a_2}) \end{bmatrix},$$

denn die Grenzwerte existieren für jede der Komponentenfunktionen.

Wir führen nun den Begriff der Stetigkeit für vektorwertige Funktionen mehrerer Variablen ein.

Definition 16.19. (stetige Funktion)

Seien $\Omega \subseteq \mathbb{R}^n$ offen und $D \subseteq \mathbb{R}^n$ so, dass $\Omega \subseteq D \subseteq \bar{\Omega}$, und sei $\vec{f} : D \rightarrow \mathbb{R}^m$.

(1) \vec{f} heißt **stetig in** $\vec{a} \in D$, wenn $\lim_{\vec{x} \rightarrow \vec{a}} \vec{f}(\vec{x})$ existiert und $\lim_{\vec{x} \rightarrow \vec{a}} \vec{f}(\vec{x}) = \vec{f}(\vec{a})$ gilt.

(2) \vec{f} heißt **stetig in** D , wenn \vec{f} in jedem $\vec{a} \in D$ stetig ist.

Auch hier können wir Stetigkeit über die Stetigkeit der Komponentenfunktionen nachweisen.

Hilfssatz 16.20. (stetige Funktion \Leftrightarrow stetige Komponentenfkten.)

Seien $\Omega \subseteq \mathbb{R}^n$ offen und $D \subseteq \mathbb{R}^n$ so, dass $\Omega \subseteq D \subseteq \bar{\Omega}$, und sei $\vec{f} : D \rightarrow \mathbb{R}^m$.

(1) \vec{f} ist **stetig in** $\vec{a} \in D$ **genau dann**, wenn alle Komponentenfunktionen f_1, f_2, \dots, f_m von \vec{f} in \vec{a} stetig sind.

(2) \vec{f} ist **stetig in D genau dann**, wenn alle Komponentenfunktionen f_1, f_2, \dots, f_m von \vec{f} in jedem $\vec{a} \in D$ stetig sind.

Betrachten wir einige Beispiele.

Beispiel 16.21. (stetige Funktion)

$$(a) \quad D = \Omega = \mathbb{R}^2, \quad f : D \rightarrow \mathbb{R}, \quad f(x, y) := \begin{cases} \frac{2xy}{x^2 + y^2} & \text{für } (x, y) \neq (0, 0), \\ 0 & \text{für } (x, y) = (0, 0). \end{cases}$$

f ist nicht stetig in $(0, 0)$, da der Grenzwert $\lim_{(x,y) \rightarrow (0,0)} f(x, y)$ nach Beispiel 16.18 (a) nicht existiert.

$$(b) \quad D = \Omega = \mathbb{R}^2, \quad f : D \rightarrow \mathbb{R}, \quad f(x, y) := \begin{cases} \frac{2xy}{\sqrt{x^2 + y^2}} & \text{für } (x, y) \neq (0, 0), \\ b & \text{für } (x, y) = (0, 0). \end{cases}$$

Da der Grenzwert $\lim_{(x,y) \rightarrow (0,0)} f(x, y) = 0$ nach Beispiel 16.18 (b) existiert, ist f stetig in $(0, 0)$ genau dann, wenn $b = 0$ ist.

Die nächste Bemerkung liefert uns ein nützliches Hilfsmittel, um Funktionen auf Stetigkeit zu untersuchen.

Bemerkung 16.22. (Summen, Produkte, Quotienten und Verkettungen stetiger Funktion)

Ähnlich wie für Funktionen in einer Veränderlichen (siehe Kapitel 6 im Skript der HM A) kann man zeigen, dass Funktionen, deren Komponentenfunktionen durch Summen, Produkte, Quotienten und/oder Verkettung aus stetigen Funktionen zusammengesetzt sind, auf ihrem maximalen Definitionsbereich stetig sind.

Beispiel 16.23. (stetige Funktion)

(a) Sei $D = \Omega := \mathbb{R}^2 \setminus \{(0, 0)\}$.

Die Funktion $f : D \rightarrow \mathbb{R}, \quad f(x, y) := \frac{2xy}{x^2 + y^2}$, ist stetig in D , denn:

f ist durch Summen, Produkte, Quotienten auf $D = \mathbb{R}^2 \setminus \{(0, 0)\}$ stetiger Funktionen gebildet und der Nenner ist auf $D = \mathbb{R}^2 \setminus \{(0, 0)\}$ immer ungleich

Null. *Genauer gilt:* Wir können f darstellen als

$$f(x, y) = \frac{g(x, y) \cdot h(x, y)}{u(x, y) + v(x, y)}$$

mit den Funktionen

$$\begin{aligned} g : \mathbb{R}^2 \setminus \{(0, 0)\} &\rightarrow \mathbb{R}, & g(x, y) &:= 2x, \\ h : \mathbb{R}^2 \setminus \{(0, 0)\} &\rightarrow \mathbb{R}, & h(x, y) &:= y, \\ u : \mathbb{R}^2 \setminus \{(0, 0)\} &\rightarrow \mathbb{R}, & u(x, y) &:= x^2, \\ v : \mathbb{R}^2 \setminus \{(0, 0)\} &\rightarrow \mathbb{R}, & v(x, y) &:= y^2. \end{aligned}$$

Diese Funktionen hängen effektiv jeweils nur von x oder nur von y (aber nicht gleichzeitig von x und y) ab und sind mit unserem Wissen aus Kapitel 6 der HM A stetig auf $\mathbb{R}^2 \setminus \{(0, 0)\}$. Damit sind der Zähler $g(x, y) \cdot h(x, y)$ und der Nenner $u(x, y) + v(x, y)$ als Produkt bzw. als Summe auf $\mathbb{R}^2 \setminus \{(0, 0)\}$ stetiger Funktionen nach Bemerkung 16.22 ebenfalls auf $\mathbb{R}^2 \setminus \{(0, 0)\}$ stetig. Weil der Nenner $u(x, y) + v(x, y) = x^2 + y^2 \neq 0$ für alle $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ ist, ist f als Quotient auf $\mathbb{R}^2 \setminus \{(0, 0)\}$ stetiger Funktionen ebenfalls stetig.

(b) Sei $D = \overline{\Omega} = \mathbb{R} \times [0, \infty[$.

Die Funktion $\vec{f} : D \rightarrow \mathbb{R}^3$, $\vec{f}(x, y) := \begin{bmatrix} x + \sqrt{y} \\ e^x - y \\ \sin(x \sqrt{y}) \end{bmatrix}$, ist stetig in D ,

denn:

Die Komponentenfunktionen sind durch Summen, Produkte und Verkettungen auf $D = \mathbb{R} \times [0, \infty[$ stetiger Funktionen gebildet. Die Begründung im Detail ist analog zu Beispiel (a).

(c) Sei $D = \overline{\Omega} = \mathbb{R}^2$.

Die Funktion $f : D \rightarrow \mathbb{R}$, $f(x, y) := \begin{cases} \frac{2xy}{\sqrt{x^2 + y^2}} & \text{für } (x, y) \neq (0, 0), \\ 0 & \text{für } (x, y) = (0, 0), \end{cases}$

ist stetig in D , denn:

f ist stetig in $(x, y) \neq (0, 0)$, da f durch Summen, Produkte, Quotienten auf $D = \mathbb{R}^2 \setminus \{(0, 0)\}$ stetiger Funktionen gebildet ist und da der Nenner auf $D = \mathbb{R}^2 \setminus \{(0, 0)\}$ immer ungleich Null ist. Die Begründung im Detail ist analog zu Beispiel (a).

f ist auch stetig in $(0, 0)$ nach Beispiel 16.21 (b).

16.3 Partielle Ableitungen

In Verallgemeinerung des Begriffs der Ableitung einer Funktion einer Variablen lernen wir nun den Begriff der partiellen Ableitungen einer reellwertigen Funktion nach den verschiedenen Variablen dieser Funktion kennen.

Definition 16.24. (partielle Ableitungen einer reellwertigen Funktion)

Sei $(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ die Standardbasis des \mathbb{R}^n . Sei $\Omega \subseteq \mathbb{R}^n$ offen, $f : \Omega \rightarrow \mathbb{R}$.

- (1) f heißt in $\vec{a} \in \Omega$ **partiell nach der i -ten Variablen differenzierbar**, falls

$$(\partial_i f)(\vec{a}) := \frac{\partial}{\partial x_i} f(\vec{a}) := \lim_{h \rightarrow 0} \frac{f(\vec{a} + h \vec{e}_i) - f(\vec{a})}{h}$$

existiert. $(\partial_i f)(\vec{a}) = \frac{\partial}{\partial x_i} f(\vec{a})$ heißt dann die **partielle Ableitung von f nach der i -ten Variablen im Punkt \vec{a}** .

- (2) Existieren in $\vec{a} \in \Omega$ alle n partiellen Ableitungen

$$(\partial_1 f)(\vec{a}), \quad (\partial_2 f)(\vec{a}), \quad \dots, \quad (\partial_n f)(\vec{a}),$$

so heißt f **partiell differenzierbar in \vec{a}** .

- (3) f heißt **partiell differenzierbar in Ω** , falls f in jedem $\vec{a} \in \Omega$ partiell differenzierbar ist.

Die nächste Bemerkung gibt einen wichtigen Hinweis für die praktische Berechnung von partiellen Ableitungen.

Bemerkung 16.25. (Berechnung partieller Ableitungen)

Partielle Ableitungen berechnen sich im Prinzip wie die „normalen“ Ableitungen einer reellwertigen Funktion mit einer Variablen. Berechnet man die partielle Ableitung $\partial_i f$ so **differenziert man die Funktion $f(x_1, x_2, \dots, x_n)$ wie aus der HM A gewohnt nach x_i und behandelt dabei alle Variablen außer x_i , also alle x_k mit $k \neq i$, als Konstanten**. Dabei gelten natürlich die üblichen Rechenregeln, insbesondere die Produktregel, die Quotientenregel und die Kettenregel.

Erklärung: Es gilt

$$\begin{aligned} (\partial_i f)(\vec{\mathbf{a}}) &= \lim_{h \rightarrow 0} \frac{f(\vec{\mathbf{a}} + h \vec{\mathbf{e}}_i) - f(\vec{\mathbf{a}})}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(a_1, \dots, a_{i-1}, a_i + h, a_{i+1}, \dots, a_n) - f(a_1, \dots, a_n)}{h}, \end{aligned}$$

d.h. wir können dieses als den üblichen Differentialquotienten der Funktion einer Variable $g(t) := f(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_n)$ in $t = a_i$ interpretieren:

$$(\partial_i f)(\vec{\mathbf{a}}) = \lim_{h \rightarrow 0} \frac{g(a_i + h) - g(a_i)}{h} = g'(a_i).$$

Betrachten wir einige Beispiele.

Beispiel 16.26. (partielle Ableitungen für reellwertige Funktionen)

- (a) $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := x^2 y^3$, ist auf ganz \mathbb{R}^2 nach x bzw. y partiell differenzierbar und hat die partiellen Ableitungen

$$(\partial_1 f)(x, y) = \frac{\partial}{\partial x} f(x, y) = 2x \cdot y^3 = 2xy^3,$$

$$(\partial_2 f)(x, y) = \frac{\partial}{\partial y} f(x, y) = x^2 \cdot 3y^2 = 3x^2 y^2.$$

- (b) $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, $f(x, y, z) := \sin(x^2 + 2y - z^3)$, ist auf ganz \mathbb{R}^3 nach x bzw. y bzw. z partiell differenzierbar und hat die partiellen Ableitungen

$$(\partial_1 f)(x, y, z) = \frac{\partial}{\partial x} f(x, y, z) = \cos(x^2 + 2y - z^3) \cdot 2x,$$

$$(\partial_2 f)(x, y, z) = \frac{\partial}{\partial y} f(x, y, z) = \cos(x^2 + 2y - z^3) \cdot 2,$$

$$(\partial_3 f)(x, y, z) = \frac{\partial}{\partial z} f(x, y, z) = \cos(x^2 + 2y - z^3) \cdot (-3z^2).$$

- (c) $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := \alpha x + \beta y + \gamma$, mit $\alpha, \beta, \gamma \in \mathbb{R}$, ist auf ganz \mathbb{R}^2 nach x bzw. y partiell differenzierbar und hat die partiellen Ableitungen

$$(\partial_1 f)(x, y) = \frac{\partial}{\partial x} f(x, y) = \alpha,$$

$$(\partial_2 f)(x, y) = \frac{\partial}{\partial y} f(x, y) = \beta.$$

- (d) $f : \mathbb{R}^4 \rightarrow \mathbb{R}$, $f(\vec{x}) = f(x_1, x_2, x_3, x_4) := x_1 x_2 x_3 x_4 = x_1 \cdot x_2 \cdot x_3 \cdot x_4$, ist auf ganz \mathbb{R}^4 nach x_i , $i = 1, 2, 3, 4$, partiell differenzierbar und hat die partiellen Ableitungen

$$(\partial_1 f)(x_1, x_2, x_3, x_4) = \frac{\partial}{\partial x_1} f(x_1, x_2, x_3, x_4) = x_2 \cdot x_3 \cdot x_4,$$

$$(\partial_2 f)(x_1, x_2, x_3, x_4) = \frac{\partial}{\partial x_2} f(x_1, x_2, x_3, x_4) = x_1 \cdot x_3 \cdot x_4,$$

$$(\partial_3 f)(x_1, x_2, x_3, x_4) = \frac{\partial}{\partial x_3} f(x_1, x_2, x_3, x_4) = x_1 \cdot x_2 \cdot x_4,$$

$$(\partial_4 f)(x_1, x_2, x_3, x_4) = \frac{\partial}{\partial x_4} f(x_1, x_2, x_3, x_4) = x_1 \cdot x_2 \cdot x_3.$$

Für vektorwertige Funktionen erklären wir die partiellen Ableitungen über die partiellen Ableitungen der Komponentenfunktionen.

Definition 16.27. (partielle Ableitungen für vektorwertige Fkten.)

Sei $\Omega \subseteq \mathbb{R}^n$ offen, und sei $\vec{f} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{bmatrix} : \Omega \rightarrow \mathbb{R}^m$.

- (1) \vec{f} heißt in $\vec{a} \in \Omega$ **partiell nach der i -ten Variablen differenzierbar**, wenn $f_1, f_2, \dots, f_m : \Omega \rightarrow \mathbb{R}$ in $\vec{a} \in \Omega$ partiell nach der i -ten Variablen differenzierbar sind, d.h. wenn

$$(\partial_i f_1)(\vec{a}), \quad (\partial_i f_2)(\vec{a}), \quad \dots, \quad (\partial_i f_m)(\vec{a})$$

existieren.

$$(\partial_i \vec{f})(\vec{a}) := \frac{\partial}{\partial x_i} \vec{f}(\vec{a}) := \begin{bmatrix} (\partial_i f_1)(\vec{a}) \\ (\partial_i f_2)(\vec{a}) \\ \vdots \\ (\partial_i f_m)(\vec{a}) \end{bmatrix} \in \mathbb{R}^m$$

heißt dann die **partielle Ableitung von \vec{f} nach der i -ten Variablen im Punkt \vec{a}** .

- (2) Existieren in $\vec{a} \in \Omega$ alle partiellen Ableitungen

$$(\partial_1 \vec{f})(\vec{a}), \quad (\partial_2 \vec{f})(\vec{a}), \quad \dots, \quad (\partial_n \vec{f})(\vec{a}),$$

so heißt $\vec{\mathbf{f}}$ in $\vec{\mathbf{a}}$ *partiell differenzierbar* und

$$\begin{aligned} (\mathbf{J}\vec{\mathbf{f}})(\vec{\mathbf{a}}) &:= \left[(\partial_1 \vec{\mathbf{f}})(\vec{\mathbf{a}}) \quad (\partial_2 \vec{\mathbf{f}})(\vec{\mathbf{a}}) \quad \cdots \quad (\partial_n \vec{\mathbf{f}})(\vec{\mathbf{a}}) \right] \\ &= \begin{bmatrix} (\partial_1 f_1)(\vec{\mathbf{a}}) & (\partial_2 f_1)(\vec{\mathbf{a}}) & \cdots & (\partial_n f_1)(\vec{\mathbf{a}}) \\ (\partial_1 f_2)(\vec{\mathbf{a}}) & (\partial_2 f_2)(\vec{\mathbf{a}}) & \cdots & (\partial_n f_2)(\vec{\mathbf{a}}) \\ \vdots & \vdots & & \vdots \\ (\partial_1 f_m)(\vec{\mathbf{a}}) & (\partial_2 f_m)(\vec{\mathbf{a}}) & \cdots & (\partial_n f_m)(\vec{\mathbf{a}}) \end{bmatrix} \in \mathbb{R}^{m \times n} \end{aligned}$$

heißt die **Jacobi-Matrix** von $\vec{\mathbf{f}}$ in $\vec{\mathbf{a}}$.

- (3) $\vec{\mathbf{f}}$ heißt *partiell differenzierbar* in Ω , falls $\vec{\mathbf{f}}$ in jedem $\vec{\mathbf{a}} \in \Omega$ partiell differenzierbar ist, d.h. wenn die Jacobi-Matrix $(\mathbf{J}\vec{\mathbf{f}})(\vec{\mathbf{a}})$ in jedem $\vec{\mathbf{a}} \in \Omega$ existiert.

Betrachten wir einige Beispiele.

Beispiel 16.28. (partielle Ableitungen für vektorwertige Funktionen)

- (a) Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := x^2 y^3$. Nach Beispiel 16.26 (a) ist f auf ganz \mathbb{R}^2 partiell differenzierbar, und die Jacobi-Matrix von f ist

$$(\mathbf{J}f)(x, y) = [(\partial_1 f)(x, y) \quad (\partial_2 f)(x, y)] = [2xy^3 \quad 3x^2 y^2] \in \mathbb{R}^{1 \times 2}.$$

- (b) Sei $\vec{\mathbf{f}} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $\vec{\mathbf{f}}(x, y) := \begin{bmatrix} e^{x+y} \\ \sin(x) + \cos(y) \\ x^2 \end{bmatrix}$.

Dann ist $\vec{\mathbf{f}}$ auf ganz \mathbb{R}^2 partiell differenzierbar, und die partiellen Ableitungen von $\vec{\mathbf{f}}$ sind

$$(\partial_1 \vec{\mathbf{f}})(x, y) = \frac{\partial}{\partial x} \vec{\mathbf{f}}(x, y) = \begin{bmatrix} e^{x+y} \\ \cos(x) \\ 2x \end{bmatrix},$$

$$(\partial_2 \vec{\mathbf{f}})(x, y) = \frac{\partial}{\partial y} \vec{\mathbf{f}}(x, y) = \begin{bmatrix} e^{x+y} \\ -\sin(y) \\ 0 \end{bmatrix},$$

und die Jacobi-Matrix von \vec{f} ist

$$(\mathbf{J}\vec{f})(x, y) = \begin{bmatrix} e^{x+y} & e^{x+y} \\ \cos(x) & -\sin(y) \\ 2x & 0 \end{bmatrix} \in \mathbb{R}^{3 \times 2}.$$

(c) Sei $\mathbf{A} \in \mathbb{R}^{m \times n}$, und sei $\vec{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ definiert durch

$$\begin{aligned} \vec{f}(\vec{x}) &= \vec{f}(x_1, x_2, \dots, x_n) := \mathbf{A} \vec{x} = \mathbf{A} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= \begin{bmatrix} a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,n}x_n \\ a_{2,1}x_1 + a_{2,2}x_2 + \dots + a_{2,n}x_n \\ \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \dots + a_{m,n}x_n \end{bmatrix}. \end{aligned}$$

Dann ist \vec{f} auf \mathbb{R}^n partiell differenzierbar, und die Jacobi-Matrix ist

$$(\mathbf{J}\vec{f})(\vec{x}) = (\mathbf{J}\vec{f})(x_1, x_2, \dots, x_n) = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} = \mathbf{A}.$$

Also: Ist $\vec{f}(\vec{x}) = \mathbf{A} \cdot \vec{x}$, so ist \vec{f} auf ganz \mathbb{R}^n partiell differenzierbar, und die Jacobi-Matrix der partiellen Ableitungen ist $(\mathbf{J}\vec{f})(\vec{x}) = \mathbf{A}$.

Das nächste Beispiel zeigt ein überraschendes Phänomen.

Beispiel 16.29. (Aus partiell differenzierbar folgt nicht stetig!)

Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definiert durch $f(x, y) := \begin{cases} \frac{2xy}{x^2 + y^2} & \text{für } (x, y) \neq (0, 0), \\ 0 & \text{für } (x, y) = (0, 0). \end{cases}$

Dann ist f in $(0, 0)$ partiell differenzierbar mit der Jacobi-Matrix

$$(\mathbf{J}f)(0, 0) = [(\partial_1 f)(0, 0) \quad (\partial_2 f)(0, 0)] = [0 \quad 0],$$

denn

$$(\partial_1 f)(0, 0) = \frac{\partial}{\partial x} f(0, 0) = \lim_{h \rightarrow 0} \frac{f(0 + h, 0) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0 - 0}{h} = 0,$$

$$(\partial_2 f)(0,0) = \frac{\partial}{\partial y} f(0,0) = \lim_{h \rightarrow 0} \frac{f(0,0+h) - f(0,0)}{h} = \lim_{h \rightarrow 0} \frac{0-0}{h} = 0.$$

Aber f ist nicht stetig in $(0,0)$ nach Beispiel 16.21 (a).

Fazit aus dem vorigen Beispiel: Aus partieller Differenzierbarkeit in $\vec{\mathbf{a}} \in \Omega$ folgt **nicht** automatisch Stetigkeit im Punkt $\vec{\mathbf{a}}$!

Als letztes Konzept in diesem Teilkapitel lernen wir die Richtungsableitungen kennen.

Definition 16.30. (Richtungsableitungen)

Sei $\Omega \subseteq \mathbb{R}^n$ offen, $\vec{\mathbf{a}} \in \Omega$ und $\vec{\mathbf{v}} \in \mathbb{R}^n \setminus \{\vec{\mathbf{0}}\}$.

(1) Sei $f : \Omega \rightarrow \mathbb{R}$. Existiert

$$(D_{\vec{\mathbf{v}}} f)(\vec{\mathbf{a}}) := \lim_{t \rightarrow 0} \frac{f(\vec{\mathbf{a}} + t \vec{\mathbf{v}}) - f(\vec{\mathbf{a}})}{t},$$

so heißt $(D_{\vec{\mathbf{v}}} f)(\vec{\mathbf{a}})$ die **(Richtungs-)Ableitung von f im Punkt $\vec{\mathbf{a}}$ in Richtung $\vec{\mathbf{v}}$** .

(2) Sei $\vec{\mathbf{f}} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{bmatrix} : \Omega \rightarrow \mathbb{R}^m$.

Existieren $(D_{\vec{\mathbf{v}}} f_1)(\vec{\mathbf{a}})$, $(D_{\vec{\mathbf{v}}} f_2)(\vec{\mathbf{a}})$, \dots , $(D_{\vec{\mathbf{v}}} f_m)(\vec{\mathbf{a}})$, so heißt

$$(D_{\vec{\mathbf{v}}} \vec{\mathbf{f}})(\vec{\mathbf{a}}) := \begin{bmatrix} (D_{\vec{\mathbf{v}}} f_1)(\vec{\mathbf{a}}) \\ (D_{\vec{\mathbf{v}}} f_2)(\vec{\mathbf{a}}) \\ \vdots \\ (D_{\vec{\mathbf{v}}} f_m)(\vec{\mathbf{a}}) \end{bmatrix} \in \mathbb{R}^m$$

die **(Richtungs-)Ableitung von $\vec{\mathbf{f}}$ im Punkt $\vec{\mathbf{a}}$ in Richtung $\vec{\mathbf{v}}$** .

Bemerkung 16.31. (partielle Ableitungen als Richtungsableitungen)

(1) **Geometrische Anschauung der Richtungsableitung:** Die Funktion $g] - \varepsilon, \varepsilon[\rightarrow \mathbb{R}$, $g(t) := f(\vec{\mathbf{a}} + t \vec{\mathbf{v}})$, wobei $\varepsilon > 0$ so gewählt ist, dass $\vec{\mathbf{a}} + t \vec{\mathbf{v}} \in \Omega$ für alle $t \in] - \varepsilon, \varepsilon[$, ist die Funktion f eingeschränkt auf

das Geradenstück $\vec{\mathbf{a}} + t \vec{\mathbf{v}}$, $t \in]-\varepsilon, \varepsilon[$, in \mathbb{R}^n . Längs dieser Geraden ist f , oder genauer g , also eine Funktion einer Variablen, wie wir sie aus der HM A kennen. Existiert die Richtungsableitung von $\vec{\mathbf{f}}$ im Punkt $\vec{\mathbf{a}}$ in Richtung $\vec{\mathbf{v}}$, so ist g in $t = 0$ differenzierbar, und wir finden

$$g'(0) = \lim_{h \rightarrow 0} \frac{g(h) - g(0)}{h} = \lim_{h \rightarrow 0} \frac{f(\vec{\mathbf{a}} + h \vec{\mathbf{v}}) - f(\vec{\mathbf{a}})}{h} = (D_{\vec{\mathbf{v}}} f)(\vec{\mathbf{a}}).$$

Also können wir $g'(0) = (D_{\vec{\mathbf{v}}} f)(\vec{\mathbf{a}})$ als Steigung von g in $t = 0$ und damit als **Steigung von f eingeschränkt auf $\vec{\mathbf{a}} + t \vec{\mathbf{v}}$, $t \in]-\varepsilon, \varepsilon[$, im Punkt $\vec{\mathbf{a}}$** interpretieren.

- (2) Bei den partiellen Ableitungen einer auf einer offenen Menge $\Omega \subseteq \mathbb{R}^n$ partiell differenzierbaren reellwertigen bzw. vektorwertigen Funktion $f : \Omega \rightarrow \mathbb{R}$ bzw. $\vec{\mathbf{f}} : \Omega \rightarrow \mathbb{R}^m$ handelt es sich um die **Richtungsableitungen in Richtung der Vektoren der Standardbasis** $(\vec{\mathbf{e}}_1, \vec{\mathbf{e}}_2, \dots, \vec{\mathbf{e}}_n)$ von \mathbb{R}^n , also

$$\begin{aligned} (D_{\vec{\mathbf{e}}_i} f)(\vec{\mathbf{a}}) &= (\partial_i f)(\vec{\mathbf{a}}) && \text{für alle } i = 1, 2, \dots, n, && \text{bzw.} \\ (D_{\vec{\mathbf{e}}_i} \vec{\mathbf{f}})(\vec{\mathbf{a}}) &= (\partial_i \vec{\mathbf{f}})(\vec{\mathbf{a}}) && \text{für alle } i = 1, 2, \dots, n. \end{aligned}$$

Betrachten wir einige Beispiele für Richtungsableitungen.

Beispiel 16.32. (Richtungsableitungen)

- (a) Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := x^2 y^3$. Dann existiert für jedes $\vec{\mathbf{a}} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \in \mathbb{R}^2$

und jedes $\vec{\mathbf{v}} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{R}^2 \setminus \{\vec{\mathbf{0}}\}$ die Richtungsableitung von f in $\vec{\mathbf{a}}$ in Richtung $\vec{\mathbf{v}}$, denn

$$\begin{aligned} (D_{\vec{\mathbf{v}}} f)(\vec{\mathbf{a}}) &= \lim_{t \rightarrow 0} \frac{f(\vec{\mathbf{a}} + t \vec{\mathbf{v}}) - f(\vec{\mathbf{a}})}{t} \\ &= \lim_{t \rightarrow 0} \frac{f(a_1 + t v_1, a_2 + t v_2) - f(a_1, a_2)}{t} \\ &= \lim_{t \rightarrow 0} \frac{(a_1 + t v_1)^2 (a_2 + t v_2)^3 - a_1^2 a_2^3}{t} \\ &= \lim_{t \rightarrow 0} \frac{(a_1^2 + 2 t a_1 v_1 + t^2 v_1^2) (a_2^3 + 3 t a_2^2 v_2 + 3 t^2 a_2 v_2^2 + t^3 v_2^3) - a_1^2 a_2^3}{t} \end{aligned}$$

$$\begin{aligned}
&= \lim_{t \rightarrow 0} \frac{(a_1^2 + t(2a_1v_1 + tv_1^2))(a_2^3 + t(3a_2^2v_2 + 3ta_2v_2^2 + t^2v_2^3)) - a_1^2a_2^3}{t} \\
&= \lim_{t \rightarrow 0} \left(\frac{a_1^2a_2^3 + t(2a_1v_1 + tv_1^2)a_2^3 + ta_1^2(3a_2^2v_2 + 3ta_2v_2^2 + t^2v_2^3)}{t} \right. \\
&\quad \left. + \frac{t^2(2a_1v_1 + tv_1^2)(3a_2^2v_2 + 3ta_2v_2^2 + t^2v_2^3) - a_1^2a_2^3}{t} \right) \\
&= \lim_{t \rightarrow 0} \left(\frac{t(2a_1v_1 + tv_1^2)a_2^3 + ta_1^2(3a_2^2v_2 + 3ta_2v_2^2 + t^2v_2^3)}{t} \right. \\
&\quad \left. + \frac{t^2(2a_1v_1 + tv_1^2)(3a_2^2v_2 + 3ta_2v_2^2 + t^2v_2^3)}{t} \right) \\
&= \lim_{t \rightarrow 0} \left((2a_1v_1 + tv_1^2)a_2^3 + a_1^2(3a_2^2v_2 + 3ta_2v_2^2 + t^2v_2^3) \right. \\
&\quad \left. + t(2a_1v_1 + tv_1^2)(3a_2^2v_2 + 3ta_2v_2^2 + t^2v_2^3) \right) \\
&= 2a_1a_2^3v_1 + 3a_1^2a_2^2v_2.
\end{aligned}$$

(b) Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := \begin{cases} \frac{2xy}{x^2 + y^2} & \text{für } (x, y) \neq (0, 0), \\ 0 & \text{für } (x, y) = (0, 0). \end{cases}$

Für $\vec{\mathbf{a}} := (0, 0)$ und $\vec{\mathbf{v}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ existiert $(D_{\vec{\mathbf{v}}}f)(\vec{\mathbf{a}})$ nicht, denn:

$$\begin{aligned}
\frac{f(\vec{\mathbf{a}} + t\vec{\mathbf{v}}) - f(\vec{\mathbf{a}})}{t} &= \frac{f(0+t, 0+t) - f(0, 0)}{t} = \frac{f(t, t) - f(0, 0)}{t} \\
&= \frac{\frac{2t^2}{t^2 + t^2} - 0}{t} = \frac{1 - 0}{t} = \frac{1}{t},
\end{aligned}$$

und $\lim_{t \rightarrow 0} \frac{1}{t}$ existiert nicht.

(c) Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := \begin{cases} \frac{xy^2}{x^2 + y^4} & \text{für } (x, y) \neq (0, 0), \\ 0 & \text{für } (x, y) = (0, 0). \end{cases}$

Sei $\vec{\mathbf{a}} := \vec{\mathbf{0}}$. Dann existiert die Richtungsableitung $(D_{\vec{\mathbf{v}}}f)(\vec{\mathbf{a}})$ in $\vec{\mathbf{a}}$ für jede Richtung $\vec{\mathbf{v}} \in \mathbb{R}^2 \setminus \{\vec{\mathbf{0}}\}$, aber f ist nicht stetig in $\vec{\mathbf{a}} = \vec{\mathbf{0}}$. Wir zeigen dieses in einer Übungsaufgabe.

16.4 Totale Differenzierbarkeit

Bis jetzt haben wir nur partielle Ableitungen und Richtungsableitungen definiert, aber wir haben noch nicht erklärt, wann eine (reell- oder vektorwertige) Funktion mehrerer Variablen differenzierbar ist und somit eine Ableitung besitzt. Zur Motivation der weiter unten folgenden Definition von (totaler) Differenzierbarkeit erinnern wir uns zunächst an Differenzierbarkeit von reellwertigen Funktionen einer Variablen (siehe Kapitel 7 im Skript der HM A):

Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $f : I \rightarrow \mathbb{R}$ differenzierbar und $a \in I$. Die **Tangente an den Graphen von f im Punkt $(a, f(a))$** ist gegeben durch

$$g(x) = f(a) + f'(a)(x - a), \quad (16.1)$$

und die **Ableitung $f'(a)$** gibt die **Steigung der Tangente** an den Graphen von f im Punkt $(a, f(a))$ an.

Im nachfolgenden Satz lernen wir als Vorbereitung auf den Begriff der (totalen) Differenzierbarkeit einer Funktion mehrerer Variablen ein Analogon zu (16.1) kennen.

Satz 16.33. (Tangentialebene an den Graphen)

Sei $\Omega \subseteq \mathbb{R}^n$ offen und $f : \Omega \rightarrow \mathbb{R}$ partiell differenzierbar. Zusätzlich seien die partiellen Ableitungen $\partial_1 f, \partial_2 f, \dots, \partial_n f$ stetig in $\vec{a} \in \Omega$. Für

$$\begin{aligned} g(\vec{x}) &:= f(\vec{a}) + (\mathbf{J}f)(\vec{a}) \cdot (\vec{x} - \vec{a}) \\ &= f(\vec{a}) + [(\partial_1 f)(\vec{a}) \quad (\partial_2 f)(\vec{a}) \quad \cdots \quad (\partial_n f)(\vec{a})] \cdot \begin{bmatrix} x_1 - a_1 \\ x_2 - a_2 \\ \vdots \\ x_n - a_n \end{bmatrix} \\ &= f(\vec{a}) + \sum_{i=1}^n (\partial_i f)(\vec{a}) (x_i - a_i) \end{aligned}$$

gilt dann:

$$\lim_{\vec{x} \rightarrow \vec{a}} \frac{f(\vec{x}) - g(\vec{x})}{\|\vec{x} - \vec{a}\|} = 0.$$

Geometrische Interpretation:

- (1) Die Graphen von f und g berühren sich im Punkt $(a_1, a_2, \dots, a_n, f(\vec{a})) \in \mathbb{R}^{n+1}$.

(2) Wenn $n = 2$ ist, so beschreibt der Graph von g eine Ebene in \mathbb{R}^3 , die **Tangentialebene an den Graphen von f in dem Punkt $(a_1, a_2, f(\vec{\mathbf{a}})) \in \mathbb{R}^3$** .

Wir beweisen den Satz für den Fall $n = 2$.

Beweis des Satzes 16.33 für $n = 2$: Da $\Omega \subseteq \mathbb{R}^n$ offen ist, existiert zu $\vec{\mathbf{a}} \in \Omega$ ein $\varepsilon > 0$ mit $U_\varepsilon(\vec{\mathbf{a}}) \subseteq \Omega$.

Für $\vec{\mathbf{x}} \in U_\varepsilon(\vec{\mathbf{a}})$ gilt jeweils nach dem Mittelwertsatz der Differentialrechnung (vgl. Satz 7.19 im Skript der HM A) für die Funktion $f(x_1, x_2)$ als Funktion der Variablen x_1 bzw. als Funktion der Variablen x_2 : Es existiert ein ξ_1 zwischen a_1 und x_1 , und es existiert ein ξ_2 zwischen a_2 und x_2 , so dass

$$\begin{aligned} f(\vec{\mathbf{x}}) - f(\vec{\mathbf{a}}) &= f(x_1, x_2) - f(a_1, a_2) \\ &= (f(x_1, x_2) - f(a_1, x_2)) + (f(a_1, x_2) - f(a_1, a_2)) \\ &= (\partial_1 f)(\xi_1, x_2) \cdot (x_1 - a_1) + (\partial_2 f)(a_1, \xi_2) \cdot (x_2 - a_2). \end{aligned} \quad (16.2)$$

Somit gilt

$$\begin{aligned} f(\vec{\mathbf{x}}) - g(\vec{\mathbf{x}}) &= f(\vec{\mathbf{x}}) - \left(f(\vec{\mathbf{a}}) + (\partial_1 f)(\vec{\mathbf{a}}) \cdot (x_1 - a_1) + (\partial_2 f)(\vec{\mathbf{a}}) \cdot (x_2 - a_2) \right) \\ &= f(\vec{\mathbf{x}}) - f(\vec{\mathbf{a}}) - (\partial_1 f)(a_1, a_1) \cdot (x_1 - a_1) - (\partial_2 f)(a_1, a_2) \cdot (x_2 - a_2) \\ &\stackrel{(16.2)}{=} (\partial_1 f)(\xi_1, x_2) \cdot (x_1 - a_1) + (\partial_2 f)(a_1, \xi_2) \cdot (x_2 - a_2) \\ &\quad - (\partial_1 f)(a_1, a_2) \cdot (x_1 - a_1) - (\partial_2 f)(a_1, a_2) \cdot (x_2 - a_2) \\ &= \left((\partial_1 f)(\xi_1, x_2) - (\partial_1 f)(a_1, a_2) \right) \cdot (x_1 - a_1) \\ &\quad + \left((\partial_2 f)(a_1, \xi_2) - (\partial_2 f)(a_1, a_2) \right) \cdot (x_2 - a_2). \end{aligned}$$

Daraus folgt

$$\begin{aligned} &\frac{|f(\vec{\mathbf{x}}) - g(\vec{\mathbf{x}})|}{\|\vec{\mathbf{x}} - \vec{\mathbf{a}}\|} \\ &= \frac{|((\partial_1 f)(\xi_1, x_2) - (\partial_1 f)(a_1, a_2)) (x_1 - a_1) + ((\partial_2 f)(a_1, \xi_2) - (\partial_2 f)(a_1, a_2)) (x_2 - a_2)|}{\|\vec{\mathbf{x}} - \vec{\mathbf{a}}\|} \\ &\leq \frac{|(\partial_1 f)(\xi_1, x_2) - (\partial_1 f)(a_1, a_2)| \cdot |x_1 - a_1| + |(\partial_2 f)(a_1, \xi_2) - (\partial_2 f)(a_1, a_2)| \cdot |x_2 - a_2|}{\|\vec{\mathbf{x}} - \vec{\mathbf{a}}\|} \\ &\leq \underbrace{|(\partial_1 f)(\xi_1, x_2) - (\partial_1 f)(a_1, a_2)|}_{\leq 1} \cdot \underbrace{\frac{|x_1 - a_1|}{\|\vec{\mathbf{x}} - \vec{\mathbf{a}}\|}}_{\leq 1} + \underbrace{|(\partial_2 f)(a_1, \xi_2) - (\partial_2 f)(a_1, a_2)|}_{\leq 1} \cdot \underbrace{\frac{|x_2 - a_2|}{\|\vec{\mathbf{x}} - \vec{\mathbf{a}}\|}}_{\leq 1} \end{aligned}$$

$$\begin{aligned} &\leq |(\partial_1 f)(\xi_1, x_2) - (\partial_1 f)(a_1, a_2)| + |(\partial_2 f)(a_1, \xi_2) - (\partial_2 f)(a_1, a_2)| \\ &\rightarrow 0 \quad \text{für } \vec{x} \rightarrow \vec{a}, \end{aligned}$$

da die partiellen Ableitungen $\partial_1 f$ und $\partial_2 f$ stetig in \vec{a} sind und da aus $\vec{x} \rightarrow \vec{a}$ auch $(\xi_1, x_2) \rightarrow (a_1, a_2)$ und $(a_1, \xi_2) \rightarrow (a_1, a_2)$ folgen. \square

Nun definieren wir totale Differenzierbarkeit.

Definition 16.34. (total differenzierbar für reellwertige Funktionen)

Sei $\Omega \subseteq \mathbb{R}^n$ offen.

- (1) $f : \Omega \rightarrow \mathbb{R}$ heißt **(total) differenzierbar** in $\vec{a} \in \Omega$, wenn es eine Matrix $\mathbf{A} \in \mathbb{R}^{1 \times n}$ gibt mit

$$\lim_{\vec{x} \rightarrow \vec{a}} \frac{f(\vec{x}) - f(\vec{a}) - \mathbf{A}(\vec{x} - \vec{a})}{\|\vec{x} - \vec{a}\|} = 0.$$

$f'(\vec{a}) := \mathbf{A}$ heißt dann die **Ableitung von f im Punkt \vec{a}** .

- (2) $f : \Omega \rightarrow \mathbb{R}$ heißt **(total) differenzierbar (in Ω)**, wenn f in jedem $\vec{a} \in \Omega$ (total) differenzierbar ist.

Bemerkung 16.35. (total differenzierbar und Ableitung)

- (1) Mit $\vec{h} := \vec{x} - \vec{a} \iff \vec{x} = \vec{a} + \vec{h}$ gilt

$$\begin{aligned} &\lim_{\vec{x} \rightarrow \vec{a}} \frac{f(\vec{x}) - f(\vec{a}) - f'(\vec{a})(\vec{x} - \vec{a})}{\|\vec{x} - \vec{a}\|} \\ &= \lim_{\vec{h} \rightarrow \vec{0}} \frac{f(\vec{a} + \vec{h}) - f(\vec{a}) - f'(\vec{a})\vec{h}}{\|\vec{h}\|}. \end{aligned}$$

- (2) Falls die Ableitung $f'(\vec{a})$ existiert, ist diese **eindeutig** bestimmt, wie der nachfolgende Satz zeigt.

Satz 16.36. (total differenzierbar \implies partiell differenzierbar)

Sei $\Omega \subseteq \mathbb{R}^n$ offen, und seien $f : \Omega \rightarrow \mathbb{R}$ und $\vec{a} \in \Omega$. Ist f (total) differenzierbar in \vec{a} , so existieren alle Richtungsableitungen in \vec{a} und es gilt

für jedes $\vec{v} \in \mathbb{R}^n \setminus \{\vec{0}\}$

$$(D_{\vec{v}}f)(\vec{a}) = \underbrace{f'(\vec{a})}_{\in \mathbb{R}^{1 \times n}} \cdot \underbrace{\vec{v}}_{\in \mathbb{R}^n}.$$

Insbesondere ist f dann **partiell differenzierbar** in \vec{a} , und es gilt

$$f'(\vec{a}) = (\mathbf{J}f)(\vec{a}),$$

d.h. die Ableitung von f im Punkt \vec{a} ist die **Jacobi-Matrix** $(\mathbf{J}f)(\vec{a})$ der partiellen Ableitungen von f im Punkt \vec{a} .

Beweis von Satz 16.36: Sei $\vec{v} \in \mathbb{R}^n \setminus \{\vec{0}\}$. Falls die Richtungsableitung $(D_{\vec{v}}f)(\vec{a})$ existiert, so ist diese durch

$$(D_{\vec{v}}f)(\vec{a}) = \lim_{t \rightarrow 0} \frac{f(\vec{a} + t\vec{v}) - f(\vec{a})}{t}.$$

gegeben. – Es gilt für jede in \vec{a} (total) differenzierbare Funktion f

$$\begin{aligned} \left| \frac{f(\vec{a} + t\vec{v}) - f(\vec{a})}{t} - f'(\vec{a}) \cdot \vec{v} \right| &= \left| \frac{f(\vec{a} + t\vec{v}) - f(\vec{a}) - f'(\vec{a}) \cdot (t\vec{v})}{t} \right| \\ &= \frac{|f(\vec{a} + t\vec{v}) - f(\vec{a}) - f'(\vec{a}) \cdot (t\vec{v})|}{|t|} \\ &= \frac{|f(\vec{a} + t\vec{v}) - f(\vec{a}) - f'(\vec{a}) \cdot (t\vec{v})|}{|t| \cdot \|\vec{v}\|} \cdot \|\vec{v}\| \\ &= \frac{|f(\vec{a} + t\vec{v}) - f(\vec{a}) - f'(\vec{a}) \cdot (t\vec{v})|}{\|t\vec{v}\|} \cdot \|\vec{v}\| \\ &\xrightarrow{t \rightarrow 0} 0 \cdot \|\vec{v}\| = 0, \end{aligned}$$

wobei wir die Definition der (totalen) Differenzierbarkeit in \vec{a} mit $\vec{h} = t\vec{v} \rightarrow \vec{0}$ wenn $t \rightarrow 0$ genutzt haben. Damit folgt für jedes $\vec{v} \in \mathbb{R}^n \setminus \{\vec{0}\}$, dass $(D_{\vec{v}}f)(\vec{a})$ existiert und dass

$$(D_{\vec{v}}f)(\vec{a}) = \lim_{t \rightarrow 0} \frac{f(\vec{a} + t\vec{v}) - f(\vec{a})}{t} = f'(\vec{a}) \cdot \vec{v}.$$

Mit $\vec{v} = \vec{e}_i$, $i = 1, 2, \dots, n$, folgt die partielle Differenzierbarkeit und

$$(\partial_i f)(\vec{a}) = (D_{\vec{e}_i} f)(\vec{a}) = f'(\vec{a}) \cdot \vec{e}_i = i\text{-ter Eintrag in der } 1 \times n\text{-Matrix } f'(\vec{a}).$$

Also gilt $f'(\vec{\mathbf{a}}) = [(\partial_1 f)(\vec{\mathbf{a}}) \quad (\partial_2 f)(\vec{\mathbf{a}}) \quad \cdots \quad (\partial_n f)(\vec{\mathbf{a}})] = (\mathbf{J}f)(\vec{\mathbf{a}})$. \square

Es ist ein bisschen mühselig mit Definition 16.34 zu überprüfen, ob eine Funktion total differenzierbar ist. Daher lernen wir zunächst noch einige Resultate über (total) differenzierbare Funktionen kennen, mit deren Hilfe wir dann leichter überprüfen können, ob eine Funktion total differenzierbar ist.

Satz 16.37. (total differenzierbar \implies stetig)

Sei $\Omega \subseteq \mathbb{R}^n$ offen, und sei $f : \Omega \rightarrow \mathbb{R}$. Ist f (total) differenzierbar in $\vec{\mathbf{a}} \in \Omega$, dann ist f stetig in $\vec{\mathbf{a}}$.

Beweisidee von Satz 16.37: Mit der Nulladdition von $-f'(\vec{\mathbf{a}}) \cdot (\vec{\mathbf{x}} - \vec{\mathbf{a}})$ und $+f'(\vec{\mathbf{a}}) \cdot (\vec{\mathbf{x}} - \vec{\mathbf{a}})$ folgt

$$\begin{aligned} f(\vec{\mathbf{x}}) - f(\vec{\mathbf{a}}) &= f(\vec{\mathbf{x}}) - f(\vec{\mathbf{a}}) - f'(\vec{\mathbf{a}}) \cdot (\vec{\mathbf{x}} - \vec{\mathbf{a}}) + f'(\vec{\mathbf{a}}) \cdot (\vec{\mathbf{x}} - \vec{\mathbf{a}}) \\ &= \underbrace{\|\vec{\mathbf{x}} - \vec{\mathbf{a}}\|}_{\rightarrow 0 \text{ wenn } \vec{\mathbf{x}} \rightarrow \vec{\mathbf{a}}} \cdot \underbrace{\frac{f(\vec{\mathbf{x}}) - f(\vec{\mathbf{a}}) - f'(\vec{\mathbf{a}}) \cdot (\vec{\mathbf{x}} - \vec{\mathbf{a}})}{\|\vec{\mathbf{x}} - \vec{\mathbf{a}}\|}}_{\rightarrow 0 \text{ wenn } \vec{\mathbf{x}} \rightarrow \vec{\mathbf{a}}} + f'(\vec{\mathbf{a}}) \cdot \underbrace{(\vec{\mathbf{x}} - \vec{\mathbf{a}})}_{\rightarrow 0 \text{ wenn } \vec{\mathbf{x}} \rightarrow \vec{\mathbf{a}}} \\ &\xrightarrow{\vec{\mathbf{x}} \rightarrow \vec{\mathbf{a}}} 0. \end{aligned}$$

Also ist f stetig in $\vec{\mathbf{a}}$. \square

Satz 16.38. (Kriterium für totale Differenzierbarkeit)

Sei $\Omega \subseteq \mathbb{R}^n$ offen und $f : \Omega \rightarrow \mathbb{R}$.

- (1) Ist f partiell differenzierbar in Ω und sind alle partiellen Ableitungen $\partial_1 f, \partial_2 f, \dots, \partial_n f$ stetig in $\vec{\mathbf{a}} \in \Omega$, dann ist f (total) differenzierbar in $\vec{\mathbf{a}}$ und

$$f'(\vec{\mathbf{a}}) = (\mathbf{J}f)(\vec{\mathbf{a}}) = [(\partial_1 f)(\vec{\mathbf{a}}) \quad (\partial_2 f)(\vec{\mathbf{a}}) \quad \cdots \quad (\partial_n f)(\vec{\mathbf{a}})].$$

- (2) Existieren alle partiellen Ableitungen von f in ganz Ω und sind sie dort stetig, dann ist f (total) differenzierbar in ganz Ω . f heißt dann stetig differenzierbar in Ω , und wir schreiben $f \in \mathcal{C}^1(\Omega, \mathbb{R})$.

Dabei ist $\mathcal{C}^1(\Omega, \mathbb{R})$ der \mathbb{R} -Vektorraum der reellwertigen auf ganz Ω stetig differenzierbaren Funktionen.

Beweis von Satz 16.38: Satz 16.38 (1) folgt direkt aus Satz 16.33, und Satz 16.38 (2) ist eine Schlussfolgerung aus Satz 16.38 (1). \square

Mit Hilfe von Satz 16.38 und Satz 16.37 können wir nun für einige unserer vorigen Beispiele leicht feststellen, ob diese in bestimmten Punkten \vec{a} oder auch auf ganz Ω (total) differenzierbar sind.

Beispiel 16.39. (total differenzierbare reellwertige Funktionen)

- (a) $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := x^2 y^3$, ist auf ganz \mathbb{R}^2 nach x bzw. y partiell differenzierbar (vgl. Beispiel 16.26 (a)) und hat die auf \mathbb{R}^2 stetigen partiellen Ableitungen

$$(\partial_1 f)(x, y) = \frac{\partial}{\partial x} f(x, y) = 2 x y^3,$$

$$(\partial_2 f)(x, y) = \frac{\partial}{\partial y} f(x, y) = 3 x^2 y^2.$$

Nach Satz 16.38 ist f auf ganz \mathbb{R}^2 (total) differenzierbar und hat die Ableitung

$$f'(x, y) = (\mathbf{J}f)(x, y) = [2 x y^3 \quad 3 x^2 y^2].$$

Es gilt also $f \in \mathcal{C}^1(\mathbb{R}^2, \mathbb{R})$.

- (b) $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, $f(x, y, z) := \sin(x^2 + 2 y - z^3)$, ist auf ganz \mathbb{R}^3 nach x bzw. y bzw. z partiell differenzierbar (vgl. Beispiel 16.26 (b)) und hat die auf \mathbb{R}^3 stetigen partiellen Ableitungen

$$(\partial_1 f)(x, y, z) = \frac{\partial}{\partial x} f(x, y, z) = 2 x \cos(x^2 + 2 y - z^3),$$

$$(\partial_2 f)(x, y, z) = \frac{\partial}{\partial y} f(x, y, z) = 2 \cos(x^2 + 2 y - z^3),$$

$$(\partial_3 f)(x, y, z) = \frac{\partial}{\partial z} f(x, y, z) = -3 z^2 \cos(x^2 + 2 y - z^3).$$

Nach Satz 16.38 ist f auf ganz \mathbb{R}^3 (total) differenzierbar und hat die Ableitung

$$f'(x, y, z) = (\mathbf{J}f)(x, y, z) = [2 x \cos(x^2 + 2 y - z^3) \quad 2 \cos(x^2 + 2 y - z^3) \quad -3 z^2 \cos(x^2 + 2 y - z^3)].$$

Es gilt also $f \in \mathcal{C}^1(\mathbb{R}^3, \mathbb{R})$.

- (c) $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := \alpha x + \beta y + \gamma$, mit $\alpha, \beta, \gamma \in \mathbb{R}$, ist auf ganz \mathbb{R}^2 nach x bzw. y partiell differenzierbar (vgl. Beispiel 16.26 (c)) und hat die

auf \mathbb{R}^2 stetigen partiellen Ableitungen

$$(\partial_1 f)(x, y) = \frac{\partial}{\partial x} f(x, y) = \alpha,$$

$$(\partial_2 f)(x, y) = \frac{\partial}{\partial y} f(x, y) = \beta.$$

Nach Satz 16.38 ist f auf ganz \mathbb{R}^2 (total) differenzierbar und hat die Ableitung

$$f'(x, y) = (\mathbf{J}f)(x, y) = [\alpha \ \beta].$$

Es gilt also $f \in \mathcal{C}^1(\mathbb{R}^2, \mathbb{R})$.

- (d) $f : \mathbb{R}^4 \rightarrow \mathbb{R}$, $f(\vec{\mathbf{x}}) := f(x_1, x_2, x_3, x_4) = x_1 x_2 x_3 x_4$, ist auf ganz \mathbb{R}^4 nach x_i , $i = 1, 2, 3, 4$, partiell differenzierbar (vgl. Beispiel 16.26 (d)) und hat die auf \mathbb{R}^4 stetigen partiellen Ableitungen

$$(\partial_1 f)(x_1, x_2, x_3, x_4) = \frac{\partial}{\partial x_1} f(x_1, x_2, x_3, x_4) = x_2 x_3 x_4,$$

$$(\partial_2 f)(x_1, x_2, x_3, x_4) = \frac{\partial}{\partial x_2} f(x_1, x_2, x_3, x_4) = x_1 x_3 x_4,$$

$$(\partial_3 f)(x_1, x_2, x_3, x_4) = \frac{\partial}{\partial x_3} f(x_1, x_2, x_3, x_4) = x_1 x_2 x_4,$$

$$(\partial_4 f)(x_1, x_2, x_3, x_4) = \frac{\partial}{\partial x_4} f(x_1, x_2, x_3, x_4) = x_1 x_2 x_3.$$

Nach Satz 16.38 ist f auf ganz \mathbb{R}^4 (total) differenzierbar und hat die Ableitung

$$\begin{aligned} f'(x_1, x_2, x_3, x_4) &= (\mathbf{J}f)(x_1, x_2, x_3, x_4) \\ &= [x_2 x_3 x_4 \quad x_1 x_3 x_4 \quad x_1 x_2 x_4 \quad x_1 x_2 x_3]. \end{aligned}$$

Es gilt also $f \in \mathcal{C}^1(\mathbb{R}^4, \mathbb{R})$.

- (e) Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := \begin{cases} \frac{2xy}{x^2 + y^2} & \text{für } (x, y) \neq (0, 0), \\ 0 & \text{für } (x, y) = (0, 0). \end{cases}$

In $\vec{\mathbf{a}} := (0, 0)$ existiert die Richtungsableitung $(D_{\vec{\mathbf{v}}}f)(\vec{\mathbf{a}})$ in die Richtung $\vec{\mathbf{v}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ nach Beispiel 16.32 (b) nicht. Also kann f in $\vec{\mathbf{a}} = (0, 0)$ nicht (total) differenzierbar sein.

- (f) Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := \begin{cases} \frac{xy^2}{x^2 + y^4} & \text{für } (x, y) \neq (0, 0) \\ 0 & \text{für } (x, y) = (0, 0). \end{cases}$

Sei $\vec{\mathbf{a}} := (0, 0)$. Nach Beispiel 16.32 (c) existiert die Richtungsableitung $(D_{\vec{\mathbf{v}}}f)(\vec{\mathbf{a}})$ in $\vec{\mathbf{a}} = (0, 0)$ für jede Richtung $\vec{\mathbf{v}} \in \mathbb{R}^2 \setminus \{\vec{\mathbf{0}}\}$, aber f ist nicht stetig in $(0, 0)$. Nach Satz 16.37 muss eine in $(0, 0)$ (total) differenzierbare Funktion in $(0, 0)$ stetig sein. Also folgt, dass f in $(0, 0)$ nicht (total) differenzierbar sein kann.

Wir dehnen nun die Definition von (totaler) Differenzierbarkeit auf vektorwertige Funktionen mehrerer Variablen aus.

Definition 16.40. (total differenzierbar für vektorwertige Fkten.)

Sei $\Omega \subseteq \mathbb{R}^n$ offen, und sei $\vec{\mathbf{f}} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{bmatrix} : \Omega \rightarrow \mathbb{R}^m$.

- (1) $\vec{\mathbf{f}}$ heißt (total) differenzierbar in $\vec{\mathbf{a}} \in \Omega$, wenn f_1, f_2, \dots, f_m in $\vec{\mathbf{a}}$ (total) differenzierbar sind. Die Matrix

$$\vec{\mathbf{f}}'(\vec{\mathbf{a}}) := \begin{bmatrix} f_1'(\vec{\mathbf{a}}) \\ f_2'(\vec{\mathbf{a}}) \\ \vdots \\ f_m'(\vec{\mathbf{a}}) \end{bmatrix} \in \mathbb{R}^{m \times n}$$

heißt dann die **Ableitung von $\vec{\mathbf{f}}$ in $\vec{\mathbf{a}}$** .

- (2) $\vec{\mathbf{f}}$ heißt (total) differenzierbar in Ω , wenn $\vec{\mathbf{f}}$ in jedem $\vec{\mathbf{a}} \in \Omega$ (total) differenzierbar ist.

Bemerkung 16.41. (total differenzierbar für vektorwertige Fkten.)

Sei $\Omega \subseteq \mathbb{R}^n$ offen und $\vec{\mathbf{f}} : \Omega \rightarrow \mathbb{R}^m$.

- (1) Aus Satz 16.36 folgt: Ist $\vec{\mathbf{f}}$ (total) differenzierbar in $\vec{\mathbf{a}} \in \Omega$, so ist $\vec{\mathbf{f}}$ partiell differenzierbar in $\vec{\mathbf{a}}$, und es gilt

$$\vec{\mathbf{f}}'(\vec{\mathbf{a}}) = (\mathbf{J}\vec{\mathbf{f}})(\vec{\mathbf{a}}) = \begin{bmatrix} (\partial_1 f_1)(\vec{\mathbf{a}}) & (\partial_2 f_1)(\vec{\mathbf{a}}) & \cdots & (\partial_n f_1)(\vec{\mathbf{a}}) \\ (\partial_1 f_2)(\vec{\mathbf{a}}) & (\partial_2 f_2)(\vec{\mathbf{a}}) & \cdots & (\partial_n f_2)(\vec{\mathbf{a}}) \\ \vdots & \vdots & \ddots & \vdots \\ (\partial_1 f_m)(\vec{\mathbf{a}}) & (\partial_2 f_m)(\vec{\mathbf{a}}) & \cdots & (\partial_n f_m)(\vec{\mathbf{a}}) \end{bmatrix} \in \mathbb{R}^{m \times n},$$

d.h. die Ableitung ist die **Jacobi-Matrix der partiellen Ableitungen von \vec{f}** .

(2) Aus Satz 16.37 folgt: Ist \vec{f} **(total) differenzierbar in $\vec{a} \in \Omega$** , so ist \vec{f} **stetig in \vec{a}** .

(3) Aus Definition 16.34 und Bemerkung 16.35 folgt: \vec{f} ist **(total) differenzierbar in $\vec{a} \in \Omega$** genau dann, wenn es ein $\mathbf{A} \in \mathbb{R}^{m \times n}$ mit

$$\lim_{\vec{h} \rightarrow \vec{0}} \frac{1}{\|\vec{h}\|} \left(\vec{f}(\vec{a} + \vec{h}) - \vec{f}(\vec{a}) - \mathbf{A} \vec{h} \right) = \vec{0} \quad \text{in } \mathbb{R}^m$$

gibt. $\vec{f}'(\vec{a}) := \mathbf{A}$ heißt dann die Ableitung von \vec{f} in \vec{a} .

Analog zu Satz 16.38 erhalten wir den nachfolgenden Satz.

Satz 16.42. (Kriterium für totale Differenzierbarkeit)

Sei $\Omega \subseteq \mathbb{R}^n$ offen und $\vec{f} : \Omega \rightarrow \mathbb{R}^m$.

(1) Ist \vec{f} **partiell differenzierbar in Ω** und sind **alle partiellen Ableitungen $\partial_1 \vec{f}, \partial_2 \vec{f}, \dots, \partial_n \vec{f}$ stetig in $\vec{a} \in \Omega$** , dann ist \vec{f} **(total) differenzierbar in \vec{a}** und

$$\vec{f}'(\vec{a}) = (\mathbf{J}\vec{f})(\vec{a}).$$

(2) Existieren **alle partiellen Ableitungen in ganz Ω** und sind sie dort **stetig**, dann ist f **(total) differenzierbar in ganz Ω** . f heißt dann **stetig differenzierbar in Ω** , und wir schreiben $f \in \mathcal{C}^1(\Omega, \mathbb{R}^m)$.

Dabei ist $\mathcal{C}^1(\Omega, \mathbb{R}^m)$ der \mathbb{R} -Vektorraum der \mathbb{R}^m -wertigen auf ganz Ω stetig differenzierbaren Funktionen.

Betrachten wir zwei Beispiele.

Beispiel 16.43. (total diffbar., Ableitung für vektorwertige Fkten.)

(a) Sei $\vec{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $f(x, y) := \begin{bmatrix} e^{x+y} \\ \sin(x) + \cos(y) \\ x^2 \end{bmatrix}$.

Dann hat \vec{f} auf ganz \mathbb{R}^2 stetige partielle Ableitungen (siehe Beispiel 16.28 (b))

$$(\partial_1 \vec{f})(x, y) = \frac{\partial}{\partial x} \vec{f}(x, y) = \begin{bmatrix} e^{x+y} \\ \cos(x) \\ 2x \end{bmatrix},$$

$$(\partial_2 \vec{f})(x, y) = \frac{\partial}{\partial y} \vec{f}(x, y) = \begin{bmatrix} e^{x+y} \\ -\sin(y) \\ 0 \end{bmatrix},$$

und nach Satz 16.42 ist \vec{f} auf \mathbb{R}^2 (total) differenzierbar mit der Ableitung

$$\vec{f}'(x, y) = (\mathbf{J} \vec{f})(x, y) = \begin{bmatrix} e^{x+y} & e^{x+y} \\ \cos(x) & -\sin(y) \\ 2x & 0 \end{bmatrix} \in \mathbb{R}^{3 \times 2}.$$

\vec{f} ist auf \mathbb{R}^2 stetig differenzierbar, also $\vec{f} \in \mathcal{C}^1(\mathbb{R}^2, \mathbb{R}^3)$.

(b) Sei $\mathbf{A} \in \mathbb{R}^{m \times n}$, und sei $\vec{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ definiert durch

$$\vec{f}(\vec{x}) := \mathbf{A} \vec{x} = \begin{bmatrix} a_{1,1} x_1 + a_{1,2} x_2 + \cdots + a_{1,n} x_n \\ a_{2,1} x_1 + a_{2,2} x_2 + \cdots + a_{2,n} x_n \\ \vdots \\ a_{m,1} x_1 + a_{m,2} x_2 + \cdots + a_{m,n} x_n \end{bmatrix}$$

Dann hat \vec{f} auf \mathbb{R}^n stetige partielle Ableitungen (siehe Beispiel 16.28 (c)) und ist somit auf \mathbb{R}^n (total) differenzierbar. Die Ableitung von \vec{f} ist

$$\vec{f}'(\vec{x}) = (\mathbf{J} \vec{f})(\vec{x}) = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} = \mathbf{A}.$$

$\vec{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\vec{f}(\vec{x}) = \mathbf{A} \cdot \vec{x}$, ist stetig differenzierbar auf \mathbb{R}^n , also $\vec{f} \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^m)$.

Hilfssatz 16.44. (Rechenregeln für total differenzierbare Funktionen)

Sei $\Omega \subseteq \mathbb{R}^n$ offen, und seien $\vec{f} : \Omega \rightarrow \mathbb{R}^m$, $\vec{g} : \Omega \rightarrow \mathbb{R}^m$ und $\lambda \in \mathbb{R}$.

(1) Sind \vec{f} und \vec{g} (total) differenzierbar in $\vec{a} \in \Omega$, dann ist auch $\vec{f} + \vec{g}$ (total) differenzierbar in \vec{a} und

$$(\vec{f} + \vec{g})'(\vec{a}) = \vec{f}'(\vec{a}) + \vec{g}'(\vec{a}).$$

(2) Ist \vec{f} (total) differenzierbar in $\vec{a} \in \Omega$, dann ist auch $\lambda \vec{f}$ (total) differenzierbar in \vec{a} und

$$(\lambda \vec{f})'(\vec{a}) = \lambda \vec{f}'(\vec{a}).$$

Insbesondere gelten diese Rechenregeln auch für $m = 1$ also für (total) differenzierbare reellwertige Funktionen mehrerer Variablen.

16.5 Die Kettenregel

Wir erinnern uns an die **Kettenregel** für reellwertige Funktionen einer Variablen (siehe Kapitel 7 im Skript der HM A): Seien I, J offene Intervalle. Seien $f : I \rightarrow \mathbb{R}$ und $g : J \rightarrow \mathbb{R}$ reelle Funktionen mit $f(I) \subseteq J$. Ist f differenzierbar in $x_0 \in I$ und ist g differenzierbar in $f(x_0) \in J$, so ist $g \circ f$ differenzierbar in x_0 und es gilt:

$$(g \circ f)'(x_0) = \underbrace{g'(f(x_0))}_{\text{äußere Ableitung}} \cdot \underbrace{f'(x_0)}_{\text{innere Ableitung}}.$$

Wir verallgemeinern diese Kettenregel nun für vektorwertige Funktionen mehrerer Variablen, welche verkettet werden können.

Satz 16.45. (Kettenregel)

Seien $n, m, p \in \mathbb{N}$, $\Omega \subseteq \mathbb{R}^n$ offen und $U \subseteq \mathbb{R}^m$ offen. Seien $\vec{f} : \Omega \rightarrow \mathbb{R}^m$ (total) differenzierbar in $\vec{a} \in \Omega$, $\vec{f}(\Omega) \subseteq U$ und $\vec{g} : U \rightarrow \mathbb{R}^p$ (total) differenzierbar in $\vec{f}(\vec{a}) \in U$. Dann ist $\vec{g} \circ \vec{f} : \Omega \rightarrow \mathbb{R}^p$ (total) differenzierbar in \vec{a} , und die Ableitung in \vec{a} ist

$$(\vec{g} \circ \vec{f})'(\vec{a}) = \underbrace{\vec{g}'(\vec{f}(\vec{a}))}_{\in \mathbb{R}^{p \times m}} \cdot \underbrace{\vec{f}'(\vec{a})}_{\in \mathbb{R}^{m \times n}} \in \mathbb{R}^{p \times n}.$$

Man kann die Ableitung auch ohne Matrizenmultiplikation berechnen: Die ℓ -te Komponentenfunktion von $\vec{g} \circ \vec{f}$ ist

$$(\vec{g} \circ \vec{f})_{\ell}(\vec{x}) = g_{\ell}(\vec{f}(\vec{x})) = g_{\ell}(f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)),$$

und ihre partielle Ableitung nach x_i in $\vec{x} = \vec{a}$ ist

$$(\partial_i(\vec{g} \circ \vec{f})_\ell)(\vec{a}) = (\partial_i(g_\ell \circ \vec{f}))(\vec{a}) = \sum_{k=1}^m (\partial_k g_\ell)(\vec{f}(\vec{a})) (\partial_i f_k)(\vec{a}).$$

Achtung: Es genügt **nicht**, nur partielle Differenzierbarkeit zu fordern!

Beweis von Satz 16.45: Wir müssen zeigen:

$$\lim_{\vec{h} \rightarrow \vec{0}} \frac{1}{\|\vec{h}\|} \left((\vec{g} \circ \vec{f})(\vec{a} + \vec{h}) - (\vec{g} \circ \vec{f})(\vec{a}) - \vec{g}'(\vec{f}(\vec{a})) \cdot \vec{f}'(\vec{a}) \cdot \vec{h} \right) = \vec{0} \quad \text{in } \mathbb{R}^p.$$

Da \vec{f} total differenzierbar in $\vec{a} \in U$ ist, gilt

$$\vec{\Phi}(\vec{h}) := \frac{1}{\|\vec{h}\|} \left(\vec{f}(\vec{a} + \vec{h}) - \vec{f}(\vec{a}) - \vec{f}'(\vec{a}) \cdot \vec{h} \right) \xrightarrow{\vec{h} \rightarrow \vec{0}} \vec{0} \quad \text{in } \mathbb{R}^m. \quad (16.3)$$

Umstellen von (16.3) liefert

$$\vec{\Phi}(\vec{h}) \|\vec{h}\| + \vec{f}'(\vec{a}) \cdot \vec{h} = \vec{f}(\vec{a} + \vec{h}) - \vec{f}(\vec{a}). \quad (16.4)$$

Da \vec{g} total differenzierbar in $\vec{f}(\vec{a}) \in V$ ist, gilt

$$\vec{\Psi}(\vec{k}) := \frac{1}{\|\vec{k}\|} \left(\vec{g}(\vec{f}(\vec{a}) + \vec{k}) - \vec{g}(\vec{f}(\vec{a})) - \vec{g}'(\vec{f}(\vec{a})) \cdot \vec{k} \right) \xrightarrow{\vec{k} \rightarrow \vec{0}} \vec{0} \quad \text{in } \mathbb{R}^p.$$

Daraus folgt

$$\|\vec{k}\| \vec{\Psi}(\vec{k}) = \vec{g}(\vec{f}(\vec{a}) + \vec{k}) - \vec{g}(\vec{f}(\vec{a})) - \vec{g}'(\vec{f}(\vec{a})) \cdot \vec{k}. \quad (16.5)$$

Da \vec{f} total differenzierbar in \vec{a} ist, ist \vec{f} stetig in \vec{a} und es gilt

$$\vec{k} := \vec{f}(\vec{a} + \vec{h}) - \vec{f}(\vec{a}) \xrightarrow{\vec{h} \rightarrow \vec{0}} \vec{0}. \quad (16.6)$$

Durch Einsetzen von \vec{k} definiert durch (16.6) in (16.5) folgt

$$\begin{aligned} & \|\vec{k}\| \vec{\Psi}(\vec{k}) \\ &= \vec{g}(\vec{f}(\vec{a} + \vec{h})) - \vec{g}(\vec{f}(\vec{a})) - \vec{g}'(\vec{f}(\vec{a})) \cdot (\vec{f}(\vec{a} + \vec{h}) - \vec{f}(\vec{a})) \\ &= (\vec{g} \circ \vec{f})(\vec{a} + \vec{h}) - (\vec{g} \circ \vec{f})(\vec{a}) - \vec{g}'(\vec{f}(\vec{a})) \cdot (\vec{f}(\vec{a} + \vec{h}) - \vec{f}(\vec{a})). \end{aligned}$$

Anschließendes Ersetzen von $\vec{f}(\vec{a} + \vec{h}) - \vec{f}(\vec{a})$ durch (16.4) liefert

$$\begin{aligned} & \|\vec{k}\| \vec{\Psi}(\vec{k}) \\ &= (\vec{g} \circ \vec{f})(\vec{a} + \vec{h}) - (\vec{g} \circ \vec{f})(\vec{a}) - \vec{g}'(\vec{f}(\vec{a})) \cdot (\vec{\Phi}(\vec{h}) \|\vec{h}\| + \vec{f}'(\vec{a}) \cdot \vec{h}). \end{aligned}$$

Umsortieren liefert

$$\begin{aligned} & (\vec{g} \circ \vec{f})(\vec{a} + \vec{h}) - (\vec{g} \circ \vec{f})(\vec{a}) - \vec{g}'(\vec{f}(\vec{a})) \cdot \vec{f}'(\vec{a}) \cdot \vec{h} \\ &= \|\vec{k}\| \underbrace{\vec{\Psi}(\vec{k})}_{\rightarrow \vec{0} \text{ für } \vec{k} \rightarrow \vec{0}} + \vec{g}'(\vec{f}(\vec{a})) \cdot \underbrace{\vec{\Phi}(\vec{h}) \|\vec{h}\|}_{\rightarrow \vec{0} \text{ für } \vec{h} \rightarrow \vec{0}}. \end{aligned}$$

Nun teilen wir durch $\|\vec{h}\|$ und finden

$$\begin{aligned} & \frac{1}{\|\vec{h}\|} \left((\vec{g} \circ \vec{f})(\vec{a} + \vec{h}) - (\vec{g} \circ \vec{f})(\vec{a}) - \vec{g}'(\vec{f}(\vec{a})) \cdot \vec{f}'(\vec{a}) \cdot \vec{h} \right) \\ &= \frac{\|\vec{k}\|}{\|\vec{h}\|} \underbrace{\vec{\Psi}(\vec{k})}_{\rightarrow \vec{0} \text{ für } \vec{k} \rightarrow \vec{0}} + \vec{g}'(\vec{f}(\vec{a})) \cdot \underbrace{\vec{\Phi}(\vec{h})}_{\rightarrow \vec{0} \text{ für } \vec{h} \rightarrow \vec{0}}. \end{aligned} \quad (16.7)$$

Wir untersuchen nun $\|\vec{k}\|/\|\vec{h}\|$. Durch Ersetzen von \vec{k} mittels (16.6) im ersten Schritt und Ersetzen von $\vec{f}(\vec{a} + \vec{h}) - \vec{f}(\vec{a})$ mittels (16.4) im zweiten Schritt folgt

$$\begin{aligned} \frac{\|\vec{k}\|}{\|\vec{h}\|} &= \frac{1}{\|\vec{h}\|} \|\vec{f}(\vec{a} + \vec{h}) - \vec{f}(\vec{a})\| \\ &= \frac{1}{\|\vec{h}\|} \|\vec{\Phi}(\vec{h}) \|\vec{h}\| + \vec{f}'(\vec{a}) \cdot \vec{h}\| \\ &\leq \frac{1}{\|\vec{h}\|} \left(\|\vec{\Phi}(\vec{h}) \|\vec{h}\|\| + \|\vec{f}'(\vec{a}) \cdot \vec{h}\| \right) \quad (\text{mittels Dreiecksungleichung}) \\ &= \frac{1}{\|\vec{h}\|} \left(\|\vec{h}\| \cdot \|\vec{\Phi}(\vec{h})\| + \|\vec{f}'(\vec{a}) \cdot \vec{h}\| \right) \quad \left(\begin{array}{l} \|\vec{h}\| \text{ aus erster Norm} \\ \text{herausziehen} \end{array} \right) \\ &= \|\vec{\Phi}(\vec{h})\| + \frac{1}{\|\vec{h}\|} \cdot \|\vec{f}'(\vec{a}) \cdot \vec{h}\| \\ &= \|\vec{\Phi}(\vec{h})\| + \left\| \vec{f}'(\vec{a}) \cdot \frac{\vec{h}}{\|\vec{h}\|} \right\| \quad (1/\|\vec{h}\| \text{ in zweite Norm hineinziehen}) \\ &\leq \text{Konstante für kleines } \|\vec{h}\|. \end{aligned}$$

Weiter gilt wegen (16.6), dass $\vec{\mathbf{k}} \rightarrow \vec{\mathbf{0}}$ wenn $\vec{\mathbf{h}} \rightarrow \vec{\mathbf{0}}$. Damit sieht man, dass die rechte Seite in (16.7) für $\vec{\mathbf{h}} \rightarrow \vec{\mathbf{0}}$ gegen den Nullvektor $\vec{\mathbf{0}} \in \mathbb{R}^p$ strebt, und die Behauptung ist gezeigt. \square

Betrachten wir zwei Beispiele für die Anwendung der Kettenregel.

Beispiel 16.46. (Kettenregel)

Seien

$$\Omega := \{(r, \phi) : r > 0, \phi \in \mathbb{R}\}, \quad U := \mathbb{R}^2 \setminus \{(0, 0)\},$$

$$\vec{\mathbf{f}} : \Omega \rightarrow \mathbb{R}^2, \quad \vec{\mathbf{f}}(r, \phi) = \begin{bmatrix} r \cos(\phi) \\ r \sin(\phi) \end{bmatrix},$$

$$g : U \rightarrow \mathbb{R}, \quad g(x, y) = \sqrt{x^2 + y^2} = (x^2 + y^2)^{1/2}.$$

Wir bemerken zunächst, dass $\vec{\mathbf{f}}(\Omega) = U$, so dass die Verkettung $g \circ \vec{\mathbf{f}}$ auf Ω gebildet werden kann.

(a) Direkte Berechnung von $(g \circ \vec{\mathbf{f}})'$:

$$\begin{aligned} (g \circ \vec{\mathbf{f}})(r, \phi) &= g(r \cos(\phi), r \sin(\phi)) \\ &= \sqrt{r^2 \cos^2(\phi) + r^2 \sin^2(\phi)} = \sqrt{r^2} = r, \\ (g \circ \vec{\mathbf{f}})'(r, \phi) &= \left[(\partial_1(g \circ \vec{\mathbf{f}}))(r, \phi) \quad (\partial_2(g \circ \vec{\mathbf{f}}))(r, \phi) \right] = [1 \quad 0]. \end{aligned}$$

(b) Berechnung mit Hilfe der Kettenregel:

Wir berechnen zunächst die Ableitungen von $\vec{\mathbf{f}}$ und g :

$$\begin{aligned} \vec{\mathbf{f}}'(r, \phi) &= \begin{bmatrix} \cos(\phi) & -r \sin(\phi) \\ \sin(\phi) & r \cos(\phi) \end{bmatrix}, \\ g'(x, y) &= \begin{bmatrix} \frac{x}{\sqrt{x^2 + y^2}} & \frac{y}{\sqrt{x^2 + y^2}} \end{bmatrix}. \end{aligned}$$

Mit

$$\sqrt{(r \cos(\phi))^2 + (r \sin(\phi))^2} = \sqrt{r^2 (\cos^2(\phi) + \sin^2(\phi))} = \sqrt{r^2 \cdot 1} = r$$

und der Kettenregel folgt dann

$$(g \circ \vec{\mathbf{f}})'(r, \phi) = g'(r \cos(\phi), r \sin(\phi)) \cdot \vec{\mathbf{f}}'(r, \phi)$$

$$\begin{aligned}
&= \begin{bmatrix} \frac{r \cos(\phi)}{r} & \frac{r \sin(\phi)}{r} \end{bmatrix} \cdot \begin{bmatrix} \cos(\phi) & -r \sin(\phi) \\ \sin(\phi) & r \cos(\phi) \end{bmatrix} \\
&= \begin{bmatrix} \cos(\phi) & \sin(\phi) \end{bmatrix} \cdot \begin{bmatrix} \cos(\phi) & -r \sin(\phi) \\ \sin(\phi) & r \cos(\phi) \end{bmatrix} \\
&= \begin{bmatrix} \cos(\phi) \cos(\phi) + \sin(\phi) \sin(\phi) & \cos(\phi) (-r \sin(\phi)) + \sin(\phi) r \cos(\phi) \end{bmatrix} \\
&= \begin{bmatrix} \cos^2(\phi) + \sin^2(\phi) & 0 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \end{bmatrix}.
\end{aligned}$$

Beobachtung am vorigen Beispiel: Die **direkte Berechnung** der Ableitung einer Verkettung von Funktionen **ohne die Kettenregel** geht manchmal schneller!

Frage: Wozu braucht man dann überhaupt also die Kettenregel?

Diese ist erforderlich, wenn man „abstrakt“ mit der Verkettung von Funktionen rechnet, also wenn man die verketteten Funktionen nicht direkt kennt, also wenn man z.B. die Ableitung der Verkettung einer Funktion auf einer Teilmenge von \mathbb{R}^3 mit einem Weg im \mathbb{R}^3 betrachtet. Wir illustrieren dieses im nächsten Beispiel. **Hier sollte man immer die Notation der partiellen Ableitungen mit den Nummern der Variablen verwenden** (siehe Beispiel), **damit man mit den Namen der Variablen nicht durcheinander gerät!**

Beispiel 16.47. (Kettenregel)

Seien $U \subseteq \mathbb{R}^3$ offen und $g : U \rightarrow \mathbb{R}$ stetig differenzierbar. Ist $\vec{x} :]a, b[\rightarrow U$ ein \mathcal{C}^1 -Weg, so gilt für $g \circ \vec{x} :]a, b[\rightarrow \mathbb{R}$

$$(g \circ \vec{x})(t) = g(\vec{x}(t)) = g(x_1(t), x_2(t), x_3(t)),$$

und die Ableitung von $g \circ \vec{x}$ berechnet sich mit der Kettenregel wie folgt:

$$\begin{aligned}
(g \circ \vec{x})'(t) &= g'(\vec{x}(t)) \cdot \vec{x}'(t) \\
&= \begin{bmatrix} (\partial_1 g)(\vec{x}(t)) & (\partial_2 g)(\vec{x}(t)) & (\partial_3 g)(\vec{x}(t)) \end{bmatrix} \cdot \begin{bmatrix} x_1'(t) \\ x_2'(t) \\ x_3'(t) \end{bmatrix} \\
&= (\partial_1 g)(\vec{x}(t)) \cdot x_1'(t) + (\partial_2 g)(\vec{x}(t)) \cdot x_2'(t) + (\partial_3 g)(\vec{x}(t)) \cdot x_3'(t).
\end{aligned}$$

In der HM C werden wir viele „abstrakte“ Anwendungen der Kettenregel wie im vorigen Beispiel benötigen, wenn wir entlang von Wegen oder auf Oberflächen über reellwertige oder vektorwertige Funktionen integrieren.

Wir lernen nun noch den neuen Begriff eines Gebiets kennen, und können dann mit Hilfe der Kettenregel eine interessante Schlussfolgerung über stetige differenzierbare reellwertige Funktionen auf einem Gebiet ziehen.

Definition 16.48. (Gebiet)

Sei $\Omega \subseteq \mathbb{R}^n$ offen. Ω heißt ein **Gebiet**, wenn es für alle $\vec{\mathbf{a}}, \vec{\mathbf{b}} \in \Omega$ einen \mathcal{C}^1 -Weg in Ω gibt, der $\vec{\mathbf{a}}$ mit $\vec{\mathbf{b}}$ verbindet.

Die Bilder in Abbildung 16.2 illustrieren den Unterschied zwischen offenen Mengen in \mathbb{R}^2 , die Gebiete sind, und offenen Mengen in \mathbb{R}^2 , die keine Gebiete sind.

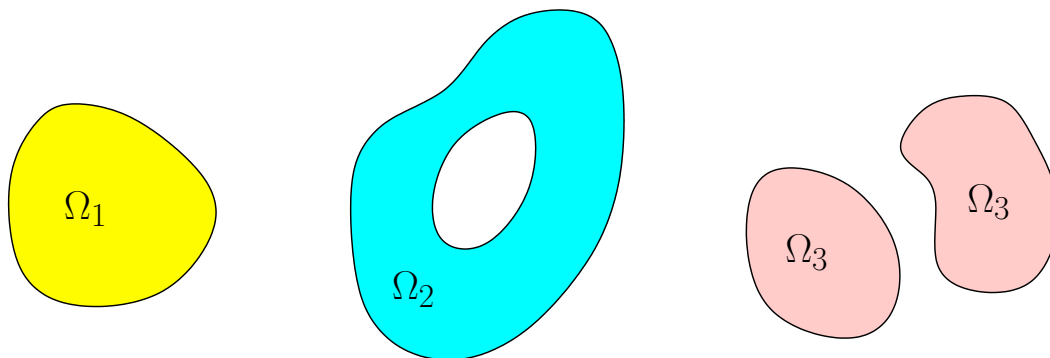


Abbildung 16.2: Bei allen drei Mengen Ω_1, Ω_2 und Ω_3 gehöre der Rand nicht mit zur Menge, so dass die Mengen offen sind. Ω_1 und Ω_2 sind jeweils ein Gebiet. Ω_3 ist dagegen kein Gebiet.

Satz 16.49. (Ableitung ist auf Gebiet Null $\implies \mathcal{C}^1$ -Funktion ist konstant)

Sei $\Omega \subseteq \mathbb{R}^n$ ein Gebiet. Sei $f \in \mathcal{C}^1(\Omega, \mathbb{R})$ mit

$$f'(\vec{\mathbf{x}}) = [(\partial_1 f)(\vec{\mathbf{x}}) \quad (\partial_2 f)(\vec{\mathbf{x}}) \quad \cdots \quad (\partial_n f)(\vec{\mathbf{x}})] = [0 \quad 0 \quad \cdots \quad 0] = \mathbf{O}_{1 \times n}$$

für alle $\vec{\mathbf{x}} \in \Omega$. Dann ist f konstant auf Ω .

Beweis von Satz 16.49: Seien $\vec{\mathbf{a}}, \vec{\mathbf{b}} \in \Omega$ beliebig. Wir zeigen, dass dann gilt $f(\vec{\mathbf{a}}) = f(\vec{\mathbf{b}})$. Da $\vec{\mathbf{a}}, \vec{\mathbf{b}} \in \Omega$ beliebig sind, hat die Funktion f dann in allen Punkten aus Ω denselben konstanten Wert.

Da Ω ein Gebiet ist, existiert ein \mathcal{C}^1 -Weg $\vec{\mathbf{x}} : I \rightarrow \Omega$ auf einem offenen Intervall $I \subseteq \mathbb{R}$ mit $[\alpha, \beta] \subseteq I$ und

$$\vec{\mathbf{x}}(\alpha) = \vec{\mathbf{a}} \quad \text{und} \quad \vec{\mathbf{x}}(\beta) = \vec{\mathbf{b}}.$$

Wir betrachten nun die Funktion $h := f \circ \vec{x} : I \rightarrow \mathbb{R}$. Die Kettenregel liefert

$$h'(t) = \underbrace{f'(\vec{x}(t))}_{=\mathbf{O}_{1 \times n}} \cdot \vec{x}'(t) = \mathbf{O}_{1 \times n} \cdot \vec{x}'(t) = 0 \quad \text{für alle } t \in I.$$

Eine reellwertige Funktion, deren Ableitung auf einem offenen Intervall Null ist, ist auf dem offenen Intervall konstant (vgl. Kapitel 7 im Skript der HM A). Somit ist $h = f \circ \vec{x}$ konstant auf I . Insbesondere folgt daraus, dass

$$f(\vec{\mathbf{a}}) = f(\vec{x}(\alpha)) = (f \circ \vec{x})(\alpha) = h(\alpha) = h(\beta) = (f \circ \vec{x})(\beta) = f(\vec{x}(\beta)) = f(\vec{\mathbf{b}})$$

gilt. □

16.6 Höhere Ableitungen

Als Vorbereitung auf die Einführung von höheren Ableitungen von Funktionen mehrerer Variablen machen wir uns noch einmal klar, was es bedeutet, wenn f in $\mathcal{C}^1(\Omega, \mathbb{R}^m)$ mit $\Omega \subseteq \mathbb{R}^n$ offen ist: Dann existieren alle partiellen Ableitungen $\partial_1 \vec{\mathbf{f}}, \partial_2 \vec{\mathbf{f}}, \dots, \partial_n \vec{\mathbf{f}} : \Omega \rightarrow \mathbb{R}^m$ und sind stetige Funktionen in Ω .

Definition 16.50. (zweimal stetig differenzierbare Funktion)

Sei $\Omega \subseteq \mathbb{R}^n$ offen. Eine Funktion $\vec{\mathbf{f}} \in \mathcal{C}^1(\Omega, \mathbb{R}^m)$ heißt **zweimal stetig differenzierbar auf Ω** , wenn die partiellen Ableitungen $\partial_1 \vec{\mathbf{f}}, \partial_2 \vec{\mathbf{f}}, \dots, \partial_n \vec{\mathbf{f}}$ existieren und in $\mathcal{C}^1(\Omega, \mathbb{R}^m)$ sind, d.h. wenn

$$\partial_k(\partial_i \vec{\mathbf{f}}) : \Omega \rightarrow \mathbb{R}^m$$

für alle $k, i \in \{1, 2, \dots, n\}$ existieren und stetig sind. Wir schreiben dann $\vec{\mathbf{f}} \in \mathcal{C}^2(\Omega, \mathbb{R}^m)$ und

$$\partial_k \partial_i \vec{\mathbf{f}} := \partial_k(\partial_i \vec{\mathbf{f}}) = \frac{\partial}{\partial x_k} \left(\frac{\partial}{\partial x_i} \vec{\mathbf{f}} \right) = \frac{\partial^2}{\partial x_k \partial x_i} \vec{\mathbf{f}}.$$

Beispiel 16.51. (zweimal stetig differenzierbare Funktion)

Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := x^2 y^3$. Dann existieren die partiellen Ableitungen erster und zweiter Ordnung auf ganz \mathbb{R}^2 ,

$$(\partial_1 f)(x, y) = 2xy^3, \quad (\partial_2 f)(x, y) = 3x^2 y^2,$$

$$\begin{aligned}(\partial_1^2 f)(x, y) &= 2y^3, & (\partial_1 \partial_2 f)(x, y) &= 6xy^2, \\(\partial_2 \partial_1 f)(x, y) &= 6xy^2, & (\partial_2^2 f)(x, y) &= 6x^2y,\end{aligned}$$

und diese sind jeweils stetig auf \mathbb{R}^2 . Es gilt also $f \in \mathcal{C}^2(\mathbb{R}^2, \mathbb{R})$

Wir beobachten, dass im vorigen Beispiel gilt

$$(\partial_2 \partial_1 f)(x, y) = (\partial_1 \partial_2 f)(x, y).$$

Allgemein gilt die folgende Aussage über „gemischte“ partielle Ableitungen zweiter Ordnung.

Satz 16.52. (Satz von Schwarz)

Sei $\Omega \subseteq \mathbb{R}^n$ offen. Für $\vec{f} \in \mathcal{C}^2(\Omega, \mathbb{R}^m)$ gilt

$$\partial_i \partial_k \vec{f} = \partial_k \partial_i \vec{f} \quad \text{für alle } i, k \in \{1, 2, \dots, n\}.$$

Beweis von Satz 16.52 (für $n = 2$ und $m = 1$): Sei $(x_1, x_2) \in \Omega$ beliebig. Wähle $h > 0$ so klein, dass das Quadrat mit den Eckpunkten $(x_1, x_2), (x_1 + h, x_2), (x_1, x_2 + h), (x_1 + h, x_2 + h)$ ganz in Ω liegt. Das geht immer, da Ω offen ist.

Wir definieren nun die Funktionen

$$\begin{aligned}g(x_1) &:= f(x_1, x_2 + h) - f(x_1, x_2), & \text{wobei } x_2 \text{ fest ist,} \\k(x_2) &:= f(x_1 + h, x_2) - f(x_1, x_2), & \text{wobei } x_1 \text{ fest ist.}\end{aligned}$$

Dann gilt

$$\begin{aligned}g(x_1 + h) - g(x_1) &= f(x_1 + h, x_2 + h) - f(x_1 + h, x_2) - (f(x_1, x_2 + h) - f(x_1, x_2)) \\&= f(x_1 + h, x_2 + h) - f(x_1, x_2 + h) - (f(x_1 + h, x_2) - f(x_1, x_2)) \\&= k(x_2 + h) - k(x_2),\end{aligned}$$

also

$$g(x_1 + h) - g(x_1) = k(x_2 + h) - k(x_2). \quad (16.8)$$

Nach dem Mittelwertsatz der Differentialrechnung (siehe Satz 7.19 im Kapitel 7 im Skript der HM A) existieren jeweils $\xi_1 \in]x_1, x_1 + h[$ und $\xi_2 \in]x_2, x_2 + h[$ mit

$$g(x_1 + h) - g(x_1) = g'(\xi_1) h,$$

$$k(x_2 + h) - k(x_2) = k'(\xi_2) h.$$

Einsetzen in (16.8) liefert

$$g'(\xi_1) h = k'(\xi_2) h \iff g'(\xi_1) = k'(\xi_2). \quad (16.9)$$

Das Berechnen der Ableitungen liefert

$$g'(x_1) := \frac{\partial}{\partial x_1} f(x_1, x_2 + h) - \frac{\partial}{\partial x_1} f(x_1, x_2) = (\partial_1 f)(x_1, x_2 + h) - (\partial_1 f)(x_1, x_2),$$

$$k'(x_2) := \frac{\partial}{\partial x_2} f(x_1 + h, x_2) - \frac{\partial}{\partial x_2} f(x_1, x_2) = (\partial_2 f)(x_1 + h, x_2) - (\partial_2 f)(x_1, x_2),$$

und Einsetzen in (16.9) liefert

$$(\partial_1 f)(\xi_1, x_2 + h) - (\partial_1 f)(\xi_1, x_2) = (\partial_2 f)(x_1 + h, \xi_2) - (\partial_2 f)(x_1, \xi_2). \quad (16.10)$$

Wir betrachten nun die stetig differenzierbaren partiellen Ableitungen $\partial_1 f$ und $\partial_2 f$ jeweils als Funktion der anderen Variablen, also

$$m(x_2) := (\partial_1 f)(\xi_1, x_2) \quad \text{und} \quad \ell(x_1) := (\partial_2 f)(x_1, \xi_2),$$

und nutzen den Mittelwertsatz der Differentialrechnung erneut: Nach dem Mittelwertsatz der Differentialrechnung existieren jeweils $\eta_1 \in]x_1, x_1 + h[$ und $\eta_2 \in]x_2, x_2 + h[$ mit

$$\begin{aligned} (\partial_1 f)(\xi_1, x_2 + h) - (\partial_1 f)(\xi_1, x_2) &= m(x_2 + h) - m(x_2) \\ &= m'(\eta_2) h = (\partial_2 \partial_1 f)(\xi_1, \eta_2) h, \\ (\partial_2 f)(x_1 + h, \xi_2) - (\partial_2 f)(x_1, \xi_2) &= \ell(x_1 + h) - \ell(x_1) \\ &= \ell'(\eta_1) h = (\partial_1 \partial_2 f)(\eta_1, \xi_2) h. \end{aligned}$$

Einsetzen in (16.10) liefert

$$(\partial_2 \partial_1 f)(\xi_1, \eta_2) h = (\partial_1 \partial_2 f)(\eta_1, \xi_2) h \iff (\partial_2 \partial_1 f)(\xi_1, \eta_2) = (\partial_1 \partial_2 f)(\eta_1, \xi_2).$$

Für $h \rightarrow 0$ gilt $(\xi_1, \eta_2) \rightarrow (x_1, x_2)$ und $(\eta_1, \xi_2) \rightarrow (x_1, x_2)$. Also folgt aus der Stetigkeit der „gemischten“ partiellen Ableitungen zweiter Ordnung $\partial_1 \partial_2 f$ und $\partial_2 \partial_1 f$, dass

$$(\partial_2 \partial_1 f)(x_1, x_2) = (\partial_1 \partial_2 f)(x_1, x_2)$$

gilt. □

Bemerkung 16.53. (Satz von Schwarz)

Sind die zweiten partiellen Ableitungen **nicht stetig**, so gilt der Satz von Schwarz im Allgemeinen nicht. Dieses zeigt das folgende Beispiel:

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) := \begin{cases} xy \frac{x^2 - y^2}{x^2 + y^2} & \text{wenn } (x, y) \neq (0, 0), \\ 0 & \text{wenn } (x, y) = (0, 0). \end{cases}$$

Man kann zeigen, dass f in $\mathcal{C}^1(\mathbb{R}^2, \mathbb{R})$ ist und dass f sogar zweimal differenzierbar ist. Die zweiten partiellen Ableitungen sind aber nicht mehr auf ganz \mathbb{R}^2 stetig. Die gemischten zweiten partiellen Ableitungen dieser Funktion im Punkt $(0, 0)$ sind aber

$$(\partial_1 \partial_2 f)(0, 0) = -1 \quad \text{und} \quad (\partial_2 \partial_1 f)(0, 0) = 1;$$

sie stimmen also hier nicht überein.

Definition 16.54. (höhere Ableitungen)

Seien $\Omega \subseteq \mathbb{R}^n$ offen und $s \in \mathbb{N}_0$. Wir sagen $\vec{f} : \Omega \rightarrow \mathbb{R}^m$ ist in $\mathcal{C}^{s+1}(\Omega, \mathbb{R}^m)$, wenn \vec{f} in Ω partiell differenzierbar ist und wenn gilt $\partial_1 \vec{f}, \partial_2 \vec{f}, \dots, \partial_n \vec{f} \in \mathcal{C}^s(\Omega, \mathbb{R}^m)$. \vec{f} heißt dann **$(s + 1)$ -mal stetig differenzierbar in Ω** .

16.7 Der Satz von Taylor

Zunächst erinnern wir uns an den **Satz von Taylor für reellwertige Funktionen einer Variablen** (siehe Satz 7.32 in Kapitel 7 im Skript der HM A), also für den Sonderfall $n = m = 1$:

Seien $I \subseteq \mathbb{R}$ ein offenes Intervall, $t_0 \in I$ und $f \in \mathcal{C}^{s+1}(I, \mathbb{R})$. Zu jedem $t \in I$ gibt es dann ein τ zwischen t und t_0 (welches in der Regel von t und t_0 abhängt), so dass gilt:

$$f(t) = \underbrace{\sum_{k=0}^s \frac{1}{k!} f^{(k)}(t_0) (t - t_0)^k}_{\text{Taylorpolynom vom Grad } s} + \underbrace{\frac{1}{(s+1)!} f^{(s+1)}(\tau) (t - t_0)^{s+1}}_{\text{Restglied}} \quad (16.11)$$

Um den Satz von Taylor verallgemeinern zu können, benötigen wir zunächst die sogenannte Multiindex-Schreibweise.

Notation 16.55. (Multiindex-Schreibweise)

Seien $\alpha_1, \dots, \alpha_n \in \mathbb{N}_0$. Dann heißt $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ ein **Multiindex**. Wir verwenden folgende die Notation:

$$\begin{aligned} |\alpha| &:= \alpha_1 + \alpha_2 + \dots + \alpha_n \\ \alpha! &:= \alpha_1! \cdot \alpha_2! \cdot \dots \cdot \alpha_n! \\ \vec{x}^\alpha &:= x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdot \dots \cdot x_n^{\alpha_n} \quad \text{für } \vec{x} = (x_1, \dots, x_n) \in \mathbb{R}^n. \\ \partial^\alpha \vec{f} &:= \partial_1^{\alpha_1} \partial_2^{\alpha_2} \dots \partial_n^{\alpha_n} \vec{f} \quad \text{für } \vec{f} \in \mathcal{C}^s(\Omega, \mathbb{R}^n), \text{ und } |\alpha| \leq s. \end{aligned}$$

Betrachten wir ein Beispiel, um uns mit der Multiindex-Schreibweise vertraut zu machen.

Beispiel 16.56. (Multiindexschreibweise)

Sei $n = 3$ und $\alpha = (3, 0, 1)$. Dann gelten

$$\begin{aligned} |\alpha| &= 3 + 0 + 1 = 4, \\ \alpha! &= 3! \cdot 0! \cdot 1! = 6 \cdot 1 \cdot 1 = 6, \\ \vec{x}^\alpha &= x_1^3 \cdot x_2^0 \cdot x_3^1 = x_1^3 x_3, \end{aligned}$$

und für die auf ganz \mathbb{R}^3 beliebig oft stetig differenzierbare Funktion

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}, \quad f(x_1, x_2, x_3) := e^{3x_1 + 2x_2^2 + x_3^3},$$

gilt

$$(\partial^\alpha f)(x_1, x_2, x_3) = (\partial_1^3 \partial_2^0 \partial_3^1 f)(x_1, x_2, x_3) = (\partial_1^3 \partial_3 f)(x_1, x_2, x_3) = 81 x_3^2 e^{3x_1 + 2x_2^2 + x_3^3}.$$

Als weitere Vorbereitung des Satzes von Taylor brauchen wir noch den Begriff einer konvexen Menge in \mathbb{R}^n .

Definition 16.57. (konvexe Menge in \mathbb{R}^n)

Eine Teilmenge M von \mathbb{R}^n heißt **konvex**, wenn für alle $\vec{x}, \vec{y} \in M$ die Verbindungsstrecke (Verbindungsgerade) von \vec{x} nach \vec{y} ganz in M verläuft.

Beispiel 16.58. (konvexe Menge)

(a) Die offene Kugel $U_r(\vec{\mathbf{a}})$ um $\vec{\mathbf{a}}$ mit Radius r in \mathbb{R}^n ist konvex, denn:

Seien $\vec{\mathbf{x}}$ und $\vec{\mathbf{y}}$ zwei beliebige Punkte aus $U_r(\vec{\mathbf{a}})$. Dann müssen wir zeigen, dass die Verbindungsstrecke

$$\vec{\mathbf{x}} + t(\vec{\mathbf{y}} - \vec{\mathbf{x}}) = (1-t)\vec{\mathbf{x}} + t\vec{\mathbf{y}}, \quad t \in [0, 1],$$

ganz in $U_r(\vec{\mathbf{a}})$ liegt. (Dieses ist anschaulich klar, wenn man sich ein Bild im \mathbb{R}^2 malt.) Für $t \in [0, 1]$ gilt mit der Dreiecksungleichung, dass

$$\begin{aligned} \|(1-t)\vec{\mathbf{x}} + t\vec{\mathbf{y}} - \vec{\mathbf{a}}\| &= \|(1-t)(\vec{\mathbf{x}} - \vec{\mathbf{a}}) + t(\vec{\mathbf{y}} - \vec{\mathbf{a}})\| \\ &\leq \|(1-t)(\vec{\mathbf{x}} - \vec{\mathbf{a}})\| + \|t(\vec{\mathbf{y}} - \vec{\mathbf{a}})\| \\ &\leq (1-t) \underbrace{\|\vec{\mathbf{x}} - \vec{\mathbf{a}}\|}_{< r} + t \underbrace{\|\vec{\mathbf{y}} - \vec{\mathbf{a}}\|}_{< r} \\ &< (1-t)r + tr = r, \end{aligned}$$

d.h. für jedes $t \in [0, 1]$ liegt $(1-t)\vec{\mathbf{x}} + t\vec{\mathbf{y}}$ in $U_r(\vec{\mathbf{a}})$. Also ist $U_r(\vec{\mathbf{a}})$ konvex.

(b) Mit einer analogen Rechnung zeigt man, dass $\overline{U_r(\vec{\mathbf{a}})}$ konvex ist.

(c) Der Rand $\partial U_r(\vec{\mathbf{a}}) = \{\vec{\mathbf{x}} \in \mathbb{R}^n : \|\vec{\mathbf{x}} - \vec{\mathbf{a}}\| = r\}$ ist **nicht** konvex, denn:

Wir geben den Nachweis der Einfachheit halber für den Fall $\vec{\mathbf{a}} = \vec{\mathbf{0}}$. Der Fall $\vec{\mathbf{a}} \neq \vec{\mathbf{0}}$ kann analog gezeigt werden (gleiche Idee), aber er wird technisch komplizierter.

Wir betrachten $\vec{\mathbf{x}} \in \partial U_r(\vec{\mathbf{0}})$. Dann ist auch $-\vec{\mathbf{x}} \in \partial U_r(\vec{\mathbf{0}})$, denn $\|-\vec{\mathbf{x}}\| = \|\vec{\mathbf{x}}\| = \|\vec{\mathbf{x}} - \vec{\mathbf{0}}\| = r$. Der Punkt

$$\vec{\mathbf{x}} + \frac{1}{2}((-\vec{\mathbf{x}}) - \vec{\mathbf{x}}) = \vec{\mathbf{x}} + \frac{1}{2}(-2\vec{\mathbf{x}}) = \vec{\mathbf{0}}$$

liegt auf der Verbindungsstrecke von $\vec{\mathbf{x}}$ und $-\vec{\mathbf{x}}$, aber er ist nicht in $\partial U_r(\vec{\mathbf{0}})$. Also ist $\partial U_r(\vec{\mathbf{0}})$ nicht konvex.

(d) Das L-förmige Gebiet

$$\begin{aligned} L &:= (]0, 1[\times]0, 2[) \cup (]0, 2[\times]0, 1[) \\ &= \{(x, y) \in \mathbb{R}^2 : 0 < x < 1, 0 < y < 2\} \\ &\quad \cup \{(x, y) \in \mathbb{R}^2 : 0 < x < 2, 0 < y < 1\} \end{aligned}$$

ist **nicht** konvex, denn man kann zeigen die Verbindungsstrecke von $(\frac{1}{2}, \frac{3}{2})$ und $(\frac{3}{2}, \frac{1}{2})$ aus L den Punkt $(1, 1)$ enthält, der nicht zu L gehört.

Ist Ω offen und konvex, so ist Ω ein Gebiet. Umgekehrt gibt es aber viele Gebiete, die nicht konvex sind!

Nach diesen Vorbereitungen können wir den Satz von Taylor für reellwertige Funktionen mehrerer Variablen formulieren.

Satz 16.59. (Satz von Taylor/Taylor-Formel in n Variablen)

Sei $\Omega \subseteq \mathbb{R}^n$ offen und konvex, und seien $f \in C^{s+1}(\Omega, \mathbb{R})$ und $\vec{a}, \vec{x} \in \Omega$. Dann gibt es $\vec{z} \in \Omega$ auf der Verbindungsstrecke zwischen \vec{a} und \vec{x} mit

$$f(\vec{x}) = \underbrace{\sum_{|\alpha| \leq s} \frac{1}{\alpha!} (\partial^\alpha f)(\vec{a}) (\vec{x} - \vec{a})^\alpha}_{=: T_s(\vec{x}; \vec{a})} + \underbrace{\sum_{|\alpha|=s+1} \frac{1}{\alpha!} (\partial^\alpha f)(\vec{z}) (\vec{x} - \vec{a})^\alpha}_{\text{Restglied}}.$$

$T_s(\vec{x}; \vec{a})$ heißt das **Taylorpolynom von f vom Grad s um den Entwicklungspunkt \vec{a}** . Der Punkt \vec{z} hängt normalerweise von \vec{x} und \vec{a} ab.

Beweis von Satz 16.59: Sei $I \subseteq \mathbb{R}$ ein maximales offenes Intervall mit

$$\vec{y}(t) := \vec{a} + t(\vec{x} - \vec{a}) \in \Omega \quad \text{für alle } t \in I.$$

Dann ist $\vec{y} : I \rightarrow \mathbb{R}^n$ ein beliebig oft stetig differenzierbarer Weg, also $\vec{y} \in C^\infty(I, \mathbb{R}^n)$. Insbesondere gilt $[0, 1] \subseteq I$, da Ω konvex ist und somit die Verbindungsstrecke von \vec{a} und \vec{x} ganz in Ω liegt.

Wir betrachten nun die Hilfsfunktion

$$h : I \rightarrow \mathbb{R}, \quad h(t) := f(\vec{y}(t)) = f(\vec{a} + t(\vec{x} - \vec{a})).$$

Nach der Kettenregel ist $h \in C^{s+1}(I, \mathbb{R})$, weil $f \in C^{s+1}(\Omega, \mathbb{R})$ und $\vec{y} \in C^\infty(I, \mathbb{R}^n)$ gilt und weil $h(t) = f(\vec{y}(t)) = (f \circ \vec{y})(t)$ ist. Weiter gilt $h(0) = f(\vec{a})$ und $h(1) = f(\vec{x})$.

Mit der Kettenregel erhält man

$$\begin{aligned} h'(t) &= f'(\vec{y}(t)) \cdot \vec{y}'(t) = f'(\vec{a} + t(\vec{x} - \vec{a})) \cdot (\vec{x} - \vec{a}) \\ &= \sum_{k=1}^n (\partial_k f)(\vec{a} + t(\vec{x} - \vec{a})) (x_k - a_k) \\ &= \sum_{|\alpha|=1} (\partial^\alpha f)(\vec{a} + t(\vec{x} - \vec{a})) (\vec{x} - \vec{a})^\alpha \end{aligned}$$

$$\begin{aligned}
h''(t) &= \sum_{k=1}^n \sum_{j=1}^n (\partial_j \partial_k f)(\vec{a} + t(\vec{x} - \vec{a})) (x_j - a_j) (x_k - a_k) \\
&= \sum_{|\alpha|=2} \frac{2!}{\alpha!} (\partial^\alpha f)(\vec{a} + t(\vec{x} - \vec{a})) (\vec{x} - \vec{a})^\alpha.
\end{aligned}$$

Setzen wir den Prozess fort, so finden wir allgemein für die k -te Ableitung:

$$h^{(k)}(t) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} (\partial^\alpha f)(\vec{a} + t(\vec{x} - \vec{a})) (\vec{x} - \vec{a})^\alpha, \quad k = 0, 1, 2, \dots, s+1.$$

(Um sich die Faktoren $k!/\alpha!$ zu überlegen braucht man etwas Kombinatorik.) Der Satz von Taylor (16.11) mit $t = 1$ und $t_0 = 0$ angewendet auf die Funktion h liefert nun

$$\begin{aligned}
f(\vec{x}) = h(1) &= \sum_{k=0}^s \frac{1}{k!} h^{(k)}(0) (1-0)^k + \frac{1}{(s+1)!} h^{(s+1)}(\tau) (1-0)^{s+1} \\
&= \sum_{k=0}^s \frac{1}{k!} \sum_{|\alpha|=k} \frac{k!}{\alpha!} (\partial^\alpha f)(\vec{a}) (\vec{x} - \vec{a})^\alpha \\
&\quad + \frac{1}{(s+1)!} \sum_{|\alpha|=s+1} \frac{(s+1)!}{\alpha!} (\partial^\alpha f)(\vec{a} + \tau(\vec{x} - \vec{a})) (\vec{x} - \vec{a})^\alpha \\
&= \sum_{|\alpha| \leq s} \frac{1}{\alpha!} (\partial^\alpha f)(\vec{a}) (\vec{x} - \vec{a})^\alpha + \sum_{|\alpha|=s+1} \frac{1}{\alpha!} (\partial^\alpha f)(\vec{z}) (\vec{x} - \vec{a})^\alpha,
\end{aligned}$$

wobei $\tau \in]0, 1[$ ist und damit $\vec{z} := \vec{a} + \tau(\vec{x} - \vec{a})$ auf der Strecke zwischen \vec{a} und \vec{x} liegt. \square

Beispiel 16.60. (Satz von Taylor)

Seien $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := \sin(3x + y^2)$ und $\vec{a} = (0, 0)$ sowie $s = 3$.

Aufgabe: Bestimme das Taylorpolynom dritten Grades $T_3((x, y); (0, 0))$ um den Entwicklungspunkt $\vec{a} = (0, 0)$.

Lösung: Wir berechnen die partielle Ableitungen für $|\alpha| \leq 3$:

$$\begin{aligned}
\alpha = (1, 0) : & \quad (\partial_1 f)(x, y) = 3 \cos(3x + y^2) \\
\alpha = (0, 1) : & \quad (\partial_2 f)(x, y) = 2y \cos(3x + y^2) \\
\alpha = (2, 0) : & \quad (\partial_1^2 f)(x, y) = -9 \sin(3x + y^2) \\
\alpha = (1, 1) : & \quad (\partial_1 \partial_2 f)(x, y) = -6y \sin(3x + y^2)
\end{aligned}$$

$$\begin{aligned} \alpha = (0, 2) : \quad & (\partial_2^2 f)(x, y) = 2 \cos(3x + y^2) - 4y^2 \sin(3x + y^2) \\ \alpha = (3, 0) : \quad & (\partial_1^3 f)(x, y) = -27 \cos(3x + y^2) \\ \alpha = (2, 1) : \quad & (\partial_1^2 \partial_2 f)(x, y) = -18y \cos(3x + y^2) \\ \alpha = (1, 2) : \quad & (\partial_1 \partial_2^2 f)(x, y) = -6 \sin(3x + y^2) - 12y^2 \cos(3x + y^2) \\ \alpha = (0, 3) : \quad & (\partial_2^3 f)(x, y) = -12y \sin(3x + y^2) - 8y^3 \cos(3x + y^2) \end{aligned}$$

Also lautet das Taylorpolynom vom Grad 3 um $\vec{\mathbf{a}} = (0, 0)$

$$\begin{aligned} T_3((x, y); (0, 0)) &= \sum_{|\alpha| \leq 3} \frac{(\partial^\alpha f)(0, 0)}{\alpha!} ((x, y) - (0, 0))^\alpha = \sum_{|\alpha| \leq 3} \frac{(\partial^\alpha f)(0, 0)}{\alpha!} (x, y)^\alpha \\ &= f(0, 0) + \frac{1}{1!0!} (\partial_1 f)(0, 0) x^1 y^0 + \frac{1}{0!1!} (\partial_2 f)(0, 0) x^0 y^1 \\ &\quad + \frac{1}{2!0!} (\partial_1^2 f)(0, 0) x^2 y^0 + \frac{1}{1!1!} (\partial_1 \partial_2 f)(0, 0) x^1 y^1 + \frac{1}{0!2!} (\partial_2^2 f)(0, 0) x^0 y^2 \\ &\quad + \frac{1}{3!0!} (\partial_1^3 f)(0, 0) x^3 y^0 + \frac{1}{2!1!} (\partial_1^2 \partial_2 f)(0, 0) x^2 y^1 \\ &\quad + \frac{1}{1!2!} (\partial_1 \partial_2^2 f)(0, 0) x^1 y^2 + \frac{1}{0!3!} (\partial_2^3 f)(0, 0) x^0 y^3 \\ &= 0 + 3x + 0 \cdot y + \frac{1}{2} \cdot 0 \cdot x^2 + 0 \cdot xy + \frac{1}{2} \cdot 2y^2 \\ &\quad + \frac{1}{6} \cdot (-27)x^3 + \frac{1}{2} \cdot 0 \cdot x^2 y + \frac{1}{2} \cdot 0 \cdot xy^2 + \frac{1}{6} \cdot 0 \cdot y^3 \\ &= 3x + y^2 - \frac{9}{2}x^3. \end{aligned}$$

Im Kontext des Satzes von Taylor lernen wir als Letztes noch die Hesse-Matrix der zweiten Ableitungen kennen.

Bemerkung 16.61. (Hesse-Matrix)

Sei $\Omega \subseteq \mathbb{R}^n$ offen und konvex, und sei $f \in \mathcal{C}^2(\Omega, \mathbb{R})$. Für $\vec{\mathbf{a}}, \vec{\mathbf{x}} \in \Omega$ ist das Taylorpolynom ersten Grades von f um $\vec{\mathbf{a}}$ durch

$$\begin{aligned} T_1(\vec{\mathbf{x}}; \vec{\mathbf{a}}) &= f(\vec{\mathbf{a}}) + \sum_{|\alpha|=1} \frac{1}{\alpha!} (\partial^\alpha f)(\vec{\mathbf{a}}) (\vec{\mathbf{x}} - \vec{\mathbf{a}})^\alpha \\ &= f(\vec{\mathbf{a}}) + \sum_{k=1}^n (\partial_k f)(\vec{\mathbf{a}}) (x_k - a_k) \\ &= f(\vec{\mathbf{a}}) + f'(\vec{\mathbf{a}}) \cdot (\vec{\mathbf{x}} - \vec{\mathbf{a}}) \end{aligned}$$

gegeben. Nach dem Satz von Taylor mit $s = 1$ existiert \vec{z} auf der Verbindungsstrecke von \vec{a} und \vec{x} mit

$$\begin{aligned} f(\vec{x}) - T_1(\vec{x}; \vec{a}) &= \sum_{|\alpha|=2} \frac{1}{\alpha!} (\partial^\alpha f)(\vec{z}) (\vec{x} - \vec{a})^\alpha \\ &= \frac{1}{2} \sum_{i=1}^n \sum_{k=1}^n (\partial_i \partial_k f)(\vec{z}) (x_i - a_i) (x_k - a_k) \\ &= \frac{1}{2} \langle \vec{x} - \vec{a} | \mathbf{A} (\vec{x} - \vec{a}) \rangle \end{aligned}$$

wobei

$$\mathbf{A} = \begin{bmatrix} (\partial_1 \partial_1 f)(\vec{z}) & (\partial_1 \partial_2 f)(\vec{z}) & \cdots & (\partial_1 \partial_n f)(\vec{z}) \\ (\partial_2 \partial_1 f)(\vec{z}) & (\partial_2 \partial_2 f)(\vec{z}) & \cdots & (\partial_2 \partial_n f)(\vec{z}) \\ \vdots & \vdots & & \vdots \\ (\partial_n \partial_1 f)(\vec{z}) & (\partial_n \partial_2 f)(\vec{z}) & \cdots & (\partial_n \partial_n f)(\vec{z}) \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

$(\mathbf{H}f)(\vec{z}) := \mathbf{A}$ heißt die **Hesse-Matrix von f in \vec{z}** .

Berechnen wir die Hesse-Matrix für ein Beispiel.

Beispiel 16.62. (Hesse-Matrix)

Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = x^2 y^3$. Nach Beispiel 16.51 sind dann die partiellen Ableitungen zweiter Ordnung

$$\begin{aligned} (\partial_1^2 f)(x, y) &= 2y^3, & (\partial_1 \partial_2 f)(x, y) &= 6xy^2, \\ (\partial_2 \partial_1 f)(x, y) &= 6xy^2, & (\partial_2^2 f)(x, y) &= 6x^2y. \end{aligned}$$

Also ist die Hesse-Matrix von f

$$(\mathbf{H}f)(x, y) = \begin{bmatrix} 2y^3 & 6xy^2 \\ 6xy^2 & 6x^2y \end{bmatrix}.$$

Extremwertaufgaben

In diesem Kapitel interessieren wir uns für lokale und globale Extremalstellen einer Funktion, d.h. Punkte in der Definitionsmenge, an denen die Funktion lokal oder global maximale bzw. minimale Funktionswerte annimmt. Handelt es sich um globale Extremalstellen, so nennen wir die zugehörigen maximalen bzw. minimalen Funktionswerte Extrema der Funktion, genauer das Maximum bzw. das Minimum der Funktion.

Wie findet man die lokalen und globalen Extremalstellen? – Wir werden diese Frage zunächst für zweimal stetig differenzierbare Funktionen auf einer offenen Definitionsmenge mit Hilfe der ersten partiellen Ableitungen und der Hesse-Matrix untersuchen. Danach lernen wir, wie man mit Hilfe von Nebenbedingungen dieses Problem für Funktionen auf sogenannten kompakten Mengen untersuchen kann. Dabei lernen wir die Methode der Lagrange-Multiplikatoren kennen.

17.1 Lokale Extremalstellen

Wir starten mit der Definition von lokalen Extremalstellen von reellwertigen Funktionen auf einer offenen Teilmenge von \mathbb{R}^n . Diese Definition ist analog zu derjenigen für Funktionen einer Variablen (siehe Kapitel 7 im Skript der HM A).

Definition 17.1. (lokale Extremalstellen und kritische Punkte)

Seien $\Omega \subseteq \mathbb{R}^n$ offen und $f \in \mathcal{C}^1(\Omega, \mathbb{R})$.

- (1) $\vec{\mathbf{a}} \in \Omega$ heißt eine **lokale Maximalstelle** von f , wenn ein $r > 0$ existiert mit $U_r(\vec{\mathbf{a}}) \subseteq \Omega$ und $f(\vec{\mathbf{x}}) \leq f(\vec{\mathbf{a}})$ für alle $\vec{\mathbf{x}} \in U_r(\vec{\mathbf{a}})$.
- (2) $\vec{\mathbf{a}} \in \Omega$ heißt eine **lokale Minimalstelle** von f , wenn ein $r > 0$ existiert mit $U_r(\vec{\mathbf{a}}) \subseteq \Omega$ und $f(\vec{\mathbf{x}}) \geq f(\vec{\mathbf{a}})$ für alle $\vec{\mathbf{x}} \in U_r(\vec{\mathbf{a}})$.
- (3) $\vec{\mathbf{a}} \in \Omega$ heißt eine **lokale Extremalstelle** von f , wenn $\vec{\mathbf{a}}$ eine lokale Maximalstelle oder eine lokale Minimalstelle von f ist.
- (4) $\vec{\mathbf{a}} \in \Omega$ heißt ein **kritischer Punkt** von f , wenn $f'(\vec{\mathbf{a}}) = \mathbf{O}_{1 \times n}$ ist.

Der nächste Satz stellt einen Zusammenhang zwischen Extremalstellen und kritischen Punkten einer \mathcal{C}^1 -Funktion her.

Satz 17.2. (lokale Extremalstelle von \mathcal{C}^1 -Fkt. \implies kritischer Punkt)

Lokale Extremalstellen von \mathcal{C}^1 -Funktionen sind kritische Punkte.

Satz 17.2 liefert also eine **notwendige Bedingung** für lokale Extremalstellen $\vec{\mathbf{a}}$ einer \mathcal{C}^1 -Funktion f : Es muss gelten $f'(\vec{\mathbf{a}}) = \mathbf{O}_{1 \times n}$. – Allerdings folgt aus der Bedingung $f'(\vec{\mathbf{a}}) = \mathbf{O}_{1 \times n}$ noch **nicht**, dass $\vec{\mathbf{a}}$ eine lokale Extremalstelle ist. Wir haben lediglich einen möglichen Kandidaten für eine solche identifiziert.

Beweis von Satz 17.2: Sei $\vec{\mathbf{a}}$ eine lokale Extremalstelle von $f \in \mathcal{C}^1(\Omega, \mathbb{R})$. Wir wählen $r > 0$ mit $U_r(\vec{\mathbf{a}}) \subseteq \Omega$ und mit $f(\vec{\mathbf{x}}) \leq f(\vec{\mathbf{a}})$ bzw. $f(\vec{\mathbf{x}}) \geq f(\vec{\mathbf{a}})$ für $\vec{\mathbf{x}} \in U_r(\vec{\mathbf{a}})$. Für $\vec{\mathbf{v}} \in \mathbb{R}^n$ mit $\|\vec{\mathbf{v}}\| = 1$ betrachten wir die \mathcal{C}^1 -Funktion einer Variablen

$$h(t) := f(\vec{\mathbf{a}} + t\vec{\mathbf{v}}), \quad t \in]-r, r[.$$

Dann hat $h :]-r, r[\rightarrow \mathbb{R}$ in $t = 0$ eine lokale Extremalstelle, also ist $h'(0) = 0$ (siehe Kapitel 7 im Skript der HM A). Mit der Kettenregel folgt dann

$$h'(t) = f'(\vec{\mathbf{a}} + t\vec{\mathbf{v}}) \cdot \vec{\mathbf{v}}, \quad t \in]-r, r[.$$

Einsetzen von $t = 0$ liefert also:

$$0 = h'(0) = f'(\vec{\mathbf{a}}) \cdot \vec{\mathbf{v}} \quad \text{für jedes } \vec{\mathbf{v}} \in \mathbb{R}^n \text{ mit } \|\vec{\mathbf{v}}\| = 1.$$

Insbesondere gilt für $\vec{\mathbf{v}} = \vec{\mathbf{e}}_k$, $k = 1, 2, \dots, n$, dass

$$0 = f'(\vec{\mathbf{a}}) \cdot \vec{\mathbf{e}}_k = (\partial_k f)(\vec{\mathbf{a}}),$$

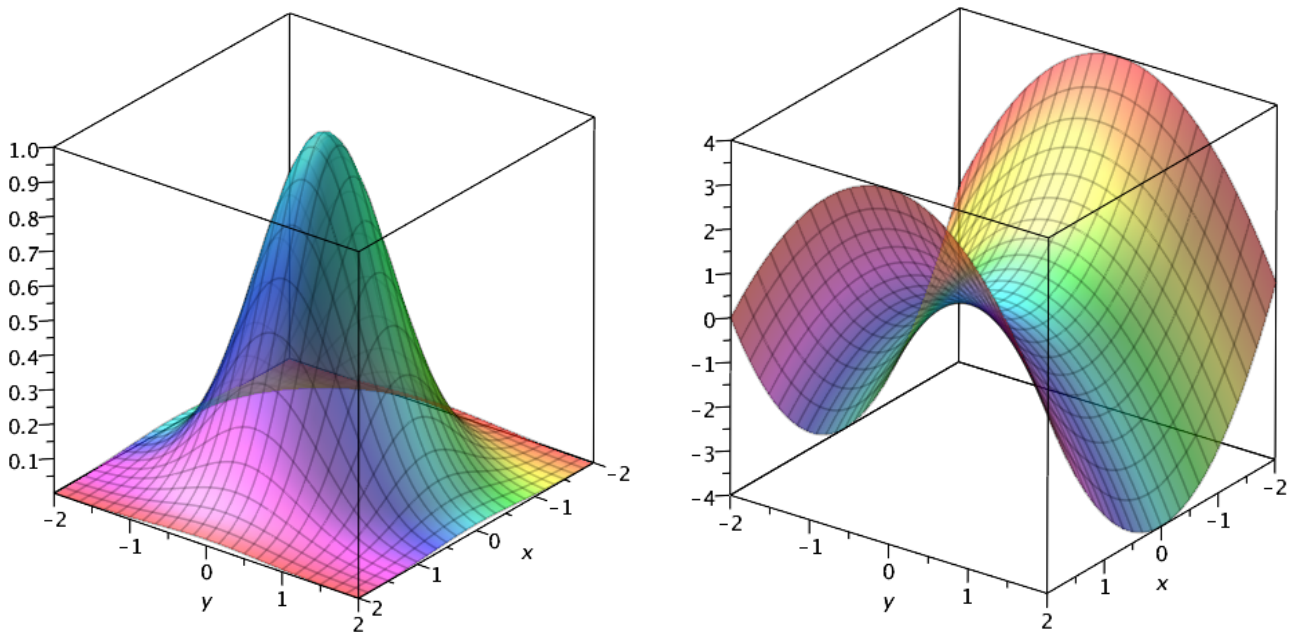


Abbildung 17.1: Die Graphen der Funktionen $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := e^{-(x^2+y^2)}$, (linkes Bild) und $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := x^2 - y^2$, (rechtes Bild).

d.h. alle partiellen Ableitungen in $\vec{\mathbf{a}}$ sind Null. Also ist

$$f'(\vec{\mathbf{a}}) = [(\partial_1 f)(\vec{\mathbf{a}}) \quad (\partial_2 f)(\vec{\mathbf{a}}) \quad \cdots \quad (\partial_n f)(\vec{\mathbf{a}})] = [0 \quad 0 \quad \cdots \quad 0] = \mathbf{O}_{1 \times n}. \quad \square$$

Betrachten wir einige Beispiele. Wenn wir hierbei von einer \mathcal{C}^∞ -Funktion sprechen, genauer einer Funktion in $\mathcal{C}^\infty(\Omega, \mathbb{R})$, so meinen wir hiermit analog zu $\mathcal{C}^k(\Omega, \mathbb{R})$ eine reellwertige Funktion, die beliebig oft (also unendlich oft) auf Ω stetig differenzierbar ist.

Beispiel 17.3. (lokale Extremalstellen und kritische Punkte)

(a) Die \mathcal{C}^∞ -Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := e^{-(x^2+y^2)}$, hat die Ableitung

$$f'(x, y) = [-2x e^{-(x^2+y^2)} \quad -2y e^{-(x^2+y^2)}].$$

Die nach Satz 17.2 notwendige Bedingung für eine lokale Extremalstelle ist

$$f'(x, y) = [-2x e^{-(x^2+y^2)} \quad -2y e^{-(x^2+y^2)}] = [0 \quad 0],$$

und diese liefert

$$(-2x e^{-(x^2+y^2)} = 0 \text{ und } -2y e^{-(x^2+y^2)} = 0) \iff (x = 0 \text{ und } y = 0),$$

d.h. $(0, 0)$ ist der einzige kritische Punkt von f .

Frage: Ist $(0, 0)$ eine lokale Extremalstelle von f ?

Für $(x, y) \in \mathbb{R}^2$ gilt

$$f(x, y) = e^{-(x^2+y^2)} \leq 1 = e^0 = f(0, 0), \quad \text{weil } -(x^2 + y^2) \leq 0 \text{ ist.}$$

Also ist $(0, 0)$ eine lokale (und auch globale) Maximalstelle von f .

Der Graph der Funktion ist im linken Bild in Abbildung 17.1 gezeichnet.

(b) Die \mathcal{C}^∞ -Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := x^2 - y^2$, hat die Ableitung

$$f'(x, y) = [2x \quad -2y].$$

Die nach Satz 17.2 notwendige Bedingung für eine lokale Extremalstelle ist

$$f'(x, y) = [2x \quad -2y] = [0 \quad 0],$$

und diese liefert

$$(2x = 0 \quad \text{und} \quad -2y = 0) \quad \iff \quad x = y = 0,$$

d.h. $(0, 0)$ ist der einzige kritische Punkt von f .

Frage: Ist $(0, 0)$ eine lokale Extremalstelle von f ?

Für $x \neq 0$ gilt: $f(x, 0) = x^2 > 0 = f(0, 0)$.

Für $y \neq 0$ gilt: $f(0, y) = -y^2 < 0 = f(0, 0)$.

Also ist $(0, 0)$ weder eine lokale Maximalstelle noch eine lokale Minimalstelle.

Der Graph der Funktion ist im rechten Bild in Abbildung 17.1 gezeichnet.

(c) Die \mathcal{C}^∞ -Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := xy$, hat die Ableitung

$$f'(x, y) = [y \quad x].$$

Die nach Satz 17.2 notwendige Bedingung für eine lokale Extremalstelle

$$f'(x, y) = [y \quad x] = [0 \quad 0]$$

liefert den einzigen kritischen Punkt $(0, 0)$ von f .

Frage: Ist $(0, 0)$ eine lokale Extremalstelle von f ?

Für $x > 0$ und $y < 0$ gilt: $f(x, y) = xy < 0 = f(0, 0)$.

Für $x > 0$ und $y > 0$ gilt: $f(x, y) = xy > 0 = f(0, 0)$.

Also ist $(0, 0)$ keine lokale Extremalstelle von f .

Der Graph der Funktion ist im linken Bild in Abbildung 17.2 gezeichnet.

- (d) Die \mathcal{C}^∞ -Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := -\cos((x-1)^2 + (y-1)^2)$, hat die Ableitung

$$\begin{aligned} f'(x, y) &= \left[2(x-1) \sin((x-1)^2 + (y-1)^2) \quad 2(y-1) \sin((x-1)^2 + (y-1)^2) \right]. \end{aligned}$$

Die nach Satz 17.2 notwendige Bedingung für eine lokale Extremalstelle

$$f'(x, y) = [0 \quad 0]$$

liefert

$$\begin{aligned} &\left((x=1 \text{ oder } \sin((x-1)^2 + (y-1)^2) = 0) \right. \\ &\quad \left. \text{und } (y=1 \text{ oder } \sin((x-1)^2 + (y-1)^2) = 0) \right) \\ \iff &\left((x=1 \text{ oder } (x-1)^2 + (y-1)^2 = k\pi \text{ mit } k \in \mathbb{N}_0) \right. \\ &\quad \left. \text{und } (y=1 \text{ oder } (x-1)^2 + (y-1)^2 = k\pi \text{ mit } k \in \mathbb{N}_0) \right). \end{aligned}$$

Also sind $(1, 1)$ und alle Punkte auf $\partial U_{\sqrt{k\pi}}(1, 1)$ mit $k \in \mathbb{N}$, d.h. alle Punkte auf den Kreisen mit Radius $\sqrt{k\pi}$, $k \in \mathbb{N}$, um $(1, 1)$, kritische Punkte von f .

Frage: Welche dieser kritischen Punkte sind lokale Extremalstellen von f ?

Der Punkt $(1, 1)$ ist eine lokale Minimalstelle, denn für alle $(x, y) \in \mathbb{R}^2$ gilt

$$\begin{aligned} f(x, y) &= -\cos((x-1)^2 + (y-1)^2) \\ &\geq -1 = -\cos(0) = -\cos((1-1)^2 + (1-1)^2). \end{aligned}$$

Betrachten wir ein $(x_k, y_k) \in \partial U_{\sqrt{k\pi}}(1, 1)$ mit $k \in \mathbb{N}$, so gilt

$$f(x_k, y_k) = -\cos((x_k-1)^2 + (y_k-1)^2) = -\cos(k\pi) = (-1)^{k+1}.$$

Also gilt für alle $(x, y) \in \mathbb{R}^2$ für gerades $k \in \mathbb{N}$

$$\begin{aligned} f(x, y) &= -\cos((x-1)^2 + (y-1)^2) \\ &\geq -1 = (-1)^{k+1} = -\cos((x_k-1)^2 + (y_k-1)^2). \end{aligned}$$

Also gilt für alle $(x, y) \in \mathbb{R}^2$ für ungerades $k \in \mathbb{N}$

$$\begin{aligned} f(x, y) &= -\cos((x-1)^2 + (y-1)^2) \\ &\leq 1 = (-1)^{k+1} = -\cos((x_k-1)^2 + (y_k-1)^2). \end{aligned}$$

Also sind alle Punkte auf $\partial U_{\sqrt{k\pi}}(1, 1)$ mit $k \in \mathbb{N}$ ungerade lokale Maximalstellen und alle Punkte auf $\partial U_{\sqrt{k\pi}}(1, 1)$ mit $k \in \mathbb{N}$ gerade lokale Minimalstellen von f .

Der Graph der Funktion ist im rechten Bild in Abbildung 17.2 gezeichnet.

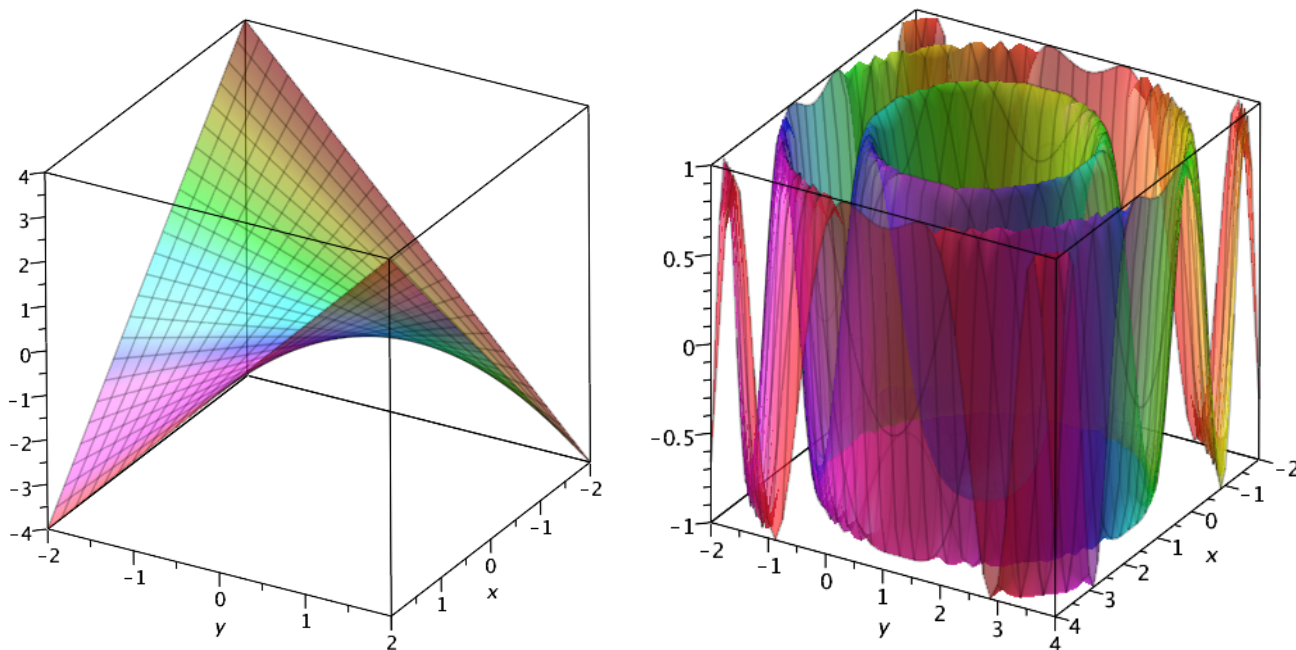


Abbildung 17.2: Die Graphen der Funktionen $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := xy$, (linkes Bild) und $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := -\cos((x-1)^2 + (y-1)^2)$, (rechtes Bild).

Die folgende Bemerkung liefert die Motivation für das nachfolgende Teilkapitel.

Bemerkung 17.4. (lokale Extremalstellen und die Hesse-Matrix)

Seien $\Omega \subseteq \mathbb{R}^n$ offen und $f \in \mathcal{C}^2(\Omega, \mathbb{R})$. Sei $\vec{a} \in \Omega$ ein kritischer Punkt von f , d.h. $f'(\vec{a}) = \mathbf{O}_{1 \times n}$. Weiter sei $r > 0$ so gewählt, dass $U_r(\vec{a}) \subseteq \Omega$ ist.

- (1) Für $\vec{x} \in U_r(\vec{a})$ existiert nach dem Satz von Taylor (siehe Satz 16.59 und Bemerkung 16.61) ein \vec{z} auf der Verbindungsstrecke von \vec{a} und \vec{x} , so dass gilt

$$\begin{aligned} f(\vec{x}) &= f(\vec{a}) + \underbrace{f'(\vec{a})}_{=\mathbf{O}_{1 \times n}} (\vec{x} - \vec{a}) + \frac{1}{2} \langle \vec{x} - \vec{a} | (\mathbf{H}f)(\vec{z}) (\vec{x} - \vec{a}) \rangle \\ &= f(\vec{a}) + \frac{1}{2} \langle \vec{x} - \vec{a} | (\mathbf{H}f)(\vec{z}) (\vec{x} - \vec{a}) \rangle, \end{aligned} \quad (17.1)$$

wobei $(\mathbf{H}f)(\vec{z})$ die **Hesse-Matrix** der zweiten partiellen Ableitungen (siehe Bemerkung 16.61) im Punkt \vec{z} ist. Falls $\varepsilon > 0$ existiert mit

$$\langle \vec{x} - \vec{a} | (\mathbf{H}f)(\vec{z}) (\vec{x} - \vec{a}) \rangle > 0 \quad \text{für alle } \vec{x} \in U_\varepsilon(\vec{a}) \setminus \{\vec{a}\},$$

dann folgt aus (17.1), dass $f(\vec{x}) > f(\vec{a})$ für alle $\vec{x} \in U_\varepsilon(\vec{a}) \setminus \{\vec{a}\}$ ist, d.h. \vec{a} ist eine **lokale Minimalstelle**.

Falls $\varepsilon > 0$ existiert mit

$$\langle \vec{x} - \vec{a} | (\mathbf{H}f)(\vec{z})(\vec{x} - \vec{a}) \rangle < 0 \quad \text{für alle } \vec{x} \in U_\varepsilon(\vec{a}) \setminus \{\vec{a}\},$$

dann folgt aus (17.1), dass $f(\vec{x}) < f(\vec{a})$ für alle $\vec{x} \in U_\varepsilon(\vec{a}) \setminus \{\vec{a}\}$ ist, d.h. \vec{a} ist eine **lokale Maximalstelle**.

- (2) Da $f \in \mathcal{C}^2(\Omega, \mathbb{R})$ ist, gilt $\partial_i \partial_k f = \partial_k \partial_i f$ für alle $i, k = 1, 2, \dots, n$ nach dem Satz von Schwarz (siehe Satz 16.52). Daher gilt ist $((\mathbf{H}f)(\vec{x}))^T = (\mathbf{H}f)(\vec{x})$, d.h. die Hesse-Matrix $(\mathbf{H}f)(\vec{x})$ ist eine **symmetrische Matrix**. Deshalb betrachten wir im nachfolgenden Teilkapitel symmetrische Matrizen.

17.2 Symmetrische Matrizen

In diesem Teilkapitel lernen wir symmetrische positiv definite, negativ definite und indefinite Matrizen kennen.

Definition 17.5. (symmetrische Matrix)

Eine Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ heißt **symmetrisch**, wenn $\mathbf{A}^T = \mathbf{A}$ gilt.

In Worten: Eine reelle Matrix ist symmetrisch, wenn sie quadratisch ist und bei Spiegelung an der Diagonalen (von links oben nach rechts unten) wieder sich selbst ergibt. – Betrachten wir einige Beispiele für symmetrische Matrizen.

Beispiel 17.6. (symmetrische Matrix)

(a) $\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix}$ ist symmetrisch.

(b) $\mathbf{B} = \begin{bmatrix} 1 & 3 \\ -3 & 2 \end{bmatrix}$ ist nicht symmetrisch.

(c) $\mathbf{C} = \begin{bmatrix} 1 & -2 & -3 \\ -2 & 4 & 2 \\ -3 & 2 & -5 \end{bmatrix}$ ist symmetrisch.

(d) Jede Diagonalmatrix ist symmetrisch.

Die nächste Bemerkung hält eine interessante Eigenschaft symmetrischer Matrizen fest.

Bemerkung 17.7. (wichtige Eigenschaft symmetrischer Matrizen)

Ist $\mathbf{A} \in \mathbb{R}^{n \times n}$ symmetrisch, dann gilt für alle $\vec{x}, \vec{y} \in \mathbb{C}^n$

$$\langle \mathbf{A} \vec{x} | \vec{y} \rangle = \langle \vec{x} | \mathbf{A} \vec{y} \rangle.$$

Beweis: Das Standardskalarprodukt zweier Vektoren $\vec{x}, \vec{y} \in \mathbb{C}^n$ können wir mit der Matrizenmultiplikation auch schreiben als

$$\langle \vec{x} | \vec{y} \rangle = \overline{x_1} y_1 + \overline{x_2} y_2 + \dots + \overline{x_n} y_n = \begin{bmatrix} \overline{x_1} & \overline{x_2} & \dots & \overline{x_n} \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \overline{\vec{x}}^T \cdot \vec{y}.$$

Damit folgt

$$\langle \mathbf{A} \vec{x} | \vec{y} \rangle = (\overline{\mathbf{A} \vec{x}})^T \vec{y} = (\overline{\mathbf{A}} \overline{\vec{x}})^T \vec{y} = \overline{\vec{x}}^T \overline{\mathbf{A}}^T \vec{y} = \overline{\vec{x}}^T (\mathbf{A} \vec{y}) = \langle \vec{x} | \mathbf{A} \vec{y} \rangle,$$

wobei wir im vierten Schritt $\overline{\mathbf{A}} = \mathbf{A}$ und damit $\overline{\mathbf{A}}^T = \mathbf{A}^T = \mathbf{A}$ (weil $\mathbf{A} \in \mathbb{R}^{n \times n}$ symmetrisch ist) genutzt haben.

Nun lernen wir zwei zentrale Sätze über symmetrische Matrizen kennen.

Satz 17.8. (Eigenwerte und Eigenvektoren einer symm. Matrix)

Sei $\mathbf{A} \in \mathbb{R}^{n \times n}$ *symmetrisch*. Dann gelten:

- (1) Alle Eigenwerte von \mathbf{A} sind *reell*.
- (2) Eigenvektoren zu *verschiedenen* Eigenwerten von \mathbf{A} sind *orthogonal*.

Wir beweisen diesen Satz, weil der Beweis einfach und illustrativ ist.

Beweis von Satz 17.8:

- (1) Sei $\lambda \in \mathbb{C}$ ein Eigenwert von \mathbf{A} mit einem Eigenvektor $\vec{x} \in \mathbb{C}^n \setminus \{\vec{0}\}$, also $\mathbf{A} \vec{x} = \lambda \vec{x}$. Da $\mathbf{A} \in \mathbb{R}^{n \times n}$ symmetrisch ist, gilt nach Bemerkung 17.7

$$\langle \mathbf{A} \vec{x} | \vec{x} \rangle = \langle \vec{x} | \mathbf{A} \vec{x} \rangle. \quad (17.2)$$

Weil \vec{x} ein Eigenvektor zum Eigenwert $\lambda \in \mathbb{C}$ ist, gilt

$$\begin{aligned}\langle \mathbf{A} \vec{x} | \vec{x} \rangle &= \langle \lambda \vec{x} | \vec{x} \rangle = \bar{\lambda} \langle \vec{x} | \vec{x} \rangle = \bar{\lambda} \|\vec{x}\|^2, \\ \langle \vec{x} | \mathbf{A} \vec{x} \rangle &= \langle \vec{x} | \lambda \vec{x} \rangle = \lambda \langle \vec{x} | \vec{x} \rangle = \lambda \|\vec{x}\|^2.\end{aligned}$$

Einsetzen in (17.2) liefert

$$\bar{\lambda} \|\vec{x}\|^2 = \lambda \|\vec{x}\|^2.$$

Da $\vec{x} \neq \vec{0}$ und damit $\|\vec{x}\| \neq 0$ ist, folgt $\lambda = \bar{\lambda}$, d.h. $\lambda \in \mathbb{R}$.

- (2) Die symmetrische Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ habe (mindestens) zwei verschiedene Eigenwerte λ, μ mit $\lambda \neq \mu$. Nach (1) sind λ und μ reell. Sei $\vec{x} \in \mathbb{C}^n \setminus \{\vec{0}\}$ ein Eigenvektor zu λ , und sei $\vec{y} \in \mathbb{C}^n \setminus \{\vec{0}\}$ ein Eigenvektor zu μ . Dann gilt nach Bemerkung 17.7

$$\langle \mathbf{A} \vec{x} | \vec{y} \rangle = \langle \vec{x} | \mathbf{A} \vec{y} \rangle, \quad (17.3)$$

und wir finden

$$\begin{aligned}\langle \mathbf{A} \vec{x} | \vec{y} \rangle &= \langle \lambda \vec{x} | \vec{y} \rangle = \bar{\lambda} \langle \vec{x} | \vec{y} \rangle = \lambda \langle \vec{x} | \vec{y} \rangle, \\ \langle \vec{x} | \mathbf{A} \vec{y} \rangle &= \langle \vec{x} | \mu \vec{y} \rangle = \mu \langle \vec{x} | \vec{y} \rangle.\end{aligned}$$

Einsetzen in (17.3) liefert

$$\lambda \langle \vec{x} | \vec{y} \rangle = \mu \langle \vec{x} | \vec{y} \rangle \iff (\lambda - \mu) \langle \vec{x} | \vec{y} \rangle = 0.$$

Da $\lambda \neq \mu$ gilt, folgt $\langle \vec{x} | \vec{y} \rangle = 0$. □

Satz 17.9. (ONB aus den EVen einer symmetrischen Matrix)

Sei $\mathbf{A} \in \mathbb{R}^{n \times n}$ symmetrisch. Dann hat \mathbb{R}^n eine **Orthonormalbasis aus Eigenvektoren von \mathbf{A}** .

Beweis von Satz 17.9 für $n = 2$: Sei $\vec{x} \in \mathbb{R}^2$ ein Eigenvektor zum Eigenwert $\lambda \in \mathbb{R}$. Für $\vec{y} \perp \vec{x}$ mit $\vec{y} \neq \vec{0}$ gilt dann nach Bemerkung 17.7 (da A symmetrisch ist):

$$\langle \mathbf{A} \vec{y} | \vec{x} \rangle = \langle \vec{y} | \mathbf{A} \vec{x} \rangle = \langle \vec{y} | \lambda \vec{x} \rangle = \lambda \langle \vec{y} | \vec{x} \rangle = 0.$$

Also ist $\mathbf{A} \vec{y} \perp \vec{x}$. Da $n = 2$ ist, sind alle zu \vec{x} in \mathbb{R}^2 orthogonalen Vektoren parallel, und somit existiert ein Faktor $\mu \in \mathbb{R}$ mit $\mathbf{A} \vec{y} = \mu \vec{y}$. Daher ist \vec{y} ein Eigenvektor zum Eigenwert μ , und

$$\left(\frac{\vec{x}}{\|\vec{x}\|}, \frac{\vec{y}}{\|\vec{y}\|} \right)$$

ist eine Orthonormalbasis von \mathbb{R}^2 . □

Betrachten wir hierzu ein Beispiel.

Beispiel 17.10. (ONB aus den EVen einer symmetrischen Matrix)

Wir betrachten die symmetrische Matrix

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

Das charakteristische Polynom ist

$$\begin{aligned} p_{\mathbf{A}}(\lambda) &= \det(\mathbf{A} - \lambda \mathbf{E}_3) = \det \left(\begin{bmatrix} -\lambda & 0 & -1 \\ 0 & 1 - \lambda & 0 \\ -1 & 0 & -\lambda \end{bmatrix} \right) \\ &= \lambda^2 (1 - \lambda) + 0 + 0 - (1 - \lambda) - 0 - 0 \\ &= (\lambda^2 - 1)(1 - \lambda) = (\lambda - 1)(\lambda + 1)(1 - \lambda) = -(\lambda - 1)^2 (\lambda + 1), \end{aligned}$$

d.h. wir finden die Eigenwerte $\lambda_1 = 1$ (mit der algebraischen Vielfachheit 2) und $\lambda_2 = -1$.

Wir berechnen die zugehörigen Eigenräume:

$$\lambda_1 = 1 : \left[\begin{array}{ccc|c} -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 \end{array} \right] \begin{array}{c} \xrightarrow{Z_3 \rightarrow Z_3 - Z_1} \\ \xrightarrow{Z_1 \rightarrow -Z_1} \\ \iff \\ \downarrow \end{array} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

$$\implies (u_1 = -u_3, \quad u_2, u_3 \text{ bel.}) \implies E_{\mathbf{A}}(1) = \text{LH} \left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \right),$$

wobei wir die Basis des Eigenraums gerade so gewählt haben, dass die beiden Eigenvektoren orthogonal und normiert sind.

$$\lambda_2 = -1 : \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 2 & 0 & 0 \\ -1 & 0 & 1 & 0 \end{array} \right] \begin{array}{c} \xrightarrow{Z_3 \rightarrow Z_3 + Z_1} \\ \xrightarrow{Z_2 \rightarrow \frac{1}{2} Z_2} \\ \iff \\ \downarrow \end{array} \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

$$\implies (u_1 = u_3, \quad u_2 = 0, \quad u_3 \text{ bel.}) \implies E_{\mathbf{A}}(-1) = \text{LH} \left(\begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} \right),$$

wobei wir die Basis des Eigenraums gerade so gewählt haben, dass der einzige Eigenvektor normiert ist.

Nach Satz 17.9 ist damit

$$\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} \right)$$

eine Orthonormalbasis von \mathbb{R}^3 , die aus Eigenvektoren von \mathbf{A} besteht.

Nun lernen wir drei wichtige mögliche Zusatzeigenschaften symmetrischer Matrizen kennen.

Definition 17.11. (positiv definite, negativ definite bzw. indefinite symmetrische Matrix)

Sei $\mathbf{A} \in \mathbb{R}^{n \times n}$ *symmetrisch*.

- (1) \mathbf{A} heißt **positiv definit**, wenn alle Eigenwerte von \mathbf{A} **positiv** (d.h. > 0) sind.
- (2) \mathbf{A} heißt **negativ definit**, wenn alle Eigenwerte von \mathbf{A} **negativ** (d.h. < 0) sind.
- (3) \mathbf{A} heißt **indefinit**, wenn \mathbf{A} **sowohl positive als auch negative Eigenwerte hat**.

Wir halten fest, dass eine symmetrische Matrix keine der drei Eigenschaften aus Definition 17.11 haben muss. Beispielsweise kann eine symmetrische Matrix neben positiven Eigenwerten noch den Eigenwert 0 haben.

Betrachten wir einige Beispiele für positiv definite, negativ definite bzw. indefinite Matrizen.

Beispiel 17.12. (positive definite, negativ definite bzw. indefinite symmetrische Matrix)

(a) Die symmetrische Matrix $\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \\ 0 & 1 & 3 \end{bmatrix}$ ist positiv definit, denn:

Das charakteristische Polynom ist

$$\begin{aligned} p_{\mathbf{A}}(\lambda) &= \det(\mathbf{A} - \lambda \mathbf{E}_3) = \det \left(\begin{bmatrix} 1 - \lambda & -1 & 0 \\ -1 & 2 - \lambda & 1 \\ 0 & 1 & 3 - \lambda \end{bmatrix} \right) \\ &= (1 - \lambda)(2 - \lambda)(3 - \lambda) + 0 + 0 - 0 - (1 - \lambda) - (3 - \lambda) \\ &= (1 - \lambda)(2 - \lambda)(3 - \lambda) - 2(2 - \lambda) \\ &= (2 - \lambda)(\lambda^2 - 4\lambda + 3 - 2) \\ &= (2 - \lambda)(\lambda^2 - 4\lambda + 1) \\ &= (2 - \lambda)((\lambda - 2)^2 - 3) \\ &= (2 - \lambda)(\lambda - 2 - \sqrt{3})(\lambda - 2 + \sqrt{3}). \end{aligned}$$

Also sind die Eigenwerte $\lambda_1 = 2$, $\lambda_2 = 2 + \sqrt{3}$, und $\lambda_3 = 2 - \sqrt{3}$. Da alle Eigenwerte positiv sind, ist die Matrix \mathbf{A} positiv definit.

(b) Die symmetrische Matrix $\mathbf{A} = \begin{bmatrix} -1 & -1 & 0 \\ -1 & -2 & 1 \\ 0 & 1 & -3 \end{bmatrix}$ ist negativ definit, denn:

Das charakteristische Polynom ist

$$\begin{aligned} p_{\mathbf{A}}(\lambda) &= \det(\mathbf{A} - \lambda \mathbf{E}_3) = \det \left(\begin{bmatrix} -1 - \lambda & -1 & 0 \\ -1 & -2 - \lambda & 1 \\ 0 & 1 & -3 - \lambda \end{bmatrix} \right) \\ &= (-1 - \lambda)(-2 - \lambda)(-3 - \lambda) + 0 + 0 - 0 - (-1 - \lambda) - (-3 - \lambda) \\ &= (-1 - \lambda)(-2 - \lambda)(-3 - \lambda) - 2(-2 - \lambda) \\ &= (-2 - \lambda)(\lambda^2 + 4\lambda + 3 - 2) \\ &= (-2 - \lambda)(\lambda^2 + 4\lambda + 1) \\ &= (-2 - \lambda)((\lambda + 2)^2 - 3) \\ &= (-2 - \lambda)(\lambda + 2 - \sqrt{3})(\lambda + 2 + \sqrt{3}). \end{aligned}$$

Also sind die Eigenwerte $\lambda_1 = -2$, $\lambda_2 = -2 + \sqrt{3}$, und $\lambda_3 = -2 - \sqrt{3}$. Da alle Eigenwerte negativ sind, ist die Matrix \mathbf{A} negativ definit.

(c) Die symmetrische Matrix $\mathbf{A} = \begin{bmatrix} -1 & -1 & 0 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{bmatrix}$ ist indefinit, denn:

Das charakteristische Polynom ist

$$\begin{aligned} p_{\mathbf{A}}(\lambda) &= \det(\mathbf{A} - \lambda \mathbf{E}_3) = \det \left(\begin{bmatrix} -1 - \lambda & -1 & 0 \\ -1 & 2 - \lambda & 1 \\ 0 & 1 & -1 - \lambda \end{bmatrix} \right) \\ &= (-1 - \lambda)(2 - \lambda)(-1 - \lambda) + 0 + 0 - 0 - (-1 - \lambda) - (-1 - \lambda) \\ &= (-1 - \lambda)(2 - \lambda)(-1 - \lambda) - 2(-1 - \lambda) \\ &= (-1 - \lambda)(\lambda^2 - \lambda - 2 - 2) \\ &= (-1 - \lambda)(\lambda^2 - \lambda - 4) \\ &= (-1 - \lambda) \left(\left(\lambda - \frac{1}{2} \right)^2 - \frac{17}{4} \right) \\ &= (-1 - \lambda) \left(\lambda - \frac{1}{2} - \frac{\sqrt{17}}{2} \right) \left(\lambda - \frac{1}{2} + \frac{\sqrt{17}}{2} \right) \\ &= (-1 - \lambda) \left(\lambda - \frac{1 + \sqrt{17}}{2} \right) \left(\lambda - \frac{1 - \sqrt{17}}{2} \right). \end{aligned}$$

Also sind die Eigenwerte $\lambda_1 = -1$, $\lambda_2 = \frac{1 + \sqrt{17}}{2}$, und $\lambda_3 = \frac{1 - \sqrt{17}}{2}$. Da \mathbf{A} positive und negative Eigenwerte hat, ist die Matrix \mathbf{A} indefinit.

(d) Die symmetrische Matrix $\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ ist weder positiv definit noch

negativ definit noch indefinit, denn sie hat die Eigenwerte $\lambda_1 = 1$, $\lambda_2 = 2$ und $\lambda_3 = 0$.

Als Letztes lernen wir eine alternative Charakterisierung für eine positiv definite bzw. eine negativ definite Matrix kennen, die wir im nächsten Teilkapitel benötigen.

Satz 17.13. (Kriterium für positiv definit bzw. negativ definit)

Sei $\mathbf{A} \in \mathbb{R}^{n \times n}$ *symmetrisch*.

- (1) \mathbf{A} ist **positiv definit** genau dann, wenn gilt $\langle \vec{x} | \mathbf{A} \vec{x} \rangle > 0$ für alle $\vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\}$.
- (2) \mathbf{A} ist **negativ definit** genau dann, wenn gilt $\langle \vec{x} | \mathbf{A} \vec{x} \rangle < 0$ für alle $\vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\}$.

Wir beweisen nur die erste Aussage dieses Satzes; die zweite zeigt man analog.

Beweis von Satz 17.13 (1): Wir müssen zwei Richtungen zeigen:

\implies : Sei \mathbf{A} positiv definit. Nach Satz 17.8 und Definition 17.11 sind dann alle Eigenwerte von \mathbf{A} reell und positiv.

Ist \vec{u} ein Eigenvektor zum Eigenwert $\lambda > 0$, so gilt

$$\langle \vec{u} | \mathbf{A} \vec{u} \rangle = \langle \vec{u} | \lambda \vec{u} \rangle = \lambda \langle \vec{u} | \vec{u} \rangle = \underbrace{\lambda}_{>0} \underbrace{\|\vec{u}\|^2}_{>0} > 0.$$

Nun wollen wir beliebiges $\vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\}$ betrachten. Da \mathbf{A} symmetrisch ist, hat \mathbb{R}^n nach Satz 17.9 eine Orthonormalbasis $(\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n)$ aus Eigenvektoren von \mathbf{A} . Seien $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$ die zugehörigen (nicht notwendigerweise verschiedenen) Eigenwerte, d.h. $\mathbf{A} \vec{b}_k = \lambda_k \vec{b}_k$ für $k = 1, 2, \dots, n$. Der Vektor \vec{x} hat eine Darstellung

$$\vec{x} = \sum_{k=1}^n c_k \vec{b}_k$$

bzgl. der Basis $(\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n)$ mit eindeutig bestimmten Koeffizienten $c_1, c_2, \dots, c_n \in \mathbb{R}$. Damit gilt

$$\mathbf{A} \vec{x} = \mathbf{A} \left(\sum_{k=1}^n c_k \vec{b}_k \right) = \sum_{k=1}^n c_k \mathbf{A} \vec{b}_k = \sum_{k=1}^n c_k \lambda_k \vec{b}_k$$

und somit

$$\begin{aligned} \langle \vec{x} | \mathbf{A} \vec{x} \rangle &= \left\langle \sum_{k=1}^n c_k \vec{b}_k \left| \sum_{\ell=1}^n c_\ell \lambda_\ell \vec{b}_\ell \right. \right\rangle = \sum_{k=1}^n \sum_{\ell=1}^n \lambda_\ell c_k c_\ell \underbrace{\langle \vec{b}_k | \vec{b}_\ell \rangle}_{=0 \text{ wenn } k \neq \ell} \\ &= \sum_{k=1}^n \lambda_k c_k^2 \underbrace{\langle \vec{b}_k | \vec{b}_k \rangle}_{=1} = \sum_{k=1}^n \underbrace{\lambda_k}_{>0} \underbrace{c_k^2}_{>0} > 0. \end{aligned}$$

Damit ist die Hinrichtung bewiesen.

\Leftarrow : Es gelte $\langle \vec{x} | \mathbf{A} \vec{x} \rangle > 0$ für alle $\vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\}$. Da \mathbf{A} symmetrisch ist, wissen wir nach Satz 17.9, dass es eine Orthonormalbasis $(\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n)$ aus Eigenvektoren von \mathbf{A} gibt. Seien $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$ die zugehörigen (nicht notwendigerweise verschiedenen) reellen Eigenwerte, d.h. $\mathbf{A} \vec{b}_k = \lambda_k \vec{b}_k$ für $k = 1, 2, \dots, n$. (Wir wissen nach Satz 17.8, dass diese reell sind, weil \mathbf{A} symmetrisch ist.) Dann gilt für $k = 1, 2, \dots, n$

$$0 < \langle \vec{b}_k | \mathbf{A} \vec{b}_k \rangle = \langle \vec{b}_k | \lambda_k \vec{b}_k \rangle = \lambda_k \langle \vec{b}_k | \vec{b}_k \rangle = \lambda_k \underbrace{\|\vec{b}_k\|^2}_{>0} \implies 0 < \lambda_k.$$

Also sind die n Eigenwerte $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$ von \mathbf{A} alle positiv, d.h. die Matrix \mathbf{A} ist positiv definit. \square

17.3 Hinreichende Bedingung für lokale Extremalstellen

Mit dem Wissen aus dem vorigen Teilkapitel können wir nun eine hinreichende Bedingung für lokale Extremalstellen angeben. Das Wort „hinreichend“ bedeutet dabei, dass immer eine lokale Extremalstelle vorliegt, wenn die Bedingung erfüllt ist. Es kann aber auch eine lokale Extremalstelle vorliegen, ohne dass die Bedingung erfüllt ist (d.h. aus der Tatsache, dass die eine lokale Extremalstelle vorliegt, folgt noch nicht, dass die Bedingung immer erfüllt ist).

Satz 17.14. (hinreichende Bedingung für lokale Extremalstellen)

Seien $\Omega \subseteq \mathbb{R}^n$ offen und $f \in \mathcal{C}^2(\Omega, \mathbb{R})$. Sei $\vec{a} \in \Omega$ ein kritischer Punkt von f , und sei $\mathbf{H}f$ die Hesse-Matrix der zweiten partiellen Ableitungen von f .

- (1) Ist $(\mathbf{H}f)(\vec{a})$ **positiv definit**, dann hat f in \vec{a} eine **lokale Minimalstelle**.
- (2) Ist $(\mathbf{H}f)(\vec{a})$ **negativ definit**, dann hat f in \vec{a} eine **lokale Maximalstelle**.
- (3) Ist $(\mathbf{H}f)(\vec{a})$ **indefinit**, dann hat f in \vec{a} einen sogenannten **Sattelpunkt**.
- (4) Ist $(\mathbf{H}f)(\vec{a})$ weder positiv definit noch negativ definit noch indefinit, so kann man **keine** Aussage darüber treffen, ob eine lokale Minimalstelle vorliegt.

le, eine lokale Maximalstelle oder ein Sattelpunkt oder keines von allen dreien vorliegt.

Satz 17.14 (3) kann als **Definition eines Sattelpunkts** betrachtet werden. Der Name ist ganz anschaulich, denn für $f \in \mathcal{C}^2(\Omega, \mathbb{R})$ mit $\Omega \subseteq \mathbb{R}^2$ offen werden wir an den Beispielen sehen, dass die Fläche des Graphen der Funktion in der Nähe eines Sattelpunkts einem Reitsattel ähnelt.

Wir beweisen nur die erste Aussage. Die zweite Aussage zeigt man analog.

Beweis von Satz 17.14 (1): Da $(\mathbf{H}f)(\vec{\mathbf{a}})$ positiv definit ist und da die partiellen Ableitungen zweiter Ordnung stetig sind (wegen $f \in \mathcal{C}^2(\Omega, \mathbb{R})$), existiert ein $\varepsilon > 0$, so dass $(\mathbf{H}f)(\vec{\mathbf{z}})$ positiv definit ist für alle $\vec{\mathbf{z}} \in U_\varepsilon(\vec{\mathbf{a}})$. (*Erklärung:* Wegen der Stetigkeit der partiellen Ableitungen zweiter Ordnung folgt aus $\langle \vec{\mathbf{x}} | (\mathbf{H}f)(\vec{\mathbf{a}}) \vec{\mathbf{x}} \rangle > 0$ für alle $\vec{\mathbf{x}} \in \mathbb{R}^n \setminus \{\vec{\mathbf{0}}\}$ (alternative Charakterisierung für positiv definit, siehe Satz 17.13), dass auf einer kleinen Kugel $U_\varepsilon(\vec{\mathbf{a}})$ um $\vec{\mathbf{a}}$ ebenfalls $\langle \vec{\mathbf{x}} | (\mathbf{H}f)(\vec{\mathbf{z}}) \vec{\mathbf{x}} \rangle > 0$ für alle $\vec{\mathbf{x}} \in \mathbb{R}^n \setminus \{\vec{\mathbf{0}}\}$ für alle $\vec{\mathbf{z}} \in U_\varepsilon(\vec{\mathbf{a}})$ gelten muss.)

Für $\vec{\mathbf{y}} \in U_\varepsilon(\vec{\mathbf{a}}) \setminus \{\vec{\mathbf{a}}\}$ existiert nach dem Satz von Taylor (siehe Satz 16.59 und Bemerkung 16.61) ein $\vec{\mathbf{z}}$ auf der Verbindungsstrecke von $\vec{\mathbf{a}}$ und $\vec{\mathbf{y}}$, so dass gilt

$$\begin{aligned} f(\vec{\mathbf{y}}) &= f(\vec{\mathbf{a}}) + \underbrace{f'(\vec{\mathbf{a}})}_{=0}(\vec{\mathbf{y}} - \vec{\mathbf{a}}) + \frac{1}{2} \langle \vec{\mathbf{y}} - \vec{\mathbf{a}} | (\mathbf{H}f)(\vec{\mathbf{z}}) \cdot (\vec{\mathbf{y}} - \vec{\mathbf{a}}) \rangle \\ &= f(\vec{\mathbf{a}}) + \frac{1}{2} \langle \vec{\mathbf{y}} - \vec{\mathbf{a}} | (\mathbf{H}f)(\vec{\mathbf{z}}) (\vec{\mathbf{y}} - \vec{\mathbf{a}}) \rangle, \end{aligned}$$

wobei $f'(\vec{\mathbf{a}}) = 0$ ist, da $\vec{\mathbf{a}}$ ein kritischer Punkt ist. Da $U_\varepsilon(\vec{\mathbf{a}})$ konvex ist, liegt $\vec{\mathbf{z}}$ in $U_\varepsilon(\vec{\mathbf{a}})$, und somit ist $(\mathbf{H}f)(\vec{\mathbf{z}})$ positiv definit. Also gilt nach Satz 17.13

$$f(\vec{\mathbf{y}}) = f(\vec{\mathbf{a}}) + \frac{1}{2} \underbrace{\langle \vec{\mathbf{y}} - \vec{\mathbf{a}} | (\mathbf{H}f)(\vec{\mathbf{z}}) \cdot (\vec{\mathbf{y}} - \vec{\mathbf{a}}) \rangle}_{>0},$$

d.h. $f(\vec{\mathbf{a}}) < f(\vec{\mathbf{y}})$. Da $\vec{\mathbf{y}} \in U_\varepsilon(\vec{\mathbf{a}}) \setminus \{\vec{\mathbf{a}}\}$ beliebig war, ist $\vec{\mathbf{a}}$ eine lokale Minimalstelle. \square

Betrachten wir einige Beispiele.

Beispiel 17.15. (lokale Extremalstellen und kritische Punkte)

- (a) Die \mathcal{C}^∞ -Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := e^{-(x^2+y^2)}$, aus Beispiel 17.3 (a) hat die Ableitung

$$f'(x, y) = \left[-2x e^{-(x^2+y^2)} \quad -2y e^{-(x^2+y^2)} \right],$$

und wir fanden den einzigen kritischen Punkt $(0, 0)$. Die Hesse-Matrix ist

$$(\mathbf{H}f)(x, y) = \begin{bmatrix} (-2 + 4x^2) e^{-(x^2+y^2)} & 4xy e^{-(x^2+y^2)} \\ 4xy e^{-(x^2+y^2)} & (-2 + 4y^2) e^{-(x^2+y^2)} \end{bmatrix},$$

und für $(x, y) = (0, 0)$ finden wir

$$(\mathbf{H}f)(0, 0) = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}.$$

Es folgt

$$\begin{aligned} p_{(\mathbf{H}f)(0,0)}(\lambda) &= \det((\mathbf{H}f)(0, 0) - \lambda \mathbf{E}_2) = \det\left(\begin{bmatrix} -2 - \lambda & 0 \\ 0 & -2 - \lambda \end{bmatrix}\right) \\ &= (-2 - \lambda)^2 \quad \implies \quad \lambda_1 = \lambda_2 = -2 < 0. \end{aligned}$$

Also ist $(\mathbf{H}f)(0, 0)$ negativ definit, und in $(0, 0)$ liegt eine lokale Maximalstelle vor.

Der Graph der Funktion ist im linken Bild in Abbildung 17.1 gezeichnet.

- (b) Die \mathcal{C}^∞ -Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := x^2 - y^2$, aus Beispiel 17.3 (b) hat die Ableitung

$$f'(x, y) = [2x \quad -2y],$$

und wir fanden den einzigen kritischen Punkt $(0, 0)$. Die Hesse-Matrix ist

$$(\mathbf{H}f)(x, y) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix},$$

und für $(x, y) = (0, 0)$ finden wir

$$(\mathbf{H}f)(0, 0) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}.$$

Es folgt

$$\begin{aligned} p_{(\mathbf{H}f)(0,0)}(\lambda) &= \det((\mathbf{H}f)(0, 0) - \lambda \mathbf{E}_2) = \det\left(\begin{bmatrix} 2 - \lambda & 0 \\ 0 & -2 - \lambda \end{bmatrix}\right) \\ &= (2 - \lambda)(-2 - \lambda) \quad \implies \quad (\lambda_1 = 2 > 0 \text{ und } \lambda_2 = -2 < 0). \end{aligned}$$

Also ist $(\mathbf{H}f)(0, 0)$ indefinit, und in $(0, 0)$ liegt ein Sattelpunkt vor. Dieses kann man an dem Graphen von f im rechten Bild in Abbildung 17.1 sehr gut sehen.

- (c) Die \mathcal{C}^∞ -Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := xy$, aus Beispiel 17.3 (c) hat die Ableitung

$$f'(x, y) = [y \ x],$$

und wir fanden den einzigen kritischen Punkt $(0, 0)$. Die Hesse-Matrix ist

$$(\mathbf{H}f)(x, y) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

und für $(x, y) = (0, 0)$ finden wir

$$(\mathbf{H}f)(0, 0) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Es folgt

$$\begin{aligned} p_{(\mathbf{H}f)(0,0)}(\lambda) &= \det((\mathbf{H}f)(0, 0) - \lambda \mathbf{E}_2) = \det\left(\begin{bmatrix} -\lambda & 1 \\ 1 & -\lambda \end{bmatrix}\right) = \lambda^2 - 1 \\ &= (\lambda - 1)(\lambda + 1) \quad \implies \quad \lambda_1 = 1 \quad \text{und} \quad \lambda_2 = -1. \end{aligned}$$

Also ist $(\mathbf{H}f)(0, 0)$ indefinit, und in $(0, 0)$ liegt ein Sattelpunkt vor. Dieses kann man an dem Graphen von f im linken Bild in Abbildung 17.2 gut sehen.

- (d) Die \mathcal{C}^∞ -Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := -\cos((x-1)^2 + (y-1)^2)$, aus Beispiel 17.3 (d) hat die Ableitung

$$\begin{aligned} f'(x, y) &= [2(x-1) \sin((x-1)^2 + (y-1)^2) \quad 2(y-1) \sin((x-1)^2 + (y-1)^2)], \end{aligned}$$

und wir fanden, dass $(1, 1)$ und alle Punkte auf $\partial U_{\sqrt{k\pi}}(1, 1)$ mit $k \in \mathbb{N}$ (also alle Punkte auf den Kreisen mit Radius $\sqrt{k\pi}$ um $(1, 1)$) kritische Punkte von f sind. Die Hesse-Matrix von f ist

$$\begin{aligned} (\mathbf{H}f)(x, y) &= \\ &= \begin{bmatrix} 2 \sin((x-1)^2 + (y-1)^2) + & 4(x-1)(y-1) \cdot \\ 4(x-1)^2 \cos((x-1)^2 + (y-1)^2) & \cos((x-1)^2 + (y-1)^2) \\ & 4(x-1)(y-1) \cdot \\ \cos((x-1)^2 + (y-1)^2) & 2 \sin((x-1)^2 + (y-1)^2) + \\ & 4(y-1)^2 \cos((x-1)^2 + (y-1)^2) \end{bmatrix}. \end{aligned}$$

Für den Punkt $(1, 1)$ finden wir die Hesse-Matrix

$$(\mathbf{H}f)(1, 1) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

mit dem einzigen Eigenwert $\lambda_1 = \lambda_2 = 0$. In $(1, 1)$ liegt die Situation in Satz 17.14 (4) vor, und wir können mit Hilfe der Hesse-Matrix keine Aussage über die Existenz einer lokalen Extremalstelle oder eines Sattelpunktes treffen.

Für (x_k, y_k) in $\partial U_{\sqrt{k\pi}}(1, 1)$ mit $k \in \mathbb{N}$ gilt

$$\begin{aligned}\cos((x_k - 1)^2 + (y_k - 1)^2) &= \cos(k\pi) = (-1)^k, \\ \sin((x_k - 1)^2 + (y_k - 1)^2) &= \sin(k\pi) = 0,\end{aligned}$$

und wir finden die Hesse-Matrix

$$(\mathbf{H}f)(x_k, y_k) = \begin{bmatrix} 4(x_k - 1)^2(-1)^k & 4(x_k - 1)(y_k - 1)(-1)^k \\ 4(x_k - 1)(y_k - 1)(-1)^k & 4(y_k - 1)^2(-1)^k \end{bmatrix},$$

deren charakteristisches Polynom wie folgt lautet:

$$\begin{aligned}p_{(\mathbf{H}f)(x_k, y_k)}(\lambda) &= \det((\mathbf{H}f)(x_k, y_k) - \lambda \mathbf{E}_2) \\ &= \det \left(\begin{bmatrix} 4(x_k - 1)^2(-1)^k - \lambda & 4(x_k - 1)(y_k - 1)(-1)^k \\ 4(x_k - 1)(y_k - 1)(-1)^k & 4(y_k - 1)^2(-1)^k - \lambda \end{bmatrix} \right) \\ &= \left(4(x_k - 1)^2(-1)^k - \lambda \right) \left(4(y_k - 1)^2(-1)^k - \lambda \right) \\ &\quad - \left(4(x_k - 1)(y_k - 1)(-1)^k \right)^2 \\ &= \lambda^2 - 4(x_k - 1)^2(-1)^k \lambda - 4(y_k - 1)^2(-1)^k \lambda \\ &= \lambda \left(\lambda - 4(-1)^k ((x_k - 1)^2 + (y_k - 1)^2) \right) \\ &= \lambda (\lambda - 4(-1)^k k \pi).\end{aligned}$$

Die Hesse-Matrix hat also die Eigenwerte $\lambda_1 = 0$ und $\lambda_2 = 4(-1)^k k \pi$, d.h. ein Eigenwert ist Null und der andere positiv, wenn k gerade ist, und negativ, wenn k ungerade ist. Hier liegt die Situation in Satz 17.14 (4) vor, und wir können mit Hilfe der Hesse-Matrix keine Aussage über die Existenz lokaler Extremalstellen oder Sattelpunkte treffen.

Der Graph der Funktion ist im rechten Bild in Abbildung 17.2 gezeichnet, und wir wissen aus Beispiel 17.3 (d) bereits, dass die Funktion f in allen kritischen Punkten lokale Extremalstellen hat.

(e) Die \mathcal{C}^∞ -Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) := y^2(x - 1) + x^2(x + 1) = y^2(x - 1) + x^3 + x^2,$$

hat die Ableitung

$$f'(x, y) = [y^2 + 3x^2 + 2x \quad 2y(x - 1)]$$

und die Hesse-Matrix

$$(\mathbf{H}f)(x, y) = \begin{bmatrix} 6x + 2 & 2y \\ 2y & 2(x - 1) \end{bmatrix}.$$

Die kritischen Punkte sind durch die Bedingung

$$f'(x, y) = [y^2 + 3x^2 + 2x \quad 2y(x - 1)] = [0 \quad 0]$$

gegeben. Die zweite Bedingung $2y(x - 1) = 0$ liefert $y = 0$ oder $x = 1$. Wir setzen jede dieser beiden Bedingungen jeweils in $0 = y^2 + 3x^2 + 2x$ ein, und finden

$$\begin{aligned} y = 0 : \quad 0 &= y^2 + 3x^2 + 2x = 0^2 + 3x^2 + 2x \\ &\implies 0 = 3x \left(x + \frac{2}{3} \right) \implies x = 0 \text{ oder } x = -\frac{2}{3}, \end{aligned}$$

$$\begin{aligned} x = 1 : \quad 0 &= y^2 + 3x^2 + 2x = y^2 + 3 \cdot 1^2 + 2 \cdot 1 \\ &\implies 0 = y^2 + 5 \quad \text{und diese Gleichung hat keine reelle Lösung.} \end{aligned}$$

Also finden wir die beiden kritischen Punkte $(x, y) = (0, 0)$ und $(x, y) = (-\frac{2}{3}, 0)$.

In $(0, 0)$ ist die Hesse-Matrix

$$(\mathbf{H}f)(0, 0) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix},$$

und diese hat die Eigenwerte $\lambda_1 = 2$ und $\lambda_2 = -2$, d.h. $(\mathbf{H}f)(0, 0)$ ist indefinit. Also ist $(0, 0)$ ein Sattelpunkt.

In $(-\frac{2}{3}, 0)$ ist die Hesse-Matrix

$$(\mathbf{H}f) \left(-\frac{2}{3}, 0 \right) = \begin{bmatrix} -2 & 0 \\ 0 & -\frac{10}{3} \end{bmatrix},$$

und diese Matrix hat die Eigenwerte $\lambda_1 = -2$ und $\lambda_2 = -\frac{10}{3}$, d.h. die Hesse-Matrix $(\mathbf{H}f) \left(-\frac{2}{3}, 0 \right)$ ist negativ definit. Also hat f in $(-\frac{2}{3}, 0)$ eine lokale Maximalstelle.

Der Graph der Funktion ist im linken Bild in Abbildung 17.3 gezeichnet, und man kann sowohl den Sattelpunkt in $(0, 0)$ wie auch die lokale Maximalstelle in $(-\frac{2}{3}, 0)$ erkennen.

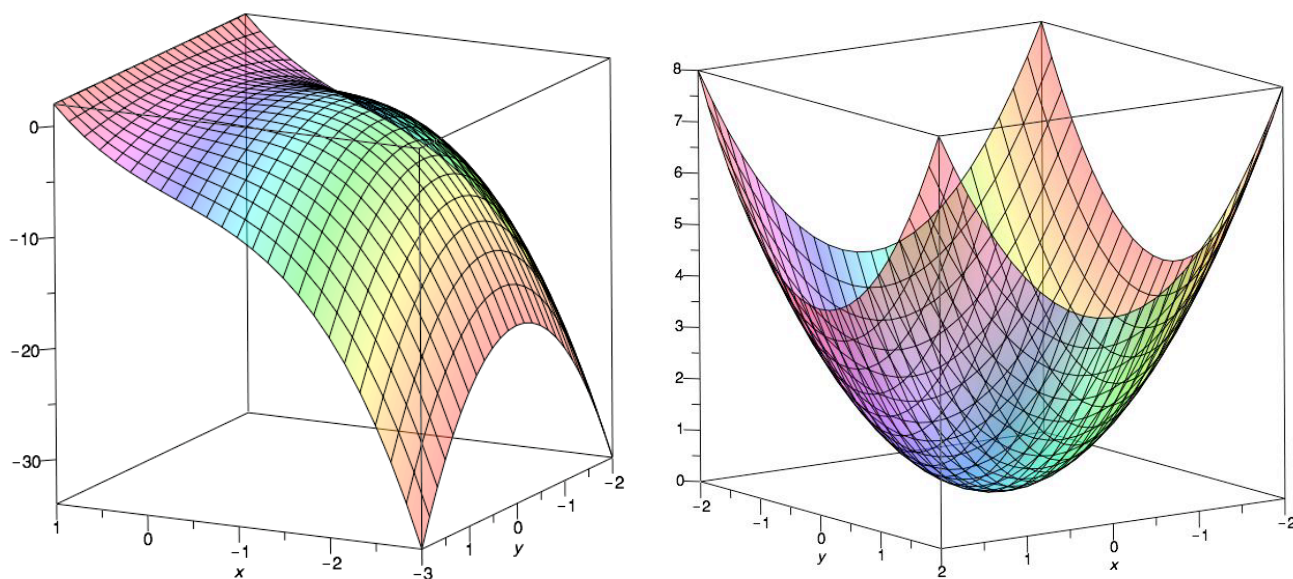


Abbildung 17.3: Die Graphen der Funktionen $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := y^2(x - 1) + x^2(x + 1)$, (linkes Bild) bzw. $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := x^2 + y^2$, (rechtes Bild). Dabei ist im linken Bild das Koordinatenkreuz der (x, y) -Ebene um 270° gegen den Uhrzeigersinn gedreht worden, damit die lokale Maximalstelle sichtbar ist.

Wir halten noch eine Beobachtung allgemein fest, die wir in den vorigen Beispielen bei der Berechnung der Eigenwerte mehrfach benutzt haben.

Bemerkung 17.16. (Eigenwerte einer Diagonalmatrix)

Ist \mathbf{D} eine Diagonalmatrix, also $\mathbf{D} = [d_{i,k}] \in \mathbb{C}^{n \times n}$ mit $d_{i,k} = 0$ wenn $i \neq k$ ist, dann sind die Diagonaleinträge von \mathbf{D} gerade die Eigenwerte von \mathbf{D} , denn:

$$\begin{aligned}
 p_{\mathbf{D}}(\lambda) &= \det(\mathbf{D} - \lambda \mathbf{E}_n) = \det \left(\begin{bmatrix} d_{1,1} - \lambda & 0 & \cdots & 0 \\ 0 & d_{2,2} - \lambda & \cdots & \vdots \\ \vdots & \cdots & \cdots & 0 \\ 0 & \cdots & 0 & d_{n,n} - \lambda \end{bmatrix} \right) \\
 &= (d_{1,1} - \lambda) (d_{2,2} - \lambda) \cdots (d_{n,n} - \lambda).
 \end{aligned}$$

17.4 Globale Extrema

Nun betrachten wir globale Extremalstellen, also Punkte, deren Funktionswert größer oder gleich bzw. kleiner oder gleich als alle anderen Funktionswerte ist. Genauer interessieren wir uns vor allem für die Funktionswerte in diesen Punkten, also für das Maximum der Funktion (also den größten Funktionswert) und das Minimum der Funktion (also den kleinsten Funktionswert), sofern solche Werte existieren. Man bezeichnet solche Werte als (globale) Extrema der Funktion.

Definition 17.17. (globales Maximum bzw. globales Minimum)

Sei D eine nichtleere Teilmenge von \mathbb{R}^n und $f : D \rightarrow \mathbb{R}$.

(1) Ein Wert $M \in \mathbb{R}$ heißt ein **(globales) Maximum** von f auf D , wenn

(i) $M = f(\vec{\mathbf{a}})$ für ein $\vec{\mathbf{a}} \in D$ und

(ii) $f(\vec{\mathbf{x}}) \leq M$ für alle $\vec{\mathbf{x}} \in D$ gilt.

Schreibweise: $\max_{\vec{\mathbf{x}} \in D} f(\vec{\mathbf{x}}) := M$.

(2) Ein Wert $m \in \mathbb{R}$ heißt ein **(globales) Minimum** von f auf D , wenn

(i) $m = f(\vec{\mathbf{a}})$ für ein $\vec{\mathbf{a}} \in D$ und

(ii) $f(\vec{\mathbf{x}}) \geq m$ für alle $\vec{\mathbf{x}} \in D$ gilt.

Schreibweise: $\min_{\vec{\mathbf{x}} \in D} f(\vec{\mathbf{x}}) := m$

Beispiel 17.18. (globale Extremalstellen)

(a) Wir betrachten die \mathcal{C}^∞ -Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := x^2 + y^2$. Dann ist $m = 0$ eine (globales) Minimum von f auf \mathbb{R}^2 , denn

(i) $f(0, 0) = 0$ und

(ii) $f(x, y) = x^2 + y^2 \geq 0$ für alle $(x, y) \in \mathbb{R}^2$.

Also gilt

$$\min_{(x,y) \in \mathbb{R}^2} f(x, y) = \min_{(x,y) \in \mathbb{R}^2} (x^2 + y^2) = 0.$$

f hat aber kein globales Maximum, weil die Funktionswerte von f beliebig groß werden.

Der Graph der Funktion ist im rechten Bild in Abbildung 17.3 gezeichnet.

(b) Die \mathcal{C}^∞ -Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := e^{-(x^2+y^2)}$, hat nach Beispiel 17.3 (a) die lokale Extremalstelle $(0, 0)$. In diesem Punkt tritt das (globale) Maximum $M = 1$ auf \mathbb{R}^2 auf, denn

(i) $f(0, 0) = 1$ und

(ii) $f(x, y) = e^{-(x^2+y^2)} \leq 1$ für alle $(x, y) \in \mathbb{R}^2$.

Also gilt

$$\max_{(x,y) \in \mathbb{R}^2} f(x, y) = \max_{(x,y) \in \mathbb{R}^2} e^{-(x^2+y^2)} = 1.$$

f hat aber kein globales Minimum, weil die Funktionswerte von f alle $f(x, y) > 0$ erfüllen, aber sich dem Wert Null beliebig dicht annähern. Der Wert 0 ist das Infimum aller Funktionswerte, also

$$0 = \inf_{(x,y) \in \mathbb{R}^2} e^{-(x^2+y^2)}.$$

Der Graph der Funktion ist im linken Bild in Abbildung 17.1 gezeichnet.

(c) Die \mathcal{C}^∞ -Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := -\cos((x-1)^2 + (y-1)^2)$, aus Beispiel 17.3 (d) hat im Punkt $(1, 1)$ eine lokale Minimalstelle, denn für alle $(x, y) \in \mathbb{R}^2$ gilt

$$\begin{aligned} f(x, y) &= -\cos((x-1)^2 + (y-1)^2) \\ &\geq -1 = -\cos(0) = -\cos((1-1)^2 + (1-1)^2). \end{aligned}$$

Wir sehen an dieser Rechnung direkt, dass $m = -1$ ein (globales) Minimum von f ist. Dieses (globale) Minimum $m = -1$ wird auch in allen Punkten (x_k, y_k) aus $\partial U_{\sqrt{k\pi}}(1, 1)$ mit $k \in \mathbb{N}$ gerade angenommen, denn für alle $(x, y) \in \mathbb{R}^2$ gilt für gerades $k \in \mathbb{N}$

$$\begin{aligned} f(x, y) &= -\cos((x-1)^2 + (y-1)^2) \\ &\geq -1 = (-1)^{k+1} = -\cos((x_k-1)^2 + (y_k-1)^2). \end{aligned}$$

In allen Punkten (x_k, y_k) aus $\partial U_{\sqrt{k\pi}}(1, 1)$ mit $k \in \mathbb{N}$ ungerade gilt für alle $(x, y) \in \mathbb{R}^2$

$$\begin{aligned} f(x, y) &= -\cos((x-1)^2 + (y-1)^2) \\ &\leq 1 = (-1)^{k+1} = -\cos((x_k-1)^2 + (y_k-1)^2). \end{aligned}$$

Also ist $M = 1$ ein (globales) Maximum von f , dass in allen Punkten aus $\partial U_{\sqrt{k\pi}}(1, 1)$ mit $k \in \mathbb{N}$ ungerade angenommen wird.

Wir haben also gefunden

$$\min_{(x,y) \in \mathbb{R}^2} f(x, y) = \min_{(x,y) \in \mathbb{R}^2} -\cos((x-1)^2 + (y-1)^2) = -1,$$

$$\max_{(x,y) \in \mathbb{R}^2} f(x, y) = \max_{(x,y) \in \mathbb{R}^2} -\cos((x-1)^2 + (y-1)^2) = 1.$$

Der Graph der Funktion ist im rechten Bild in Abbildung 17.2 gezeichnet.

Was ist mit Extremalstellen und dem (globalen) Maximum und Minimum einer Funktion, die nicht auf ganz \mathbb{R}^n sondern auf einer Menge $\bar{\Omega} \subseteq \mathbb{R}^n$ mit $\bar{\Omega} \neq \mathbb{R}^n$ definiert ist?

Hier müssen wir bei der Untersuchung auf ein (globales) Maximum bzw. Minimum an den Randpunkten von $\bar{\Omega}$ aufpassen: Am Rand $\partial\Omega = \partial\bar{\Omega}$ greift Satz 17.14 nicht, denn dieser Satz setzt einen **offenen** Definitionsbereich der \mathcal{C}^2 -Funktion voraus und kann daher nur (lokal) für Punkte aus $\bar{\Omega}$ angewendet werden, um die man eine kleine Kugel legen kann, die noch ganz in $\bar{\Omega} \setminus \partial\Omega$ liegt.

Um diese Situation einfacher analysieren und beschreiben zu können, führen wir noch weitere Begriffe für Mengen in \mathbb{R}^n ein.

Definition 17.19. (beschränkte, abgeschlossene bzw. kompakte Menge in \mathbb{R}^n)

- (1) $K \subseteq \mathbb{R}^n$ heißt **beschränkt**, wenn ein $C > 0$ existiert mit $\|\vec{x}\| \leq C$ für alle $\vec{x} \in K$.
- (2) $K \subseteq \mathbb{R}^n$ heißt **abgeschlossen**, wenn der Rand ∂K von K ganz zu K gehört, d.h. wenn $\partial K \subseteq K$.
- (3) $K \subseteq \mathbb{R}^n$ heißt **kompakt**, wenn K beschränkt und abgeschlossen ist.

Beispiel 17.20. (beschränkte, abgeschlossene, kompakte Menge in \mathbb{R}^n)

- (a) $K_1 := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ ist kompakt, denn K_1 ist beschränkt, da $\|(x, y)\| \leq 1$ für alle $(x, y) \in K_1$, und abgeschlossen, da

$$\partial K_1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\} \subseteq K_1.$$

Analog ist $\{\vec{x} \in \mathbb{R}^n : \|\vec{x}\| \leq 1\}$ kompakt.

- (b) $K_2 := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$ ist nicht kompakt, denn der Rand

$$\partial K_2 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$$

ist nicht in K_2 enthalten. K_2 ist beschränkt, da $\|(x, y)\| \leq 1$ für alle $(x, y) \in K_2$ ist und offen, da $K_2 = U_1(0, 0)$.

Analog ist $\{\vec{x} \in \mathbb{R}^n : \|\vec{x}\| < 1\}$ nicht kompakt.

- (c) $K_3 := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ ist kompakt, denn K_3 ist beschränkt, da $\|(x, y)\| \leq 1$ für alle $(x, y) \in K_3$, und abgeschlossen, da und $\partial K_3 = K_3$.

Analog ist $\{\vec{x} \in \mathbb{R}^n : \|\vec{x}\| = 1\}$ kompakt.

- (d) \mathbb{R}^n ist nicht kompakt, denn \mathbb{R}^n ist nicht beschränkt.

- (e) Der Würfel $W_1 := \{\vec{x} \in \mathbb{R}^n : \max\{|x_1|, |x_2|, \dots, |x_n|\} \leq 1\}$ ist kompakt, denn er ist beschränkt, da für alle $\vec{x} \in W_1$ gilt

$$\begin{aligned} \|\vec{x}\| &= \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} = \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2} \\ &\leq \sqrt{1 + 1 + \dots + 1} = \sqrt{n}, \end{aligned}$$

und abgeschlossen, da sein Rand

$$\partial W_1 := \{\vec{x} \in \mathbb{R}^n : \max\{|x_1|, |x_2|, \dots, |x_n|\} = 1\}$$

zu dem Würfel dazugehört.

Nun können wir den Satz über das (globale) Maximum bzw. Minimum einer stetigen Funktion mit einer kompakten Definitionsmenge in \mathbb{R}^n formulieren.

Satz 17.21. (stetige Funktion auf kompakter Menge hat globales Minimum und Maximum)

Sei $K \subseteq \mathbb{R}^n$ **kompakt** und nicht leer. Ist $f : K \rightarrow \mathbb{R}$ stetig, so hat f sowohl ein **globales Minimum** als auch ein **globales Maximum**.

Bemerkung 17.22. (Warum ist Satz 17.21 wichtig?)

Sei $f : K \rightarrow \mathbb{R}$ eine stetige Funktion auf der kompakten Menge $K \subseteq \mathbb{R}^n$. Nach Satz 17.21 wissen wir, dass f auf K ein globales Minimum und ein globales Maximum hat, d.h. es gibt \vec{a}_1 und \vec{a}_2 in K mit

$$\begin{aligned} f(\vec{a}_1) &\leq f(\vec{x}) && \text{für alle } \vec{x} \in K, \\ f(\vec{a}_2) &\geq f(\vec{x}) && \text{für alle } \vec{x} \in K. \end{aligned}$$

Dabei ist \vec{a}_1 bzw. \vec{a}_2 jeweils entweder in der offenen Menge $K \setminus \partial K$ oder im Rand ∂K von K .

Ist f eingeschränkt auf $K \setminus \partial K$ in $\mathcal{C}^2(K \setminus \partial K, \mathbb{R})$, dann können wir Kandidaten für lokale Extremalstellen in $K \setminus \partial K$ mit der Hilfe von Satz 17.2 und Satz 17.14 identifizieren und untersuchen. Wir benötigen allerdings noch ein Verfahren, um die Punkte in ∂K zu untersuchen.

Betrachten zwei Beispiele.

Beispiel 17.23. (Minimum bzw. Maximum einer Funktion auf einer kompakten Menge)

Betrachten wir die Funktion \mathcal{C}^∞ -Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = e^{-(x^2+y^2)},$$

aus Beispiel 17.3 (a) und Beispiel 17.15 (a). Wir interessieren uns für die globalen Extrema von f auf der kompakten Einheitskreisscheibe

$$K := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}.$$

Nutzen wir ebene Polarkoordinaten $x = r \cos(\phi)$, $y = r \sin(\phi)$, so sehen wir, dass

$$\begin{aligned} f(r \cos(\phi), r \sin(\phi)) &= \exp\left(-\left[(r \cos(\phi))^2 + (r \sin(\phi))^2\right]\right) \\ &= \exp\left(-r^2 [\cos^2(\phi) + \sin^2(\phi)]\right) = \exp(-r^2) = e^{-r^2}, \end{aligned}$$

d.h. wir haben eine Funktion, die nur vom Radius r , also vom Abstand zum Nullpunkt $(0, 0)$, abhängt (vgl. auch das linke Bild in Abbildung 17.1). Die kompakte Einheitskreisscheibe besteht in Polarkoordinaten aus allen Punkte mit $x^2 + y^2 = \|(x, y)\|^2 = r^2 \leq 1$ also mit $r \leq 1$. Aus dem Verhalten von $g(r) := e^{-r^2}$ auf $[0, 1]$ können wir nun alles ablesen: Diese Funktion g hat ihr Maximum $M = 1 = e^0$ in $r = 0$, und dieses entspricht dem (globalen) Maximum $M = 1$ von f in $(0, 0)$ (vgl. Beispiele 17.18 (b) und 17.3 (a)). g hat ihr Minimum $m = e^{-1}$ in $r = 1$, und dieses entspricht allen Punkten auf dem Rand

$$\partial K := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}.$$

der Einheitskreisscheibe.

Also hat $f : K \rightarrow \mathbb{R}$, $f(x, y) = e^{-(x^2+y^2)}$, das Maximum $M = 1$, welches in $(0, 0)$ angenommen wird, und das Minimum $m = e^{-1}$, welches in allen Punkten auf ∂K angenommen wird.

Im vorigen Beispiel hatten wir eine um $(0, 0)$ rotationssymmetrische Funktion auf einer um $(0, 0)$ rotationssymmetrischen kompakten Menge und konnten daher das Problem der Bestimmung des (globalen) Maximums bzw. Minimums auf das Problem der Bestimmung des Maximums bzw. Minimums einer Funktion einer Variablen reduzieren. Leider ist die Situation nicht immer so einfach, wie das nächste Beispiel zeigt.

Beispiel 17.24. (Minimum bzw. Maximum einer Funktion auf einer kompakten Menge)

Betrachten wir die Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) := xy.$$

Gesucht sind die globalen Extrema von f auf der kompakten Einheitskreisscheibe

$$K := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}.$$

Da diese Funktion nicht rotationssymmetrisch um $(0, 0)$ ist, können wir hier nicht so bequem wie im vorigen Beispiel vorgehen.

Lösungsweg:

- (1) Wir bestimmen zunächst die lokalen Extremalstellen von f , die in der offenen Menge $K \setminus \partial K$ liegen, und berechnen die zugehörigen Funktionswerte. Diese sind Kandidaten für die globalen Extrema.

In Beispiel 17.15 (c) haben wir gesehen, dass f keine lokalen Extrema in \mathbb{R}^2 und damit auch in $K \setminus \partial K$ hat. Daher erhalten wir auf diese Weise keine Kandidaten für globale Extrema.

- (2) Wir betrachten nun den Rand ∂K von K . Dieser ist hier die Einheitskreislinie

$$\partial K = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}.$$

Wir suchen nun lokale Extrema von $f(x, y) = xy$ unter der Nebenbedingung

$$x^2 + y^2 = 1 \quad \Longleftrightarrow \quad x^2 + y^2 - 1 = 0.$$

Wie man Kandidaten (x, y) für lokale Extrema unter Nebenbedingungen berechnet, ist das Thema des nachfolgenden Teilkapitels. Wenn wir die Kandidaten $(x, y) \in \partial K$ der lokalen Extrema gefunden haben, bestimmen wir deren Funktionswerte, um zu sehen, ob in einem dieser Punkte ein Minimum bzw. Maximum auftritt.

17.5 Extrema unter Nebenbedingungen

In Beispiel 17.24 haben wir gesehen, dass man zur Bestimmung der Extrema einer \mathcal{C}^2 -Funktion f auf einer kompakten Menge $K \subseteq \mathbb{R}^n$ die Punkte auf dem Rand ∂K daraufhin untersuchen muss, ob in ihnen f ein Extremum annimmt. Wie bereits im Beispiel 17.24 angekündigt, machen wir dieses, indem wir $f(\vec{x})$ mit der Nebenbedingung $\vec{x} \in \partial K$ betrachten. Diese Nebenbedingung lässt sich in der

Regel mit einer \mathcal{C}^1 -Funktion $g : \mathbb{R}^n \rightarrow \mathbb{R}$ als $g(\vec{x}) = 0$ beschreiben. In Beispiel 17.24 war $g(x, y) = x^2 + y^2 - 1$.

Wir formulieren unsere neue Aufgabe zunächst mathematisch und noch etwas allgemeiner.

Definition 17.25. (lokale Extrema unter Nebenbedingungen)

Seien $n, p \in \mathbb{N}$ mit $p < n$, $\Omega \subseteq \mathbb{R}^n$ offen, $f \in \mathcal{C}^1(\Omega, \mathbb{R})$ und $\vec{g} \in \mathcal{C}^1(\Omega, \mathbb{R}^p)$.

- (1) f hat in $\vec{a} \in \Omega$ ein **lokales Minimum** unter der Nebenbedingung $\vec{g}(\vec{x}) = \vec{0}$, wenn
- (i) $\vec{g}(\vec{a}) = \vec{0}$ ist und
 - (ii) ein $\varepsilon > 0$ existiert mit $f(\vec{x}) \geq f(\vec{a})$ für alle $\vec{x} \in U_\varepsilon(\vec{a})$ mit $\vec{g}(\vec{x}) = \vec{0}$.
- (2) f hat in $\vec{a} \in \Omega$ ein **lokales Maximum** unter der Nebenbedingung $\vec{g}(\vec{x}) = \vec{0}$, wenn
- (i) $\vec{g}(\vec{a}) = \vec{0}$ ist und
 - (ii) ein $\varepsilon > 0$ existiert mit $f(\vec{x}) \leq f(\vec{a})$ für alle $\vec{x} \in U_\varepsilon(\vec{a})$ mit $\vec{g}(\vec{x}) = \vec{0}$.

Der nächste Satz sagt uns indirekt, wie wir Kandidaten für lokale Extrema unter Nebenbedingungen finden können.

Satz 17.26. (Lagrangesche Multiplikatorenregel)

Sei $\Omega \subseteq \mathbb{R}^n$ offen, $f \in \mathcal{C}^1(\Omega, \mathbb{R})$ und $\vec{g} \in \mathcal{C}^1(\Omega, \mathbb{R}^p)$ mit $p < n$. Die Ableitung $\vec{g}'(\vec{x}) \in \mathbb{R}^{p \times n}$ von \vec{g} habe in allen $\vec{x} \in \Omega$, die $\vec{g}(\vec{x}) = \vec{0}$ erfüllen, den Rang p . Hat f in \vec{a} ein lokales Extremum unter der Nebenbedingung $\vec{g}(\vec{x}) = \vec{0}$, dann existieren $\lambda_1, \lambda_2, \dots, \lambda_p \in \mathbb{R}$, so dass der Punkt \vec{a} für die Funktion

$$F : \Omega \rightarrow \mathbb{R}, \quad F(x_1, x_2, \dots, x_n) := f(x_1, x_2, \dots, x_n) + \sum_{k=1}^p \lambda_k g_k(x_1, x_2, \dots, x_n),$$

ein **kritischer Punkt** ist. Die Zahlen $\lambda_1, \lambda_2, \dots, \lambda_p$ heißen **Lagrange-Multiplikatoren**.

Die Existenzaussage des vorigen Satzes liefert uns ein Berechnungsverfahren für

Punkte, die Kandidaten für Extrema unter der gegebenen Nebenbedingung sind. Dieses wird in der nachfolgenden Methode erklärt.

Methode 17.27. (Verfahren der Lagrangeschen Multiplikatoren)

Es gelten die Bezeichnungen und Voraussetzungen aus Satz 17.26. Nach Satz 17.26 sind Punkte $\vec{\mathbf{a}}$, in denen $F'(\vec{\mathbf{a}}) = \mathbf{O}_{1 \times n}$ gilt, Kandidaten für Punkte in denen lokale Extrema von f unter der Nebenbedingung $\vec{\mathbf{g}}(\vec{\mathbf{x}}) = \vec{\mathbf{0}}$ auftreten. $F'(\vec{\mathbf{x}}) = \mathbf{O}_{1 \times n}$ liefert

$$F'(\vec{\mathbf{x}}) = f'(\vec{\mathbf{x}}) + \sum_{k=1}^p \lambda_k g'_k(\vec{\mathbf{x}}) = \mathbf{O}_{1 \times n}$$

oder ausführlicher

$$(\partial_i f)(\vec{\mathbf{x}}) + \lambda_1 (\partial_i g_1)(\vec{\mathbf{x}}) + \lambda_2 (\partial_i g_2)(\vec{\mathbf{x}}) + \dots + \lambda_p (\partial_i g_p)(\vec{\mathbf{x}}) = 0, \quad i = 1, 2, \dots, n. \quad (17.4)$$

Zusätzlich zu diesen n Gleichungen muss man noch die p Gleichungen berücksichtigen, die man aus der Nebenbedingung $\vec{\mathbf{g}}(\vec{\mathbf{x}}) = \vec{\mathbf{0}}$ erhält, also

$$g_k(\vec{\mathbf{x}}) = 0, \quad k = 1, 2, \dots, p, \quad (17.5)$$

wobei g_1, g_2, \dots, g_p die p Komponentenfunktionen von $\vec{\mathbf{g}}$ sind.

Aus den $n + p$ Gleichungen (17.4) und (17.5) kann man nun die $n + p$ Unbekannten $a_1, a_2, \dots, a_n \in \mathbb{R}$ von $\vec{\mathbf{x}} = \vec{\mathbf{a}} = (a_1, a_2, \dots, a_n)$ und $\lambda_1, \lambda_2, \dots, \lambda_p \in \mathbb{R}$ bestimmen. (Achtung: Diese Gleichungen ergeben **nicht immer ein lineares** Gleichungssystem, denn die Gleichungen sind oft **nicht-linear** in x_1, x_2, \dots, x_n und $\lambda_1, \lambda_2, \dots, \lambda_p$.)

Bemerkung 17.28. (Verfahren der Lagrangeschen Multiplikatoren)

Es gelten die Bezeichnungen und Voraussetzungen aus Satz 17.26. Man erhält alle $n + p$ Gleichungen (17.4) und (17.5), wenn man die kritischen Punkte der Funktion $\Phi : \Omega \times \mathbb{R}^p \rightarrow \mathbb{R}$,

$$\Phi(x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_p) := f(x_1, x_2, \dots, x_n) + \sum_{k=1}^p \lambda_k g_k(x_1, x_2, \dots, x_n),$$

sucht. Hier haben wir in der Funktion F nun auch noch $\lambda_1, \lambda_2, \dots, \lambda_p$ als zusätzliche Variablen aufgefasst. Die partiellen Ableitungen von Φ nach den

Variablen x_1, x_2, \dots, x_n sind dieselben wie die von F , und aus den partiellen Ableitungen nach $\lambda_1, \lambda_2, \dots, \lambda_p$ erhält man jeweils die Nebenbedingungen $g_1(\vec{x}) = 0, g_2(\vec{x}) = 0, \dots, g_p(\vec{x}) = 0$.

Betrachten wir mehrere Beispiele.

Beispiel 17.29. (Extrema unter Nebenbedingungen)

(a) Bestimmen Sie die globalen Extrema der Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) := xy,$$

auf der Einheitskreislinie $S := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$.

Hier ist also $p = 1$ und die Nebenbedingung lautet

$$g(x, y) := x^2 + y^2 - 1 = 0.$$

Schritt 1: Überprüfen der Rangbedingung

$$g'(x, y) = [2x \quad 2y],$$

und für alle (x, y) mit $x^2 + y^2 = 1$ ist $(x, y) \neq (0, 0)$ und daher $g'(x, y) \neq \mathbf{0}_{1 \times 2}$ und somit $\text{Rang}(g'(x, y)) = 1$. Also ist die Rangbedingung erfüllt.

Schritt 2: Anwenden der Lagrange-Multiplikatorenregel

$$\Phi(x, y, \lambda) := f(x, y) + \lambda g(x, y) = xy + \lambda(x^2 + y^2 - 1)$$

Die partiellen Ableitungen sind

$$\begin{aligned} \frac{\partial}{\partial x} \Phi(x, y, \lambda) &= y + 2\lambda x, \\ \frac{\partial}{\partial y} \Phi(x, y, \lambda) &= x + 2\lambda y, \\ \frac{\partial}{\partial \lambda} \Phi(x, y, \lambda) &= x^2 + y^2 - 1. \end{aligned}$$

Die partiellen Ableitungen gleich null setzen liefert die folgenden drei Gleichungen:

$$y + 2\lambda x = 0 \quad (\text{I})$$

$$x + 2\lambda y = 0 \quad (\text{II})$$

$$x^2 + y^2 - 1 = 0 \quad (\text{III})$$

Schritt 3: Lösen des Gleichungssystems

$$(I) \iff y = -2\lambda x$$

$$\text{Einsetzen in (II): } x - 4\lambda^2 x = 0 \iff x(1 - 4\lambda^2) = 0$$

$$\iff \left(x = 0 \text{ oder } \lambda^2 = \frac{1}{4} \right)$$

Wäre $x = 0$, so folgt aus (I) $y = 0$. Diese ist ein Widerspruch zu (III). Also kann $x = 0$ nicht auftreten.

$$\text{Daher gilt } \lambda^2 = \frac{1}{4}, \text{ d.h. } \lambda = \frac{1}{2} \text{ oder } \lambda = -\frac{1}{2}.$$

$$\text{Einsetzen von } \lambda = -\frac{1}{2} \text{ bzw. } \lambda = \frac{1}{2} \text{ in (I) liefert: } y = x \text{ bzw. } y = -x$$

Einsetzen von $y = x$ bzw. $y = -x$ in (III) liefert:

$$2x^2 = 1 \iff x^2 = \frac{1}{2} \iff \left(\boxed{x = \frac{1}{\sqrt{2}}} \text{ oder } \boxed{x = -\frac{1}{\sqrt{2}}} \right)$$

Also sind die Kandidaten für lokale Extrema auf der Einheitskreislinie durch

$$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right), \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right), \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right)$$

gegeben.

Schritt 4: Berechnen der Funktionswerte

$$f\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = f\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = \frac{1}{2}$$

$$f\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = f\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = -\frac{1}{2}.$$

Ergebnis: Da die Einheitskreislinie S kompakt ist und f auf S stetig ist, nimmt f auf S nach Satz 17.21 ein Minimum und ein Maximum an. Nach Satz 17.26 und Methode 17.27 müssen die Punkte, in denen das Maximum bzw. Minimum auftreten, als mögliche Kandidaten von dem Verfahren der Lagrange-Multiplikatoren erfasst sein. Daher können wir schließen, dass gilt

$$\text{Maximum: } \max_{(x,y) \in S} f(x,y) = \max_{(x,y) \in S} xy = \frac{1}{2},$$

$$\text{Minimum: } \min_{(x,y) \in S} f(x,y) = \min_{(x,y) \in S} xy = -\frac{1}{2}.$$

- (b) Mit dem Ergebnis aus (a) können wir auch Beispiel 17.24 abschließen: Hier waren die Extrema von $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := xy$, auf der kompakten Einheitskreisscheibe $K := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ gesucht.

Aus Beispiel 17.15 (c) wissen wir bereits, dass f keine lokalen Extremalstellen auf \mathbb{R}^2 und damit insbesondere keine lokalen Extremalstellen auf der offenen Menge $K \setminus \partial K$ hat.

Nach (a) hat f auf $S = \partial K$ die Extrema $-\frac{1}{2}$ und $\frac{1}{2}$.

Damit folgt, dass die Extrema von f auf K die folgenden sind:

$$\begin{aligned} \text{Maximum:} \quad & \max_{(x,y) \in K} f(x, y) = \max_{(x,y) \in K} xy = \frac{1}{2}, \\ \text{Minimum:} \quad & \min_{(x,y) \in K} f(x, y) = \min_{(x,y) \in K} xy = -\frac{1}{2}. \end{aligned}$$

Beispiel 17.30. (Extrema unter Nebenbedingungen)

Gesucht sind die globalen Extrema von $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := xy^2$, auf der Einheitskreislinie $x^2 + y^2 = 1$, also auf

$$S := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}.$$

Hier ist also $p = 1$ und $g(x, y) := x^2 + y^2 - 1$ (wie in vorigen Beispiel).

Schritt 1: Überprüfen der Rangbedingung

$$g'(x, y) = [2x \quad 2y],$$

und für alle (x, y) mit $x^2 + y^2 = 1$ ist $(x, y) \neq (0, 0)$ und daher $g'(x, y) \neq \mathbf{0}_{1 \times 2}$ und somit $\text{Rang}(g'(x, y)) = 1$. Also ist die Rangbedingung erfüllt.

Schritt 2: Anwenden der Lagrange-Multiplikatorenregel

$$\Phi(x, y, \lambda) := xy^2 + \lambda(x^2 + y^2 - 1)$$

Die partiellen Ableitungen sind

$$\begin{aligned} \frac{\partial}{\partial x} \Phi(x, y, \lambda) &= y^2 + 2\lambda x, \\ \frac{\partial}{\partial y} \Phi(x, y, \lambda) &= 2xy + 2\lambda y, \\ \frac{\partial}{\partial \lambda} \Phi(x, y, \lambda) &= x^2 + y^2 - 1. \end{aligned}$$

Die partiellen Ableitungen gleich null setzen liefert die folgenden drei Gleichungen:

$$y^2 + 2\lambda x = 0 \quad (\text{I})$$

$$2xy + 2\lambda y = 0 \quad (\text{II})$$

$$x^2 + y^2 - 1 = 0 \quad (\text{III})$$

Schritt 3: Lösen des Gleichungssystems

$$(\text{II}) \iff 2y(x + \lambda) = 0 \iff (\boxed{y = 0} \text{ oder } \boxed{\lambda = -x})$$

- Falls $y = 0$ ist, folgt aus (III): $x^2 = 1 \iff (\boxed{x = 1} \text{ oder } \boxed{x = -1})$,
und aus (I) folgt: $\lambda = 0$

Für diesen Fall erhalten wir also zwei Kandidaten für lokale Extremalstellen:

$$\boxed{(1, 0), (-1, 0)}$$

- Falls $\boxed{\lambda = -x}$, ist folgt aus (I): $y^2 - 2x^2 = 0 \iff y^2 = 2x^2$,

und Einsetzen von $y^2 = 2x^2$ in (III) liefert: $3x^2 = 1 \iff x^2 = \frac{1}{3}$

$$\iff \boxed{x = \pm\sqrt{\frac{1}{3}}}$$

Einsetzen von $x^2 = \frac{1}{3}$ in (I) liefert: $y^2 = 2x^2 = \frac{2}{3} \iff \boxed{y = \pm\sqrt{\frac{2}{3}}}$

Für diesen Fall erhalten wir also die folgenden Kandidaten für lokale Extremalstellen:

$$\left(\sqrt{\frac{1}{3}}, \sqrt{\frac{2}{3}}\right), \left(\sqrt{\frac{1}{3}}, -\sqrt{\frac{2}{3}}\right), \left(-\sqrt{\frac{1}{3}}, \sqrt{\frac{2}{3}}\right), \left(-\sqrt{\frac{1}{3}}, -\sqrt{\frac{2}{3}}\right).$$

Schritt 4: Berechnen der Funktionswerte

$$f(1, 0) = 0,$$

$$f(-1, 0) = 0,$$

$$f\left(\sqrt{\frac{1}{3}}, \sqrt{\frac{2}{3}}\right) = \sqrt{\frac{1}{3}} \cdot \frac{2}{3} = \frac{2}{3\sqrt{3}},$$

$$f\left(\sqrt{\frac{1}{3}}, -\sqrt{\frac{2}{3}}\right) = \sqrt{\frac{1}{3}} \cdot \frac{2}{3} = \frac{2}{3\sqrt{3}},$$

$$f\left(-\sqrt{\frac{1}{3}}, \sqrt{\frac{2}{3}}\right) = -\sqrt{\frac{1}{3}} \cdot \frac{2}{3} = -\frac{2}{3\sqrt{3}},$$

$$f\left(-\sqrt{\frac{1}{3}}, -\sqrt{\frac{2}{3}}\right) = -\sqrt{\frac{1}{3}} \cdot \frac{2}{3} = -\frac{2}{3\sqrt{3}}.$$

Ergebnis: Da die Einheitskreislinie S kompakt ist und f auf S stetig ist, nimmt f auf S nach Satz 17.21 ein Minimum und ein Maximum an. Nach Satz 17.26 und Methode 17.27 müssen die Punkte, in denen das Maximum bzw. Minimum auftreten, als Kandidaten für lokale Extrema vom Verfahren der Lagrange-Multiplikatoren erfasst sein. Daher können wir schließen, dass gilt

$$\begin{aligned} \text{Maximum:} \quad & \max_{(x,y) \in S} f(x,y) = \max_{(x,y) \in S} x y^2 = \frac{2}{3\sqrt{3}}, \\ \text{Minimum:} \quad & \min_{(x,y) \in S} f(x,y) = \min_{(x,y) \in S} x y^2 = -\frac{2}{3\sqrt{3}}. \end{aligned}$$

In unserem dritten Beispiel haben wir eine vektorwertige Nebenbedingung, d.h. konkret zwei reellwertige Nebenbedingungen.

Beispiel 17.31. (Extrema unter Nebenbedingungen)

Bestimmen Sie die Extrema von $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, $f(x, y, z) := 5x + y - 3z$, auf dem Schnitt der Ebene $x + y + z = 0$ mit der Kugeloberfläche $x^2 + y^2 + z^2 = 1$.

Hier ist also $p = 2$, $K := \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0 \text{ und } x^2 + y^2 + z^2 = 1\}$, und die Nebenbedingungen werden durch die Funktion

$$\vec{g}(x, y, z) := \begin{bmatrix} g_1(x, y, z) \\ g_2(x, y, z) \end{bmatrix} = \begin{bmatrix} x + y + z \\ x^2 + y^2 + z^2 - 1 \end{bmatrix}$$

beschrieben. Da die Ebene $x + y + z = 0$ durch den Nullpunkt $(0, 0, 0)$ geht, ist K eine Kreislinie im \mathbb{R}^3 mit Radius 1 und Mittelpunkt in $(0, 0, 0)$.

Schritt 1: Überprüfen der Rangbedingung

$$\vec{g}'(x, y, z) = \begin{bmatrix} 1 & 1 & 1 \\ 2x & 2y & 2z \end{bmatrix}$$

Für die Methode der Lagrange-Multiplikatoren muss $\text{Rang}(\vec{g}'(x, y, z)) = 2$ für alle $(x, y, z) \in K$ gelten. Wegen $x + y + z = 0$ kann in K die Gleichung $x = y = z$ nicht gelten. Dann muss $\vec{g}'(x, y, z)$ zwei linear unabhängige Zeilenvektoren haben, und es folgt $\text{Rang}(\vec{g}'(x, y, z)) = 2$ für alle $(x, y, z) \in K$.

Schritt 2: Anwenden der Lagrange-Multiplikatorenregel

$$\Phi(x, y, z, \lambda, \mu) := 5x + y - 3z + \lambda(x + y + z) + \mu(x^2 + y^2 + z^2 - 1)$$

Die partiellen Ableitungen sind

$$\frac{\partial}{\partial x} \Phi(x, y, z, \lambda, \mu) = 5 + \lambda + 2\mu x,$$

$$\frac{\partial}{\partial y} \Phi(x, y, z, \lambda, \mu) = 1 + \lambda + 2\mu y,$$

$$\frac{\partial}{\partial z} \Phi(x, y, z, \lambda, \mu) = -3 + \lambda + 2\mu z,$$

$$\frac{\partial}{\partial \lambda} \Phi(x, y, z, \lambda, \mu) = x + y + z,$$

$$\frac{\partial}{\partial \mu} \Phi(x, y, z, \lambda, \mu) = x^2 + y^2 + z^2 - 1.$$

Die partiellen Ableitungen gleich null setzen liefert die fünf Gleichungen

$$5 + \lambda + 2\mu x = 0 \quad (\text{I})$$

$$1 + \lambda + 2\mu y = 0 \quad (\text{II})$$

$$-3 + \lambda + 2\mu z = 0 \quad (\text{III})$$

$$x + y + z = 0 \quad (\text{IV})$$

$$x^2 + y^2 + z^2 - 1 = 0 \quad (\text{V})$$

Schritt 3: Lösen des Gleichungssystems

$$(\text{I}) + (\text{II}) + (\text{III}) \text{ liefert: } 3 + 3\lambda + 2\mu \underbrace{(x + y + z)}_{=0 \text{ nach (IV)}} = 0 \quad \Longrightarrow \quad 3(1 + \lambda) = 0$$

$$\Longrightarrow \quad \lambda = -1$$

Einsetzen von $\lambda = -1$ in (I), (II) und (III) liefert:

$$\text{Aus (I): } 4 + 2\mu x = 0 \quad \Longleftrightarrow \quad \mu x = -2 \quad \Longleftrightarrow \quad \left(x = -\frac{2}{\mu} \text{ und } \mu \neq 0 \right)$$

$$\text{Aus (II): } 2\mu y = 0 \quad \Longleftrightarrow \quad \mu y = 0 \quad \Longleftrightarrow \quad (y = 0 \text{ oder } \mu = 0)$$

$$\text{Aus (III): } -4 + 2\mu z = 0 \quad \Longleftrightarrow \quad \mu z = 2 \quad \Longleftrightarrow \quad \left(z = \frac{2}{\mu} \text{ und } \mu \neq 0 \right)$$

Wir sehen also aus (I) und (III), dass $\mu = 0$ nicht gelten kann, und finden somit aus (I), (II) und (III), dass

$$\boxed{x = -\frac{2}{\mu}, \quad y = 0, \quad z = \frac{2}{\mu}.} \quad (17.6)$$

Einsetzen dieser Bedingungen in (V) liefert

$$\begin{aligned} \left(-\frac{2}{\mu}\right)^2 + 0^2 + \left(\frac{2}{\mu}\right)^2 - 1 = 0 &\iff \frac{8}{\mu^2} = 1 \\ \iff \mu^2 = 8 &\iff \left(\boxed{\mu = 2\sqrt{2}} \text{ oder } \boxed{\mu = -2\sqrt{2}}\right). \end{aligned}$$

Einsetzen von $\mu = 2\sqrt{2}$ bzw. $\mu = -2\sqrt{2}$ in (17.6) liefert die folgenden Kandidaten für lokale Extrema

$$\left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), \quad \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right).$$

Diese erfüllen auch Gleichung (IV).

Schritt 4: Berechnen der Funktionswerte an möglichen lokalen Extremalstellen

$$\begin{aligned} f\left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) &= -\frac{5}{\sqrt{2}} - \frac{3}{\sqrt{2}} = -\frac{8}{\sqrt{2}} = -4\sqrt{2}, \\ f\left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right) &= \frac{5}{\sqrt{2}} + \frac{3}{\sqrt{2}} = \frac{8}{\sqrt{2}} = 4\sqrt{2}. \end{aligned}$$

Ergebnis: Da die Kreislinie K im \mathbb{R}^3 kompakt ist und f auf K stetig ist, nimmt f auf K nach Satz 17.21 ein Minimum und ein Maximum an. Nach Satz 17.26 und Methode 17.27 müssen die Punkte, in denen das Maximum bzw. Minimum auftreten, als Kandidaten für lokale Extrema vom Verfahren der Lagrange-Multiplikatoren erfasst sein. Daher können wir schließen, dass gilt

$$\text{Maximum:} \quad \max_{(x,y,z) \in K} f(x, y, z) = \max_{(x,y,z) \in K} (5x + y - 3z) = 4\sqrt{2},$$

$$\text{Minimum:} \quad \min_{(x,y,z) \in K} f(x, y, z) = \min_{(x,y,z) \in K} (5x + y - 3z) = -4\sqrt{2}.$$