$$\det\left(\frac{\partial f_i}{\partial x_j}(a,b)\right)_{1\le i,j\le 2} \neq 0$$

LEHRBUCH

Christoph Ableitinger
Angela Herrmann

Lernen aus Musterlösungen zur Analysis und Linearen Algebra

Ein Arbeits- und Übungsbuch

2. Auflage

Springer Spektrum

Lernen aus Musterlösungen zur Analysis und Linearen Algebra

Christoph Ableitinger · Angela Herrmann

Lernen aus Musterlösungen zur Analysis und Linearen Algebra

Ein Arbeits- und Übungsbuch

2., verbesserte Auflage

Springer Spektrum

Dr. Christoph Ableitinger
Universität Wien, Österreich
christoph.ableitinger@univie.ac.at

Angela Herrmann
Universität Duisburg-Essen, Deutschland
angela.herrmann@uni-due.de

ISBN 978-3-658-02352-2
ISBN 978-3-658-02353-9 (eBook)
DOI 10.1007/978-3-658-02353-9

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Springer Spektrum

Planung und Lektorat: Ulrike Schmickler-Hirzebruch | Barbara Gerlach

Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier.

Springer Spektrum ist eine Marke von Springer DE. Springer DE ist Teil der Fachverlagsgruppe Springer Science+Business Media
www.springer-spektrum.de

Vorwort

Liebe Studentin, lieber Student,

schön, dass Sie sich dazu entschieden haben, mit diesem Buch zu arbeiten. Sie haben damit eine in gewisser Hinsicht einzigartige Lektüre vor sich. Es ist das erste und bis zum Erscheinen dieser Neuauflage auch das einzige Buch, das seinen Fokus auf die Arbeit mit ausführlichen Musterlösungen setzt. Es soll dazu beitragen, die Lücke zu schließen zwischen dem, was von Studierenden in den Mathematikveranstaltungen des ersten Studienjahres erwartet wird und dem, was sie zu diesem Zeitpunkt zu leisten im Stande sind.

Der Beginn Ihres Mathematikstudiums steckt voller Herausforderungen. Sie müssen viele neue Konzepte, Begriffe sowie Sprech- und Denkweisen erlernen. Sie werden sehen, bei der Bearbeitung der Übungsaufgaben laufen all diese Anforderungen zusammen. Wenig verwunderlich also, dass Jahr für Jahr viele Studierende[1] große Schwierigkeiten beim Lösen ihrer Haus- und Klausuraufgaben haben. Aber gerade das erfolgreiche Lösen von Aufgaben ist entscheidend für den Studienerfolg im ersten Jahr.

Durch die großzügige Förderung der *Deutsche Telekom-Stiftung* konnte an der Universität Duisburg-Essen das auf drei Jahre angelegte Projekt *Mathematik besser verstehen* durchgeführt werden. Es sollte damit den Überforderungen des ersten Studienjahres und der Hürde zwischen Schul- und Hochschulmathematik begegnet werden. Ziel dieses Projekts war unter anderem, die Studierenden beim Bearbeiten ihrer wöchentlichen Übungsaufgaben zu unterstützen. Ein wesentlicher Strang dieser Begleitung war die Anfertigung ausführlicher Musterlösungen. Es hat sich nämlich gezeigt, dass Studierende mit den üblicherweise recht kurz gefassten Musterlösungen nur schlecht zurechtkommen. Die positive Resonanz auf diese Projektaktivität hat uns dazu ermutigt, das vorliegende Buch zu verfassen.

Im *ersten Kapitel* gehen wir noch einmal genauer darauf ein, wie es zur Entstehung des Buches gekommen ist, welche Erwartungen wir erfüllen können bzw. enttäuschen müssen und welche inhaltliche Ausrichtung wir verfolgen.

Wir wollten nicht nur eine bloße Ansammlung von Musterlösungen schreiben, sondern das Thema *Lernen aus Musterlösungen* auch aus lern- und kognitionspsychologischer Sicht beleuchten. Dies geschieht im *zweiten Kapitel*, in dem wir die Konzepte der *Cognitive Load Theory*, des *Example-based Learning* und des *Cognitive Apprenticeship* vorstellen. Sie bilden das theoretische Fundament des sonst sehr praxisorientierten Buches.

Im *dritten Kapitel* benennen und reflektieren wir typische Teilprozesse beim Lösen klassischer Aufgaben aus der Analysis und der Linearen Algebra, die wir in unserer Forschungsarbeit herausgearbeitet haben. Die Analyse von Teillösungsprozessen er-

[1] Es wird nach Möglichkeit versucht, geschlechtsneutrale Personenbezeichnungen zu verwenden. Aus Gründen der Lesbarkeit werden wir im Folgenden allerdings teilweise auf Formulierungen wie „Studienanfängerinnen und -anfänger" verzichten und lediglich die allgemeine Bezeichnung „Studienanfänger" benutzen. Alle personenbezogenen Aussagen gelten – sofern nicht explizit anders formuliert – stets für Frauen und Männer gleichermaßen.

möglicht eine Meta-Sicht auf die Musterlösungen, die auch für künftige eigenständige Aufgabenlösungsprozesse von Bedeutung sein wird. Es werden außerdem zusätzliche Anforderungen des Aufgabenlösens diskutiert, die nebenbei im ersten Studienjahr mitgelernt werden müssen.

Der zweite Teil, der sich aus den *Kapiteln 4* bis *9* zusammensetzt, stellt den Kern des Buches dar. Hier finden Sie einen großen Fundus an typischen Aufgaben aus den Lehrveranstaltungen Analysis und Lineare Algebra und die dazu passenden ausführlichen Musterlösungen. In der vorliegenden Neuauflage des Buches haben wir außerdem ein Kapitel mit Musterlösungen eingeschoben, die sich mit Grundlagen mathematischen Arbeitens beschäftigen. Dadurch soll ein noch leichterer Einstieg für Sie als Leserin bzw. Leser geschaffen werden.

Es handelt sich in diesem zweiten Teil um authentische Aufgaben, die in dieser Formulierung auch im Übungsbetrieb am Campus Essen eingesetzt wurden. Verständnisfragen, Übungen und Arbeitsaufträge am Ende der jeweiligen Musterlösungen sollen Sie zum Nachdenken und zum intensiven Durch*arbeiten* der Lösungen motivieren. Denn durch bloßes Durch*lesen* der Lösungen können Sie sicher nicht das volle Potenzial des vorliegenden Buches ausschöpfen.

Auf welche Art und Weise Studierende tatsächlich mit den Musterlösungen arbeiten, versuchen wir übrigens in Begleitforschungsprojekten herauszufinden. Dabei hat sich u. a. gezeigt, dass das Arbeiten mit ausführlichen Musterlösungen neben vielen beobachtbaren positiven Wirkungen auch einen negativen Effekt haben kann – nämlich dann, wenn selbst einfache Passagen aus den Lösungen einfach unverstanden auf analoge Übungsaufgaben übertragen werden. Der Lerneffekt ist dabei natürlich vernachlässigbar! Versuchen Sie also – trotz der in diesem Buch abgedruckten „Vorlagen" – möglichst eigenständig an Ihren Übungsaufgaben zu arbeiten. Dozenten, die Musterlösungen aus diesem Buch in der Lehre einsetzen möchten, könnten je nach Lernstand ihrer Studierenden einzelne Passagen auch entfernen und die entstehenden Lücken zur Übung schließen lassen.

Im anschließenden *Übungsteil* können Sie sich selbst am Verfassen ausführlicher Musterlösungen zu vorgegebenen komprimierten Musterlösungen versuchen. Dazu müssen Sie zuerst die bewusst knapp gehaltenen Musterlösungen verstehen und sie anschließend in eine gut nachvollziehbare Form bringen. Sie werden sehen – das bringt Ihnen sowohl Übung im Durchdringen fremder als auch im Verfassen eigener mathematischer Texte.

Im abschließenden vierten Teil des Buches finden Sie Antwortvorschläge zu den Verständnisfragen und Arbeitsaufträgen des zweiten Teils und Lösungsvorschläge zu den Übungen des zweiten und dritten Teils.

Unser ausdrücklicher Dank gilt *Prof. Lisa Hefendehl-Hebeker* und *Prof. Gebhard Böckle* für die wunderbare Zusammenarbeit im Projekt *Mathematik besser verstehen* und für die befruchtenden Diskussionen dieses Buch betreffend. Außerdem danken wir *Prof. Wolfgang Lempken*, *Prof. Patrizio Neff*, *Prof. Petra Wittbold*, *Dr. Waldemar Pompe*, *Dr. Reiner Staszewski* und *Dr. Aleksandra Zimmermann* für die Erlaubnis, die von ihnen im Lehrbetrieb verwendeten Aufgaben in diesem Buch abzudrucken. Für die reibungslose und motivierende Zusammenarbeit mit dem Verlag Vieweg+Teubner (jetzt

Springer Spektrum) bedanken wir uns stellvertretend bei *Ulrike Schmickler-Hirzebruch*. *Karen Lippert* möchten wir für die Gestaltung des Layouts unseren Dank aussprechen. Der *Deutsche Telekom-Stiftung* gilt unser Dank für die großzügige Unterstützung des Projektes *Mathematik besser verstehen*, aus dem dieses Buch hervorgehen konnte. Herzlich bedanken wollen wir uns schließlich bei unseren Korrekturlesern *Johannes Lankeit* und *Stephan Lehmich* für ihre sorgfältige Arbeit.

Wien und Essen im Juni 2013 Christoph Ableitinger und Angela Herrmann

Inhaltsverzeichnis

Teil I

Einleitung, Theorie

1 Entstehung und inhaltliche Ausrichtung des Buches

Diskontinuität. Der Übergang von der Schul- zur Hochschulmathematik fällt erfahrungsgemäß vielen Studierenden sehr schwer. Sie erkennen das Schulfach, das in aller Regel mit guten Noten absolviert wurde, an der Hochschule nicht wieder. Stattdessen werden sie mit einer neuen Art des Denkens, mit neuen, andersartigen Fragestellungen und einem hohen Maß an Abstraktheit konfrontiert. Man spricht in diesem Zusammenhang zuweilen von einer *Diskontinuität* zwischen Schule und Hochschule, die erst überwunden werden müsse. Die hohen Abbrecherquoten in den ersten beiden Semestern des Mathematikstudiums unterstreichen deutlich, dass die Mathematik-Fachbereiche dieses Problem noch nicht in zufriedenstellender Weise lösen konnten. Ein Ansatz läge sicher darin, Studieninteressenten schon vorab besser darüber zu informieren, was sie in einem Mathematikstudium erwartet. Dadurch könnten falsche Vorstellungen schon im Keim erstickt und so die Entscheidungen für bzw. gegen die Studienwahl erleichtert werden.

Lehr- und Lernstrukturen verbessern. Eine zweite Möglichkeit bestünde darin, die Lehr- und Lernstrukturen besser an die individuellen Bedürfnisse der Studierenden anzupassen. Damit ist einerseits die Anschlussfähigkeit des Lernstoffs an das Vorwissen aus der Schule gemeint, andererseits auch die didaktische Aufbereitung des Lehrstoffs sowie die Gestaltung der Lernumgebungen, damit produktives, kognitiv aktivierendes Mathematiktreiben ermöglicht wird. Leider stehen für umfassende Reformen in diese Richtung den wenigsten Fachbereichen ausreichende Personal-, Raum- und Finanzressourcen zur Verfügung. Meist hängen Innovationen im Lehrbetrieb am Engagement einzelner Dozenten und Lehrpersonen. Nachhaltige, über den eigenen Fachbereich hinausgehende Veränderungen werden auf diese Weise selten auf den Weg gebracht.

Begleitende Unterstützung. Das vorliegende Buch möchte einen dritten Weg verfolgen. Entstanden ist die Idee im Rahmen des Projektes *Mathematik besser verstehen* an der Universität Duisburg-Essen, das begleitend zu den Anfängervorlesungen Analysis und Lineare Algebra durchgeführt wurde. Der Anspruch dieser von der *Deutsche Telekom-Stiftung* unterstützten Unternehmung war nicht, Vorlesungs- und Lehrstrukturen grundlegend zu verändern. Es ging vielmehr darum, den Studierenden bei den bestehenden Lehrveranstaltungen unterstützend unter die Arme zu greifen. Eine zumeist große Herausforderung für Studierende der Eingangsphase stellen die Übungsaufgaben dar, die Woche für Woche schriftlich zu bearbeiten sind. Eine der vielen Projektaktivitäten von *Mathematik besser verstehen* war das Anfertigen *ausführlicher* Musterlösungen, die effektive Lernprozesse bei den Studierenden einleiten und begleiten sollen.

C. Ableitinger, A. Herrmann, *Lernen aus Musterlösungen zur Analysis und Linearen Algebra*,
DOI 10.1007/978-3-658-02353-9_1, © Springer Fachmedien Wiesbaden 2013

Dazu möchten wir an dieser Stelle anhand einer typischen Aufgabe aus der Analysis 1 motivieren, wie es zu dieser Projektaktivität überhaupt gekommen ist:

Aufgabe 1.1 Sei $f : [0,1] \to [0,1]$ stetig. Zeigen Sie, dass es ein $c \in [0,1]$ mit $f(c) = c$ gibt, dass f also einen Fixpunkt besitzt.

Vielleicht ist Ihnen diese Aufgabe in Ihrer eigenen Analysis-Veranstaltung schon begegnet. Falls nicht, dann erkennen Sie aber bestimmt das Aufgabenformat wieder. Aufgaben des ersten Studienjahres ähneln einander oft hinsichtlich ihrer knappen Formulierung. Die meisten beinhalten keine Begründung, warum man gerade diese Aufgabe lösen soll, was man an ihr lernen kann, usw. Und die wenigsten werden durch geeignete graphische Veranschaulichungen begleitet. Das führt für Studierende, die wöchentlich mit solchen Aufgaben konfrontiert werden, häufig zu Schwierigkeiten. Wie soll eine Lösung gefunden werden, wenn nicht einmal die Aufgabenstellung aufgrund ihres dichten und abstrakten Informationsgehaltes verstanden wird?

Komprimierte Musterlösungen. Falls Dozenten ihren Studierenden überhaupt Musterlösungen zur Verfügung stellen, dann sind diese Musterlösungen meist ebenso knapp formuliert wie die Aufgabenstellung selbst. Sie sind auf die wesentlichen Informationen verdichtet, die zur erfolgreichen Lösung der jeweiligen Aufgabe nötig sind. Es wird dabei in keiner Weise der Aufgabenlösungs*prozess* abgebildet, der hinter der akkurat aufgeschriebenen fertigen Lösung steckt. Selbst wenn also Studierende die Lösungen im Nachhinein nachvollziehen können, lernen sie dabei nur bedingt etwas über das Aufgabenlösen an sich. Das Fach Mathematik tendiert dazu, seine Erkenntnisse so zu formulieren, dass die Spuren ihrer Entdeckung kaum noch sichtbar sind. Es zeichnet sich durch Prägnanz, Eleganz und die Vermeidung von Redundanzen aus. Sie werden im Laufe dieses Buches noch Gelegenheit bekommen, das Verfassen komprimierter Musterlösungen zu trainieren. Zunächst werden wir uns aber mit den Schwierigkeiten auseinandersetzen, die solche Musterlösungen für den Lernprozess mit sich bringen und wir werden Vorschläge für das Verfassen ausführlicher Musterlösungen machen. Eine zur Aufgabe 1.1 passende *komprimierte* Musterlösung könnte etwa wie folgt aussehen:

Komprimierte Lösung 1.1 Wenn $f(0) = 0$ bzw. $f(1) = 1$, dann ist $c = 0$ bzw. $c = 1$. Nehmen wir also an, dass $f(0) \neq 0$ und $f(1) \neq 1$, d. h. $f(0) > 0$, $f(1) < 1$. Sei g definiert durch $g(x) := f(x) - x$. Dann ist g auf $[0,1]$ stetig und

$$g(0) = f(0) > 0, g(1) = f(1) - 1 < 0.$$

Aus dem Zwischenwertsatz folgt: $\exists\, c \in (0,1)$ mit $g(c) = 0$, d. h. $f(c) = c$.

Sie können gerne versuchen, diese Lösung schrittweise nachzuvollziehen. Vermutlich wird Ihnen das auch gelingen. Und doch kostet es viel Anstrengung zu erfassen, warum gleich zu Beginn der Lösung zwei Sonderfälle betrachtet werden. Auch wird nicht geklärt, warum denn der Zwischenwertsatz das geeignete Hilfsmittel auf dem Weg zur Lösung ist. Der Transfer der Fixpunktaussage in eine dem Zwischenwertsatz zugängliche Nullstellenaussage wird ebenfalls nicht explizit gemacht.

Ausführliche Musterlösungen. Wir haben uns daher im Projekt *Mathematik besser verstehen* dazu entschlossen, zu einigen typischen Aufgaben aus der Analysis und der Linearen Algebra ausführliche Musterlösungen zu verfassen und diese unseren Studierenden im Anschluss an ihre eigenen Aufgabenlösungsprozesse zur Verfügung zu stellen. Einerseits sollten diese Lösungen als Kontrollmöglichkeit für jene Studierenden dienen, die die Aufgabe erfolgreich bearbeitet haben. Andererseits erhielten weniger erfolgreiche Studierende die Möglichkeit, wenigstens im Nachhinein die Aufgabenstellung, die Entwicklung von Lösungsideen sowie die komplette Genese der Aufgabenlösung zu verstehen. Eine zur Aufgabe 1.1 passende *ausführliche* Musterlösung sieht z. B. so aus:

Ausführliche Lösung 1.1 Wir sollen also in dieser Aufgabe zeigen, dass f unter den gegebenen Voraussetzungen einen Fixpunkt besitzt. Demnach soll lediglich die Existenz gezeigt werden, also weder Eindeutigkeit bewiesen noch eine konkrete Stelle angegeben werden.

Ein Satz aus der Vorlesung, der hier helfen könnte, ist der Zwischenwertsatz[1]. Dieser macht eine Existenzaussage, allerdings über eine Nullstelle und nicht über einen Fixpunkt. Wir sollten also versuchen, die Fixpunktaussage in eine Nullstellenaussage zu transformieren:

$$f(c) = c \quad \Leftrightarrow \quad f(c) - c = 0.$$

Wir definieren jetzt eine neue Funktion g durch $g(x) := f(x) - x$. Wenn wir zeigen können, dass diese Funktion g in $[0,1]$ eine Nullstelle besitzt, sind wir fertig. Das versuchen wir durch Anwendung des Zwischenwertsatzes.

Überprüfen der Voraussetzungen: g ist stetig, da f stetig ist und die Funktion h mit $h(x) := x$ stetig ist. $g(0) = f(0) - 0 \geq 0$, $g(1) = f(1) - 1 \leq 0$, da f laut Voraussetzung nur Werte in $[0,1]$ annimmt.

Die Voraussetzungen für den Zwischenwertsatz sind so nicht vollständig erfüllt, wir brauchen eine Fallunterscheidung:

1. Fall: $g(0) = 0$. Dann gilt aber $f(0) - 0 = 0$ und somit $f(0) = 0$. Demnach ist also $c = 0$ ein Fixpunkt von f.

2. Fall: $g(1) = 0$. Dann gilt aber $f(1) - 1 = 0$ und somit $f(1) = 1$. Demnach ist also $c = 1$ ein Fixpunkt von f.

3. Fall: $g(0) > 0$ und $g(1) < 0$. Dann folgt aus dem Zwischenwertsatz, dass es eine Stelle $c \in (0,1)$ mit $g(c) = 0$ gibt. Damit gilt aber $f(c) - c = 0$ und somit $f(c) = c$. Folglich ist diese Stelle c ein Fixpunkt von f.

Auslöser für dieses Buch. Die große Nachfrage und die vielen positiven Rückmeldungen unserer Studierenden haben dazu geführt, das Thema „Ausführliche Musterlösungen" auch aus fachdidaktischer, lernpsychologischer und bildungswissenschaftlicher Perspektive zu beleuchten. Eine ausführliche Darstellung finden Sie in den Kapiteln 2 und 3. Die positive Resonanz auf unsere Arbeit hat uns dazu ermutigt, die

[1] Wir legen folgende Formulierung des Zwischenwertsatzes zugrunde: Sei $g : [a,b] \to \mathbb{R}$ stetig und es gelte $g(a) < g(b)$. Dann gibt es zu jeder Zahl y mit $g(a) < y < g(b)$ eine Zahl $c \in (a,b)$ mit $g(c) = y$.

im Projekt entstandenen ausführlichen Musterlösungen einer breiteren Öffentlichkeit zugänglich zu machen. Das vorliegende Buch ist diesem Anliegen entsprungen.

Intention der Musterlösungen. Die Musterlösungen in diesem Buch haben begleitenden Charakter. Das Durcharbeiten der präsentierten Aufgaben kann selbstverständlich kein Lehrbuch und keine Vorlesung ersetzen. Auch ist dieses Buch kein Übungsbuch im klassischen Sinne, es sind nicht alle Themen der Anfängervorlesungen umfassend repräsentiert. Und doch ist es uns ein Anliegen gewesen, eine große Sammlung idealtypischer Aufgaben aus der Analysis und der Linearen Algebra zusammenzustellen. Die ausgewählten Musterlösungen sind ein authentisches Abbild des Lehrbetriebes, wie er an den meisten deutschen Hochschulen ablaufen könnte. Wir haben keine eigens für das Buch „durchgestylten" Aufgaben erfunden, sondern sie so übernommen, wie sie im Lehrbetrieb der ersten beiden Semester am Campus Essen tatsächlich vorgekommen sind.

Die vorgestellten Musterlösungen sind nicht als Vorbilder zum Abschreiben gedacht. Falls Sie das Buch aus diesem Grund gekauft haben, müssen wir Sie leider enttäuschen. Nur selten werden die Musterlösungen genau zu den Aufgaben passen, die Sie selbst zu bearbeiten haben. Dennoch wird es häufig möglich sein, Ideen aus unseren ausführlichen Musterlösungen auf Ihre Aufgaben zu übertragen. Und das ist auch durchaus legitim. Vor allem sollen Sie aber durch das Arbeiten mit diesem Buch dazu befähigt werden, wichtige Konzepte aus Ihren Vorlesungen besser zu verstehen und zu durchdringen. Das funktioniert meist gerade anhand von Aufgaben sehr gut. Viele Begriffe entfalten erst in Beispielen und Aufgaben ihre vollständige Bedeutung und Ausschärfung. Sie sollen die Vorlesungsinhalte direkt auf Aufgaben angewandt sehen, um Sie dann auch selbst in Ihren Übungs- und Klausuraufgaben anwenden zu können.

Insofern ist es von immenser Bedeutung, dass Sie die ausführlichen Musterlösungen auch wirklich durcharbeiten und nicht nur überfliegen. Die im Anschluss an die ausführlichen Musterlösungen gestellten Verständnisfragen sollen Sie dazu ermuntern, sich die Lösungen und einzelne Argumentationsschritte noch einmal selbst zu erklären. Nutzen Sie diese Chance unbedingt!

2 Lerntheoretische Grundlagen

Lernen aus Lösungen. Das Lernen aus Musterlösungen ist keine neue Idee. Sowohl die Kognitionspsycholgie wie auch die Bildungswissenschaften haben sich damit auseinandergesetzt. In diesem Abschnitt möchten wir gerne je einen wissenschaftlichen Ansatz aus diesen beiden Disziplinen vorstellen. Sie fungieren als theoretisches Fundament unserer Arbeit. Jeweils am Ende der Vorstellung finden Sie Literaturhinweise zum weiteren Vertiefen in die Thematik.

2.1 Cognitive Load Theory und Example-based Learning

Kognitive Überforderung. Beim Bearbeiten klassischer Übungsaufgaben muss der Aufgabenlöser eine Vielzahl unterschiedlicher kognitiver Aktivitäten gleichzeitig ausführen. Zum einen müssen die theoretischen Konzepte und Methoden aus den Vorlesungen verstanden werden. Des Weiteren müssen neue Begriffe erlernt, mit den bisherigen Begriffen in Verbindung gebracht und auf die konkreten Kontexte der zu bearbeitenden Aufgaben angewandt werden. Dazu kommen noch heuristische Fähigkeiten wie das Generieren guter Lösungsideen, das Auswählen geeigneter Lösungsstrategien und das flexible Hantieren mit den in der Aufgabe vorkommenden Objekten. Nicht zuletzt müssen auch zum Verstehen der zumeist in sehr hoher Informationsdichte formulierten Aufgabenstellungen sowie zum passenden Darstellen und Aufschreiben der eigenen Lösung anspruchsvolle Denkleistungen vollbracht werden.

Sie müssen sich also nicht wundern, wenn Sie sich beim Bearbeiten Ihrer Übungsaufgaben überfordert fühlen. Sie sind mit diesem Problem nicht alleine. Klarerweise sollen Sie in Ihrem Studium dazu befähigt werden, eigenständig komplexe Aufgaben zu lösen. Denn gerade die Fähigkeit zum schnellen Einarbeiten in vielschichtige, abstrakte Sachverhalte und zum erfolgreichen Finden entsprechender Lösungen zeichnet gut ausgebildete Mathematiker aus. Dieses Ziel kann man allerdings auf unterschiedlichen Wegen erreichen. Das Durcharbeiten von Musterlösungen und das damit implizit verbundene Lernen aus authentischen Vorbildern stellt eine aus didaktischer und lerntheoretischer Sicht effektive Möglichkeit dar.

Cognitive Load Theory. Die Cognitive Load Theory ist eine in den 1980er und 1990er Jahren entwickelte Theorie zur Beschreibung unterschiedlicher kognitiver Belastungen. Sie versucht, die Leistungsfähigkeit beim Erwerb kognitiver Fähigkeiten mit Kapazitätsbeschränkungen des menschlichen Arbeitsgedächtnisses in Verbindung zu bringen (siehe z. B. Renkl et al. 2003). Das Informationsverarbeitungssystem eines Menschen unterliegt natürlichen Beschränkungen. Es muss demnach mit den zur

C. Ableitinger, A. Herrmann, *Lernen aus Musterlösungen zur Analysis und Linearen Algebra*,
DOI 10.1007/978-3-658-02353-9_2, © Springer Fachmedien Wiesbaden 2013

Verfügung stehenden Kapazitäten sinnvoll umgegangen werden, um ein erfolgreiches und effektives Funktionieren des Systems zu ermöglichen.

Die Cognitive Load Theory unterscheidet drei unterschiedliche Formen kognitiver Belastungen (vgl. Sweller et al. 1998):

- **Intrinsic Load (innewohnende Belastung):** Damit ist die dem Lerngegenstand innewohnende Komplexität gemeint. In unserem Fall bezieht sich dies also auf die mathematischen Konzepte sowie auf die neu zu erlernenden und zu verstehenden Begriffe. Die daraus erwachsende kognitive Belastung kann nicht unmittelbar durch eine Veränderung der Lernumgebung, der Aufgabenformate, etc. verringert werden. Die Mathematik an sich bleibt eben ein anspruchsvolles Lernfeld.
- **Extraneous Load (äußere Belastung):** Wie der Name schon andeutet, handelt es sich hierbei um eine kognitive Belastung, die von außen durch die Gestaltung der Lernumgebung bedingt wird. Sie hat nicht unmittelbar mit dem Lernstoff zu tun, sondern vielmehr mit der Art und Weise, wie der Lernstoff präsentiert wird bzw. wie die zu bearbeitenden Aufgabenstellungen angeordnet, formuliert und dargestellt werden. Hier kann also didaktisch interveniert werden, indem man versucht, Extraneous Load weitestgehend zu verringern. Das ist insbesondere für Lernsituationen mit hoher intrinsischer Belastung von Relevanz.
- **Germane Load (lernbezogene Belastung):** Diese Art der kognitiven Belastung resultiert aus dem Lernprozess, also aus Phasen der Schemakonstruktion. Sie tritt immer dann auf, wenn beim Lernenden Verstehensprozesse ablaufen, wenn z. B. neue Begriffe mit bereits ausgebildeten internen Repräsentationen des Lerngegenstandes verwoben werden müssen. Es ist also von großer Bedeutung, die verfügbaren kognitiven Kapazitäten gerade für diese Art der Belastung zu reservieren. Nimmt dagegen die äußere Belastung zu viel Kapazität des Informationsverarbeitungssystems in Anspruch, so wird das Lernen in aller Regel ineffektiv bleiben. Durch das Arbeiten mit Musterlösungen können äußere Belastungen zugunsten der lernbezogenen Belastung reduziert werden, indem die besonders beanspruchenden Prozesse des Aufgabenlösens wegfallen.

Example-based Learning. Das Lernen aus vorgelösten Aufgaben ist Forschungsgegenstand der Kognitionspsychologie. Es konnte in zahlreichen Studien nachgewiesen werden, dass diese Methode gerade am Beginn eines Lernprozesses effektiver ist, als das sofortige eigenständige Problemlösen (siehe z. B. Renkl et al. 2003, Wittwer et al. 2010). So soll sich an die Phase des Vorstellens und Erklärens eines Prinzips oder einer Methode eine durchaus lange Phase des Lernens aus Lösungsbeispielen anschließen. Erst danach sollen sich Lernende selbstständig an der Lösung von Aufgaben versuchen. Erklärt wird dieser Effekt durch die oben beschriebene Cognitive Load Theory.

Das Bereitstellen von Musterlösungen an Lernende kann aber auch einen negativen Effekt haben. Sie sollten sich als Studentin bzw. Student niemals in Sicherheit wiegen, nur weil Sie ausgegebene Musterlösungen ordentlich abheften und bis zur Klausurvorbereitung sicher verwahren. Musterlösungen entfalten nur dann ihr volles Potential für den Lernprozess, wenn sie auch gründlich durchgearbeitet und verstanden werden. Unsere eigenen Lehr- und Lernerfahrungen zeigen, dass dazu ein großes

Maß an Selbstdisziplin erforderlich ist. Um den Einsatz von Fleiß und Ausdauer sowie um die Entwicklung einer guten Portion an Frustrationstoleranz kommen Sie also nicht herum, weder mit noch ohne Musterlösungen. Wir versuchen Ihnen bei der Überwindung dieser inneren Hürde zu helfen und Sie beim Durcharbeiten der Lösungen durch gezielte Anregungen zu unterstützen. Wir folgen dabei Empfehlungen aus der Kognitionspsychologie. Diese hat nämlich herausgefunden, dass vorgelöste Aufgaben nur dann effektiv sind, wenn beim Arbeiten mit ihnen Prozesse des Self-Explaining stattfinden. Diese wiederum können von außen durch Denkanstöße und Verständnisfragen eingeleitet werden (siehe z. B. Renkl et al. 2009).

Umsetzung im Buch. Im vorliegenden Buch versuchen wir der lernbezogenen Belastung dadurch Raum zu geben, dass wir die äußeren Belastungen nach Möglichkeit reduzieren. Durch das Arbeiten mit vorgelösten Aufgaben fallen die kognitiv sehr belastenden Prozesse des Aufgabenlösens komplett weg. Sie können sich also voll und ganz auf das Verstehen der Lösungen konzentrieren und Ihren Fokus auf das Verstehen und Durchdringen der mathematischen Sachverhalte richten. Auch die Aufgabenstellungen werden bei den meisten Musterlösungen nochmal anders formuliert. Es wird die Motivation geklärt, warum man sich für die Aufgabenstellung überhaupt interessiert und welchen Platz die Aufgabe im Gesamtkontext der Analysis bzw. der Linearen Algebra einnimmt.

Erneut sei betont, dass es nicht beim ausschließlichen Arbeiten mit Musterlösungen bleiben soll und darf. Selbstverständlich sollen Sie später auch Tätigkeiten wie das eigenständige Aufgabenlösen und das Einordnen der Aufgabe in einen größeren Zusammenhang lernen. Am Beginn des Lernprozesses ist es jedoch förderlich, sich zunächst auf das Verstehen der Inhalte zu konzentrieren.

Literatur zum Weiterlesen:

Renkl, A., Gruber, H., Weber, S., Lerche, T., & Schweizer, K. (2003). Cognitive Load beim Lernen aus Lösungsbeispielen. *Zeitschrift für Pädagogische Psychologie, 17*, 93-101.

Renkl, A., Hilbert, T., & Schworm, S. (2009). Example-based Learning in Heuristic Domains: A Cognitive Load Theory Account. *Educational Psychology Review, 21*, 67-78.

Sweller, J., van Merriënboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251-296.

Wittwer, J., & Renkl, A. (2010). How Effective are Instructional Explanations in Example-Based Learning? A Meta-Analytic Review. *Educational Psychology Review, 22*, 393-409.

2.2 Cognitive Apprenticeship

Zu wenige Vorbilder. Üblicherweise werden in Vorlesungen aus den Gebieten der Analysis und der Linearen Algebra im Anschluss an neue Definitionen und gerade bewiesene Sätze kurze, idealtypische Beispiele vorgestellt, an denen das eben Gelernte ausgeschärft und exemplifiziert wird. Aufgrund des Zeitmangels kann diesen Beispielen allerdings nicht der gebührende Zeitrahmen eingeräumt werden. Weitere

Beispiele und Aufgaben werden in den Übungsbetrieb ausgelagert. Dieser wiederum läuft häufig so ab, dass Studierende selbstständig entweder an sogenannten Präsenzaufgaben während der Übungszeit bzw. überhaupt sofort an Hausaufgaben arbeiten müssen. Nachdem die Aufgaben für die Studierenden meist neuartig und in vielen Fällen schwer zugänglich sind, fällt der Beginn meist recht schwer. In kleinen Lerngruppen, mit Hilfe des Internets und einschlägiger Lehrbücher wird unter Zeitdruck versucht, möglichst alle Aufgaben bis zum Abgabetermin zu lösen. Es mangelt vielerorts an professioneller Unterstützung und an authentischen Vorbildern im Aufgabenlösen. Warum benutzen wir das Wort *authentisch*? Zwar werden nach Abgabe die Hausaufgaben in den Übungsgruppen besprochen. Sehr oft werden dabei aber nur komprimierte Musterlösungen an die Tafel geschrieben und bestenfalls von Expertenseite kommentiert. Der Aufgabenlösungsprozess bleibt dabei meist versteckt.

Lernen im Handwerk. Eine Gruppe amerikanischer Bildungswissenschaftler hat in den späten 1980er Jahren die Idee geboren, das Lernen in handwerklichen Berufen auch auf das Erlernen kognitiver Fähigkeiten zu übertragen (siehe Collins et al. 1989). Grundprinzip ist dort das Meister-Schüler-Schema. Der Spezialist führt eine Tätigkeit vor, der Lehrling ahmt ihn nach. Erst danach wird die neu erlernte Aktivität in regelmäßiger Übung gefestigt und variiert. Collins, Brown und Newman haben an dieses Grundschema noch weitere Phasen angefügt und ihr Lernmodell *Cognitive Apprenticeship* genannt. Das lässt sich mit „Kognitives Handwerk“ oder „Kognitive Anlehre“ übersetzen. Das Cognitive Apprenticeship umfasst insgesamt fünf Phasen:

Modelling: Der Lernprozess wird durch ein modellhaftes Vorstellen einer neu zu erlernenden Fähigkeit eingeleitet. Der Experte führt z. B. vor, wie eine neue Rechenmethode oder ein neuer Algorithmus angewandt wird. Dabei ist wichtig, dass der Experte keine vorbereitete, von heuristischen Überlegungen und Gedankengängen „gesäuberte“ Vorstellung abliefert. Im Gegenteil kommt es hier gerade darauf an, dass er ein echtes Vorbild für einen Aufgabenlösungsprozess gibt. Er soll dabei – so gut es geht – implizites Expertenwissen explizieren. Die Lernenden sollen an seinen Gedankengängen teilhaben dürfen, sollen über alternative Lösungswege informiert werden, dürfen den Spezialisten auch einmal beim vorläufigen Scheitern sehen.

Coaching: An das Modelling schließt sich eine Phase des betreuten Nachahmens an. Die Lernenden sollen ähnlich strukturierte Probleme lösen und werden dabei vom Experten mit Hinweisen und motivierenden Fragen begleitet. Diese Unterstützung kann zu Beginn noch sehr intensiv sein, soll jedoch im Laufe des Übungsprozesses immer mehr zurückgenommen werden. Schließlich sollen die Lernenden dazu befähigt werden, mit dem neu Erlernten sicher und selbstständig umzugehen.

Articulation: Können die Lernenden mit ihrer neu erworbenen Fähigkeit verständig umgehen, so sollen sie das auch nach außen transportieren. Indem sie über ihr Problemlösen sprechen, führen sie selbst eine Art Modelling durch. Das Ausformulieren von Gedankengängen hilft dabei, Sachverhalte zu strukturieren und sie sich auch selbst nochmal klar zu machen.

Reflection: Ein interessantes Detail beim Cognitive Apprenticeship ist, dass die Phase der Reflexion nicht am Beginn des Lernprozesses steht, sondern beinahe an seinem Ende. Warum eine Methode funktioniert, wie sie funktioniert und in welchem Bereich sie funktioniert, kann man erst erkunden und begründen, nachdem man die Ausführung der Methode beherrscht. Das Reflektieren soll zusammen mit dem Experten erfolgen.

Exploration: Die neu erlernte Fähigkeit soll natürlich so flexibel und vielseitig wie möglich angewandt werden können. Das Übertragen auf neuartige Kontexte und Situationen sowie das Variieren der Problemstellung komplettieren somit den Lernprozess. Dabei kann natürlich neuerliche Unterstützung durch den Spezialisten nötig werden.

Umsetzung im Buch: Wir versuchen zumindest die erste, die dritte und die vierte Phase des Cognitive Apprenticeship durch das vorliegende Buch abzudecken. In den Kapiteln 5 bis 9 geben wir im Sinne des Modellings authentische Vorbilder für das Aufgabenlösen im Bereich mathematischer Grundlagen und in den Gebieten Analysis und Lineare Algebra. Wir haben dabei versucht, nicht nur Lösungen anzubieten, sondern auch deutlich zu machen, wie man diese Lösungen entwickelt. Danach müsste sich eine Phase des Coachings anschließen. Sie müssen also selbst Aufgaben lösen, am besten mit Unterstützung Ihrer Tutorinnen und Tutoren. Das kann leider nicht durch ein Buch geleistet werden. Zum Artikulieren regen wir Sie dann im Übungsteil an, indem wir Sie einladen, *ausführliche* Musterlösungen zu vorgegebenen *komprimierten* Musterlösungen zu verfassen. Zwar explizieren Sie dabei nicht Ihre eigenen Gedankengänge, aber immerhin müssen Sie überlegen, welche Gedankengänge zur vorliegenden Lösung gemacht werden mussten. Durch die Verständnisfragen im Anschluss an die einzelnen ausführlichen Musterlösungen im Teil II werden Sie laufend zum Reflektieren angeregt. Auch die Kennzeichnung der einzelnen Teilprozesse des Aufgabenlösens am Seitenrand neben den Musterlösungen (siehe dazu auch Kapitel 3) soll Sie dazu ermutigen, über den Zweck der einzelnen Passagen der Musterlösungen nachzudenken.

Literatur zum Weiterlesen:

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18*, S. 32-42.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive Apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Hrsg.), *Knowing, learning and instruction* (S. 453-494). Hillsdale: Lawrence Erlbaum Associates.

Kaiser, H. (2010). *Rechnen und Mathematik anwendungsbezogen unterrichten*. Zürich: Swissmem.

3 Teilprozesse beim Aufgabenlösen

Teilprozesse beim Aufgabenlösen. In diesem Kapitel stellen wir ein Werkzeug zur Aufgabenanalyse vor, das wir in unserer fachdidaktischen Forschung entwickelt haben (siehe Ableitinger 2012). Es stellt die beim mathematischen Aufgabenlösen ablaufenden Prozesse dar und erlaubt es demzufolge, die entscheidenden Phasen in Aufgabenlösungen aufzudecken. Dabei gewinnt man Einsicht in die Schwierigkeiten, mit denen Studierende üblicherweise beim Bearbeiten ihrer Aufgaben konfrontiert werden. Wir nutzen dieses Analysewerkzeug in den Kapiteln 5 bis 9, indem wir die ablaufenden Teilprozesse in den ausführlichen Musterlösungen am Seitenrand markieren.

Die in den folgenden Abschnitten beschriebenen Prozesse laufen dabei nicht unbedingt in der dargebotenen Reihenfolge ab. Es kann z. B. sein, dass ein Teilprozess auch mehrmals im gesamten Aufgabenlösungsprozess auftaucht, während ein anderer vielleicht gar nicht nötig ist. Sie werden das in den Musterlösungen in den Kapiteln 5 bis 9 sehen können.

Änderung des Problemzustandes. Ein Aufgabenlösungsprozess hat im Wesentlichen drei Zustände. Zunächst ist da der Ausgangszustand – die Aufgabenstellung. Ziel ist es, eine Lösung zur Aufgabe zu finden. Das ist der Endzustand. Und dazwischen muss der Ausgangszustand irgendwie in den Endzustand übergeführt werden. Das ist der Weg zwischen den beiden Zuständen, der während des Aufgabenlösens beschritten werden muss. Normalerweise gibt es einige unterschiedliche Möglichkeiten, wie man vom Ausgangs- zum Endzustand gelangen kann. Manchmal muss man auf dem Weg auch anhalten, innehalten, einen Blick auf die Karte werfen (also überlegen, wie man weiter vorgehen könnte) und sich klarmachen, wie weit man schon gekommen ist bzw. welches Stück Weg noch vor einem liegt. Es gibt also während des Aufgabenlösens Teilprozesse, in denen der Problemzustand nicht verändert wird (siehe die Abschnitte 3.1 und 3.7). Dann gibt es einen Teilprozess, durch den die Änderung des Problemzustandes schon angebahnt und manchmal sogar festgelegt wird (siehe Abschnitt 3.2). Und die wohl produktivsten Teilprozesse des Aufgabenlösens sind jene, in denen der Problemzustand tatsächlich verändert wird (siehe die Abschnitte 3.3 bis 3.6). Die letztgenannten Teilprozesse sind es übrigens, die man normalerweise auch in komprimierten Musterlösungen vorfindet.

Kennzeichnung der Teilprozesse in den Musterlösungen. In den Musterlösungen in den Kapiteln 5 bis 9 haben wir versucht, *allen sieben Teilprozessen* den ihnen gebührenden Platz zu geben. Die Teilprozesse sind in diesen Musterlösungen am Seitenrand gekennzeichnet. Sie werden feststellen, dass es in einigen Fällen nicht eindeutig ist, welchem Teilprozess eine Stelle in der Aufgabenlösung zuzuordnen ist. Manchmal

C. Ableitinger, A. Herrmann, *Lernen aus Musterlösungen zur Analysis und Linearen Algebra*,
DOI 10.1007/978-3-658-02353-9_3, © Springer Fachmedien Wiesbaden 2013

laufen Prozesse nämlich auch gleichzeitig ab bzw. lässt sich in einigen Fällen nur interpretieren, was sich der Aufgabenlöser an einer bestimmten Stelle wohl gedacht haben wird. Im Allgemeinen werden Sie jedoch die nachfolgend beschriebenen Teilprozesse problemlos in den ausführlichen Musterlösungen wiederfinden können.

Aufbau der folgenden Abschnitte. Eingeleitet werden die Abschnitte 3.1 bis 3.7 jeweils durch Fragen und Bemerkungen, die Studierende häufig äußern, während sie Aufgaben lösen bzw. nachdem sie sich im Nachhinein Musterlösungen angesehen haben. Wahrscheinlich erkennen Sie sich selbst in einigen dieser Fragen wieder. Es folgen eine Beschreibung des jeweiligen Teilprozesses sowie typische Beispiele, die diese Beschreibungen unterstützen sollen. Auch wenn in der Schulmathematik nicht alle genannten Teilprozesse gleichberechtigt im Unterricht repräsentiert sind, geben wir jeweils *ein* Beispiel aus dem Schulstoff. Gerade für Leserinnen und Leser, die noch am Beginn Ihres Studiums stehen, wollen wir somit die Teilprozesse an ihnen vertrauten Kontexten erklären. Leichter fällt es uns natürlich, Beispiele aus dem Hochschulstoff zu liefern. Zur erfolgreichen Lösung klassischer Aufgaben aus der Analysis und der Linearen Algebra braucht man nämlich in aller Regel eine Verzahnung aller sieben Teilprozesse. Das ist mitunter ein Grund, warum das Aufgabenlösen an der Hochschule meist anspruchsvoller ist als jenes an der Schule.

3.1 P: Problembewusstsein schaffen

Typische Studierendenäußerungen:

- „Ich sehe den Wald vor lauter Bäumen nicht!“
- „Was ist eigentlich das Problem bei dieser Aufgabe? Die ist doch trivial zu beantworten, oder?“
- „Was ist denn der springende Punkt?“

Beschreibung: Es gibt zwei unterschiedliche Situationen, in denen die Entwicklung von Problembewusstsein wichtig ist:

a) Bevor Sie eine Aufgabe sinnvoll bearbeiten können, müssen Sie zunächst einordnen können, zu welchem Themenfeld sie überhaupt gehört. Das ist im Übungsbetrieb üblicherweise nicht schwierig, denn die Übungsaufgaben passen meistens zum Vorlesungsstoff, der gerade behandelt wurde. Bei Klausuren müssen Sie diese Fähigkeit allerdings unter Beweis stellen. Sie müssen dann aber auch noch erkennen, worin das eigentliche Problem der Aufgabenstellung besteht und inwiefern die Aufgabe nicht trivial zu beantworten ist. Erst danach öffnen sich Handlungsspielräume und Sie können über mögliche Strategien zur Lösungsfindung nachdenken.

b) Zum Problembewusstsein gehört aber auch, dass Sie während des Aufgabenlösungsprozesses bzw. an seinem Ende den springenden Punkt der Aufgabe erkennen und benennen können oder rückblickend die Rolle der Aufgabe im Kontext des Themengebietes identifizieren können. Das ist wichtig für die Bearbeitung künftiger, ähnlicher Aufgabenstellungen. Nur wenn Sie sich später an die wesentlichen Ideen aus den Aufgabenlösungen erinnern können, können Sie dieses Erfahrungswissen für das effektive Bearbeiten neuer Aufgaben nutzbar machen.

In beiden Kategorien wird also der Problemzustand analysiert – einmal vorausschauend, einmal rückblickend – er wird jedoch durch diesen Teilprozess nicht verändert.

Beispiel aus der Schulmathematik:

Beispiel 3.1 Wollen Sie den größtmöglichen Definitionsbereich der Funktion f mit dem Funktionsterm $f(x) = \frac{\sqrt{x+3}}{x^2-5x+6}$ bestimmen, so sollten Sie sich Folgendes klar machen: Sie müssen gerade jene Stellen $x \in \mathbb{R}$ suchen, für die der Nenner Null wird bzw. für die der Ausdruck unter der Wurzel negativ wird. All diese Stellen dürfen Sie nicht in die Definitionsmenge aufnehmen. Es geht also darum zu erkennen, für welche Werte von x die vorkommenden Terme nicht definiert sind.

Beispiele aus der Hochschulmathematik:

Beispiel 3.2 Es sei festzustellen, in welchen Punkten die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x,y) = |x \cdot y|$ differenzierbar ist. Sie müssen dann erfassen, dass lediglich die Punkte (x,y) mit $x \cdot y = 0$ Schwierigkeiten bereiten.

Beispiel 3.3 Es sei der Grenzwert $\lim\limits_{n\to\infty}(1^3 + 2^3 + \cdots + n^3)^{\frac{1}{n}}$ zu berechnen. Hier müssen Sie erkennen, dass Sie nicht zuerst den Grenzwert von $\frac{1}{n}$ auswerten und danach $(1^3 + 2^3 + \cdots + n^3)^0 = 1$ schreiben dürfen.

Beispiel 3.4 Wenn Sie gerade die Funktion $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^3$ auf Stetigkeit bzw. gleichmäßige Stetigkeit untersucht haben, können Sie sie anschließend im Gedächtnis als ein Paradebeispiel einer Funktion abspeichern, die eben stetig, aber nicht gleichmäßig stetig ist.

Beispiel 3.5 Sie können nach der Untersuchung von $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^3$ für sich die Erkenntnis ziehen, dass eine Funktion an einer Stelle kein lokales Extremum besitzen muss, obwohl die erste Ableitung an dieser Stelle gleich Null ist.

3.2 K: Klärung der Handlungsoptionen

Typische Studierendenäußerungen:

- „Welche Lösungsstrategien stehen mir überhaupt zur Verfügung? Welche davon soll ich wählen?“
- „Wie kann ich schon im Vorhinein erahnen, welche Methode zum Ziel führen wird?“

Beschreibung: In diesem Teilprozess müssen Sie Handlungsoptionen sichten, die zur Aufgabenlösung beitragen könnten, ohne dass Sie bereits die Detailausführung in Angriff nehmen. Sie müssen abwägen, welche Art von Strategie, welche Theorie, welcher Satz, welche Definition, welcher Kalkül zum Einsatz kommen könnte. In der Folge müssen Sie die Entscheidung für eine der Optionen treffen bzw. fehlgeschlagene Handlungsoptionen während des Lösungsprozesses verwerfen und neue aufnehmen.

Dieses Klären von Handlungsoptionen begleitet den gesamten Aufgabenlösungsprozess, ohne dass Sie dabei schon den Problemzustand verändern. Die Veränderung des Problemzustandes wird allerdings schon angebahnt und durch Ihre Entscheidung zugunsten einer der Handlungsoptionen beeinflusst bzw. sogar festgelegt.

Beispiel aus der Schulmathematik:

Beispiel 3.6 Wenn Sie eine Stammfunktion zu einer Funktion bestimmen sollen, so müssen Sie dazu üblicherweise überlegen, welche Integrationsmethode Sie verwenden können. Während bei $\int x \sin(x)\ dx$ die partielle Integration funktioniert, müssen Sie bei $\int x \sin(x^2)\ dx$ substituieren. Das Abwägen unterschiedlicher Möglichkeiten und das Treffen einer Wahl gehören in diese Kategorie, das eigentliche Anwenden der gewählten Integrationsmethode allerdings noch nicht.

Beispiele aus der Hochschulmathematik:

Beispiel 3.7 Sollen Sie über die Konvergenz einer Reihe entscheiden, müssen Sie die zur Verfügung stehenden Konvergenzkriterien gegeneinander abwägen und sich für eines entscheiden. Liefert dann beispielsweise das Quotientenkriterium kein Ergebnis, müssen Sie es abwählen und ein anderes Kriterium verwenden.

Beispiel 3.8 Wollen Sie mit dem Eigenwertkriterium feststellen, dass für die Funktion $f : \mathbb{R}^n \to \mathbb{R}$ an einer kritischen Stelle x_0 ein lokales Extremum vorliegt, und besitzt die Hesse-Matrix an der Stelle x_0 lauter positive Eigenwerte und einen Eigenwert 0, so müssen Sie das Kriterium verwerfen und anders argumentieren.

Beispiel 3.9 Müssen Sie einen Beweis führen, so müssen Sie sich anfangs für ein Beweisformat entscheiden (direkt, indirekt, Induktion).

Beispiel 3.10 Sollen Sie entscheiden, ob ein Grenzwert der Form $\lim_{(x,y)\to(a,b)} f(x,y)$ existiert, müssen Sie sich vorher darüber klar werden, was es heißt, dass so ein Grenzwert existiert bzw. nicht existiert (das gehört noch zum Teilprozess *Problembewusstsein schaffen*). Schließlich müssen Sie sich aber für eine der beiden Vorgehensweisen (Existenz beweisen oder widerlegen) entscheiden und gegebenenfalls an einem Punkt der Aufgabenlösung auf die andere Handlungsoption umschwenken.

3.3 Z: Einen Zugriff herstellen, die Aufgabe handhabbar machen

Typische Studierendenäußerungen:

- „Wie soll ich denn bei der Aufgabe anfangen?“
- „Ich komme hier nicht mehr weiter!“
- „Mir ist das zu abstrakt – was bedeutet das denn konkret?“
- „Wie bekomme ich bei der Aufgabe einen Fuß in die Tür?“

Beschreibung: Bei vielen Aufgaben müssen Sie sich als Aufgabenlöser – um die gewählte Handlungsoption durchführen zu können – zuerst einen Ankerpunkt verschaffen, an dem Sie mit der Bearbeitung ansetzen können. Sie müssen gewissermaßen die äußere Schale der Aufgabe aufbrechen und die Aufgabenstellung so repräsentieren, dass sie für die zur Verfügung stehenden Methoden zugänglich wird. Das kann manchmal durch einen Sichtweisenwechsel erfolgen, manchmal aber auch durch einen Übergang zu geeigneten Objekten bzw. zu einer geeigneten Darstellung der *vorliegenden* mathematischen Objekte. Dieser Teilprozess hat *öffnenden* Charakter, insofern sich danach meist neue Handlungsoptionen bieten (die Sie vorher evtl. noch gar nicht gesehen haben).

Beispiel aus der Schulmathematik:

Beispiel 3.11 Angenommen, Sie sollen die Gleichung jener Geraden bestimmen, die durch die Punkte $(3, 4)$ und $(7, 2)$ geht. Sie können sich diese Aufgabe zugänglich machen, indem Sie $y = kx + d$ ansetzen. Danach brauchen Sie die Aufgabe nur noch abzuarbeiten, indem Sie die Koordinaten der beiden Punkte einsetzen und das entstehende Gleichungssystem lösen.

Beispiele aus der Hochschulmathematik:

Beispiel 3.12 Es sei eine Mengeninklusion zu zeigen. Es lohnt sich dann meist, ein beliebiges Element aus der einen Menge herauszugreifen und zu zeigen, dass es auch in der anderen Menge liegt. Das Herunterbrechen der Aufgabe auf die Ebene der Elemente macht die Problemstellung zugänglich. Das Hantieren mit dem „fassbaren" Element eröffnet Ihnen neue Handlungsmöglichkeiten.

Beispiel 3.13 Sollen Sie die Differenzierbarkeit einer Funktion $f : \mathbb{R} \to \mathbb{R}$ an einer Stelle nachweisen, so erlaubt Ihnen das Aufstellen des Differenzenquotienten einen Zugriff auf das Problem. Den Term können Sie dann manipulieren, seinen Grenzwert berechnen usw.

Beispiel 3.14 Es sei $\sqrt{2 + 3i}$ zu berechnen. Diese Aufgabenstellung können Sie öffnen, indem Sie den Ansatz $\sqrt{2 + 3i} = a + bi$ machen. Danach brauchen Sie nur noch zu quadrieren und einen Koeffizientenvergleich durchzuführen. Das Ansetzen einer unbekannten komplexen Zahl als $a + bi$ öffnet den Handlungsspielraum.

Beispiel 3.15 Wissen Sie schon, dass 2 eine obere Schranke einer Menge ist und wollen Sie beweisen, dass 2 sogar die kleinste obere Schranke ist, so können Sie sich Zugriff verschaffen, indem Sie für beliebiges $\varepsilon > 0$ die Zahl $2 - \varepsilon$ betrachten. Können Sie beweisen, dass $2 - \varepsilon$ für kein $\varepsilon > 0$ eine obere Schranke ist, dann ist 2 die kleinste obere Schranke.

Beispiel 3.16 Möchten Sie beweisen, dass $0.\bar{9}$ gleich 1 ist, so können Sie indirekt annehmen, dass $0.\bar{9} = 1 - \varepsilon$ für ein $\varepsilon > 0$ ist und diese Annahme zu einem Widerspruch führen. Das konkrete Benennen des vermeintlichen Abstandes zwischen $0.\bar{9}$ und 1 macht die Problemstellung abarbeitbar.

3.4 A: Anpassen oder Prüfen der Passung

Typische Studierendenäußerungen:

- „Wie kann ich herausfinden, ob ich einen bestimmten Satz anwenden darf?“
- „Die Aufgabe sieht ganz anders aus als der Satz, den ich laut Hinweis anwenden soll – wie passen denn die beiden zusammen?“

Beschreibung: Zielen Sie auf die Verwendung eines bestimmten Werkzeuges (z. B. Satz, Methode, Kalkül) ab, so müssen Sie die Problemstellung an dieses Werkzeug anpassen. Sie müssen dazu die vorliegende Situation eventuell modifizieren. In manchen Fällen müssen Notationen angeglichen werden oder man muss erkennen, dass das Werkzeug abgesehen von der Notation schon zur vorliegenden Situation passt. Auch das Überprüfen von Voraussetzungen eines Satzes und das gezielte Umstrukturieren eines Ausdrucks, um ihn in eine gewünschte Form zu bringen, gehören in diese Kategorie. Im Gegensatz zur Kategorie Z handelt es sich hierbei um *fokussierende* Tätigkeiten. An den Teilprozess A schließt sich klarerweise meist das Anwenden des vorbereiteten Werkzeugs an.

Beispiel aus der Schulmathematik:

Beispiel 3.17 Angenommen, Sie möchten die quadratische Gleichung $3x^2 + 4x - 9 = 0$ mit Hilfe der p-q-Formel lösen. Dann müssen Sie zunächst durch 3 dividieren, um die Gleichung an die Methode anzupassen.

Beispiele aus der Hochschulmathematik:

Beispiel 3.18 Haben Sie einen Funktionengrenzwert zu berechnen und kommt dabei die unbestimmte Form „$\infty - \infty$“ heraus, können Sie manchmal den Ausdruck so umformen, dass die Regeln von de l'Hospital Anwendung finden. Das wäre dann eine gezielte Umformung, um eine bestimmte Form zu erhalten bzw. um eine bestimmte Methode anwenden zu können (z. B. bei: $\lim\limits_{x\to 0}(\dfrac{1}{\sin(x)} - \dfrac{1}{x})$). Danach müssen Sie noch überprüfen, ob die neue Situation tatsächlich die Voraussetzungen für die Anwendbarkeit der Methode erfüllt.

Beispiel 3.19 Es sei mit Hilfe des Mittelwertsatzes der Differentialrechnung zu zeigen, dass $e^x > 1 + x$ für $x > 0$ gilt.
Lösung: Sei $f(x) := e^x$. Aus dem Mittelwertsatz folgt: $\frac{e^x - e^0}{x-0} = f'(c)$ für ein $c \in (0, x)$. Nachdem $f'(c) = e^c > 1$ gilt, folgt die Behauptung nach kurzer Umformung.
Es musste hier zunächst Kompatibilität zum Mittelwertsatz hergestellt werden, d. h., es musste der Differenzenquotient für e^x mit der zu beweisenden Ungleichung in Verbindung bzw. zur Passung gebracht werden.

3.5 H: Handwerk

Typische Studierendenäußerungen:

- „Das sieht doch eigentlich ganz einfach aus! Da wird ja nur gerechnet!"
- „Wie komme ich denn von dem einen Ausdruck zu dem anderen?"

Beschreibung: Mit Handwerk ist das Ausführen von Verfahren und Techniken gemeint, die zur Routine geworden sind. Selbstverständlich ist es subjektiv unterschiedlich, welche Arbeitsphasen eine Person als Handwerk bezeichnen würde. Eine Tätigkeit, die für einen Experten Routine ist, kann für einen Anfänger noch eine große kognitive Herausforderung darstellen. Legen wir das (normativ zu erwartende) Kompetenzniveau von Studienanfängern zugrunde, so können wir Tätigkeiten wie das Manipulieren von Termen, das Abarbeiten von Kalkülen oder das Anwenden von einfachen bekannten Resultaten als Handwerk bezeichnen.

Beispiel aus der Schulmathematik:

Beispiel 3.20 Bruch- und Potenzrechenregeln

Beispiele aus der Hochschulmathematik:

Beispiel 3.21 Multiplizieren zweier vorgegebener Matrizen

Beispiel 3.22 Berechnen partieller Ableitungen einer Funktion in mehreren Variablen

Beispiel 3.23 Ziehen einfacher Schlussfolgerungen

Beispiel 3.24 Anwenden eines Satzes (nachdem schon seine Anwendbarkeit überprüft wurde)

3.6 T: Tricks

Typische Studierendenäußerungen:

- „Woher weiß ich denn, dass ich das so machen soll?"
- „Ich verstehe das zwar im Nachhinein, selber wäre ich da aber nie drauf gekommen!"

Beschreibung: An manchen Stellen der Aufgabenbearbeitung hilft nur noch ein Trick, eine spezielle Repräsentation oder Umdeutung des Problems. Diese Tricks sind meist lokale, „vom Himmel fallende" Ideen, die sehr speziell auf eine bestimmte Situation passen und nur erfahrenen Aufgabenlösern geläufig sind. In manchen Fällen können Sie diese Tricks allerdings auch durch das Anwenden heuristischer Strategien finden. Die Funktion von Tricks ist meist das Vorantreiben des Lösungsprozesses, wenn dieser ins Stocken geraten ist, bzw. das elegante Abkürzen der Aufgabenlösung.

Letztlich öffnen auch Tricks Handlungsspielräume und machen die weitere Bearbeitung der Aufgabe zugänglich. Insofern könnte man sie der Kategorie Z zuordnen.

Wir haben uns trotzdem dazu entschieden, Tricks hier separat anzuführen (im Gegensatz zur Publikation Ableitinger 2012). Viele Studierende beruhigt es nämlich zu wissen, dass an manchen Stellen Tricks nötig sind, auf die man nicht ohne Weiteres selbst kommt. Dadurch relativiert sich gegebenenfalls das eigene Steckenbleiben auf dem Weg zur Lösung.

Beispiel aus der Schulmathematik:

Beispiel 3.25 Wollen Sie die Funktion $\ln(x)$ integrieren, so schreiben Sie am besten $1 \cdot \ln(x)$ und integrieren partiell. Das Hinzufügen des Faktors 1 ist ein Trick, der ganz lokal in dieser Situation hilft.

Beispiele aus der Hochschulmathematik:

Beispiel 3.26 Ist der Grenzwert

$$\lim_{(x,y)\to(0,0)} x^2 \cdot y^2 \cdot \ln(x^2 + y^2)$$

zu berechnen, lohnt es sich, in Polarkoordinaten zu transformieren. Danach braucht man nur noch r gegen Null laufen zu lassen (und nicht mehr zwei Variablen gleichzeitig).

Beispiel 3.27 Sollen Sie

$$\lim_{x\to 0} \frac{e^x - 1}{x}$$

berechnen und haben Sie dazu noch nicht die Regeln von de l'Hospital zur Verfügung, können Sie den Ausdruck als Differentialquotienten auffassen:

$$\lim_{x\to 0} \frac{e^x - 1}{x} = \lim_{x\to 0} \frac{e^x - e^0}{x - 0} = f'(0),$$

wobei $f(x) := e^x$.

3.7 B: Begleitende, strukturierende Kommentare und Erläuterungen

Typische Studierendenäußerungen:

- „Wie detailliert muss ich die Lösung aufschreiben?“
- „Wie kann ich meine Lösung besser strukturieren? In welcher Reihenfolge soll ich die einzelnen Teile aufschreiben?“
- „Wozu dient denn dieser Argumentationsschritt in der Musterlösung?“

Beschreibung: Zwischen einzelnen Teilschritten der Aufgabenbearbeitung wird Resümee gezogen. Es wird über schon Erreichtes reflektiert und festgehalten, was noch zu zeigen ist. Durch begleitende Kommentare kann auch nochmal der entscheidende Punkt in der Aufgabenbearbeitung akzentuiert werden. Schließlich fällt auch das Beurteilen des Erfolgs von zuvor durchgeführten Handlungen in diese Kategorie.

Beispiele aus der Schul- und Hochschulmathematik:

Beispiel 3.28 „Offenbar hat diese Methode zu einem Punkt geführt, an dem wir nicht mehr weiterkommen. Die Strategie hat sich also nicht bewährt."

Beispiel 3.29 „Wir haben bis jetzt gezeigt, dass die Funktion in allen Punkten $x \neq 0$ differenzierbar ist. Bleibt noch die Stelle $x = 0$ zu untersuchen."

Beispiel 3.30 „Man beachte an dieser Stelle, dass sich der Ausdruck aus einem linearen und einem quadratischen Term zusammensetzt."

Beispiel 3.31 „Zu zeigen: ..."

3.8 Die Teilprozesse in einer vollständigen Musterlösung

Abgrenzungen. In kompletten Aufgabenlösungen laufen die sieben vorgestellten Teilprozesse in unterschiedlichen Reihenfolgen ab. Manchmal sind an einem Bearbeitungsschritt auch zwei oder sogar drei Teilprozesse gleichzeitig beteiligt. Außerdem sind ab und zu die Grenzen gar nicht klar, an denen ein Teilprozess endet und ein neuer beginnt. Und doch lassen sich typische Aufgaben aus dem ersten Studienjahr vollständig durch die sieben Teilprozesse beschreiben. Sie können das im zweiten Teil des Buches sehen. Wir haben den dort präsentierten Musterlösungen die jeweils ablaufenden Prozesse zugeordnet. Sie finden daher eine Vielzahl an Musterlösungen, bei denen die sieben Teilprozesse am Seitenrand markiert sind.

Zuordnen der Teilprozesse. Bei einigen Musterlösungen bekommen Sie dann selbst die Gelegenheit, die Teilprozesse den einzelnen Bearbeitungsschritten zuzuordnen. Sie werden sehen, dass das gar nicht so einfach ist. Dazu müssen Sie nämlich die Aufgabenlösungen gut verstanden haben und den Zweck der einzelnen Bearbeitungsschritte erkennen können. Manchmal hängt die Entscheidung für oder gegen einen Teilprozess auch davon ab, auf welchem Lernstand man sich selbst gerade befindet. Was für einen Experten bloßes *Handwerk* ist, kann Ihnen selbst vielleicht noch sehr schwer fallen. Sie würden dann einen entsprechenden Bearbeitungsschritt vielleicht als *Trick* oder als *Zugriff herstellen* interpretieren.

Exemplarisch möchten wir Ihnen das Zuordnen der Teilprozesse an jener Aufgabenlösung vorstellen, die wir schon in Kapitel 1 präsentiert haben. Die drei unterschiedlichen Farbnuancen der Kästchen (hell-, mittel- und dunkelblau) werden im Abschnitt 4.1 erklärt. Sie sind an dieser Stelle noch nicht wichtig.

Teilprozesse in Aufgabe 1.1:

B Wir sollen also in dieser Aufgabe zeigen, dass f unter den gegebenen Voraussetzungen einen Fixpunkt besitzt.

P Demnach soll lediglich die Existenz gezeigt werden, also weder Eindeutigkeit beweisen noch eine konkrete Stelle angegeben werden

K Ein Satz aus der Vorlesung, der hier helfen könnte, ist der Zwischenwertsatz.

B Dieser macht eine Existenzaussage, allerdings über eine Nullstelle und nicht über einen Fixpunkt.

K Wir sollten also versuchen, die Fixpunktaussage in eine Nullstellenaussage zu transformieren:

A
$$f(c) = c \quad \Leftrightarrow \quad f(c) - c = 0$$

Z Wir definieren jetzt eine neue Funktion g durch $g(x) := f(x) - x$.

K Wenn wir zeigen können, dass diese Funktion g in $[0,1]$ eine Nullstelle besitzt, sind wir fertig. Das versuchen wir durch Anwendung des Zwischenwertsatzes.

B Überprüfen der Voraussetzungen:

A g ist stetig, da f stetig ist und die Funktion h mit $h(x) := x$ stetig ist. $g(0) = f(0) - 0 \geq 0$, $g(1) = f(1) - 1 \leq 0$, da f laut Voraussetzung nur Werte in $[0,1]$ annimmt.

B Die Voraussetzungen für den Zwischenwertsatz sind so nicht vollständig erfüllt, wir brauchen eine Fallunterscheidung:

H 1. Fall: $g(0) = 0$. Dann gilt aber $f(0) - 0 = 0$ und somit $f(0) = 0$.

B Demnach ist also $c = 0$ ein Fixpunkt von f.

H 2. Fall: $g(1) = 0$. Dann gilt aber $f(1) - 1 = 0$ und somit $f(1) = 1$.

B Demnach ist also $c = 1$ ein Fixpunkt von f.

H 3. Fall: $g(0) > 0$ und $g(1) < 0$. Dann folgt aus dem Zwischenwertsatz, dass es eine Stelle $c \in (0,1)$ mit $g(c) = 0$ gibt. Damit gilt aber $f(c) - c = 0$ und somit $f(c) = c$.

B Folglich ist diese Stelle c ein Fixpunkt von f.

3.9 Quer zu den Teilprozessen liegende Kompetenzen und Dispositionen

Im vorigen Abschnitt haben Sie gesehen, dass man typische Musterlösungen vollständig durch die sieben Teilprozesse beschreiben kann. Wenn Sie jetzt denken, dass Sie dann auch nur diese Teilprozesse zu üben brauchen, um eine professionelle Mathematikerin bzw. ein professioneller Mathematiker zu werden, dann liegen Sie aber leider falsch. Quer zu diesen Prozessen gibt es noch viele weitere Denkleistungen und Haltungen, die Sie zum erfolgreichen Aufgabenlösen benötigen. Ohne sie können Sie auch die sieben beschriebenen Teilprozesse nicht angemessen durchführen. Die folgenden intellektuellen, sozialen und affektiven Kompetenzen befähigen Sie dazu, in gewissen Situationen zu erkennen, welchen Schritt Sie in einem Aufgabenlösungsprozess als Nächstes setzen müssen.

Structure Sense. Sie haben schon in der Schule damit begonnen, Termstrukturen „richtig" lesen zu lernen. Es ist eine wichtige Fähigkeit, einem Term schon seinen nächsten Bearbeitungsschritt ansehen zu können. Haben Sie etwa den Ausdruck

$$\frac{1}{\sqrt{x}-\sqrt{x_0}}$$

vor sich, so wird es in den meisten Fällen (natürlich kommt das im Einzelfall auf den Kontext an) sinnvoll sein, den Ausdruck mit $\sqrt{x}+\sqrt{x_0}$ zu erweitern. Haben Sie also das Gespür für die Struktur des Terms, so führt Sie das automatisch zum Ausführen von Handwerk. Structure Sense kann aber auch andere Teilprozesse in Gang bringen. Beispielsweise können Sie manchmal schon der Gestalt einer auf Konvergenz zu untersuchenden Reihe ansehen, welches Konvergenzkriterium helfen könnte. Hier hilft Ihnen der Structure Sense also bei der Klärung von Handlungsoptionen.

Symbol Sense. Eine der größten Schwierigkeiten für Studienanfänger ist der flexible Umgang mit Variablen. Sie müssen z. B. bei gegebener Abbildung $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x,y) = 2x + xy$ erfassen können, dass beim Ausdruck $f(a+h)$ die Symbole a und h jeweils für Vektoren mit zwei Komponenten stehen müssen. Erst diese Erkenntnis ermöglicht es Ihnen, sich einen Zugriff auf die Situation zu verschaffen. Sie können dann nämlich schreiben: $a = (a_1, a_2)$ und $h = (h_1, h_2)$. Danach lässt sich der Funktionswert für den Vektor $a+h$ einfach berechnen. Es gibt im ersten Studienjahr eine Vielzahl von Situationen und Aufgaben, wo Symbol Sense eine wichtige Rolle spielt. Sie können diese Fähigkeit trainieren, indem Sie sich jedes Mal klar machen, mit welcher Art von Objekt Sie es bei einem vorliegenden Symbol zu tun haben.

Anwenden von Methoden. Sie lernen in Ihrem Studium eine Vielzahl von Methoden und Kalkülen kennen. Doch eine Methode zu kennen reicht im Allgemeinen nicht aus, um sie auch korrekt und zielführend anwenden zu können. Selten bekommen Sie eine explizite Gebrauchsanweisung mitgeliefert. Es muss Ihnen vielmehr beigebracht werden, wie man die Methode auf einen Kontext bezieht. Indem Sie andere Leute beim Anwenden einer Regel beobachten (oder es in einem Buch lesen), entwickelt sich implizit die Fähigkeit mit, die Methode eigenständig auf eine Aufgabe anzuwenden.

Befolgen impliziter Konventionen. Es gibt in einem Mathematikstudium viele Dinge, die *nebenbei* mitgelernt werden, ohne jemals explizit gemacht zu werden. Das Befolgen impliziter Normen und Gepflogenheiten spielt z. B. eine wichtige Rolle bei der Entscheidung, wie detailliert man eine Aufgabenlösung aufschreibt oder mit welcher Argumentationstiefe man einen Beweis führt. Auch gehört das Wählen einer geeigneten Notation (f, g für Abbildungen, u, v für Vektoren, Großbuchstaben für Mengen, etc.) in diese Kategorie.

Mut zum intellektuellen Risiko. Manchmal muss man sich überwinden, eine Aufgabe überhaupt anzupacken. Die äußere Gestalt, die vorkommenden Objekte oder das ungewohnte Aufgabenformat können einen demotivieren, mit dem Lösungsprozess überhaupt einmal zu beginnen. Vielleicht würden Sie sich selbst auch durch folgende

Aufgabenstellung abgeschreckt fühlen: Untersuchen Sie die Funktion $f : (-1, 1) \to \mathbb{R}$ mit

$$f(x) = \begin{cases} \arctan(x^2) & -1 < x < 0 \\ \sin(x^5)\sqrt{1 - x^2} & 0 \leq x < 1 \end{cases}$$

auf Differenzierbarkeit!

Aufgabenlösen im Mathematikstudium erfordert Mut, sich auf neuartige Probleme einzulassen! Oft stellt sich dann erst in der Bearbeitung heraus, dass die Aufgabe gar nicht so schwierig ist, wie sie aussieht.

Flexibles Uminterpretieren. In der Aufgabenlösung aus Abschnitt 3.8 muss an einer Stelle die Fixpunktaussage in eine Nullstellenaussage übergeführt werden. Dazu müssen Sie flexibel und beweglich denken können. Diese Fähigkeit benötigen Sie in sehr vielen Aufgabenlösungen. Manchmal müssen Sie in einer bestimmten Situation eine *Reihe* als *Folge* von Partialsummen interpretieren, manchmal eine *Folge* als *Abbildung* mit speziellem Definitionsbereich oder eine *Abbildung* als *Vektor* auffassen. Diese Kompetenz entwickelt sich im Laufe der Zeit von alleine mit, sofern Sie selbst regelmäßig Aufgaben lösen bzw. Musterlösungen durcharbeiten.

Literatur zum Weiterlesen:

Ableitinger, Ch. (2012): Typische Teilprozesse beim Lösen hochschulmathematischer Aufgaben: Kategorienbildung und Ankerbeispiele. *Journal für Mathematik-Didaktik, 33*(1), 87-111.

Arcavi, A. (1994). Symbol Sense: Informal Sense-making in Formal Mathematics. *For the Learning of Mathematics, 14*(3), 24-35.

Hoch, M., & Dreyfus, T. (2005). Students' difficulties with applying a familiar formula in an unfamiliar context. In H. L. Chick, & J. L. Vincent (Hrsg.), *Proceedings of the 29th Conference of the International Group for Psychology of Mathematics Education* (Vol. 3, S. 145-152). Melbourne: PME.

Wittmann, E. (1975). *Grundfragen des Mathematikunterrichts*. 3. Aufl. Braunschweig: Vieweg.

Teil II

Ausführliche Musterlösungen

4 Allgemeines zu Musterlösungen

In den Kapiteln 5 bis 9 finden Sie eine Sammlung an ausführlichen Musterlösungen zu Aufgaben aus den Themengebieten mathematische Grundlagen, Analysis 1 und 2, Lineare Algebra 1 und 2. Bevor Sie sich mit diesen ausführlichen Musterlösungen beschäftigen, sollten Sie sich die Zeit nehmen, ein paar Bemerkungen von uns diesbezüglich zu lesen.

4.1 Bemerkungen zu den ausführlichen Musterlösungen

Anspruch an die Musterlösungen. Im Vergleich zu den üblicherweise im Lehrbetrieb oder in Lehrbüchern eingesetzten *komprimierten Musterlösungen* möchten die *ausführlichen Musterlösungen* in diesem Buch zusätzliche Informationen bereitstellen, die für den Lernprozess wichtig sein können. Wir möchten nicht bloß die Lösung der Aufgabe darstellen, sondern auch ein paar Worte darüber verlieren, was eigentlich das Interessante an der Aufgabe ist. Wir möchten klären, inwiefern es gewinnbringend sein kann, sich mit der Aufgabe zu beschäftigen. Außerdem wird auch auf das Entwickeln und Ausformulieren von Lösungsideen Wert gelegt. Wir versuchen zu motivieren, wie man auf eine Lösungsstrategie überhaupt kommt. Manchmal passiert es beim Aufgabenlösen, dass man an irgendeiner Stelle feststeckt und es auf eine andere Art und Weise versuchen muss. Auch solche Situationen haben wir in unseren Musterlösungen abgebildet, sofern der Ablauf dadurch nicht verwirrend wird und zu sehr vom funktionierenden Lösungsweg wegführt. Jedenfalls werden aber die entscheidenden Punkte in der Lösung hervorgehoben und detailliert beschrieben. Nach Möglichkeit sind alle Argumentationsschritte aufgeführt, sodass Sie dem Lösungsstrang in aller Regel gut folgen können sollten.

Nur eine Lösung pro Aufgabe. Klarerweise sind Lösungswege selten eindeutig. Es gäbe zu den meisten Aufgaben, die wir in diesem Teil des Buches betrachten, viele unterschiedliche Lösungsmöglichkeiten. Und es wäre bestimmt auch interessant, verschiedene Wege miteinander zu vergleichen oder aus der Vielfalt an Möglichkeiten jene zu suchen, die einem am besten liegt. Dennoch haben wir uns dazu entschieden, jeweils nur eine Lösung pro Aufgabe aufzuschreiben und dafür im Gegenzug die Anzahl der Aufgaben zu erhöhen. Außerdem bergen die ohnehin schon sehr ausführlichen Musterlösungen die Gefahr, nicht in allen Details gelesen und intensiv

C. Ableitinger, A. Herrmann, *Lernen aus Musterlösungen zur Analysis und Linearen Algebra*,
DOI 10.1007/978-3-658-02353-9_4, © Springer Fachmedien Wiesbaden 2013

durchgearbeitet zu werden. Würden wir also zu jeder Aufgabe viele unterschiedliche Lösungswege darstellen, würde das diese Überforderung nur noch verstärken. In einigen Musterlösungen werden aber zumindest andere Lösungswege angedeutet oder sogar in Ansätzen durchgeführt.

Behandelte Themen. In unserem Buch greifen wir sehr viele zentrale Begriffe und Konzepte aus den mathematischen Grundlagen, der Analysis und der Linearen Algebra auf. Die hier präsentierte Aufgabenpalette deckt allerdings keineswegs den gesamten möglichen Stoff entsprechender Vorlesungen lückenlos ab. Das Buch kann also keine Vorlesung und kein Lehrbuch ersetzen. Umgekehrt finden Sie hier aber womöglich auch Aufgaben zu Themen, die in Ihren eigenen Vorlesungen gar nicht behandelt wurden. Gerade in der Linearen Algebra 2 gibt es nämlich keinen strikt festgelegten Stoffkanon, an den sich die Dozentinnen und Dozenten halten. Lassen Sie sich dadurch also nicht irritieren!

Wichtige Sätze und Definitionen. Zur Lösung mancher Aufgaben muss man einen Satz oder eine Definition verwenden. Aus Platzgründen können wir aber nicht jeden Satz und jede Definition aufschreiben. Wir überlassen es Ihnen, geläufige Sätze, Methoden, usw. in Ihrer Vorlesungsmitschrift, einem einschlägigen Lehrbuch oder im Internet nachzulesen. Sätze, bei denen es auf eine bestimmte Notation ankommt oder zu denen es in der Literatur unterschiedliche Versionen gibt, sind jedoch an Ort und Stelle formuliert. Damit wollen wir Missverständnissen vorbeugen.

Markierung der Teilprozesse. Am Seitenrand der meisten Musterlösungen sind neben den einzelnen Bearbeitungsschritten die gerade ablaufenden Teilprozesse markiert (siehe dazu Kapitel 3). Zeichenerklärung:

P *Problembewusstsein schaffen*

K *Klärung der Handlungsoptionen*

Z *Einen Zugriff herstellen*

A *Anpassen oder Prüfen der Passung*

H *Handwerk*

T *Trick*

B *Begleitende, strukturierende Kommentare und Erläuterungen*

Die einzelnen Teilprozesse werden nach Farbnuancen unterschieden. Die hellblauen Teilprozesse verändern den Problemzustand nicht, der mittelblaue Teilprozess bereitet die Veränderung schon konkret vor und die dunkelblauen Teilprozesse verändern tatsächlich den Zustand des Problems.

Nichteindeutigkeit der Teilprozesse. Sie werden sich vielleicht an manchen Stellen fragen, warum gerade dieser oder jener Teilprozess neben einem Bearbeitungsschritt markiert ist. Vielleicht hätten Sie selbst einen anderen Teilprozess zugeordnet. Das ist gar kein Problem. Gerade wenn man an manchen Stellen über die Auswahl eines Teilprozesses streiten kann und Interpretationsspielräme offen sind, werden dadurch metakognitive Prozesse angeregt. Es ist jedenfalls gewinnbringend, von einer Außensicht auf die Aufgabenlösungen zu blicken, über die Funktion der einzelnen Bearbeitungsschritte nachzudenken und so schließlich auch über das eigene Denken zu reflektieren. Haben Sie die unterschiedlichen Teilprozesse irgendwann verinnerlicht, so können sie Ihnen auch bei selbstständigen Problemlöseprozessen als Steuerungswerkzeug bzw. als Lösungsplan dienen. Manchmal werden Sie vermutlich auch denken, dass die angegebenen Teilprozesse zu grob eingeteilt sind. Einige der Bearbeitungsphasen könnte man sicher noch in kürzere Teilpassagen untergliedern. Darauf haben wir aus Gründen eines übersichtlicheren Layouts verzichtet.

Verständnisfragen. Im Anschluss an die jeweilige Musterlösung folgen Verständnisfragen zu ganz speziellen Aspekten der Lösung. Diese sollen Sie dazu anregen, sich über die Aufgabe Gedanken zu machen bzw. die Aufgabenlösung genau zu studieren. Die Fragen sind dabei nicht nach ihrem Schwierigkeitsgrad sortiert, sondern verlaufen meist chronologisch entlang der Musterlösung. Sie werden sehen – manche Fragen sind sehr einfach, andere werden Sie sicher herausfordern! Im Lösungsteil des Buches (Kapitel 11) finden Sie mögliche Antworten zu den Verständnisfragen. Manche Fragen sind jedoch so offen formuliert, dass sich Ihre eigenen Antworten teilweise von den vorgegebenen Antwortvorschlägen unterscheiden werden. Dennoch sollten Sie diese Fragen unbedingt bearbeiten, denn auch bei mündlichen Prüfungen werden häufig Fragen dieser Art gestellt. Überdies sollten Sie durch die Verständnisfragen ermuntert werden, sich auch selbst zusätzliche Fragen zu stellen bzw. sich diese Fragehaltung auch für das Arbeiten mit anderen mathematischen Texten (Lehrbuch, Vorlesungsmitschrift, etc.) anzugewöhnen. Die dabei eingeleiteten Prozesse des *Sich-Selbst-Erklärens* eines mathematischen Sachverhaltes sind für das Verstehen von entscheidender Bedeutung.

Übungen. Bei einigen der Musterlösungen sind seitlich *keine* Teilprozesse markiert. Dies sollen Sie zum Anlass nehmen, selbst darüber nachzudenken, welche Teilprozesse aus Kapitel 3 zu den einzelnen Bearbeitungsschritten passen könnten. Wir fordern Sie dazu an entsprechender Stelle konkret durch eine Übung auf, die Sie im Anschluss an die Verständnisfragen finden. Entsprechende Lösungsvorschläge sind im vierten Teil des Buches (Kapitel 11) abgedruckt. Wundern Sie sich jedoch nicht, wenn nicht alle Ihre Zuordnungen mit jenen im Lösungsvorschlag übereinstimmen. Wie oben schon erwähnt gibt es dabei manchmal Interpretationsspielräume. Wichtig erscheint uns vielmehr, dass Sie selbst über die Bedeutung einzelner Bearbeitungsschritte nachdenken und so ein tieferes Verständnis erlangen.

Arbeitsaufträge. In den Arbeitsaufträgen bekommen Sie die Gelegenheit, komprimierte Musterlösungen zu den Aufgaben aus den Kapiteln 5 bis 9 zu verfassen. Wichtig ist, dass Sie die ausführlichen Musterlösungen zunächst gründlich durcharbeiten. Dadurch bekommen Sie einen Einblick, welche Teile der Musterlösung wesentlich sind und welche Passagen Sie einsparen können. Im vierten Teil des Buches (Kapitel 11) finden Sie zu jeder dieser ausführlichen Lösungen einen Vorschlag für eine entsprechende komprimierte Musterlösung. Selbstverständlich muss Ihre komprimierte Musterlösung nicht genauso aussehen wie unser Lösungsvorschlag. Er soll Ihnen nur *eine* Möglichkeit präsentieren und somit ein Anhaltspunkt für Ihre eigene Musterlösung sein, falls Sie unsicher sind. Wie genau wir uns das Verfassen komprimierter Musterlösungen vorstellen und was Sie dabei beachten sollten, stellen wir im nächsten Abschnitt genauer vor.

Empfehlung. Konzentrieren Sie sich beim Arbeiten mit den Musterlösungen zunächst auf den Lösungstext selbst. Erst wenn Sie glauben, alles verstanden zu haben, kümmern Sie sich um die Beantwortung der Verständnisfragen. Schließlich sollten Sie Ihren Fokus auf die markierten Teilprozesse richten. Versuchen Sie nachzuvollziehen, inwiefern die einzelnen Teilprozesse zu den jeweiligen Bearbeitungsschritten passen. Dadurch bekommen Sie einen Einblick in die Struktur der Musterlösung und erkennen, welche Teile der Lösung besonders wichtig sind und welche Funktionen die einzelnen Passagen haben. Das wird Ihnen dann auch beim Verfassen komprimierter Musterlösungen helfen.

4.2 Hinweise zum Verfassen komprimierter Musterlösungen

Ausführliche Musterlösungen als Ausgangspunkt. Die in den Kapiteln 5 bis 9 folgenden ausführlichen Musterlösungen sollen Ihre Grundlage für das Verfassen komprimierter Musterlösungen sein. Das prägnante Aufschreiben mathematischer Texte ist eine wichtige Fähigkeit, die Sie in Ihrem Studium unbedingt erlernen sollten. Um etwas griffig und treffend aufschreiben zu können, muss man es aber zunächst einmal gedanklich durchdrungen haben. Erst dann können Sie nämlich entscheiden, welche Aspekte Sie in den Text aufnehmen und welche Sie getrost weglassen können. Insofern halten wir es für einen guten Weg, wenn Sie zunächst mit einer ausführlichen Musterlösung arbeiten und sie zu verstehen versuchen. Danach sollen Sie abwägen, welche Informationen für die Darstellung der Lösung unverzichtbar sind. Diese Informationen sollen Sie schließlich zu einer komprimierten Musterlösung vereinen.

Prägnanz und Eleganz. Ästhetik ist eine subjektive Frage. Und doch gibt es in der mathematischen Community häufig Konsens darüber, wann ein mathematischer Text elegant ist und wann nicht. Diese Einschätzungsfähigkeit wird in einem Mathematikstudium implizit mitgelernt. Sie kommt beim Verfassen komprimierter Lösungen jedenfalls zum Tragen. Es reicht also nicht, einfach aus einer ausführlichen Musterlösung scheinbar unwichtige Passagen herauszustreichen. Es geht auch darum, Bearbeitungsteile eventuell anders anzureihen, mehrere Fälle einer Fallunterscheidung gleichzeitig abzuarbeiten bzw. die Entstehungsgeschichte einer Lösung ihrer übersichtlichen und prägnanten Darstellung zu opfern.

Vollständigkeit. Achten Sie beim Verfassen komprimierter Musterlösungen darauf, dass sie vollständig bleiben, dass also keine essentiellen Informationen verlorengehen! Gerade bei Klausuren ist es wichtig, dass nachvollziehbar bleibt, wie Sie von der Aufgabenstellung Schritt für Schritt zur Lösung gelangen. Selbstverständlich hat dieses Auswählen relevanter Informationen und die Entscheidung für ein bestimmtes Argumentationsniveau eine subjektive Komponente. Was für den einen ausführlich genug ist, ist für den anderen lückenhaft und unvollständig. Am besten orientieren Sie sich dabei an Ihren Dozenten, um bei Klausuren keine bösen Überraschungen zu erleben.

Teilprozesse als Hilfe. Am Seitenrand neben den meisten Musterlösungen der Kapitel 5 bis 9 finden Sie die zugehörigen Teilprozesse aus Kapitel 3. Auch diese Markierungen können Ihnen eventuell dabei helfen zu entscheiden, welche Bearbeitungsschritte unbedingt auch in die komprimierte Musterlösung gehören. Das werden in den meisten Fällen nämlich gerade die dunkelblau markierten Passagen sein. Während dieser Teilprozesse wird ja der Problemzustand tatsächlich verändert, während in den anderen Teilprozessen entweder eine Meta-Sicht auf die Aufgabe eingenommen wird (hellblau) oder eine Veränderung des Problemzustandes erst angebahnt wird (mittelblau).

Ein eigener Anspruch. Sie dürfen das Schreiben komprimierter Musterlösungen nicht unterschätzen! Auch wenn das Verfassen ausführlicher Musterlösungen die vermutlich anspruchsvollere Tätigkeit ist, so gibt es dennoch viele wichtige Aspekte zu beachten. Wenn Sie das nicht glauben, so lesen Sie nochmal die eben genannten Punkte durch!

5 Musterlösungen zu mathematischen Grundlagen

5.1 Summen- und Produktzeichen

Aufgabe 5.1 a) Berechnen Sie die Summe $\sum_{k=2}^{5}\left((-1)^k \cdot \frac{k-3}{k+7}\right)$.

b) Berechnen Sie das Produkt $\prod_{j=-3}^{5} m^j$ (wobei $m \neq 0$ gilt).

c) Schreiben Sie das Produkt $\frac{2}{4} \cdot \frac{5}{7} \cdot \frac{10}{10} \cdot \frac{17}{13} \cdot \frac{26}{16}$ mit Hilfe des Produktzeichens.

d) Führen Sie eine Indexverschiebung von $\sum_{k=1}^{n} \frac{(k-2)^2}{3^{k+1}}$ zu $\sum_{k=0}^{\dots} \dots$ durch.

e) Führen Sie eine Indexverschiebung von $\prod_{l=3}^{6} \frac{l+3}{l^2}$ zu $\prod_{l=\dots}^{\dots} \frac{l+8}{\dots}$ durch.

Ausführliche Lösung 5.1 a) In diesem Aufgabenteil soll geübt werden, wie die Summennotation gelesen werden muss. B

Der Summationsindex k läuft in unserem Fall von 2 bis 5 (deswegen heißt k auch Laufindex), d. h., es müssen nacheinander die Zahlen 2, 3, 4 und 5 für k in den Term $(-1)^k \cdot \frac{k-3}{k+7}$ eingesetzt werden. Die entstehenden Terme müssen dann summiert werden – das sagt uns das Summenzeichen $\sum$. P

Setzen wir exemplarisch für $k = 2$ in den Term ein. Es ergibt sich $(-1)^2 \cdot \frac{(2-3)}{2+7} = -\frac{1}{9}$. Analog geht man für die anderen Werte vor, die für k eingesetzt werden sollen. Z

Insgesamt erhält man: H

$$\begin{aligned}\sum_{k=2}^{5}\left((-1)^k \cdot \frac{k-3}{k+7}\right) &= -\frac{1}{9} + (-1)^3 \cdot \frac{3-3}{3+7} + (-1)^4 \cdot \frac{4-3}{4+7} + (-1)^5 \cdot \frac{5-3}{5+7} \\ &= -\frac{1}{9} - 0 + \frac{1}{11} - \frac{1}{6} = -\frac{37}{198}.\end{aligned}$$

b) Beim Produkt $\prod_{j=-3}^{5} m^j$ kommen im Term m^j zwei Variablen vor. Man muss also aufpassen, dass man die Zahlen -3 bis 5 für die richtige Variable, nämlich für den P

C. Ableitinger, A. Herrmann, *Lernen aus Musterlösungen zur Analysis und Linearen Algebra*,
DOI 10.1007/978-3-658-02353-9_5, © Springer Fachmedien Wiesbaden 2013

Laufindex j einsetzt. m hingegen bleibt variabel. Die einzelnen Terme werden diesmal miteinander multipliziert – das erkennt man am Produktzeichen $\prod$.

H

Es ergibt sich

$$\begin{aligned}\prod_{j=-3}^{5} m^j &= m^{-3} \cdot m^{-2} \cdot m^{-1} \cdot m^0 \cdot m^1 \cdot m^2 \cdot m^3 \cdot m^4 \cdot m^5 \\ &= \frac{m^1 \cdot m^2 \cdot m^3 \cdot m^4 \cdot m^5}{m^1 \cdot m^2 \cdot m^3} = m^9,\end{aligned}$$

wobei die Potenzrechenregeln $a^{-n} = \frac{1}{a^n}$, $a^n \cdot a^m = a^{n+m}$ und $\frac{a^n}{a^m} = a^{n-m}$ verwendet wurden.

P

c) Bei dieser Aufgabe muss man zunächst die Struktur hinter dem Produkt $\frac{2}{4} \cdot \frac{5}{7} \cdot \frac{10}{10} \cdot \frac{17}{13} \cdot \frac{26}{16}$ erkennen. Man muss gleichsam die Entstehungsgeschichte der einzelnen Faktoren rekonstruieren. Welcher Term führt – wenn man für einen noch festzulegenden Laufindex nacheinander bestimmte Werte einsetzt – zu den gegebenen Faktoren?

Z

Etwas leichter gelingt eine solche Rekonstruktion bei den Nennern. An der Zahlenfolge 4, 7, 10, 13, 16 erkennt man, dass in jedem Schritt die Zahl 3 addiert wird.

H

Erhöht sich der Laufindex um 1, so soll sich das Folgenglied also um 3 erhöhen. Das leistet der lineare Ausdruck $1 + 3n$. Wenn man für den Laufindex die Zahlen 1 bis 5 einsetzt, erhält man gerade die Zahlenfolge der Nenner.

K

Für den Zähler braucht man etwas mehr Struktursinn. Die Folge 2, 5, 10, 17, 26 offenbart nicht sofort, welcher Term hier gesucht ist. Untersucht man wie vorher die Abstände zwischen je zwei aufeinanderfolgenden Zahlen, erkennt man aber auch hier ein System: Die Abstände vergrößern sich von Schritt zu Schritt um 2. Allerdings ist es gar nicht so einfach, einen Term zu finden, der das leistet. Versuchen Sie es ruhig einmal!

T

Hier hilft ein Sichtweisenwechsel: Die Zahlen 2, 5, 10, 17, 26 kann man sich auch aus den Quadratzahlen 1, 4, 9, 16, 25 entstanden denken. Sie unterscheiden sich ja jeweils nur um 1. Ein passender Term für den Zähler ist also $n^2 + 1$, wobei der Index n wieder von 1 bis 5 läuft.

P

Das ist ein entscheidender Punkt an dieser Stelle. Die Terme in Zähler und Nenner sollen am Ende denselben Laufindex mit demselben Laufbereich haben! In unserem Fall passt das glücklicherweise zusammen. Hätten die Laufindizes in Zähler und Nenner allerdings unterschiedliche Laufbereiche, müsste man entweder im Zähler oder im Nenner eine Indexverschiebung durchführen. Wie man das macht, sehen Sie in der nächsten Teilaufgabe.

B

Wir erhalten:

$$\frac{2}{4} \cdot \frac{5}{7} \cdot \frac{10}{10} \cdot \frac{17}{13} \cdot \frac{26}{16} = \prod_{n=1}^{5} \frac{n^2 + 1}{1 + 3n}.$$

d) Der Laufbereich des Index soll um 1 nach unten „verschoben“ werden. Das erkennt man daran, dass k anstatt bei 1 nun bei 0 beginnen soll.

P

Nachdem die Summe insgesamt aber wieder gleich viele Summanden haben soll (nämlich n Stück), darf k nur noch bis $n-1$ laufen.

H

Am Term hinter dem Summenzeichen verändert sich durch diese Indexverschiebung auch etwas. Sehen wir uns dazu an, wie der erste Summand der ursprünglichen Summe ausgesehen hat: $\frac{(1-2)^2}{3^{1+1}} = \frac{1}{9}$. In der neuen Summe wird zur Berechnung des ersten Summanden für k die Zahl 0 eingesetzt.

Z

Zähler und Nenner sollen sich dadurch aber natürlich nicht ändern. Die Abnahme des Werts von k um 1 lässt sich kompensieren, indem man im Zähler $(k-1)^2$ statt $(k-2)^2$ und im Nenner 3^{k+2} statt 3^{k+1} schreibt. Dadurch erhält man nämlich als ersten Summanden $\frac{(0-1)^2}{3^{0+2}} = \frac{1}{9}$, also genau das Gewünschte.

A

Man kann sich das auch etwas formaler so überlegen: Wenn die Werte des Laufbereichs von k um 1 vermindert werden, kann man diese Verminderung ausgleichen, indem man im Summanden jedes k durch $k+1$ ersetzt. Im Zähler würde das also bedeuten:

K

$$((k+1)-2)^2 = (k-1)^2.$$

Die Indexverschiebung sieht dann wie folgt aus:

B

$$\sum_{k=1}^{n} \frac{(k-2)^2}{3^{k+1}} = \sum_{k=0}^{n-1} \frac{(k-1)^2}{3^{k+2}}.$$

e) Bei diesem Produkt erkennt man nicht sofort am Laufbereich, um welche Zahl der Laufindex verschoben wird.

P

Die einzige Vergleichsmöglichkeit bietet sich bei der Betrachtung des Zählers.

K

Nachdem bei der Indexverschiebung der Zähler um 5 vergrößert wird (von $l+3$ auf $l+8$), muss diesmal der Laufbereich des Index l um 5 nach unten verschoben werden. l muss also von -2 bis 1 laufen.

H

Im Nenner müssen wir entsprechend l durch $l+5$ ersetzen und erhalten insgesamt:

A

$$\prod_{l=3}^{6} \frac{l+3}{l^2} = \prod_{l=-2}^{1} \frac{l+8}{(l+5)^2}.$$

Verständnisfragen 5.1

1. Wohin ist in Teil b) der Faktor m^0 verschwunden? Er steht weder im Zähler noch im Nenner des entstandenen Bruchs.
2. Hätte man bei Aufgabe c) im Nenner auch den Term $4+3n$ verwenden können? Falls ja, was müsste man am Laufbereich bzw. am Zählerterm verändern? Falls nein, begründen Sie, warum dies nicht möglich ist.

3. Stellen Sie sich vor, der Zähler des Summanden in d) wäre vor der Indexverschiebung $3k+5$ gewesen. Wie hätte dieser Zähler bei der Indexverschiebung verändert werden müssen?

Arbeitsauftrag 5.1 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.1.

5.2 Vollständige Induktion

Aufgabe 5.2 Beweisen Sie, dass für alle natürlichen Zahlen $n \geq 2$ die Ungleichung

$$2^{n+1} \geq 3n+2$$

gilt.

K **Ausführliche Lösung 5.2** Wir beweisen mit vollständiger Induktion, da es sich hierbei um eine Aussage über die *natürlichen Zahlen* handelt!

B Beweis durch Induktion nach n:

A *Induktionsanfang* für $n=2$:

H Da für $n=2$

$$2^{n+1} = 2^{2+1} = 2^3 = 8 \geq 8 = 3 \cdot 2 + 2 = 3n+2$$

gilt, ist die Aussage für $n=2$ bewiesen und der Induktionsanfang vollbracht.

Z *Induktionsschritt*: Wir wollen den Induktionsschritt von n nach $n+1$ durchführen.

Dafür nehmen wir als *Induktionsvoraussetzung* (IV) an, dass die Ungleichung

$$2^{n+1} \geq 3n+2$$

für ein beliebiges, aber festes $n \geq 2$ bereits gilt.

Beh.: Die Aussage gilt dann auch für $n+1$.

P Der Beweis dieser Behauptung ist der Hauptteil des Beweises der Ungleichung.

K Dafür müssen wir als Erstes die Aussage für $n+1$ formulieren. Dazu ersetzen wir *alle* n's in der Aussage (für n) durch $n+1$.

A Für $n+1$ müssen wir also Folgendes zeigen:

$$2^{(n+1)+1} \geq 3(n+1)+2.$$

Um die Aussage für $n+1$ zu beweisen, versuchen wir die Aussage (hier eine Ungleichung) so umzuformen, dass wir die Induktionsvoraussetzung anwenden können. Weitere Umformungen und Überlegungen sollten uns dann zum Ziel führen. Wir fangen an, die linke Seite der Ungleichung so umzuformen[1], dass wir die Induktionsvoraussetzung verwenden können:

$$2^{(n+1)+1} = 2^{n+1} \cdot 2^1.$$

Jetzt können wir die Induktionsvoraussetzung anwenden:

$$\begin{aligned} 2^{(n+1)+1} &= 2^{n+1} \cdot 2 \\ &\overset{\mathrm{IV}}{\geq} (3n+2) \cdot 2. \end{aligned}$$

Wir dürfen nun den Blick darauf nicht verlieren, was wir zeigen wollen. Nach mehreren Umformungen wollen wir schließlich bei dem Ergebnis

$$2^{(n+1)+1} \geq \ldots \geq 3(n+1) + 2$$

ankommen.

Man hat jetzt zwei Möglichkeiten (vielleicht kennen Sie sogar noch mehr) vorzugehen. So können wir zum Beispiel auf beiden Seiten Umformungen machen und schauen, ob wir uns in der Mitte irgendwo treffen. Das heißt, wir formen einfach mal die Terme $(3n+2) \cdot 2$ und $3(n+1)+2$ weiter um:

$$\begin{aligned} 2^{(n+1)+1} &= 2^{n+1} \cdot 2 \\ &\overset{\mathrm{IV}}{\geq} (3n+2) \cdot 2 \\ &= 6n+4 \\ &\geq \ldots ? \\ &\geq \ldots ? \\ &\geq 3n+3+2 \\ &= 3(n+1)+2. \end{aligned}$$

Das wäre ein erster Schritt, und weiter könnte es so gehen:

$$\begin{aligned} 2^{(n+1)+1} &= 2^{n+1} \cdot 2 \\ &\overset{\mathrm{IV}}{\geq} (3n+2) \cdot 2 \\ &= 6n+4 \\ &= 3n+(3n+4) \\ &\geq \ldots ? \\ &\geq 3n+5 \\ &= 3n+3+2 \\ &= 3(n+1)+2. \end{aligned}$$

[1] Manchmal ist es auch geschickter, die Aussage „anders herum“ zu lesen, also mit der rechten Seite zu starten.

B

Da haben wir uns schon deutlich angenähert. Denn das Einzige, was jetzt noch zu zeigen bleibt, ist doch, dass $(3n+4) \geq 5$ gilt.

H

Das sieht man aber recht schnell ein, da $n \geq 2$ ist, und damit auf jeden Fall $3n \geq 6 \geq 1$ gilt. Also gilt insgesamt $3n+4 \geq 1+4 = 5$.

B

Das haben wir uns jetzt sehr ausführlich überlegt, in unserem Beweis reicht es, wenn wir das so aufschreiben:

$$\begin{aligned}
2^{(n+1)+1} &= 2^{n+1} \cdot 2 \\
&\overset{\text{IV}}{\geq} (3n+2) \cdot 2 \\
&= 6n+4 \\
&= 3n + (\underbrace{3n}_{\geq 1;\text{ da } n \geq 2} + 4) \\
&\geq 3n + (1+4) \\
&= 3n+5 \\
&= 3n+3+2 \\
&= 3(n+1)+2.
\end{aligned}$$

Und schon steht da der ganze Beweis! Natürlich würde man auch nur diesen letzten Teil aufschreiben, die Zwischenüberlegungen mit den „Pünktchen" sind nur fürs Schmierpapier!

K

Eine andere Möglichkeit ist, die Aussage (hier: Ungleichung), die noch zu beweisen bleibt, durch äquivalente Umformungen auf eine wahre Aussage zurückzuführen. In unserem Fall müssen wir ja noch

$$(3n+2) \cdot 2 \geq 3(n+1) + 2$$

zeigen. Um zu kennzeichnen, dass das noch zu zeigen ist, schreibt man oft ein „?" über dem Relationszeichen.

B

Okay, machen wir nun äquivalente Umformungen und versuchen, die Ungleichung auf eine Ungleichung zurückzuführen, von der wir wissen, dass sie wahr ist:

H

$$\begin{aligned}
& (3n+2) \cdot 2 \overset{?}{\geq} 3(n+1) + 2 \\
\Leftrightarrow \quad & 6n+4 \overset{?}{\geq} 3n+5 && |(-3n);\ |(-4) \\
\Leftrightarrow \quad & 3n \overset{?}{\geq} 1.
\end{aligned}$$

Die letzte Ungleichung gilt, da $3n$ sogar schon ≥ 6 ist, weil $n \geq 2$ nach Voraussetzung gilt.

P

Hinweis: Bei der zweiten Vorgehensweise ist es wichtig, dass es sich bei den Umformungen um Äquivalenzen handelt. Es reicht nicht, aus der Aussage, die wir beweisen sollen, eine wahre Aussage zu folgern.

Verständnisfragen 5.2

1. Warum wurde der Induktionsanfang für $n = 2$ und nicht für $n = 1$ gemacht?
2. Beim 2. Lösungsweg wurde die Wichtigkeit der *äquivalenten* Umformungen betont. Warum sind diese so wichtig? Braucht man wirklich Äquivalenzen, würden nicht auch Folgerungen reichen? Warum?

Arbeitsauftrag 5.2 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.1.

5.3 Mengeninklusion und -gleichheit bei Bild und Urbild

Aufgabe 5.3 Sei $f : X \to Y$ eine Abbildung und seien $A, B \subseteq X$ und $C, D \subseteq Y$. Zeigen Sie:

a) $f(A \cup B) = f(A) \cup f(B)$,

b) $f^{-1}(C \cap D) = f^{-1}(C) \cap f^{-1}(D)$,

c) $f(f^{-1}(C)) \subseteq C$.

d) $f(f^{-1}(C)) = C$ für alle $C \subseteq Y \Leftrightarrow f$ ist surjektiv.

Ausführliche Lösung 5.3 In dieser Aufgabe müssen Mengengleichheiten bzw. -inklusionen nachgewiesen werden. Bei den Mengen handelt es sich dabei um Bilder und Urbilder[2] einer Abbildung $f : X \to Y$. B

a) Wir sollen zeigen, dass das Bild der Vereinigung zweier Teilmengen von X gleich der Vereinigung ihrer Bilder ist. Wir müssen also eine Mengengleichheit beweisen und das macht man üblicherweise, indem man beide Inklusionsrichtungen nachweist. Unser Beweis wird sich damit in die beiden Teile „$f(A \cup B) \subseteq f(A) \cup f(B)$“ und „$f(A) \cup f(B) \subseteq f(A \cup B)$“ aufsplitten. K

Beginnen wir mit „$f(A \cup B) \subseteq f(A) \cup f(B)$“: Eine Menge ist Teilmenge einer anderen, wenn jedes ihrer Elemente auch in der anderen Menge enthalten ist (in Formeln ausgedrückt: $M \subseteq N \;\Leftrightarrow\; \forall x \in M$ gilt: $x \in N$). Und genau so werden wir den Beweis auch führen. Wir werden zeigen, dass jedes Element von $f(A \cup B)$ auch in $f(A) \cup f(B)$ enthalten ist. Und wie machen wir das? Nun, wir zeigen die Behauptung – wie bei Allaussagen üblich – für ein beliebiges, aber festes Element.

Sei also $y \in f(A \cup B)$ beliebig, aber fest. Z

[2] Falls Sie zum Einstieg in Mengeninklusionen und -gleichheiten erst einmal ohne Bild und Urbild arbeiten möchten, so werfen Sie einen Blick in Übung 10.3 im Kapitel 10.1.

P

Bevor man in den eigentlichen Beweis einsteigt, könnte es helfen, erst einmal möglichst viele Informationen zu sammeln. Z. B. könnte man sich überlegen: Was weiß ich bereits über y, welcher Art ist die Menge $f(A \cup B)$, wo wollen wir eigentlich hin?

Machen wir das für unseren Fall:

- $y \in f(A \cup B)$ – das ist schließlich unsere Voraussetzung.
- $y \in Y$ – da es im Bild einer Teilmenge von X liegt.
- $f(A \cup B) = \{f(x) | x \in A \cup B\}$
- Ziel ist es zu zeigen: $y \in f(A) = \{f(x) | x \in A\}$ oder $y \in f(B) = \{f(x) | x \in B\}$.

B

Eine zentrale Rolle spielt hier auf jeden Fall der Begriff *Bild*, d. h. die Menge $f(M)$, wobei M die verschiedenen Teilmengen von X repräsentiert, die in der Aufgabenstellung auftauchen (A, B, $A \cup B$). Deshalb schauen wir uns zunächst an, was das Bild überhaupt ist. Definiert wird es häufig auf eine der zwei folgenden Weisen:

$$f(M) := \{f(x) | x \in M\}$$

oder

$$f(M) := \{y \in Y | \exists x \in M \text{ mit } y = f(x)\}.$$

P

Die beiden Definitionen sind natürlich äquivalent zueinander, ansonsten würden sie unterschiedliche Objekte definieren.

Das Bild einer Teilmenge ist demnach zunächst eine Teilmenge von Y, und zwar sammeln wir hier alle Elemente aus Y, die von Elementen aus M durch die Abbildung f „getroffen“ werden. Diese Definition werden wir für die verschiedenen Teilmengen A, B, $A \cup B$ von X verwenden müssen, die in der Aufgabe auftauchen.

A

Versuchen wir nun mit den bis hierhin entwickelten Ideen, den Beweis zu führen: Aus $y \in f(A \cup B)$ können wir mit Hilfe der Definition des Bildes (für die Menge $A \cup B$) folgern, dass es ein x aus der Menge $A \cup B$ geben muss, sodass $y = f(x)$ gilt. $x \in A \cup B$ heißt aber, dass x in A oder dass x in B liegt.

H

Zusammen bedeutet dies, dass es ein $x \in A$ oder ein $x \in B$ gibt, für das $y = f(x)$ gilt. Laut Definition des Bildes (einmal für A und einmal für B) folgt daraus wiederum, dass y im Bild von A oder im Bild von B liegt. Somit ist y in der Vereinigung der beiden Bilder enthalten.

B

Versuchen wir abschließend diese Überlegungen etwas formaler aufzuschreiben:

$$\begin{aligned}
y \in f(A \cup B) &\Rightarrow \exists x \in A \cup B \text{ mit } y = f(x) \\
&\Rightarrow \exists x \in A \text{ oder } \exists x \in B \text{ mit } y = f(x) \\
&\Rightarrow \exists x \in A \text{ mit } y = f(x) \text{ oder } \exists x \in B \text{ mit } y = f(x) \\
&\Rightarrow y \in f(A) \text{ oder } y \in f(B) \\
&\Rightarrow y \in f(A) \cup f(B).
\end{aligned}$$

Kommen wir nun zur zweiten Inklusion „$f(A) \cup f(B) \subseteq f(A \cup B)$“: Natürlich wäre es am einfachsten, wenn wir die Folgerungskette der anderen Inklusion einfach in der anderen Richtung aufschreiben könnten.

Aber Vorsicht(!), wir müssen jeden einzelnen Folgerungspfeil unter die Lupe nehmen und entscheiden, ob seine Umkehrung ebenfalls gilt. Dies ist durchaus nicht immer der Fall, manchmal muss man andere Wege für die Rückrichtung gehen. Da wir hier aber eigentlich nur die Definition des Bildes und die Definition von „$\cup$“ verwendet haben, stehen unsere Chancen ganz gut.

Wir suchen also Begründungen für die folgenden Implikationen:

$$\begin{aligned}
y \in f(A) \cup f(B) &\Rightarrow y \in f(A) \text{ oder } y \in f(B) && (5.1)\\
&\Rightarrow \exists x \in A \text{ mit } y = f(x) \text{ oder } \exists x \in B \text{ mit } y = f(x) && (5.2)\\
&\Rightarrow \exists x \in A \text{ oder } \exists x \in B \text{ mit } y = f(x) && (5.3)\\
&\Rightarrow \exists x \in A \cup B \text{ mit } y = f(x) && (5.4)\\
&\Rightarrow y \in f(A \cup B). && (5.5)
\end{aligned}$$

Die erste Implikation in Zeile 5.1 gilt wegen der Definition der Vereinigung zweier Mengen. Die Implikation in Zeile 5.2 ist ebenfalls wahr, weil hier lediglich die Definitionen des Bildes von A und des Bildes von B eingehen. Von Zeile 5.2 gelangen wir durch Umformulierung mit Hilfe logischer Schlussfolgerungen zu Zeile 5.3. Die vierte Implikation (Zeile 5.4) benutzt die Definition der Vereinigung und die letzte Implikation (Zeile 5.5) folgt aus der Definition des Bildes von $A \cup B$.

Sicherlich hätten wir uns hier einiges an Arbeit erspart, wenn wir gleich Äquivalenzen formuliert hätten. Es ist aber besser, wenn man zu Beginn etwas vorsichtiger ist, denn man übersieht leicht einmal, dass eine Rückrichtung evtl. nicht erfüllt ist.

b) Wir wollen unsere Überlegungen nun etwas kürzer halten. Wie in Teil a) müssen wir eine Mengengleichheit nachweisen, diesmal für den Zusammenhang zwischen Urbildern einer Abbildung. Also sollten wir uns als Erstes die Frage stellen, was ein Urbild überhaupt ist. Für eine Teilmenge $N \subseteq Y$ ist es allgemein so definiert:

$$f^{-1}(N) := \{x \in X | f(x) \in N\}.$$

Es handelt sich dabei also um eine Teilmenge von X, und zwar werden hier jene Elemente gesammelt, die durch die Abbildung f auf ein Element aus N abgebildet werden.[3]

Wir werden für den Beweis diese Definition und die Eigenschaften des Schnitts ausnutzen, so wie wir das oben für Bild und Vereinigung gemacht haben.

Sei $x \in f^{-1}(C \cap D)$ beliebig. Laut Definition des Urbildes ist $x \in X$ mit der Eigenschaft $f(x) \in C \cap D$. Das heißt aber, dass x ein Element von X ist und $f(x)$ in C und in D liegt. Daraus folgt, dass x ein Element aus X ist, für das $f(x) \in C$ ist, und dass x ein Element aus X ist mit $f(x) \in D$. Aus der Definition des Urbildes erhalten wir so: $x \in f^{-1}(C)$ und $x \in f^{-1}(D)$. Schließlich dürfen wir aus der Definition des Schnitts „$\cap$“ noch folgern, dass dann $x \in f^{-1}(C) \cap f^{-1}(D)$ gilt. Damit haben wir die Inklusion „$f^{-1}(C \cap D) \subseteq f^{-1}(C) \cap f^{-1}(D)$“ bewiesen.

[3] Beachten Sie, dass dabei f^{-1} das Urbild und nicht die Umkehrfunktion von f bezeichnet.

K Die Inklusion „$f^{-1}(C) \cap f^{-1}(D) \subseteq f^{-1}(C \cap D)$“ können wir durch Umkehrung der Implikationen beweisen.

B c) Es soll für jede Abbildung $f : X \to Y$ gezeigt werden, dass für eine Teilmenge $C \subseteq Y$ die Inklusion $f(f^{-1}(C)) \subseteq C$ gilt.

H Ist y ein beliebiges Element aus $f(f^{-1}(C))$, dann dürfen wir als Erstes mit Hilfe der Definition des Bildes folgern, dass es dann ein x aus $f^{-1}(C)$ gibt, für das $y = f(x)$ gilt. $x \in f^{-1}(C)$ bedeutet, dass $f(x) \in C$ ist. Da nun aber $y = f(x)$ war, können wir folgern, dass $y \in C$ ist. Und das wollten wir gerade zeigen.

B d) Hier müssen wir eine Äquivalenz nachweisen, bei der die Begriffe Bild, Urbild und surjektiv miteinander verknüpft werden.

K Dies machen wir, indem wir die beiden Implikationen separat beweisen.

H Nehmen wir also einmal an, dass für alle $C \subseteq Y$ nicht nur die Inklusion aus c), sondern sogar die Gleichheit der Mengen $f(f^{-1}(C))$ und C gilt. Wenn dies für jede Teilmenge gelten soll, muss es insbesondere auch für Y selbst gelten (Y ist ja schließlich eine Teilmenge seiner selbst). Also gilt: $f(f^{-1}(Y)) = Y$.

K Mit einem kleinen Blick Richtung Ziel wird klar, dass wir als Nächstes $f^{-1}(Y) = X$ zeigen müssen. Denn, wenn wir das haben, gilt doch $f(X) = Y$, woraus die Surjektivität direkt folgt.

H Die Mengengleichheit $f^{-1}(Y) = X$ gilt aber nach der Definition des Urbildes.

B Damit haben wir nun die Implikation „$f(f^{-1}(C)) = C$ für alle $C \subseteq Y \Rightarrow f$ ist surjektiv“ bewiesen.

Kommen wir nun zur Umkehrung: Sei f surjektiv. Die Mengeninklusion „$\subseteq$“ gilt nach Teil c) schon allgemein (C war dort nicht weiter spezifiziert; die Inklusion gilt also für alle Teilmengen $C \subseteq Y$). Wir müssen nur noch die Inklusion „$C \subseteq f(f^{-1}(C))$“ zeigen.

H Sei $y \in C$ beliebig. Wegen der Surjektivität der Abbildung f gibt es dann ein $x \in X$ mit $f(x) = y$. Weil y nun aber aus C ist, ist x aus dem Urbild von C: $x \in f^{-1}(C)$. Daraus folgt schon, dass $f(x) \in f(f^{-1}(C))$ ist (Bilddefinition). Und weil $y = f(x)$ war, gilt dann auch $y \in f(f^{-1}(C))$ und damit die andere Mengeninklusion. Insgesamt folgt so die Mengengleichheit.

Verständnisfragen 5.3

1. Begründen Sie, warum die in Teil b) getroffene Aussage „Die Inklusion $f^{-1}(A) \cap f^{-1}(B) \subseteq f^{-1}(A \cap B)$ können wir durch Umkehrung der Implikationen beweisen.“ berechtigt ist.
2. In Teil d) wird behauptet, dass $f^{-1}(Y) = X$ aus der Definition des Urbildes folgt. Begründen Sie diese Behauptung im Detail.
3. Im Beweis der Implikation „$\Leftarrow$“ der Äquivalenz in Teil d) wurde behauptet, dass mit Hilfe der Bilddefinition aus $x \in f^{-1}(C)$ bereits $f(x) \in f(f^{-1}(C))$ folgt. Auf welche Teilmenge bezieht sich dabei die Definition?

4. Im Beweis zu d) wurden zwei verschiedene (aber äquivalente) Definitionen der Surjektivität von f verwendet. Welche sind das?
5. Gilt die Äquivalenz in d) auch dann, wenn die Mengengleichheit $f(f^{-1}(C)) = C$ nicht für jede Teilmenge von Y, sondern nur für eine gefordert wird?

Arbeitsauftrag 5.3 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.1.

5.4 Injektivität und Surjektivität

Aufgabe 5.4 Gegeben seien die beiden Abbildungen $f : X \to Y$ und $g : Y \to X$ mit der Eigenschaft

$$(g \circ f)(x) = x \qquad \forall\, x \in X. \tag{5.6}$$

Zeigen Sie, dass f injektiv und g surjektiv ist!

Ausführliche Lösung 5.4 Wir zeigen zunächst, dass f injektiv ist.

Für die Injektivität gibt es zwei äquivalente Definitionen:

1. Die Abbildung $f : X \to Y$ ist genau dann injektiv, wenn für alle $x_1, x_2 \in X$ mit der Eigenschaft $f(x_1) = f(x_2)$ stets schon $x_1 = x_2$ folgt.
2. Die Abbildung $f : X \to Y$ ist genau dann injektiv, wenn für alle $x_1, x_2 \in X$ mit der Eigenschaft $x_1 \neq x_2$ stets schon $f(x_1) \neq f(x_2)$ folgt.

Prinzipiell ist es für den Beweis der Injektivität egal, welche der beiden Definitionen man zum Nachweis nutzt. Oft bietet sich eine aber mehr an. Tendenziell wird man meistens die erste Definition verwenden, da Gleichheit einfacher handhabbar als Ungleichheit ist – aber das ist Geschmackssache.

Wir wollen hier jedenfalls mit Hilfe der ersten Definition die Injektivität von f beweisen.

Seien $x_1, x_2 \in X$ mit $f(x_1) = f(x_2)$ beliebig. [Zu zeigen ist dann: $x_1 = x_2$.][4]

Da $f : X \to Y$ eine Abbildung von X nach Y ist, sind $f(x_1), f(x_2) \in Y$.

Da $g : Y \to X$ eine Abbildung von Y nach X ist und $f(x_1) = f(x_2) \in Y$ gilt, folgt

$$g(f(x_1)) = g(f(x_2)).$$

Das heißt aber

$$(g \circ f)(x_1) = (g \circ f)(x_2)$$

[4] Das ist ein typischer Anfang eines Beweises, bei dem man etwas *für alle* Elemente mit bestimmten Eigenschaften zeigen muss. Man nimmt sich ein beliebiges, aber festes Element mit geforderten Eigenschaften heraus und zeigt, dass die Behauptung für dieses Element gilt.

und daraus folgt schließlich mit Hilfe der Voraussetzung (5.6), dass

$$x_1 = x_2$$

ist, was wir ja zeigen mussten. Da wir $x_1, x_2 \in X$ mit $f(x_1) = f(x_2)$ beliebig gewählt hatten, gilt für alle $x_1, x_2 \in X$ mit $f(x_1) = f(x_2)$, dass stets schon $x_1 = x_2$ folgt. Das entspricht aber gerade der Definition für Injektivität. Demnach ist f also tatsächlich injektiv.

Als Nächstes zeigen wir, dass g surjektiv ist.

Auch für die Surjektivität gibt es zwei äquivalente Definitionen:

1. Die Abbildung $g : Y \to X$ ist surjektiv genau dann, wenn für jedes $x \in X$ ein $y \in Y$ existiert, sodass $g(y) = x$ ist.
2. Die Abbildung $g : Y \to X$ ist surjektiv genau dann, wenn $g(Y) = X$ gilt.

In den meisten Fällen verwendet man die erste Definition, um eine Abbildung auf Surjektivität zu überprüfen. Die zweite Definition ist hilfreich, wenn man z. B. über die Anzahl der Elemente von $g(Y)$ und X schon etwas weiß. Wir verwenden hier auf jeden Fall die erste Definition.

Wir sollen also zeigen, dass für jedes $x \in X$ ein $y \in Y$ mit $g(y) = x$ existiert. Wie man etwas für alle Elemente zeigt, haben wir oben schon gelernt. Das machen wir hier genauso:

Sei $x \in X$ beliebig.

Jetzt müssen wir zu diesem x ein $y \in Y$ finden, für das $g(y) = x$ gilt.

Soll man zeigen, dass *ein* Element mit bestimmten Eigenschaften *existiert*, so bleibt einem in einigen Fällen nichts anderes übrig, als ein solches Element konkret anzugeben. Um das zu finden, braucht man aber die richtige Idee. Hat man diese, dann ist es oft nicht mehr so schwer, zu zeigen, dass dieses Element tatsächlich die gewünschten Eigenschaften besitzt.

Zurück zu unserer Aufgabe: Wir erinnern uns an die Voraussetzung (5.6)! Für x gilt:

$$(g \circ f)(x) = x.$$

Daraus folgt doch aber

$$g(f(x)) = x.$$

Und, sehen Sie schon, wie wir y wählen müssen? Nein? Dann überlegen Sie doch mal, in welcher Menge $f(x)$ liegt! Richtig, in Y, da f ja von X nach Y abbildet. Dann könnten wir doch einfach $y := f(x)$ wählen, oder? Dann gilt nämlich $y \in Y$ und auch

$$g(y) = g(f(x)) = x,$$

wie wir uns gerade schon überlegt haben. Und so ein y hatten wir schließlich gesucht.

Verständnisfragen 5.4

1. Beim Beweis der Injektivität von f wurde benutzt, dass $f : X \to Y$ und $g : Y \to X$ Abbildungen sind. Welche Eigenschaft der Abbildung wurde jeweils genau benutzt?

2. Können Sie aus dem Beweis zur Surjektivität erkennen, welche groben Schritte man *allgemein* vornimmt, um die Surjektivität einer Abbildung $f : M \to N$ zu zeigen (M, N Mengen)?

Übung 5.4 Ordnen Sie nun selbst den einzelnen Bearbeitungsschritten der Musterlösung die Phasen aus Kapitel 3 zu! Tragen Sie dazu die jeweiligen Buchstaben in die kleinen Quadrate am Seitenrand ein!

Arbeitsauftrag 5.4 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen, der Übung und dem Arbeitsauftrag finden Sie in Abschnitt 11.1.

5.5 Äquivalenzrelation und Äquivalenzklassen

Aufgabe 5.5 Auf $(\mathbb{R} \times \mathbb{R}) \setminus \{(0,0)\}$ ist die folgende Relation gegeben:

$$(x_1, y_1) \sim (x_2, y_2) :\Leftrightarrow x_1 \cdot y_2 = x_2 \cdot y_1.$$

Zeigen Sie, dass $\sim$ eine Äquivalenzrelation ist. Bestimmen Sie weiterhin die Äquivalenzklasse von (a, b).

Ausführliche Lösung 5.5 Die Aufgabe teilt sich in zwei Teile. Wir sollen zunächst überprüfen, ob die angegebene Relation eine Äquivalenzrelation ist. Danach sollen wir uns mit Äquivalenzklassen beschäftigen. B

Widmen wir uns zunächst dem ersten Teil. Dazu müssen wir natürlich wissen, was eine Äquivalenzrelation ist. Schauen wir uns dazu die Definition an:

Definition: Eine Relation $\sim$ auf der Menge M heißt genau dann *Äquivalenzrelation*, wenn die folgenden drei Eigenschaften erfüllt sind:

1. $\sim$ ist reflexiv, d. h. für alle $a \in M$ gilt: $a \sim a$,

2. $\sim$ ist symmetrisch, d. h. für alle $a, b \in M$ mit $a \sim b$ gilt auch $b \sim a$.

3. $\sim$ ist transitiv, d. h. für alle $a, b, c \in M$ mit $a \sim b$ und $b \sim c$ gilt auch $a \sim c$.

Diese allgemeine Definition müssen wir nun auf die vorliegende Relation übertragen. Wir schauen uns zunächst die Menge $(\mathbb{R}\times\mathbb{R})\setminus\{(0,0)\}$ an, auf der die Relation definiert ist. Diese besteht aus Paaren reeller Zahlen (x, y) – diese können wir uns als Punkte im kartesischen Koordinatensystem vorstellen. Der Koordinatenursprung $(0, 0)$ ist dabei allerdings nicht enthalten. Halten wir also Folgendes fest: $M = (\mathbb{R} \times \mathbb{R})\setminus\{(0,0)\}$. Ein Element a der Menge M entspricht dann einem Zahlenpaar: $a = (x, y)$. A

B

Mit diesem Wissen können wir die Definition nun an die gegebene Relation anpassen und bekommen so einen Fahrplan dafür, was zu überprüfen ist:

A

Die Relation $\sim$ auf der Menge $(\mathbb{R} \times \mathbb{R}) \setminus \{(0,0)\}$ ist genau dann eine Äquivalenzrelation, wenn die folgenden drei Eigenschaften erfüllt sind:

1. $\sim$ ist reflexiv, d. h. für alle $(x,y) \in (\mathbb{R} \times \mathbb{R}) \setminus \{(0,0)\}$ gilt: $(x,y) \sim (x,y)$,
2. $\sim$ ist symmetrisch, d. h. für alle $(x_1,y_1),(x_2,y_2) \in (\mathbb{R} \times \mathbb{R}) \setminus \{(0,0)\}$ mit $(x_1,y_1) \sim (x_2,y_2)$ gilt auch $(x_2,y_2) \sim (x_1,y_1)$.
3. $\sim$ ist transitiv, d. h. für alle $(x_1,y_1),(x_2,y_2),(x_3,y_3) \in (\mathbb{R} \times \mathbb{R}) \setminus \{(0,0)\}$ mit $(x_1,y_1) \sim (x_2,y_2)$ und $(x_2,y_2) \sim (x_3,y_3)$ gilt auch $(x_1,y_1) \sim (x_3,y_3)$.

K

Nun können wir nach und nach zeigen, dass die einzelnen Punkte tatsächlich für die gegebene Relation erfüllt sind:

Z

1. $\sim$ *ist reflexiv*: Da wir hier eine Allaussage beweisen müssen, nehmen wir uns ein beliebiges Element und zeigen, dass die Eigenschaft für dieses Element erfüllt ist. Also sei $(x,y) \in (\mathbb{R} \times \mathbb{R}) \setminus \{(0,0)\}$ beliebig.

   A

   Es gilt trivialerweise (und ja, hier dürfen wir das wohl wirklich mal sagen, weil auf beiden Seiten der Gleichung dasselbe steht):

   $$x \cdot y = x \cdot y.$$

   Wenn wir dies nun mit der Definition der Relation abgleichen, erhalten wir daraus:

   $$(x,y) \sim (x,y).$$

   Dabei entspricht das x_1 der Definition dem x der linken Seite der Gleichung, y_1 dem y der rechten Seite der Gleichung, x_2 dem x der rechten Seite der Gleichung und y_2 dem y der linken Seite der Gleichung.

Z

2. $\sim$ *ist symmetrisch*: Auch hier haben wir wieder eine Allaussage nachzuweisen. Diesmal allerdings mit einer einschränkenden Eigenschaft. Wir nehmen uns daher beliebige Elemente $(x_1,y_1),(x_2,y_2) \in (\mathbb{R} \times \mathbb{R}) \setminus \{(0,0)\}$ für die bereits $(x_1,y_1) \sim (x_2,y_2)$ gilt. Wir beweisen, dass dann auch $(x_2,y_2) \sim (x_1,y_1)$ gilt:

   H

   $$(x_1,y_1) \sim (x_2,y_2) \overset{\text{Def.}\sim}{\Rightarrow} x_1 \cdot y_2 = x_2 \cdot y_1 \tag{5.7}$$
   $$\Rightarrow x_2 \cdot y_1 = x_1 \cdot y_2 \tag{5.8}$$
   $$\overset{\text{Def.}\sim}{\Rightarrow} (x_2,y_2) \sim (x_1,y_1). \tag{5.9}$$

Z

3. $\sim$ *ist transitiv*: Seien nun $(x_1,y_1),(x_2,y_2),(x_3,y_3) \in (\mathbb{R} \times \mathbb{R}) \setminus \{(0,0)\}$ beliebige Elemente, für die $(x_1,y_1) \sim (x_2,y_2)$ und $(x_2,y_2) \sim (x_3,y_3)$ gilt.

Es gilt:

$$
\begin{aligned}
&(x_1, y_1) \sim (x_2, y_2) \text{ und } (x_2, y_2) \sim (x_3, y_3) && (5.10)\\
&\quad \overset{\text{Def.}\sim}{\Rightarrow} x_1 \cdot y_2 = x_2 \cdot y_1 \text{ und } x_2 \cdot y_3 = x_3 \cdot y_2 && (5.11)\\
&\quad \Rightarrow x_1 \cdot y_2 \cdot x_2 \cdot y_3 = x_2 \cdot y_1 \cdot x_3 \cdot y_2 && (5.12)\\
&\quad \overset{x_2 \cdot y_2 \neq 0}{\Rightarrow} x_1 \cdot y_3 = x_3 \cdot y_1 && (5.13)\\
&\quad \overset{\text{Def.}\sim}{\Rightarrow} (x_1, y_1) \sim (x_3, y_3). && (5.14)
\end{aligned}
$$

Im dritten Schritt (von 5.12 nach 5.13) brauchen wir die Einschränkung $x_2 \cdot y_2 \neq 0$ damit wir die Folgerung treffen können. Den Fall $x_2 \cdot y_2 = 0$ müssen wir also gesondert betrachten:

Da in der Menge der Ursprung $(0,0)$ nicht enthalten ist, können wir folgern, dass *entweder* x_2 *oder* y_2 gleich Null ist, nicht aber beide gleichzeitig. Damit erhalten wir aus Zeile 5.11 für den Fall $x_2 = 0$, dass $x_1 = 0$ und $x_3 = 0$ sind. Daraus lässt sich aber wiederum schließen, dass $x_1 \cdot y_3 = x_3 \cdot y_1$ gilt und somit (x_1, y_1) in Relation zu (x_3, y_3) steht (formal: $(x_1, y_1) \sim (x_3, y_3)$). Der Fall $y_2 = 0$ verläuft analog dazu.

Sicherlich wären auch andere Umformungen der Gleichungen in Zeile 5.11 denkbar, die ebenfalls das Ergebnis $(x_1, y_1) \sim (x_3, y_3)$ liefern. Die obige Kernidee, die jeweiligen Seiten der beiden Gleichungen miteinander zu multiplizieren, erscheint dem einen oder anderen sicherlich als Trick. Beispielsweise könnte man auch eine der beiden Gleichungen nach einer Variablen umformen und in die andere einsetzen. Beachten Sie, dass Sie bei anderen Umformungen ebenfalls eine Fallunterscheidung vornehmen müssen.

Wir haben nun unseren Fahrplan abgearbeitet und können daraus schließen, dass es sich bei der Relation in der Tat um eine Äquivalenzrelation handelt.

Als Nächstes beschäftigen wir uns mit der Äquivalenzklasse $[(a, b)]$. Bevor wir diese allgemein aufstellen, betrachten wir ein Beispiel. Dazu suchen wir uns ein Element der Menge $(\mathbb{R} \times \mathbb{R}) \setminus \{(0,0)\}$ aus und überlegen uns, wie dessen Äquivalenzklasse aussieht. Nehmen wir den Punkt $(1, 2)$ als Beispiel.

Bevor wir anfangen können, müssen wir natürlich erst einmal wissen, was eine Äquivalenzklasse ist. Dazu schauen wir uns auch hier die Definition an.

Definition: Die Äquivalenzklasse eines Elements $a \in M$ bzgl. einer Äquivalenzrelation $\sim$ ist wie folgt definiert:

$$[a] := \{b \in M | a \sim b\}.$$

Bei der Äquivalenzklasse handelt es sich also um eine *Menge*. Diese ist insbesondere eine *Teilmenge von M*. Außerdem soll für alle Elemente der Menge gelten, dass sie in Relation zu a stehen. In der Äquivalenzklasse von a werden also alle Elemente gesammelt, die in Relation zu a stehen.

A

Soweit so gut, was bedeutet dies nun aber für den vorliegenden Fall? Wir leisten zunächst wieder Übersetzungsarbeit:

$$[(1,2)] = \{(x,y) \in (\mathbb{R} \times \mathbb{R}) \setminus \{(0,0)\} | (1,2) \sim (x,y)\}.$$

H

Diese Menge schauen wir uns nun etwas genauer an. Als Erstes wenden wir dafür einfach die Definition der Relation an, formen etwas um und erhalten:

$$\begin{aligned}[(1,2)] &= \{(x,y) \in (\mathbb{R} \times \mathbb{R}) \setminus \{(0,0)\} | 1 \cdot y = x \cdot 2\} \\ &= \{(x,y) \in (\mathbb{R} \times \mathbb{R}) \setminus \{(0,0)\} | y = 2x\}.\end{aligned}$$

B

Erkennen Sie schon etwas? Kommt Ihnen der Ausdruck „$y = 2x$“ irgendwie bekannt vor?

P

Genau, es handelt sich dabei um nichts anderes als eine Geradengleichung. Die Äquivalenzklasse von $(1,2)$ ist also nichts anderes als die Menge der Punkte auf der Geraden mit Steigung 2 durch den Ursprung, allerdings ohne den Ursprung selbst – der Punkt $(0,0)$ ist ja nicht in unserer Menge enthalten (vgl. Abbildung 5.1).

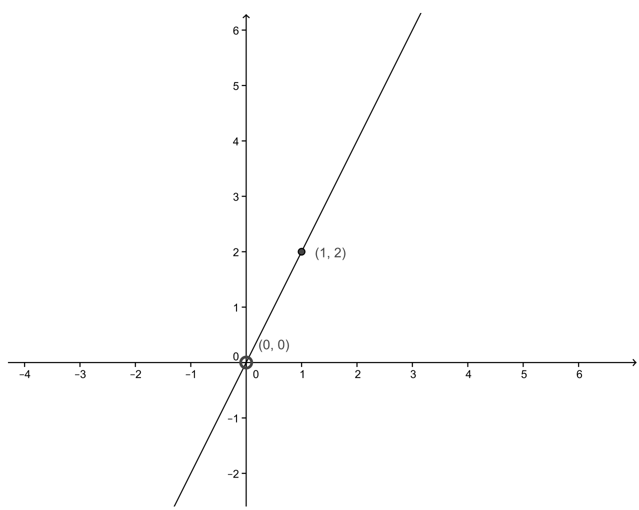

Abbildung 5.1: Äquivalenzklasse von $(1,2)$

Für ein Beispiel haben wir uns den Sachverhalt nun klar gemacht. Aber wir sollten ja die Äquivalenzklasse eines *allgemeinen* Punktes (a,b) ermitteln. Das wird nun hoffentlich recht ähnlich verlaufen.

Wir ersetzen in unserem Beispiel die Zahl 1 durch a und die Zahl 2 durch b und führen analoge Schritte durch:

K

H

$$\begin{aligned}[(a,b)] &= \{(x,y) \in (\mathbb{R} \times \mathbb{R}) \backslash \{(0,0)\} | (a,b) \sim (x,y)\} \\ &= \{(x,y) \in (\mathbb{R} \times \mathbb{R}) \backslash \{(0,0)\} | a \cdot y = x \cdot b\} \\ &= \{(x,y) \in (\mathbb{R} \times \mathbb{R}) \backslash \{(0,0)\} | y = \frac{b}{a} \cdot x\}.\end{aligned}$$

Die Äquivalenzklasse von (a,b) ist demnach die Menge aller Punkte auf der Ursprungsgeraden (wieder ohne den Ursprung selbst) mit der Steigung $\frac{b}{a}$.

Aber *Vorsicht*! Wir haben einfach durch a geteilt. Das dürfen wir natürlich nur, wenn a ungleich Null ist. Wir müssen also noch betrachten, was für $a = 0$ passiert.

P

Gesagt, getan:

H

$$\begin{aligned}[(0,b)] &= \{(x,y) \in (\mathbb{R} \times \mathbb{R}) \backslash \{(0,0)\} | 0 \cdot y = x \cdot b\} \\ &= \{(x,y) \in (\mathbb{R} \times \mathbb{R}) \backslash \{(0,0)\} | x \cdot b = 0\}.\end{aligned}$$

Da b nicht auch 0 sein kann, folgt:

$$[(0,b)] = \{(x,y) \in (\mathbb{R} \times \mathbb{R}) \backslash \{(0,0)\} | x = 0\}.$$

Das ist nicht anderes als die y-Achse ohne den Punkt $(0,0)$.

Verständnisfragen 5.5

1. Zu Beginn haben wir die allgemeine Definition der Äquivalenzrelation auf die vorliegende Relation übertragen und festgehalten, dass beispielsweise die Menge M dabei $(\mathbb{R} \times \mathbb{R}) \backslash \{(0,0)\}$ entspricht. Geben Sie auch für die Variablen a, b und c die Entsprechungen an! Also: Für $\sim$ reflexiv: $a = \ldots$, für $\sim$ symmetrisch: $a = \ldots$, $b = \ldots$, für $\sim$ transitiv: $a = \ldots$, $b = \ldots$, $c = \ldots$

2. Im Beweis der Transitivität wurde im Fall „$x_2 = 0$“ gefolgert, dass $x_1 = x_3 = 0$ gilt. Welche Überlegungen gehen dort im Detail ein? Stellen Sie die gleichen Überlegungen für den nicht ausgeführten Fall $y_2 = 0$ an!

3. Wäre $\sim$ auch eine Äquivalenzrelation, wenn man statt $(\mathbb{R} \times \mathbb{R}) \backslash \{(0,0)\}$ die Menge $\mathbb{R} \times \mathbb{R}$ zugrunde legen würde? Begründen Sie!

4. Falls in Ihrer Vorlesung Repräsentantensysteme behandelt wurden, geben Sie eines für die vorliegende Relation an!

Arbeitsauftrag 5.5 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.1.

6 Musterlösungen aus der Analysis 1

6.1 Supremum und Infimum

Aufgabe 6.1 Bestimmen Sie das Infimum und das Supremum der folgenden Menge:

$$M = \left\{ \frac{3x}{x+2} \mid x > 0 \right\}.$$

B

Ausführliche Lösung 6.1 Wir suchen die kleinste obere Schranke (Supremum) sowie die größte untere Schranke (Infimum) der Menge M. Hilfreich wäre es, wenn wir schon eine Idee hätten, welche Zahlen dafür in Frage kommen.

Z

Vorüberlegung zum Finden möglicher Kandidaten für $\sup M$ und $\inf M$: Für x dürfen wir in den Term $\frac{3x}{x+2}$ laut Definition von M nur positive Zahlen einsetzen. Machen wir das einmal für ein paar Werte: $\frac{3\cdot 1}{1+2} = 1$; $\frac{3\cdot 2}{2+2} = 1.5$; $\frac{3\cdot 0.01}{0.01+2} \approx 0.015$; $\frac{3\cdot 0.001}{0.001+2} \approx 0.0015$; $\frac{3\cdot 10}{10+2} = 2.5$; $\frac{3\cdot 100}{100+2} \approx 2.9$; $\frac{3\cdot 10000}{10000+2} \approx 2.999$; usw. Das sind also Beispiele für Zahlen, die in M liegen. Graphisch kann man diese Elemente veranschaulichen, indem man den Term $\frac{3x}{x+2}$ als Funktionsterm auffasst (siehe Abbildung 6.1).

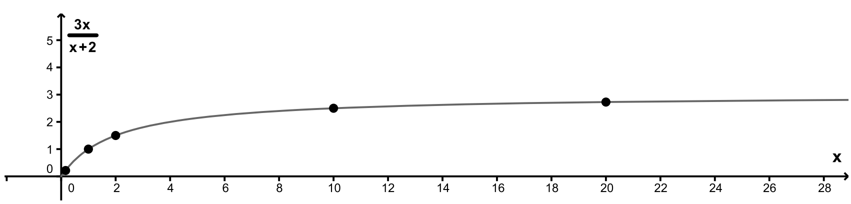

Abbildung 6.1: Der Term $\frac{3x}{x+2}$ aufgefasst als Funktionsterm

Welche Zahl wäre ein Kandidat für das Supremum? Wir suchen dabei die kleinste Zahl, die größer (oder gleich) als *alle* Zahlen in M ist. Nachdem wir bei unseren Beispielen keine Zahl herausbekommen haben, die größer als 3 ist und nachdem sich die Zahlen für größer werdendes x offenbar der Zahl 3 (von unten!) beliebig nähern, werden wir wohl 3 als Kandidaten für das Supremum versuchen.

C. Ableitinger, A. Herrmann, *Lernen aus Musterlösungen zur Analysis und Linearen Algebra*,
DOI 10.1007/978-3-658-02353-9_6, © Springer Fachmedien Wiesbaden 2013

Welche Zahl wäre ein Kandidat für das Infimum? Offenbar liegen in M nur positive Zahlen, d. h. jede negative Zahl scheint eine untere Schranke für M zu sein. Als größte untere Schranke kommt wohl nur die Zahl 0 in Frage[1].

Die beiden Vermutungen „3 ist das Supremum von M“ und „0 ist das Infimum von M“ müssen natürlich noch bewiesen werden. Es könnte ja sein, dass unsere Beispielauswahl von oben nicht sehr geschickt war!

Behauptung 1: 3 ist das Supremum von M.

Beweisstrategie: Behauptung 1 kann man in zwei Teilbehauptungen zerlegen, die wir separat beweisen werden:

1.1.) 3 ist eine obere Schranke von M (siehe auch Abbildung 6.2).

1.2.) 3 ist die *kleinste* obere Schranke von M.

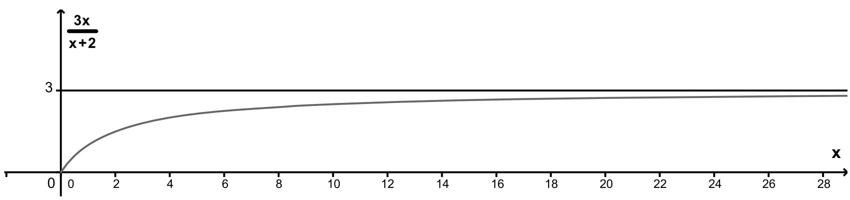

Abbildung 6.2: 3 ist eine obere Schranke von M.

Beweis von 1.1.):

$$\begin{array}{rrcll} & 3 & \geq & \frac{3x}{x+2} & \text{für alle } x > 0 \\ \Leftrightarrow & 3(x+2) & \geq & 3x & \text{für alle } x > 0 \\ \Leftrightarrow & 3x+6 & \geq & 3x & \text{für alle } x > 0 \\ \Leftrightarrow & 6 & \geq & 0 & \text{für alle } x > 0\text{, was trivialerweise erfüllt ist.} \end{array}$$

Beweis von 1.2.): Z. z.: 3 ist die *kleinste* obere Schranke. Aber wie zeigt man soetwas?

Dass 3 die *kleinste* obere Schranke ist bedeutet, dass es keine Zahl kleiner als 3 gibt, die auch noch obere Schranke ist.

Dann ist die Aussage „3 ist die kleinste obere Schranke“ aber gleichbedeutend damit, dass für alle $\varepsilon > 0$ die Zahl $3 - \varepsilon$ *keine* obere Schranke mehr ist, d. h. die Menge M dringt (mit mindestens einem ihrer Elemente) in die Lücke zwischen $3 - \varepsilon$ und 3 ein (siehe Abbildung 6.3).

Um das zu beweisen genügt es, für *ein* beliebiges[2], aber festes $\varepsilon_0 > 0$ *eine* Zahl x_0 zu

---

[1] Dass man sich bei solchen Vermutungen auch täuschen kann, zeigt die Menge $\left\{\frac{x+1}{1000x} \,|\, x > 0\right\}$, deren Infimum die Zahl 0.001 ist.

[2] *Ein beliebiges* bedeutet nicht *ein konkretes*. Es reicht also nicht, den Beweis z. B. nur für $\varepsilon_0 = 0.001$ zu führen! Beliebig meint also nicht, dass Sie sich irgendein konkretes ε_0 aussuchen dürfen, sondern dass der folgende Beweis für *jedes beliebige* $\varepsilon_0 > 0$ funktionieren muss.

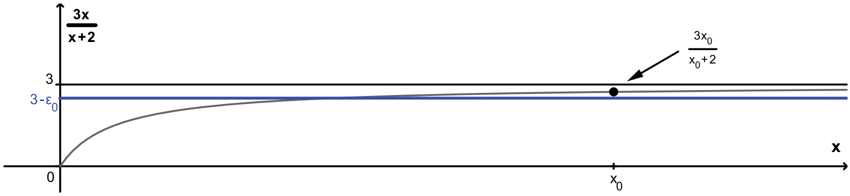

Abbildung 6.3: Für $\varepsilon_0 > 0$ ist $3 - \varepsilon_0$ keine obere Schranke von M.

finden, die uns ein Element in M liefert, das größer als $3 - \varepsilon_0$ ist[3]:

$$3 - \varepsilon_0 < \frac{3x_0}{x_0 + 2}.$$

Versuchen wir das einmal. Sei im Folgenden also $\varepsilon_0 > 0$ beliebig, aber fest.

Ein dazu passendes x_0 finden wir nach einigen Äquivalenzumformungen, mit deren Hilfe wir x_0 isolieren:

$$\begin{array}{rcl}
 & 3 - \varepsilon_0 & < & \frac{3x_0}{x_0+2} \\
\Leftrightarrow & (3 - \varepsilon_0)(x_0 + 2) & < & 3x_0 \\
\Leftrightarrow & 3x_0 - \varepsilon_0 x_0 + 6 - 2\varepsilon_0 & < & 3x_0 \\
\Leftrightarrow & 6 - 2\varepsilon_0 & < & \varepsilon_0 x_0 \\
\Leftrightarrow & \frac{6-2\varepsilon_0}{\varepsilon_0} & < & x_0.
\end{array}$$

Durch diese Umformungen haben wir x_0 durch eine Bedingung charakterisiert, die vom gewählten ε_0 abhängt. Bei der Suche nach einem passenden x_0 müssen also die beiden Bedingungen $x_0 > 0$ und $x_0 > \frac{6-2\varepsilon_0}{\varepsilon_0}$ erfüllt sein.

Je nach dem Wert von ε_0 ist eine der beiden Bedingungen schon automatisch durch die jeweils andere erfüllt. Wir führen also eine Fallunterscheidung nach ε_0 durch.

Fallunterscheidung:

1. Fall: $\varepsilon_0 > 3$. Dann ist $\frac{6-2\varepsilon_0}{\varepsilon_0}$ negativ und die Bedingung $x_0 > \frac{6-2\varepsilon_0}{\varepsilon_0}$ wird durch jede Zahl $x_0 > 0$ erfüllt. Wir können also z. B. $x_0 = 1$ wählen.

2. Fall: $\varepsilon_0 \leq 3$. Dann ist $\frac{6-2\varepsilon_0}{\varepsilon_0} \geq 0$. D. h. für $x_0 > \frac{6-2\varepsilon_0}{\varepsilon_0}$ ist die Bedingung $x_0 > 0$ automatisch erfüllt. Wir können also z. B. $x_0 = \frac{6-2\varepsilon_0}{\varepsilon_0} + 1$ wählen.

Fazit aus der Fallunterscheidung: Man kann die beiden Fälle wieder zusammenführen, indem man x_0 als das Maximum der beiden Zahlen 1 und $\frac{6-2\varepsilon_0}{\varepsilon_0} + 1$ definiert. Zusammenfassend kann man also sagen: Für jedes beliebige $\varepsilon_0 > 0$ erzeugt die Zahl

$$x_0 := \max\{1, \frac{6 - 2\varepsilon_0}{\varepsilon_0} + 1\}$$

eine Zahl $\frac{3x_0}{x_0+2}$, die größer als $3 - \varepsilon_0$ ist.

[3] Wenn einem diese Art von Beweis schon geläufig ist, kann man auch auf die Indizes bei ε_0 und x_0 verzichten. Sie sollen hier verdeutlichen, dass wir den Variablen ε und x nicht mehr ihren jeweiligen Laufbereich gestatten, sondern dass wir sie für den Moment fixieren.

Es sind beide Teilbehauptungen gezeigt und wir können folgern: 3 ist das Supremum von M.

P *Rückblickende Bemerkung:* Der Kernpunkt des Beweises von *1.2.)* war die Umformulierung der Aussage „3 ist die kleinste obere Schranke“ in die Aussage „Für alle $\varepsilon > 0$ ist die Zahl $3 - \varepsilon$ *keine* obere Schranke mehr“. Das hat es uns ermöglicht, den Beweis der Aussage auf das Umformen einer Ungleichung zu reduzieren. Wir haben gewissermaßen das Problem dem uns zur Verfügung stehenden Werkzeug zugänglich gemacht.

B *Behauptung 2:* 0 ist das Infimum von M.

K *Beweisstrategie:* Auch diese Behauptung kann man in zwei Teilbehauptungen zerlegen:

Z *2.1.)* 0 ist eine untere Schranke von M.

2.2.) 0 ist die *größte* untere Schranke von M.

H *Beweis von 2.1.)*:

$$\begin{array}{rrcll} & 0 & \leq & \frac{3x}{x+2} & \text{für alle } x > 0 \\ \Leftrightarrow & 0 & \leq & 3x & \text{für alle } x > 0 \\ \Leftrightarrow & 0 & \leq & x & \text{für alle } x > 0 \text{, was trivialerweise erfüllt ist.} \end{array}$$

B *Beweis von 2.2.)*: Z. z.: 0 ist die größte untere Schranke.

Z Das ist aber gleichbedeutend damit, dass für alle $\varepsilon > 0$ die Zahl $0 + \varepsilon$ *keine* untere Schranke mehr ist (siehe Abbildung 6.4).

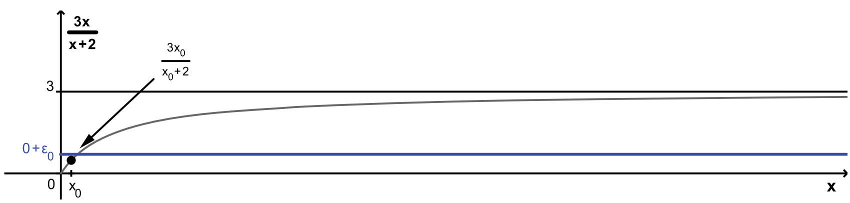

Abbildung 6.4: Für $\varepsilon_0 > 0$ ist $0 + \varepsilon_0$ keine untere Schranke von M.

K Es genügt also wieder zu zeigen, dass man zu einem beliebigen, aber festen $\varepsilon_0 > 0$ eine Zahl $x_0 > 0$ findet mit:

$$0 + \varepsilon_0 > \frac{3x_0}{x_0 + 2}.$$

Ein solches x_0 finden wir wie oben nach einigen Äquivalenzumformungen, mit deren Hilfe wir wieder x_0 isolieren:

H

$$\begin{array}{rccc} & 0 + \varepsilon_0 & > & \frac{3x_0}{x_0+2} \\ \Leftrightarrow & \varepsilon_0(x_0 + 2) & > & 3x_0 \\ \Leftrightarrow & \varepsilon_0 x_0 + 2\varepsilon_0 & > & 3x_0 \\ \Leftrightarrow & 2\varepsilon_0 & > & x_0(3 - \varepsilon_0). \end{array}$$

Wenn wir jetzt durch $(3 - \varepsilon_0)$ dividieren wollen, müssen wir eine Fallunterscheidung machen: **P**

1. Fall: $\varepsilon_0 < 3$. Dann ist die letzte Ungleichung äquivalent zu $x_0 < \frac{2\varepsilon_0}{3-\varepsilon_0}$. Wir können in diesem Fall $x_0 = \frac{\varepsilon_0}{3-\varepsilon_0}$ wählen und sind fertig. **H**

2. Fall: $\varepsilon_0 > 3$. Dann ist die letzte Ungleichung äquivalent zu $x_0 > \frac{2\varepsilon_0}{3-\varepsilon_0}$. Diese Ungleichung wird durch jedes $x_0 > 0$ erfüllt, da auf der rechten Seite eine negative Zahl steht. Auch in diesem Fall finden wir also ein passendes x_0 (z. B. $x_0 = 1$).

3. Fall: $\varepsilon_0 = 3$. Dann ist die letzte Ungleichung äquivalent zu $\varepsilon_0 > 0$, was ebenfalls durch jede Zahl $x_0 > 0$ erfüllt wird.

Es sind beide Teilbehauptungen gezeigt und wir können folgern: 0 ist das Infimum von M. **B**

Verständnisfragen 6.1

1. Hätte man die Aufgabe auch wie folgt lösen können? Wieso? Wieso nicht? $\lim\limits_{x\to+\infty} \dfrac{3x}{x+2} = 3$ und $\lim\limits_{x\searrow 0} \dfrac{3x}{x+2} = 0$. Also ist 3 das Supremum der Menge und 0 das Infimum.
2. Wieso wird im *Beweis von 1.2.)* ausgerechnet eine Fallunterscheidung nach $\varepsilon_0 > 3$ bzw. $\varepsilon_0 \leq 3$ gemacht?
3. Wieso wird im *Beweis von 2.2.)* im 1. Fall die Zahl x_0 gerade $x_0 = \frac{\varepsilon_0}{3-\varepsilon_0}$ gewählt?

Arbeitsauftrag 6.1 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.2.

6.2 Konvergenz von Folgen

Aufgabe 6.2 Untersuchen Sie die nachstehenden Folgen auf Konvergenz und bestimmen Sie gegebenenfalls ihren Grenzwert!

a) $a_n = \sum\limits_{k=1}^{n} \left(-\dfrac{1}{2}\right)^k$

b) $b_n = \dfrac{r_0 + r_1 n + \ldots + r_k n^k}{s_0 + s_1 n + \ldots + s_k n^k}$ für festes $k \in \mathbb{N}$, $k \geq 2$, gegebene r_i und $s_i \in \mathbb{R}$ mit $0 \leq i \leq k$, $s_k \neq 0$. Dabei sei der Nenner für alle $n \in \mathbb{N}$ von 0 verschieden.

c) $c_n = \dfrac{2 + \frac{1}{\sqrt{n}}}{\sqrt{n} + 5^{-n}}$, $n \in \mathbb{N}$

Ausführliche Lösung 6.2 Diese Aufgabe deckt viele Fertigkeiten ab, die Sie noch häufig in Ihrem Studium und auch später im Zusammenhang mit Folgengrenzwerten brauchen werden. Es geht darum zu erkennen, ob eine Folge konvergiert, welchen Grenzwert sie gegebenenfalls hat und wie man die Konvergenz beweist.

a) Diese Teilaufgabe ist in gewisser Hinsicht gefährlich. Man wird leicht dazu verleitet, falsch zu beginnen:

$$\sum_{k=1}^{n}\left(-\frac{1}{2}\right)^k = -\frac{1}{2}+\left(-\frac{1}{2}\right)^2+\left(-\frac{1}{2}\right)^3+\ldots+\left(-\frac{1}{2}\right)^{n-1}+\left(-\frac{1}{2}\right)^n.$$

Es würde hier *nicht* zum Ziel führen zu untersuchen, wogegen die einzelnen Summanden konvergieren und dann zu argumentieren, dass die Summe gegen die Summe der Einzelgrenzwerte konvergiert. Diese Argumentation ist für Summen mit unendlich vielen Summanden (mit wachsendem n wächst ja auch die Anzahl der Summanden in der Summe) nicht zulässig!

Vielmehr muss man hier ein geschultes Auge haben. Der Ausdruck $\sum_{k=1}^{n}\left(-\frac{1}{2}\right)^k$ ist nämlich eine endliche geometrische Summe:

$$\sum_{k=1}^{n}\left(-\frac{1}{2}\right)^k = \frac{1-\left(-\frac{1}{2}\right)^{n+1}}{1-\left(-\frac{1}{2}\right)}-1 = \frac{1-\left(-\frac{1}{2}\right)^{n+1}}{\frac{3}{2}}-1.$$

Vorsicht: In der üblichen Darstellung einer endlichen geometrischen Summe läuft der Summationsindex von 0 bis n, deshalb muss hier der Summand für $k=0$ (also die Zahl 1) abgezogen werden!

Dieser Ausdruck lässt sich jetzt viel leichter handhaben als die obige Summe. Wir können den Grenzübergang durchführen und beachten, dass $\lim_{n\to\infty}\left(-\frac{1}{2}\right)^{n+1}=0$ ist:

$$\lim_{n\to\infty}\frac{1-\left(-\frac{1}{2}\right)^{n+1}}{\frac{3}{2}}-1 = \frac{1-0}{\frac{3}{2}}-1 = \frac{2}{3}-1 = -\frac{1}{3}.$$

b) Dieser Aufgabenteil sieht ein wenig technisch und kompliziert aus – ist er aber eigentlich gar nicht. Sehen wir uns dazu ein einfaches Beispiel an:
Der Ausdruck $b_n = \frac{3-4n+2n^2}{-7+n+3n^2}$ ist genau von der in der Aufgabe angegebenen Form. (Machen Sie sich das bitte klar!) Was ist der Grenzwert der Folge b_n?
Dazu braucht man einen kleinen Trick. Wir dividieren den Zähler und den Nenner des Bruches durch die größte vorkommende Potenz von n. Das ist im vorliegenden Beispiel n^2. Wir erweitern also den Bruch mit der Zahl $\frac{1}{n^2}$, das ändert bekanntlich den Wert des Bruches nicht:

$$b_n = \frac{3-4n+2n^2}{-7+n+3n^2} = \frac{\frac{3}{n^2}-\frac{4n}{n^2}+\frac{2n^2}{n^2}}{-\frac{7}{n^2}+\frac{n}{n^2}+\frac{3n^2}{n^2}} = \frac{\frac{3}{n^2}-\frac{4}{n}+2}{-\frac{7}{n^2}+\frac{1}{n}+3}.$$

Jetzt können wir den Grenzübergang durchführen.

Jeweils die ersten beiden Terme in Zähler und Nenner sind Nullfolgen, was uns Folgendes liefert:

$$\lim_{n\to\infty} \frac{\frac{3}{n^2} - \frac{4}{n} + 2}{-\frac{7}{n^2} + \frac{1}{n} + 3} = \frac{0+0+2}{0+0+3} = \frac{2}{3}.$$

Als Ergebnis erhalten wir also genau jenen Bruch, der als Zähler und Nenner gerade die Koeffizienten von n^2 des Ausgangsbruches hat.

Dieselbe Idee können wir jetzt für den allgemeinen Bruch

$$\frac{r_0 + r_1 n + \ldots + r_k n^k}{s_0 + s_1 n + \ldots + s_k n^k}$$

verwenden. Wir erweitern dazu den Bruch mit $\frac{1}{n^k}$, da ja n^k die größte vorkommende Potenz von n ist:

$$\frac{r_0 + r_1 n + \ldots + r_k n^k}{s_0 + s_1 n + \ldots + s_k n^k} = \frac{\frac{r_0}{n^k} + \frac{r_1 n}{n^k} + \ldots + \frac{r_k n^k}{n^k}}{\frac{s_0}{n^k} + \frac{s_1 n}{n^k} + \ldots + \frac{s_k n^k}{n^k}} = \frac{\frac{r_0}{n^k} + \frac{r_1}{n^{k-1}} + \ldots + r_k}{\frac{s_0}{n^k} + \frac{s_1}{n^{k-1}} + \ldots + s_k}.$$

Jetzt stehen in Zähler und Nenner wieder jede Menge Nullfolgen, was uns folgendes Ergebnis liefert:

$$\lim_{n\to\infty} \frac{\frac{r_0}{n^k} + \frac{r_1}{n^{k-1}} + \ldots + r_k}{\frac{s_0}{n^k} + \frac{s_1}{n^{k-1}} + \ldots + s_k} = \frac{0+0+\ldots+r_k}{0+0+\ldots+s_k} = \frac{r_k}{s_k}.$$

Beachten Sie, dass $s_k \neq 0$ vorausgesetzt war. Beachten Sie außerdem, dass diese Umformungen auch für $r_k = 0$ gemacht werden können. Liegt Ihnen also ein Bruch dieser Gestalt vor, bei dem der Grad des Nenners größer oder gleich dem Grad des Zählers ist, so können Sie den Grenzwert sofort an den Koeffizienten der größten vorkommenden Potenz von n ablesen.

c) Bei diesem Aufgabenteil kann – gerade bei Anfängern – ein bestimmter Fehler häufig auftreten.

Formen wir zunächst den Ausdruck etwas um:

$$\frac{2 + \frac{1}{\sqrt{n}}}{\sqrt{n} + 5^{-n}} = \frac{2 + \frac{1}{\sqrt{n}}}{\sqrt{n} + \frac{1}{5^n}}.$$

Wozu man hier verleitet wird, ist, zuerst einmal die beiden Grenzwerte der Nullfolgen $\frac{1}{\sqrt{n}}$ und $\frac{1}{5^n}$ auszuwerten und *erst danach* den Grenzwert von $\frac{2+0}{\sqrt{n}+0}$ auszurechnen. Das darf man im Allgemeinen aber so nicht machen[4].

[4] In unserem Beispiel würden wir zwar zufälligerweise das richtige Ergebnis erhalten, der Grenzwert $\lim\limits_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e$ liefert aber ein Gegenbeispiel.

K

„Sicherer" ist es, zunächst zwei Abschätzungen (eine nach oben und eine nach unten) und ein paar Umformungen zu machen, damit wir anschließend das Einschließungskriterium anwenden können:

A

$$0 \leq \frac{2 + \frac{1}{\sqrt{n}}}{\sqrt{n} + \frac{1}{5^n}} \leq \frac{2 + \frac{1}{\sqrt{n}}}{\sqrt{n}} = \frac{\frac{2\sqrt{n}+1}{\sqrt{n}}}{\sqrt{n}} = \frac{2}{\sqrt{n}} + \frac{1}{n}.$$

Jetzt haben wir rechts eine Summe zweier Nullfolgen vorliegen, d. h.

$$\lim_{n\to\infty} \left(\frac{2}{\sqrt{n}} + \frac{1}{n}\right) = 0 + 0 = 0.$$

H

Aus dem Einschließungskriterium folgt, dass auch die Folge c_n gegen 0 konvergiert.

Verständnisfragen 6.2

1. Welchen Vorteil hat es, in Teil a) zunächst die endliche geometrische Summe zu berechnen?
2. Was würde in Teil b) passieren, wenn $s_k = 0$ und $r_k \neq 0$ wäre, wenn also der Grad des Zählers größer als der Grad des Nenners wäre?
3. Wieso wurde in Teil c) die Ungleichungskette mit dem Ausdruck $0 \leq \ldots$ begonnen?

Arbeitsauftrag 6.2 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.2.

6.3 Cauchyfolgen

Aufgabe 6.3 Zeigen Sie, dass die Folge $\{x_n\}_{n\in\mathbb{N}}$ mit

$$x_n = \frac{(-1)^n}{n^2}$$

eine Cauchyfolge ist, indem Sie direkt die Definition anwenden.

Ausführliche Lösung 6.3 Diese Aufgabe ist ein Standardbeispiel für einen Beweis, dass eine bestimmte Folge eine Cauchyfolge ist. Zunächst zur Wiederholung:

Definition Eine Zahlenfolge $\{a_n\}_{n\in\mathbb{N}}$ heißt Cauchyfolge, wenn gilt:

$$\forall\, \varepsilon > 0\ \exists\, n_0 \in \mathbb{N}: \ |a_n - a_m| < \varepsilon \text{ für alle } n,\ m > n_0.$$

Wir sollen jetzt zeigen, dass die Folge $x_n = \frac{(-1)^n}{n^2}$ eine Cauchyfolge ist. Dazu sollen wir direkt die Definition verwenden.

Wir müssen also zu jedem $\varepsilon > 0$ eine natürliche Zahl n_0 angeben können, sodass

$$|a_n - a_m| < \varepsilon \text{ für alle } n,\ m > n_0 \text{ gilt.}$$

Umformuliert heißt das: Größer als welche Zahl n_0 müssen wir n und m mindestens wählen, damit der Abstand zwischen den beiden Folgengliedern x_n und x_m kleiner als ε ist? Na gut, dann legen wir los:

Sei $\varepsilon > 0$ beliebig.

Wir starten jetzt einfach einmal bei der linken Seite der geforderten Ungleichung. Im Laufe unserer Arbeit werden wir dann (hoffentlich) an eine Stelle stoßen, an der wir ein geeignetes n_0 finden (denn das suchen wir ja!):

$$|x_n - x_m| = \left| \frac{(-1)^n}{n^2} - \frac{(-1)^m}{m^2} \right|.$$

An dieser Stelle stecken wir erstmals fest. Wir können daraus nicht sofort erkennen, wie groß wir n und m mindestens wählen müssen, damit der Ausdruck kleiner als ε wird.

Wir sollten schleunigst die Betragsstriche loswerden. Dazu machen wir eine kleine Umformung und wenden die Dreiecksungleichung an:

$$\begin{aligned} \left| \frac{(-1)^n}{n^2} - \frac{(-1)^m}{m^2} \right| &= \left| \frac{(-1)^n}{n^2} + \frac{(-1)^{m+1}}{m^2} \right| \leq \left| \frac{(-1)^n}{n^2} \right| + \left| \frac{(-1)^{m+1}}{m^2} \right| \\ &= \frac{1}{n^2} + \frac{1}{m^2} = \frac{m^2 + n^2}{n^2 m^2}. \end{aligned}$$

Jetzt stehen wir vor dem nächsten Problem. Wenn wir wüssten, welche der beiden Zahlen n und m größer ist, könnten wir weiter abschätzen – wir wissen es aber nicht.

An dieser Stelle können wir allerdings o.B.d.A. (ohne Beschränkung der Allgemeinheit) annehmen, dass $n \geq m$ gilt (falls es in Wirklichkeit gerade umgekehrt ist, kann man den Beweis analog führen – deshalb schränkt diese Annahme die Allgemeinheit des Beweises nicht ein). Damit gilt:

$$\frac{m^2 + n^2}{n^2 m^2} \leq \frac{2n^2}{n^2 m^2} = \frac{2}{m^2}.$$

Jetzt sind wir an einer Stelle, an der es sich schon gut arbeiten lässt. Wie groß müssen n und m mindestens sein, damit $\frac{2}{m^2} < \varepsilon$ ist? Umformen liefert:

$$m > \sqrt{\frac{2}{\varepsilon}}.$$

Jetzt wissen wir, wie wir n_0 wählen können, damit alles passt[5]: $n_0 := \left[\sqrt{\frac{2}{\varepsilon}}\right] + 1$.

Für $n \geq m > n_0$ gelten dann alle obigen Abschätzungen und damit $|x_n - x_m| < \varepsilon$.

[5] Die Gaußklammer [ ] rundet eine Zahl auf die nächstkleinere ganze Zahl ab, z.B.: $[3,82] = 3$. Die Wahl von $n_0 := \left[\sqrt{\frac{2}{\varepsilon}}\right] + 1$ sichert uns also, dass n_0 eine natürliche Zahl ist und dass diese größer als $\sqrt{\frac{2}{\varepsilon}}$ ist.

Verständnisfragen 6.3

1. Wieso darf man die Umformung

$$\left|\frac{(-1)^n}{n^2} - \frac{(-1)^m}{m^2}\right| = \left|\frac{(-1)^n}{n^2} + \frac{(-1)^{m+1}}{m^2}\right|$$

machen und wozu machen wir sie überhaupt?

2. Wieso ersetzt man im Ausdruck $\frac{m^2+n^2}{n^2m^2}$ nur im Zähler m^2 durch n^2, im Nenner aber nicht?

3. Könnten wir statt $n_0 := \left[\sqrt{\frac{2}{\varepsilon}}\right] + 1$ auch $n_0 := \left[\sqrt{\frac{2}{\varepsilon}}\right]$ wählen?

Übung 6.3 Ordnen Sie nun selbst den einzelnen Bearbeitungsschritten der Musterlösung die Phasen aus Kapitel 3 zu! Tragen Sie dazu die jeweiligen Buchstaben in die kleinen Quadrate am Seitenrand ein!

Arbeitsauftrag 6.3 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen, der Übung und dem Arbeitsauftrag finden Sie in Abschnitt 11.2.

6.4 Konvergenz von Reihen

Aufgabe 6.4 Untersuchen Sie, ob die Reihe konvergiert! Konvergiert sie auch absolut?

a) $\sum_{n=1}^{\infty} \left(\frac{n}{2n+1}\right)^n$

b) $\sum_{n=2}^{\infty} \frac{(-1)^n}{n-\sqrt{n}}$

K **Ausführliche Lösung 6.4** a) Welches der Konvergenzkriterien eignet sich für diese Aufgabe? Zur Auswahl stehen z. B.:

- Wurzelkriterium,
- Quotientenkriterium,
- Majorantenkriterium,
- Minorantenkriterium,
- Leibniz-Kriterium.

Wir könnten natürlich eines nach dem anderen probieren, um festzustellen, ob sich mit einem davon die Konvergenz bzw. Divergenz zeigen lässt.

Oder wir überlegen gezielt, welches der Kriterien am erfolgversprechendsten ist:

- Die Potenz n deutet darauf hin, dass das *Wurzelkriterium* hilfreich sein könnte (die n-te Wurzel und die n-te Potenz sind ja gerade zueinander inverse Operationen).
- Ein Bruch ist oftmals aber auch ein Indiz dafür, dass das *Quotientenkriterium* helfen könnte.
- Und obendrein hat die Reihe noch Ähnlichkeit mit einer geometrischen Reihe – sie ist ja von der Form $\sum\limits_{n=1}^{\infty} (\dots)^n$. Also lässt sich vielleicht auch eine geometrische Reihe $\sum\limits_{n=1}^{\infty} q^n$ als *konvergente Majorante* angeben.

Tatsächlich funktionieren prinzipiell alle drei Methoden, das Quotientenkriterium führt uns aber zu einer schwierigen Grenzwertberechnung[6]. B

1. Wurzelkriterium:

Beim Wurzelkriterium ist zu prüfen, ob

$$\limsup_{n\to\infty} \sqrt[n]{|\ a_n\ |} < 1$$

gilt.

Wir müssten also zeigen, dass der größte Häufungspunkt der Folge P

$$\{\sqrt[n]{|\ a_n\ |}\}_{n\in\mathbb{N}} = \left\{\sqrt[n]{\left|\left(\frac{n}{2n+1}\right)^n\right|}\right\}_{n\in\mathbb{N}}$$

kleiner als 1 ist.

Wie bestimmt man den größten Häufungspunkt einer Folge? Am einfachsten geht das, wenn die Folge konvergiert. Dann hat sie nämlich nur einen Häufungspunkt (der dann klarerweise auch der größte ist), und zwar den Grenzwert selbst. K

Durchführung: A

$$\sqrt[n]{\left|\left(\frac{n}{2n+1}\right)^n\right|} = \frac{n}{2n+1} = \frac{1}{2+\frac{1}{n}} \text{ und } \lim_{n\to\infty} \frac{1}{2+\frac{1}{n}} = \frac{1}{2}$$

Die Folge konvergiert also tatsächlich. Ihr größter Häufungspunkt ist demnach $\frac{1}{2}$. B

Nachdem dieser echt kleiner als 1 ist, können wir das Wurzelkriterium anwenden, woraus folgt: K

Die Reihe $\sum\limits_{n=1}^{\infty} \left(\frac{n}{2n+1}\right)^n$ konvergiert absolut (insbesondere konvergiert sie). H

[6] Die anderen Methoden funktionieren hier nicht bzw. sind hier nicht praktikabel. Das Minorantenkriterium kann man nur zum Nachweis der Divergenz einer Reihe verwenden und das Leibniz-Kriterium funktioniert nur bei alternierenden Reihen.

B *2. Quotientenkriterium:*

Beim Quotientenkriterium ist zuerst zu prüfen, ob $a_n \neq 0$ für alle $n \in \mathbb{N}$ gilt.

H Das ist aber offensichtlich erfüllt.

B Danach ist zu prüfen, ob $\limsup\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1$ gilt.

P Schon wieder müssen wir also den größten Häufungspunkt einer Folge bestimmen.

K Wir gehen wie oben vor, indem wir zunächst den Grenzwert $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|$ berechnen (wenn er überhaupt existiert und sich berechnen lässt).

A *Durchführung:*

$$\begin{aligned}
\left|\frac{a_{n+1}}{a_n}\right| &= \left|\frac{\left(\frac{n+1}{2(n+1)+1}\right)^{n+1}}{\left(\frac{n}{2n+1}\right)^n}\right| = \frac{\frac{(n+1)^{n+1}}{(2n+3)^{n+1}}}{\frac{n^n}{(2n+1)^n}} = \frac{(2n+1)^n}{(2n+3)^{n+1}} \frac{(n+1)^{n+1}}{n^n} \\
&= \left(\frac{(2n+1)(n+1)}{(2n+3)n}\right)^n \frac{n+1}{2n+3} = \left(\frac{2n^2+3n+1}{2n^2+3n}\right)^n \frac{n+1}{2n+3}
\end{aligned}$$

P Der Grenzwert dieses Ausdrucks ist leider nicht einfach zu berechnen und auch nicht leicht nach oben abzuschätzen. Sie können das ja mal versuchen.

K Wenn man bei einem Kriterium nun so feststeckt wie wir hier, dann ist es manchmal einfach am besten, man versucht ein anderes Kriterium. Vielleicht hat man damit ja mehr Glück!

3. Majorantenkriterium:

Beim Majorantenkriterium müssen wir – wie der Name schon sagt – eine konvergente *Majorante* finden. Die obige Idee war ja, eine geometrische Reihe als Majorante anzugeben.
Das klappt auch:

A
$$|a_n| = \left(\frac{n}{2n+1}\right)^n \leq \left(\frac{n}{2n}\right)^n = \left(\frac{1}{2}\right)^n.$$

H Nachdem die Reihe $\sum\limits_{n=1}^{\infty} \left(\frac{1}{2}\right)^n$ konvergiert, konvergiert die Reihe $\sum\limits_{n=1}^{\infty} \left(\frac{n}{2n+1}\right)^n$ also nach dem Majorantenkriterium absolut.

K b) Selbstverständlich kann man nun wieder versuchen, eines der oben benutzten Konvergenzkriterien anzuwenden. Sie hätten den Vorteil, dass man daraus sofort die *absolute* Konvergenz folgern könnte (woraus natürlich auch die „gewöhnliche“ Konvergenz sofort folgt).

Sie können gerne versuchen, die Aufgabe mit Quotienten- oder Wurzelkriterium zu lösen. Leider liefern aber beide Methoden als Limes superior die Zahl 1. Da weder Wurzel- noch Quotientenkriterium in diesem Fall eine Aussage über die Konvergenz bzw. Divergenz der Reihe machen, führt das zu nichts Verwertbarem.

Ein Konvergenzkriterium, das wir bis jetzt noch gar nicht verwendet haben, ist das Leibniz-Kriterium . Dieses Kriterium sagt allerdings nur etwas über die Konvergenz und nicht über die absolute Konvergenz einer Reihe. Nachdem aber die obigen Methoden keinen Erfolg v rsprechen, wären wir zumindest einmal damit zufrieden, wenn wir durch das Leibniz-Kriterium die Konvergenz eigen könnten.

Durchführung: B

Beim Leibniz-Kriterium sind zwei Dinge nachzuweisen: Erstens muss es sich um eine alternierende Reihe handeln.

Nachdem der Nenner von $\frac{(-1)^n}{n-\sqrt{n}}$ für $n \geq 2$ immer positiv und der Zähler abwechselnd positiv und negativ ist, ist diese Bedingung schon einmal erfüllt. A

Zweitens ist nachzuweisen, dass $\{|a_n|\}_{n\in\mathbb{N}} = \left\{\left|\frac{(-1)^n}{n-\sqrt{n}}\right|\right\}_{n\in\mathbb{N}}$ eine *monoton fallende* Nullfolge ist. B

Wir zeigen: $\{|a_n|\}_{n\in\mathbb{N}}$ ist monoton fallend. Zu zeigen ist also $|a_{n+1}| \leq |a_n|$ für alle $n \geq 2$. A

$$\begin{array}{ccc}
 & |a_{n+1}| & \leq & |a_n| \\
\Leftrightarrow & \dfrac{1}{n+1-\sqrt{n+1}} & \leq & \dfrac{1}{n-\sqrt{n}} \\
\Leftrightarrow & n-\sqrt{n} & \leq & n+1-\sqrt{n+1} \\
\Leftrightarrow & -\sqrt{n} & \leq & 1-\sqrt{n+1} \\
\Leftrightarrow & \sqrt{n+1} & \leq & 1+\sqrt{n} \\
\overset{n\in\mathbb{N}}{\Leftrightarrow} & n+1 & \leq & n+1+2\sqrt{n} \\
\Leftrightarrow & 0 & \leq & \sqrt{n}
\end{array}$$

Die letzte Ungleichung stimmt offensichtlich für alle $n \geq 2$. H

Aus der Äquivalenz der einzelnen Zeilen folgt dann also auch, dass $|a_{n+1}| \leq |a_n|$ gilt, was wir ja zeigen wollten. B

Wir zeigen: $\{|a_n|\}_{n\in\mathbb{N}}$ ist eine Nullfolge.

$$\lim_{n\to\infty} |a_n| = \lim_{n\to\infty} \frac{1}{n-\sqrt{n}} = \lim_{n\to\infty} \frac{\frac{1}{n}}{1-\frac{1}{\sqrt{n}}} = 0$$ A

Damit sind alle Voraussetzungen des Leibniz-Kriteriums erfüllt. Es folgt also, dass die Reihe $\sum_{n=2}^{\infty} \frac{(-1)^n}{n-\sqrt{n}}$ konvergiert. H

Absolute Konvergenz? B

Absolute Konvergenz der Reihe $\sum_{n=2}^{\infty} \frac{(-1)^n}{n-\sqrt{n}}$ bedeutet, dass die Reihe

$$\sum_{n=2}^{\infty} \left|\frac{(-1)^n}{n-\sqrt{n}}\right| = \sum_{n=2}^{\infty} \frac{1}{n-\sqrt{n}}$$

konvergiert. Das sieht aber eher schlecht aus. Warum?

P

Die einzelnen Summanden sind ja größer als jene der harmonischen Reihe.

K

Und genau das nutzen wir mit Hilfe des Minorantenkriteriums aus. Dazu zeigen wir, dass die harmonische Reihe eine divergente Minorante der zu untersuchenden Reihe ist.

H

Es gilt: $|a_n| = \frac{1}{n-\sqrt{n}} \geq \frac{1}{n}$.

P

Wir machen ja durch das Weglassen von $-\sqrt{n}$ den Nenner größer, was insgesamt den Bruch verkleinert.

H

Die harmonische Reihe $\sum\limits_{n=1}^{\infty} \frac{1}{n}$ divergiert bekanntlich, und damit auch $\sum\limits_{n=2}^{\infty} \frac{1}{n}$ (wir lassen ja nur den ersten Summanden weg).

B

Das erlaubt uns jetzt, das Minorantenkriterium anzuwenden: Die Reihe $\sum\limits_{n=2}^{\infty} \frac{(-1)^n}{n-\sqrt{n}}$ ist also nicht absolut konvergent.

Verständnisfragen 6.4

1. Muss man die Konvergenz einer Reihe mit mehreren Konvergenzkriterien nachweisen, so wie wir das in Teil a) gemacht haben?
2. Wieso ist der Grenzwert von $\left(\frac{2n^2+3n+1}{2n^2+3n}\right)^n \cdot \frac{n+1}{2n+3}$ mit $n \to \infty$ nicht einfach zu berechnen?
3. Wieso muss man beim Leibniz-Kriterium auch noch zeigen, dass die Folge eine Nullfolge ist? Ist nicht jede monoton fallende Folge eine Nullfolge?

Arbeitsauftrag 6.4 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.2.

6.5 Folgenstetigkeit

Aufgabe 6.5 Bestimmen Sie alle reellen Zahlen x_0, in denen die Funktion $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = [x](1-x)$ stetig ist.

B

Ausführliche Lösung 6.5 Es ist herauszufinden, an welchen Stellen die Funktion f stetig ist.

K

Um eine Idee dafür zu bekommen, welche Stellen überhaupt in Frage kommen, lohnt es sich, eine Skizze des Funktionsgraphen anzufertigen. Und das macht man so:

Z

Nachdem im Funktionsterm der Ausdruck $[x]$ vorkommt, sollten wir uns zunächst einmal klar machen, wie $[x]$ überhaupt definiert ist. Die eckige Klammer (auch Gaußklammer genannt) bewirkt, dass die Zahl innerhalb der Klammer auf die nächstliegende ganze Zahl abgerundet wird (z. B. gilt $[3{,}14] = 3$, $[4{,}8] = 4$ bzw. $[5] = 5$).

Betrachten wir zunächst immer nur Abschnitte auf der x-Achse, innerhalb derer $[x]$ denselben Wert liefert:

$$\begin{array}{llll}
\text{Für} & -2 \leq x < -1 & \text{gilt:} & f(x) = -2(1-x) = 2x - 2. \\
\text{Für} & -1 \leq x < 0 & \text{gilt:} & f(x) = -1(1-x) = x - 1. \\
\text{Für} & 0 \leq x < 1 & \text{gilt:} & f(x) = 0(1-x) = 0. \\
\text{Für} & 1 \leq x < 2 & \text{gilt:} & f(x) = 1(1-x) = 1 - x. \\
\text{Für} & 2 \leq x < 3 & \text{gilt:} & f(x) = 2(1-x) = 2 - 2x.
\end{array}$$

Der Graph hat dann die Gestalt aus Abbildung 6.5.

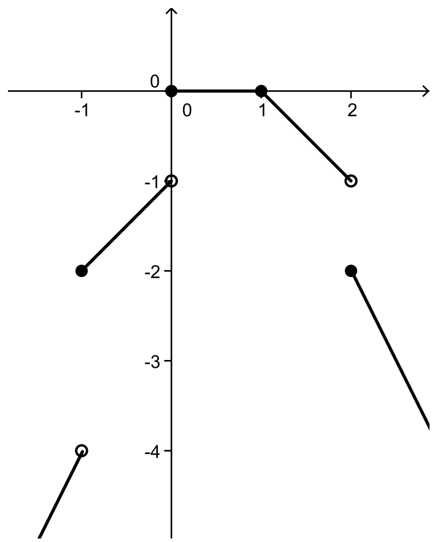

Abbildung 6.5: Graph der Funktion f aus Aufgabe 6.5

Es sieht also so aus, als wäre f an allen nichtganzzahligen Stellen und zusätzlich an der Stelle 1 stetig, an allen anderen ganzzahligen Stellen jedoch nicht.
Das beweisen wir nun.

Beginnen wir mit den Stellen $x_0 \notin \mathbb{Z}$.

Nachdem sowohl $x \mapsto [x]$ als auch $x \mapsto (1-x)$ stetig in x_0 ist, folgt: f ist stetig in x_0, weil das Produkt stetiger Funktionen stetig ist.

Jetzt kommt der schwierigere Teil: Sei also jetzt $x_0 \in \mathbb{Z}$.

Wir werden nun eine Folge $\{x_n\}_{n\in\mathbb{N}}$ mit Grenzwert x_0 angeben, für die die Stetigkeitsbedingung

$$\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n)$$

verletzt wird. Welche Folge eignet sich dafür?

Werfen wir einen Blick auf den Funktionsgraphen! Betrachten wir z. B. die Stelle 2. Wenn wir eine Folge $\{x_n\}_{n\in\mathbb{N}}$ angeben können, die sich von links dem Wert 2 auf der x-Achse nähert, dann könnte es klappen. Alle diese Folgenglieder, wenn sie nahe genug am Wert 2 sind, liefern einen Funktionswert nahe von -1. Die linke Seite der Stetigkeitsbedingung wird also vermutlich den Wert -1 annehmen. An der Stelle 2 selbst nimmt die Funktion aber den Wert -2 an – die rechte Seite der Stetigkeitsbedingung wird also den Wert -2 annehmen. Und genau das wollen wir erreichen![7]

K

So, jetzt aber formal und allgemein: Wir konstruieren zu $x_0 \in \mathbb{Z}$ eine geeignete Folge, die sich von links der Zahl x_0 nähert. Versuchen wir die Folge $\{x_0 - \frac{1}{n}\}_{n\in\mathbb{N}}$ und setzen zuerst in die linke Seite der Stetigkeitsbedingung ein:

H

$$\begin{aligned}\lim_{n\to\infty} f(x_0 - \frac{1}{n}) &= \lim_{n\to\infty} \left[x_0 - \frac{1}{n}\right](1-(x_0 - \frac{1}{n})) = \lim_{n\to\infty}(x_0-1)(1-x_0+\frac{1}{n}) \\ &= (x_0-1)(1-x_0).\end{aligned}$$

Die rechte Seite liefert:

$$f(\lim_{n\to\infty}(x_0 - \frac{1}{n})) = f(x_0) = [x_0](1-x_0) = x_0(1-x_0).$$

Jetzt vergleichen wir die linke und die rechte Seite: $(x_0-1)(1-x_0) = x_0(1-x_0)$. Wir stellen fest, dass diese Gleichung nur dann gilt, wenn $(1-x_0) = 0$ ist (ansonsten könnte man ja durch $(1-x_0)$ dividieren und würde dann einen Widerspruch erhalten).

B

Sie gilt also nur für $x_0 = 1$. Alle anderen ganzzahligen Werte für x_0 verletzen die Stetigkeitsbedingung und liefern damit den Beweis, dass f dort nicht stetig ist.

P

Vorsicht! Damit ist noch nicht gezeigt, dass f in $x_0 = 1$ stetig ist. Dazu müssen wir ja beweisen, dass die Stetigkeitsbedingung für *alle* Folgen, die gegen $x_0 = 1$ konvergieren, erfüllt ist (oben haben wir ja nur *eine spezielle* betrachtet)! Und das machen wir jetzt.

Z

Sei also $\{x_n\}_{n\in\mathbb{N}}$ eine beliebige Folge mit $\lim\limits_{n\to\infty} x_n = 1$.

K

Wir werden zeigen, dass dann die Stetigkeitsbedingung

$$\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n)$$

erfüllt ist.

H

Die rechte Seite ist leicht:

$$f(\lim_{n\to\infty} x_n) = f(1) = [1](1-1) = 0.$$

B

Es bleibt also zu zeigen, dass auch die linke Seite

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty}[x_n](1-x_n)$$

gleich 0 ist.

K

Wir werden versuchen, $|[x_n](1-x_n)|$ nach unten und oben gegen geeignete Nullfolgen abzuschätzen, um dann das Einschließungskriterium anwenden zu können.

[7] Hätten wir eine Folge betrachtet, die sich von rechts der Zahl 2 nähert, dann wären wir damit erfolglos geblieben. Dann würde die Stetigkeitsbedingung auf beiden Seiten den Wert -2 liefern.

Nach unten können wir einfach die konstante Nullfolge verwenden. Für die Abschätzung nach oben müssen wir uns überlegen, welche Werte $[x_n]$ annehmen kann: P

Laut Grenzwertdefinition existiert eine Zahl $n_0 \in \mathbb{N}$, sodass für alle $n \geq n_0$ gilt, dass $x_n \in (0, 2)$ liegt. Für diese Zahlen n gilt also: $[x_n] \in \{0, 1\}$. T

Jetzt können wir A

$$0 \leq |[x_n](1 - x_n)| = |[x_n]||1 - x_n| \leq |1 - x_n|$$

schreiben. Die rechte Seite konvergiert laut Voraussetzung gegen 0 und damit auch wie gewünscht der zweite Ausdruck in der Ungleichungskette.

Die Stetigkeitsbedingung ist also für jede beliebige gegen 1 konvergente Folge $\{x_n\}_{n \in \mathbb{N}}$ erfüllt und damit ist f in $x_0 = 1$ stetig! B

Verständnisfragen 6.5

1. Inwiefern hatten wir bei der Anfertigung der Skizze Glück?
2. Wo verwenden wir die Gleichung $\left[x_0 - \frac{1}{n}\right] = x_0 - 1$ und warum gilt sie?
3. Wo genau verwenden wir im Fall $x_0 = 1$, dass $[x_n] \in \{0, 1\}$ für $n \geq n_0$ gilt?

Arbeitsauftrag 6.5 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.2.

6.6 Stetigkeit mit Epsilon und Delta

Aufgabe 6.6 Zeigen Sie mit einem direkten ε-δ-Beweis, dass die Abbildung $x \mapsto \dfrac{1}{x^2 + 1}$ stetig auf $\mathbb{R}$ ist.

Ausführliche Lösung 6.6 Diese Aufgabe ist eine *ganz typische Aufgabe* zum Nachweis der Stetigkeit einer Funktion mit Hilfe der ε-δ-Definition. Sie sollten Aufgaben dieser Art *unbedingt* lösen können.

Wir sollen also die Stetigkeit der Funktion f mit dem Funktionsterm $f(x) := \frac{1}{x^2+1}$ auf ganz $\mathbb{R}$ nachweisen.

Seien also $x_0 \in \mathbb{R}$ und $\varepsilon > 0$ beliebig, aber fest gewählt. Wir suchen eine Zahl $\delta > 0$, sodass für alle $x \in \mathbb{R}$ mit $|x - x_0| < \delta$ gilt:

$$|f(x) - f(x_0)| < \varepsilon.$$

Wir beginnen mit:

$$|f(x) - f(x_0)| = \left| \frac{1}{x^2 + 1} - \frac{1}{x_0^2 + 1} \right| = \left| \frac{x_0^2 + 1 - (x^2 + 1)}{(x^2 + 1)(x_0^2 + 1)} \right| = \ldots$$

Unser Ziel: Wir möchten gerne (das variable) x loswerden und den ganzen Ausdruck gegen eine Konstante (nämlich ε) nach oben abschätzen. Ein paar Schritte ergeben sich dabei recht natürlich (Nenner verkleinern, Dreiecksungleichung, $|x - x_0| < \delta$ ausnutzen):

$$\begin{aligned} ... &= \left| \frac{x_0^2 - x^2}{(x^2+1)(x_0^2+1)} \right| \le \left| \frac{x_0^2 - x^2}{x_0^2+1} \right| = \frac{|x_0 - x| \cdot |x_0 + x|}{x_0^2+1} \\ &\le \frac{|x - x_0| \cdot (|x_0| + |x|)}{x_0^2+1} < \delta \cdot \frac{|x_0| + |x|}{x_0^2+1} = ... \end{aligned}$$

Jetzt kommt der schwierigste Teil: Wir möchten nun noch $|x|$ nach oben abschätzen (salopp gesprochen durch irgendeinen Ausdruck, der nur noch δ und/oder x_0 enthält).

Es bietet sich an, mit $|x - x_0| < \delta$ zu starten. Jetzt möchten wir x isolieren, dazu müssen wir den Betrag auflösen. Es wäre gut, wenn wir $|x - x_0|$ nach *unten* abschätzen könnten (nach oben bringt uns in dieser Situation leider nichts, die Dreiecksungleichung hilft hier also nicht). Dazu können wir aber die *umgekehrte Dreiecksungleichung* benutzen:

$$\begin{aligned} & ||x| - |x_0|| \le |x - x_0| < \delta \\ \Rightarrow \quad & -\delta \le |x| - |x_0| \le \delta \\ \Rightarrow \quad & |x| \le \delta + |x_0|. \end{aligned}$$

Damit ist der schwierigste Teil geschafft.

$$... \le \delta \cdot \frac{|x_0| + \delta + |x_0|}{x_0^2+1} = \frac{\delta^2 + 2\delta|x_0|}{x_0^2+1} = ...$$

Jetzt möchten wir gerne δ ausklammern, damit wir gut sehen können, wie wir δ geschickterweise wählen sollen.

Dazu legen wir vorerst $\delta \le 1$ fest. Damit gilt nämlich $\delta^2 \le \delta$:

$$... \le \frac{\delta + 2\delta|x_0|}{x_0^2+1} = \delta \cdot \frac{1 + 2|x_0|}{x_0^2+1} = ...$$

Jetzt sieht man, welche weitere Bedingung wir an δ stellen müssen:

$$\delta \le \frac{\varepsilon}{\frac{1+2|x_0|}{x_0^2+1}}.$$

Damit gilt:

$$... \le \frac{\varepsilon}{\frac{1+2|x_0|}{x_0^2+1}} \cdot \frac{1 + 2|x_0|}{x_0^2+1} = \varepsilon.$$

Wähle also

$$\delta := \min\{1, \frac{\varepsilon}{\frac{1+2|x_0|}{x_0^2+1}}\}.$$

Verständnisfragen 6.6

1. Wieso „bietet es sich an", beim Abschätzen von $|x|$ mit der Ungleichung $|x - x_0| < \delta$ zu starten?
2. Wieso kann zum Abschätzen von $|x|$ nicht die (gewöhnliche) Dreiecksungleichung verwendet werden?
3. Wieso dürfen wir einfach festlegen, dass $\delta \leq 1$ gelten soll?
4. Wieso wird δ als Minimum von 1 und $\frac{\varepsilon}{\frac{1+2|x_0|}{x_0^2+1}}$ gewählt?

Übung 6.6 Ordnen Sie nun selbst den einzelnen Bearbeitungsschritten der Musterlösung die Phasen aus Kapitel 3 zu! Tragen Sie dazu die jeweiligen Buchstaben in die kleinen Quadrate am Seitenrand ein!

Arbeitsauftrag 6.6 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen, der Übung und dem Arbeitsauftrag finden Sie in Abschnitt 11.2.

6.7 Gleichmäßige Stetigkeit und Lipschitz-Stetigkeit

Aufgabe 6.7 Es sei schon gezeigt, dass die Funktion $f : [0, \infty) \to \mathbb{R}$ mit $f(x) = \sqrt{x}$ auf ihrem Definitionsbereich stetig ist. Untersuchen Sie, ob f auf dem Intervall $[0, 1)$

a) gleichmäßig stetig,

b) Lipschitz-stetig ist.

Ausführliche Lösung 6.7 Wir sollen prüfen, ob die Funktion $f : [0, 1) \to \mathbb{R}$ mit $f(x) = \sqrt{x}$ gleichmäßig bzw. Lipschitz-stetig ist. B

Ein Satz, der in jeder Analysisvorlesung vorkommt, hilft uns schon zur Beantwortung der Frage nach der gleichmäßigen Stetigkeit. Er sagt aus, dass stetige Funktionen auf kompakten Intervallen sogar gleichmäßig stetig sind. K

a) Wäre die Funktion f auf dem abgeschlossenen Intervall $[0, 1]$ zu untersuchen, dann wären wir schon fertig. Dann hätten wir nämlich eine stetige Funktion auf einem kompakten Intervall vorliegen. Und diese ist nach dem oben formulierten Satz gleichmäßig stetig auf dem kompakten Intervall. P

Die Funktion ist allerdings auf $[0, 1)$ zu untersuchen und dieses Intervall ist *nicht* kompakt.

H Wenn wir allerdings schon wissen, dass f auf $[0,1]$ gleichmäßig stetig ist, dann muss f erst recht auf $[0,1)$ gleichmäßig stetig sein (was nämlich für alle Elemente $x, y \in [0,1]$ gilt, muss erst recht für alle Elemente $x, y \in [0,1)$ gelten).

B Damit sind wir fertig.

Bemerkung: Übrigens ist f sogar auf ganz $\mathbb{R}$ gleichmäßig stetig. Einen Beweis führen wir an dieser Stelle nicht.

K b) Wir beginnen jetzt einfach einmal so, als wollten wir die Lipschitz-Stetigkeit von f beweisen (was allerdings schlussendlich fehlschlagen wird).

B Zu zeigen ist dann, dass es eine konstante Zahl $L > 0$ gibt mit:

$$|\sqrt{x} - \sqrt{y}| \leq L \cdot |x - y|$$

für alle $x, y \in [0,1)$.

K Wir sollten also versuchen, den Ausdruck $|\sqrt{x} - \sqrt{y}|$ so umzuformen, dass darin der Ausdruck $|x - y|$ zum Vorschein kommt. Vielleicht können wir ja dann eine passende Zahl L ablesen:

A

$$|\sqrt{x} - \sqrt{y}| = \frac{|x-y|}{\sqrt{x} + \sqrt{y}} = \frac{1}{\sqrt{x} + \sqrt{y}} \cdot |x - y|.$$

P Diese Umformung dürfen wir nur dann machen, wenn x und y nicht gleichzeitig Null sind.

Z Also machen wir sicherheitshalber eine Fallunterscheidung:

H *1. Fall:* $x = 0$ und $y = 0$. Dann ist aber $|\sqrt{x} - \sqrt{y}| \leq L \cdot |x - y|$ für jede Zahl L erfüllt. Das bereitet also keine Schwierigkeiten.

2. Fall: $(x, y) \neq (0,0)$. Dann gilt wie oben gezeigt:

$$|\sqrt{x} - \sqrt{y}| = \frac{1}{\sqrt{x} + \sqrt{y}} \cdot |x - y|.$$

B Unser Ziel wäre jetzt, den Ausdruck $\frac{1}{\sqrt{x}+\sqrt{y}}$ nach oben gegen eine Konstante L abzuschätzen. Und zwar für jede beliebige Belegung von $x, y \in [0,1)$.

P Problematisch wird das offensichtlich, wenn sich x und y der Null nähern. Dann „explodiert" der Ausdruck $\frac{1}{\sqrt{x}+\sqrt{y}}$ nämlich nach $+\infty$.

Z Wie schreiben wir das auf? *Eine* der beiden Variablen können wir gleich Null setzen (unser 2. Fall verbietet ja nur, dass *beide* Variablen *gleichzeitig* Null sind). Sei also $y = 0$.

H Dann gilt:

$$\frac{1}{\sqrt{x} + \sqrt{y}} = \frac{1}{\sqrt{x}}.$$

K Nun können wir die zweite Variable gegen Null laufen lassen:

H

$$\lim_{x \searrow 0} \frac{1}{\sqrt{x}} = +\infty.$$

Das heißt, wir können keine Abschätzung des Ausdrucks $\frac{1}{\sqrt{x}+\sqrt{y}}$ nach oben durch eine Konstante L finden, die jeder Belegung von $x, y \in [0, 1)$ genügt. f ist also nicht Lipschitz-stetig.

Diese Aufgabe zeigt, dass eine Funktion auf einem Intervall gleichmäßig stetig sein kann, ohne dort auch Lipschitz-stetig zu sein. P

Verständnisfragen 6.7

1. Bevor Sie die Lösung der Aufgabe gekannt haben: Welche Stelle im Intervall $[0, 1)$ könnte denn – wenn überhaupt – am ehesten Schwierigkeiten in Bezug auf die gleichmäßige Stetigkeit der Funktion f bereiten? Begründen Sie!
2. Was passiert bei der Umformung $|\sqrt{x} - \sqrt{y}| = \dfrac{|x-y|}{\sqrt{x}+\sqrt{y}}$? Erklären Sie detailliert, wie der erste Ausdruck in den zweiten übergeführt werden kann!
3. Wieso genügt es, im 2. Fall nur für $y = 0$ nachzuweisen, dass der Ausdruck $\frac{1}{\sqrt{x}+\sqrt{y}}$ explodiert? Müsste man das nicht auch noch für alle anderen Werte von y überprüfen?

Arbeitsauftrag 6.7 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.2.

6.8 Differenzierbarkeit

Aufgabe 6.8 Bestimmen Sie mit Hilfe der Definition der Differenzierbarkeit alle Punkte $x_0 \in \mathbb{R}$, in denen die folgenden Funktionen differenzierbar sind!

a) $f(x) = x^2$

b) $g(x) = \begin{cases} x & x > 0 \\ x^2 & x \leq 0 \end{cases}$

c) $l(x) = x \cdot |x|$

Ausführliche Lösung 6.8 Das Berechnen von Ableitungen mit Hilfe der Definition der Differenzierbarkeit ist eine der Standardaufgaben, die Sie jedenfalls in Ihrem Studium lösen lernen sollen. P

a) Wir sehen uns zunächst die Definition der Differenzierbarkeit an. Sei M offen. $f : M \to \mathbb{R}$ heißt in $x_0 \in M$ differenzierbar, wenn der Grenzwert B

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert.

P Bitte beachten Sie, dass es sich bei dem Grenzwert um den *Grenzwertbegriff für Funktionen* handelt! $\lim\limits_{x \to x_0} g(x) = a$ für irgendeine Funktion g bedeutet demnach, dass $\lim\limits_{n \to \infty} g(x_n) = a$ für *alle* Folgen $\{x_n\}_{n \in \mathbb{N}}$ mit $\lim\limits_{n \to \infty} x_n = x_0$ gelten muss (*). Bei uns ist $g(x)$ also gerade der Differenzenquotient $\frac{f(x)-f(x_0)}{x-x_0}$.

Z Wir schreiben den Differentialquotienten zunächst etwas um, indem wir $x := x_0 + h$ setzen:

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Damit lässt es sich nämlich etwas besser arbeiten – das können Sie sich auch für andere Aufgaben zum Nachweis der Differenzierbarkeit merken.

B Nun aber zur Funktion f. Der zugehörige Differenzenquotient lautet:

H

$$\begin{aligned} \frac{f(x_0 + h) - f(x_0)}{h} &= \frac{(x_0 + h)^2 - x_0^2}{h} = \frac{x_0^2 + 2x_0 h + h^2 - x_0^2}{h} \\ &= \frac{2x_0 h + h^2}{h} = 2x_0 + h. \end{aligned}$$

B Wir sollen davon den Grenzwert für $h \to 0$ berechnen:

H

$$\lim_{h \to 0} (2x_0 + h) = 2x_0.$$

B Damit haben wir die Aufgabe gelöst. Für alle $x_0 \in \mathbb{R}$ gilt also: $f'(x_0) = 2x_0$ (dieses Resultat haben Sie in Ihrer Schulzeit bestimmt schon oft verwendet).

B b) Wir haben die Funktion $g(x) = \begin{cases} x & x > 0 \\ x^2 & x \leq 0 \end{cases}$ zu untersuchen.

H Klarerweise ist diese Funktion an allen Stellen $x_0 \neq 0$ differenzierbar, da auch die Funktionen $x \mapsto x$ und $x \mapsto x^2$ differenzierbar sind (für $x \mapsto x^2$ haben wir das gerade bewiesen, für $x \mapsto x$ können Sie sich das als Übung selbst überlegen).

P Nun also zur heiklen „Klebestelle" $x_0 = 0$:

H An der Stelle $x_0 = 0$ liefert der Differenzenquotient:

$$\frac{g(0 + h) - g(0)}{h} = \begin{cases} 1 & h > 0 \\ h & h < 0 \end{cases}.$$

P Es lässt sich hier schon vermuten, dass g an der Stelle 0 nicht differenzierbar ist. Wenn man sich nämlich im Differenzenquotienten mit h von links der Zahl 0 nähert, wird wohl die Zahl 0 herauskommen, nähert man sich mit h von rechts der 0, so liefert das aber 1.

K Das müssen wir natürlich formal beweisen: Wir untersuchen dazu den Grenzwert

$$\lim_{h \to 0} \frac{g(h) - g(0)}{h}$$

„entlang“ zweier unterschiedlicher Folgen $h_n := \frac{1}{n}$ und $\tilde{h}_n := -\frac{1}{n}$, die aber beide gegen 0 (die eine von rechts, die andere von links) konvergieren. Wenn g differenzierbar in 0 wäre, müsste nach (*) beide Male dasselbe Ergebnis herauskommen, tut es aber nicht:

$$\lim_{n\to\infty} \frac{g(h_n) - g(0)}{h_n} = 1, \text{ da } h_n > 0 \text{ für alle } n \text{ gilt.}$$

$$\lim_{n\to\infty} \frac{g(\tilde{h}_n) - g(0)}{\tilde{h}_n} = \lim_{n\to\infty} \tilde{h}_n = \lim_{n\to\infty} -\frac{1}{n} = 0, \text{ da } \tilde{h}_n < 0 \text{ für alle } n \text{ gilt.}$$

Hinweis: Man kann hier alternativ mit dem rechts- bzw. linksseitigen Grenzwert argumentieren.

Die Funktion g ist also nur auf $\mathbb{R} \setminus \{0\}$ differenzierbar.

c) Wir betrachten schließlich noch die Funktion $l(x) = x \cdot |x|$. Die Strategie ist dieselbe wie bei den beiden Aufgaben oben.

Zunächst ist klar, dass l in allen Punkten $x_0 \neq 0$ differenzierbar ist, da l für $x > 0$ die Funktion $x \mapsto x^2$ und für $x < 0$ die Funktion $x \mapsto -x^2$ ist.

Bleibt also wieder nur zu untersuchen, ob l differenzierbar an der „Klebestelle“ $x_0 = 0$ ist:

Der Differenzenquotient liefert diesmal $\frac{l(0+h)-l(0)}{h} = \frac{h \cdot |h|}{h} = |h|$. Dieser Ausdruck geht aber augenscheinlich für $h \to 0$ gegen 0. Wir sind fertig und können schließen, dass l auf ganz $\mathbb{R}$ differenzierbar ist.

Verständnisfragen 6.8

1. Wieso darf man in Teil a) die Ersetzung $x := x_0 + h$ vornehmen?
2. Wieso lohnt es sich in Teil a), den Ausdruck zuerst zu vereinfachen und erst danach den Grenzwert zu berechnen?
3. Wieso muss man bei der Auswertung des Differenzenquotienten an der Stelle $x_0 = 0$ in Teil b) zwei Fälle unterscheiden? Wie kommt man auf die beiden Werte 1 und h?
4. Wo kommen in Teil b) Funktionengrenzwerte vor und wo Folgengrenzwerte? Warum muss man überhaupt zu Folgengrenzwerten wechseln?
5. Wie sieht man schon dem Funktionsterm in Teil c) an, welche Stelle bei der Differenzierbarkeit Schwierigkeiten bereiten könnte?
6. Wieso verwenden wir im Teil c) (im Gegensatz zu Teil b) keine Folgengrenzwerte?
7. Muss beim Differentialquotienten immer der Wert 0 herauskommen, damit die Funktion differenzierbar ist?

Arbeitsauftrag 6.8 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.2.

6.9 Taylorpolynom

Aufgabe 6.9 Bestimmen Sie das Taylorpolynom zweiten Grades der Funktion $f : (-1, +\infty) \to \mathbb{R}$ mit $f(x) = \sqrt[4]{1+x}$ an der Stelle $x = 0$. Berechnen Sie damit den Wert von $\sqrt[4]{1.5}$ und schätzen Sie den Fehler mit Hilfe der Taylorformel ab!

P

Ausführliche Lösung 6.9 Vorbemerkung: „Warum sollte man denn $\sqrt[4]{1.5}$ überhaupt mit der Taylorformel ausrechnen wollen, so wie das in der Aufgabe gefordert ist? Wieso nimmt man nicht einfach den Taschenrechner und tippt das ein? Das geht doch viel schneller!" So oder so ähnlich könnten Ihre Gedanken sein, wenn Sie die Aufgabenstellung zum ersten Mal lesen, richtig?

Die Antwort ist ganz einfach: „Weil auch der Taschenrechner erst „lernen musste", wie man $\sqrt[4]{1.5}$ berechnet. Und wie berechnet das der Taschenrechner? Ja genau, mit Hilfe eines Taylorpolynoms!"

Die Idee ist, dass man die Funktion f an einer Stelle x durch ein geeignetes Polynom möglichst gut annähert. Will man jetzt einen Funktionswert $f(x+h)$ in der Nähe von x berechnen, so kann man genausogut den Wert des Polynoms an dieser Stelle ausrechnen. Schließlich sollte dabei eine gute Näherung für $f(x+h)$ herauskommen, weil ja die Funktion und das Polynom dort „nahe beisammen" liegen.

B

Jetzt aber zur eigentlichen Aufgabe: Wir sollen zunächst das Taylorpolynom 2. Grades an der Stelle $x = 0$ ausrechnen. Die Taylorformel lautet:

$$f(x+h) = \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!} h^k + \frac{f^{(n+1)}(x+\theta h)}{(n+1)!} h^{n+1}$$

mit $0 < \theta < 1$, für $(n+1)$-mal stetig differenzierbares f.

P

Der Term mit dem Summenzeichen beschreibt dabei das Taylorpolynom vom Grad n mit Entwicklungspunkt x (dieses wollen wir mit $T_{n,f,x}$ bezeichnen) und der letzte Summand das Restglied. In unserem Beispiel gilt also $x = 0$ und $n = 2$.

A

Wir suchen das Taylorpolynom

$$T_{2,f,0}(h) = \sum_{k=0}^{2} \frac{f^{(k)}(0)}{k!} h^k = f(0) + \frac{f'(0)}{1!} h + \frac{f''(0)}{2!} h^2.$$

B

Wir brauchen dafür also $f(0)$, $f'(0)$ und $f''(0)$, was wir aber leicht *ohne Taschenrechner* berechnen können:

H

$$\begin{aligned} f(x) &= (1+x)^{\frac{1}{4}} &\Rightarrow\quad f(0) &= 1, \\ f'(x) &= \frac{1}{4}(1+x)^{-\frac{3}{4}} &\Rightarrow\quad f'(0) &= \frac{1}{4}, \\ f''(x) &= \frac{1}{4}\left(-\frac{3}{4}\right)(1+x)^{-\frac{7}{4}} &\Rightarrow\quad f''(0) &= -\frac{3}{16}. \end{aligned}$$

Das Taylorpolynom 2. Grades zu f am Entwicklungspunkt 0 lautet damit:

$$T_{2,f,0}(h) = 1 + \frac{1}{4}h - \frac{3}{32}h^2.$$

P Dieses Polynom nähert also die Funktion f an der Stelle $x = 0$ an. Der Fehler, den man dabei macht, wird durch das Restglied beschrieben. Darauf kommen wir gleich zurück.

K Zunächst berechnen wir mit Hilfe des Taylorpolynoms den Wert $\sqrt[4]{1.5}$ näherungsweise:

H

$$\sqrt[4]{1.5} = \sqrt[4]{1 + \frac{1}{2}} = f(\frac{1}{2}) \approx T_{2,f,0}(\frac{1}{2}) = 1 + \frac{1}{4} \cdot \frac{1}{2} - \frac{3}{32} \cdot \left(\frac{1}{2}\right)^2 = \frac{141}{128}.$$

B Dafür haben wir keinen Taschenrechner verwenden müssen! Das Auswerten von Polynomen ist schlichtes Addieren, Subtrahieren und Multiplizieren!

Fehlerabschätzung: $\frac{141}{128}$ ist der gesuchte Näherungswert für $\sqrt[4]{1.5}$. Aber wie gut ist diese Näherung? Wie weit weicht sie vom tatsächlichen Wert von $\sqrt[4]{1.5}$ ab? Darüber gibt das Restglied Auskunft. Es gilt ja nach obiger Taylorformel:

A

$$\sqrt[4]{1.5} = \frac{141}{128} + \frac{f^{(3)}(0 + \frac{1}{2}\theta)}{3!}\left(\frac{1}{2}\right)^3$$

mit $0 < \theta < 1$.

Z Uns interessiert der Abstand zwischen $\sqrt[4]{1.5}$ und $\frac{141}{128}$, also formen wir um zu:

$$\left|\sqrt[4]{1.5} - \frac{141}{128}\right| = \left|\frac{f^{(3)}(\frac{1}{2}\theta)}{3!}\left(\frac{1}{2}\right)^3\right|.$$

B Wir möchten jetzt noch die rechte Seite berechnen (oder zumindest nach oben abschätzen), dafür brauchen wir die dritte Ableitung von f:

H

$$f^{(3)}(x) = \frac{21}{64}(1 + x)^{-\frac{11}{4}}.$$

Ausgewertet an der Stelle $\frac{\theta}{2}$ ergibt das $\frac{21}{64}(1 + \frac{\theta}{2})^{-\frac{11}{4}} = \frac{21}{64} \cdot \frac{1}{(1+\frac{\theta}{2})^{\frac{11}{4}}}$.

T Weil $0 < \theta < 1$ und damit $1 + \frac{\theta}{2} > 1$ gilt, können wir

$$\frac{21}{64} \cdot \frac{1}{(1 + \frac{\theta}{2})^{\frac{11}{4}}} < \frac{21}{64}$$

abschätzen.

H Damit erhalten wir also

$$\left|\sqrt[4]{1.5} - \frac{141}{128}\right| < \frac{21}{64} \cdot \frac{1}{3!} \cdot \left(\frac{1}{2}\right)^3 = \frac{7}{1024} < 0.007.$$

B Der Fehler, den wir bei der Approximation gemacht haben, ist folglich kleiner als 0.007. Der Näherungswert ist also schon ganz gut!

P

Bemerkung: Man könnte sich an dieser Stelle fragen, ob es nicht noch besser gewesen wäre, die Funktion direkt an der Stelle $x = \frac{1}{2}$ in ein Taylorpolynom zu entwickeln. Schließlich interessiert uns ja der Funktionswert $f(\frac{1}{2}) = \sqrt[4]{1.5}$ und das Taylorpolynom stimmt im Entwicklungspunkt ja (immer) *exakt* mit dem Funktionswert überein.

Sicher wäre das besser! Aber wenn Sie das versuchen, werden Sie bald feststellen, dass Sie zur Berechnung des Taylorpolynoms an der Stelle $x = \frac{1}{2}$ auch den Funktionswert $f(\frac{1}{2})$ benötigen. Eben den kennen wir aber nicht, den wollen wir ja gerade approximieren!

Es ist zur Berechnung eines Näherungswertes für $f(x_0)$ also immer sinnvoll, die Funktion an einer Stelle x (möglichst nahe x_0) in ein Taylorpolynom zu entwickeln, von der man leicht (also ohne Taschenrechner!) $f(x)$, $f'(x)$, $f''(x)$, usw. berechnen kann! Und genau das haben wir in dieser Aufgabe gemacht.

Verständnisfragen 6.9

1. Wieso muss man den Wert $\frac{1}{2}$ (und nicht den Wert 1.5) in das Taylorpolynom einsetzen, um $\sqrt[4]{1.5}$ zu berechnen?
2. Wie schafft es der Taschenrechner, dass sein Fehler bei der Berechnung von $\sqrt[4]{1.5}$ viel kleiner als 0.007 ist?
3. Inwiefern ist das Restglied in der Taylorformel unangenehm? Wie sind wir mit diesem Problem in unserer Aufgabe umgegangen?

Arbeitsauftrag 6.9 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.2.

6.10 Funktionenreihen

Aufgabe 6.10 Beweisen Sie, dass die folgenden Funktionenreihen gleichmäßig auf $\mathbb{R}$ konvergieren!

a) $f(x) = \sum\limits_{n=1}^{\infty} \frac{\ln(1 + |nx|)}{(n^3 + 1)(x^2 + 1)}$

b) $g(x) = \sum\limits_{n=1}^{\infty} \frac{x}{e^{|x|n} \cdot n}$

Ausführliche Lösung 6.10 Wir sollen in dieser Aufgabe beweisen, dass bestimmte Funktionenreihen gleichmäßig konvergieren. Hierfür bietet sich das Majorantenkriterium von Weierstraß an.

Das besagt: Wenn $||f_n||_\infty \leq c_n$ gilt und $\sum\limits_{n=1}^{\infty} c_n$ konvergent ist, dann konvergiert die Funktionenreihe $\sum\limits_{n=1}^{\infty} f_n(x)$ gleichmäßig.

a) Wir definieren $f_n(x) := \frac{\ln(1+|nx|)}{(n^3+1)(x^2+1)}$.

Es ist schwierig, davon die Supremumsnorm $||f_n||_\infty$ zu berechnen, deswegen werden wir sie im Folgenden einfach nach oben abschätzen. Das reicht ja auch. Uns stört vor allem der Logarithmus im Zähler, der den ganzen Ausdruck unhandlich macht.

Wir werden deshalb die Ungleichung $\ln(1+t) \leq t$ ausnutzen, die für alle $t \in \mathbb{R}$ gilt.

Damit folgt nämlich:

$$|f_n(x)| \leq \frac{|nx|}{(n^3+1)(x^2+1)} = \frac{n}{n^3+1} \cdot \frac{|x|}{x^2+1}.$$

Jetzt möchten wir gerne noch die Variable x loswerden, dann bleibt nur noch die Konvergenz einer Reihe mit Summationsindex n zu zeigen.

Wir nutzen dafür die Ungleichung $\frac{|x|}{x^2+1} < 1$, die sich leicht nachweisen lässt.

Es gilt dann:

$$|f_n(x)| \leq \frac{n}{n^3+1} \cdot \frac{|x|}{x^2+1} < \frac{n}{n^3+1} \leq \frac{1}{n^2}.$$

Auch die letzte Ungleichung in dieser Ungleichungskette lässt sich leicht zeigen!

Jetzt sind wir fast fertig.
Weil $|f_n(x)| \leq \frac{1}{n^2}$ für alle $x \in \mathbb{R}$ gilt, ist insbesondere $||f_n||_\infty := \sup\limits_{x\in\mathbb{R}} |f_n(x)| \leq \frac{1}{n^2}$.

Da die Reihe $\sum\limits_{n=1}^{\infty} \frac{1}{n^2}$ konvergiert, folgt nun nach dem Majorantenkriterium von Weierstraß die gleichmäßige Konvergenz der Reihe $\sum\limits_{n=1}^{\infty} f_n(x)$.

b) Hier definieren wir $g_n(x) := \frac{x}{e^{|x|n} \cdot n}$.
Diesmal lässt sich die Supremumsnorm wirklich ausrechnen:

Wir suchen das Supremum des Ausdrucks $|g_n(x)| = \frac{1}{n}|x|e^{-|x|n}$, wobei $x \in \mathbb{R}$. Wieder stört etwas an der Untersuchung dieses Ausdrucks, nämlich $|x|$.

Statt nach dem Supremum von $\frac{1}{n}|x|e^{-|x|n}$ mit $x \in \mathbb{R}$ zu suchen, können wir aber genausogut nach dem Supremum von $\frac{1}{n}te^{-tn}$ mit $t \in [0, \infty)$ suchen.

Definieren wir also $h(t) := \frac{1}{n}te^{-tn}$.

Wir suchen zuerst nach relativen Extremstellen der Funktion h:

$$h'(t) = \frac{1}{n}(te^{-tn}(-n) + e^{-tn}) \stackrel{!}{=} 0 \Longleftrightarrow -tn + 1 = 0 \Longleftrightarrow t = \frac{1}{n}.$$

Der Funktionswert an dieser Stelle ist $h(\frac{1}{n}) = \frac{1}{n^2}e^{-\frac{1}{n}n} = \frac{1}{en^2}$.

Gibt es am „Rand des Definitionsbereiches“ $[0, \infty)$ noch betragsmäßig größere Funktionswerte? Mal sehen:

$$h(0) = 0 \text{ und } \lim_{t\to\infty} h(t) = \lim_{t\to\infty} \frac{1}{n} te^{-tn} = \lim_{t\to\infty} \frac{t}{ne^{tn}} \overset{\text{de l'H.}}{=} \lim_{t\to\infty} \frac{1}{n^2 e^{tn}} = 0.$$

Die Antwort lautet also „Nein“. Damit erhalten wir:

$||g_n||_\infty = ||h||_\infty = \frac{1}{en^2}$. Da die Reihe $\sum_{n=1}^{\infty} \frac{1}{en^2}$ konvergiert, folgt nun nach dem Majorantenkriterium von Weierstraß die gleichmäßige Konvergenz der Reihe $\sum_{n=1}^{\infty} g_n(x)$.

Verständnisfragen 6.10

1. Wieso gelten die beiden in Teil a) verwendeten Ungleichungen $\frac{|x|}{x^2+1} < 1$ für $x \in \mathbb{R}$ bzw. $\frac{n}{n^3+1} < \frac{1}{n^2}$ für $n \in \mathbb{N}$? Beweisen Sie sie!
2. Wieso braucht man in Teil b) bei der Suche nach dem Supremum von $\frac{1}{n} \cdot te^{-tn}$ die Variable t nur im Intervall $[0, \infty)$ zu betrachten?
3. Wieso reicht es nicht, bei der Suche nach dem Supremum der Funktion h auf dem Intervall $[0, \infty)$ einfach das relative Extremum mittels der Methode $h'(t) = 0$ zu suchen?
4. Wieso wurde „Rand des Definitionsbereiches“ unter Gänsefüßchen gesetzt?

Übung 6.10 Ordnen Sie nun selbst den einzelnen Bearbeitungsschritten der Musterlösung die Phasen aus Kapitel 3 zu! Tragen Sie dazu die jeweiligen Buchstaben in die kleinen Quadrate am Seitenrand ein!

Arbeitsauftrag 6.10 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen, der Übung und dem Arbeitsauftrag finden Sie in Abschnitt 11.2.

7 Musterlösungen aus der Analysis 2

7.1 Funktionengrenzwerte

Aufgabe 7.1 Können die folgenden Funktionen so in $(0,0)$ definiert werden, dass sie auf ganz $\mathbb{R}^2$ stetig werden? Begründen Sie Ihre Antwort!

a) $f(x,y) = \dfrac{xy}{x^2+y^2}$

b) $g(x,y) = (x^2+y^2)^{x^2y^2}$

B

Ausführliche Lösung 7.1 Wir sollen in dieser Aufgabe feststellen, ob sich die Funktionen in $(0,0)$ so definieren lassen, dass sie stetig werden. Auf $\mathbb{R}^2 \setminus \{(0,0)\}$ sind die beiden Funktionen als Kompositionen stetiger Funktionen offensichtlich stetig.

P

a) Falls man f stetig in $(0,0)$ fortsetzen könnte, müsste

$$\lim_{(x,y)\to(0,0)} f(x,y) = f(0,0)$$

gelten für ein geeignet gewähltes $f(0,0)$.

B

Das soll etwas salopp formuliert bedeuten: „Egal entlang welches Nullfolgenpaares $\{(x_n,y_n)\}$ man den Grenzübergang Richtung $(0,0)$ ausführt, die zugehörigen Folgen der Funktionswerte $f(x_n,y_n)$ müssen alle zum selben Wert konvergieren".

P

Was ist das Problem in Aufgabenteil a)? Nachdem sowohl der Zähler als auch der Nenner des Bruches $\frac{xy}{x^2+y^2}$ mit $(x,y)\to(0,0)$ gegen Null konvergiert, können wir von vornherein nichts über diesen Grenzwert sagen. Leider gibt es für Funktionen in mehreren Variablen auch keine Regeln von de l'Hospital mehr.

K

Merkregel für solche Aufgaben: Solche Funktionengrenzwerte sind deshalb schwer zu berechnen, weil man zwei Variablen gleichzeitig gegen Null laufen lassen muss[1]. In unserem Fall nämlich x und y. Ziel ist es bei solchen Aufgaben in der Regel, die Berechnung des Grenzwertes auf eine gewöhnliche Grenzwertberechnung mit nur einer Variablen zurückzuführen. Diese sind meist einfacher durchzuführen. Wir werden auf diese *Merkregel* in der Folge immer wieder aufmerksam machen.

[1] Dies ist ein wesentlicher Unterschied zur Grenzwertberechnung mit nur einer Variablen. Während wir uns dort einer Stelle nur von rechts bzw. links nähern können, haben wir im Zweidimensionalen viel mehr Möglichkeiten. Wir können uns einem Punkt der Ebene entlang von Geraden, Parabeln, Spiralen usw. annähern, was die Sache komplexer macht.

C. Ableitinger, A. Herrmann, *Lernen aus Musterlösungen zur Analysis und Linearen Algebra*,
DOI 10.1007/978-3-658-02353-9_7, © Springer Fachmedien Wiesbaden 2013

Wie entscheidet man, ob eine Funktion stetig fortsetzbar ist oder nicht?

i) Entweder man findet zwei Nullfolgenpaare $\{(x_n, y_n)\}$ und $\{(u_n, v_n)\}$, entlang derer die entsprechenden Grenzwerte der Folgen der Funktionswerte zwei unterschiedliche Werte annehmen. Dann ist f nicht stetig fortsetzbar.

ii) Oder man findet eben keine zwei solchen Folgenpaare. Dann muss man allgemein beweisen, dass der Grenzwert der Folge der Funktionswerte entlang jedes beliebigen Nullfolgenpaares $\{(x_n, y_n)\}$ derselbe ist. Dann ist f stetig fortsetzbar.

B Für Teil a) findet man zwei solche Folgenpaare. Dafür braucht man ein geschultes Auge – passen Sie auf:

Z Wenn man sich entlang einer der beiden Koordinatenachsen nähern möchte (z. B. entlang der positiven y-Achse), dann kann man das z. B. entlang des Folgenpaares $\{(x_n, y_n)\} := \{(0, \frac{1}{n})\}$ tun (siehe Abbildung 7.1).

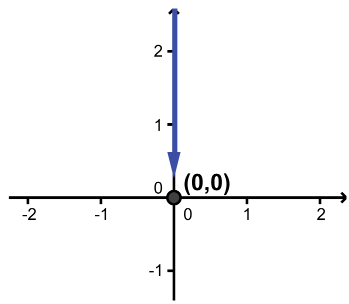

Abbildung 7.1: Annäherung an $(0, 0)$ entlang der y-Achse

H Sehen wir uns den zugehörigen Grenzwert der Folge der Funktionswerte an:

$$\lim_{n\to\infty} f(0, \frac{1}{n}) = \lim_{n\to\infty} \frac{0 \cdot \frac{1}{n}}{0^2 + \left(\frac{1}{n}\right)^2} = \lim_{n\to\infty} 0 = 0.$$

P Wir haben hier also praktischerweise ein Folgenpaar gewählt, bei dem der Nenner des Bruchs ungleich Null, der Zähler aber Null für jedes Folgenglied ist. Man müsste also demnach $f(0, 0) := 0$ definieren. Beachten Sie außerdem, dass wir der oben formulierten Merkregel genüge getan haben, indem wir den Grenzwert auf einen Grenzwert mit nur einer Variablen n zurückgeführt haben. Dieser Grenzwert war dann einfach zu berechnen.

K Wenn man sich jetzt dem Punkt $(0, 0)$ entlang eines geschickt gewählten anderen Folgenpaares nähert, dann kann man aber auch einen anderen Grenzwert der Folge der Funktionswerte erhalten. Und zwar wählen wir jetzt $\{(u_n, v_n)\}$ so, dass sich eingesetzt in den Funktionsterm die Ausdrücke mit n kürzen lassen.

Z Das lässt sich z. B. durch $\{(u_n, v_n)\} := \{(\frac{1}{n}, \frac{1}{n})\}$ erreichen (siehe Abbildung 7.2).

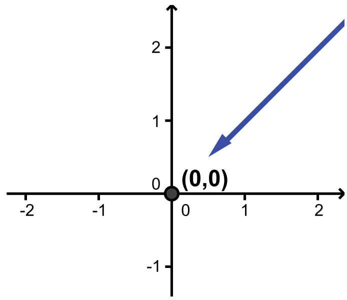

Abbildung 7.2: Annäherung an $(0,0)$ entlang der Geraden $y = x$

Eingesetzt ergibt das:

$$\lim_{n\to\infty} f\left(\frac{1}{n}, \frac{1}{n}\right) = \lim_{n\to\infty} \frac{\frac{1}{n}\cdot\frac{1}{n}}{\frac{1}{n^2}+\frac{1}{n^2}} = \lim_{n\to\infty} \frac{\frac{1}{n^2}}{\frac{2}{n^2}} = \lim_{n\to\infty} \frac{1}{2} = \frac{1}{2}.$$

Hier müssten wir also $f(0,0) := \frac{1}{2}$ definieren, um eine Chance auf Stetigkeit zu haben. Das steht aber im Widerspruch zu der Wahl von $f(0,0) = 0$ vorher.

Diese Funktion lässt sich also nicht stetig in $(0,0)$ fortsetzen.

b) Wir bringen die Funktion $g(x,y) = (x^2+y^2)^{x^2y^2}$ zunächst in eine Gestalt, für die wir bessere Bearbeitungsinstrumente besitzen: $g(x,y) = e^{x^2y^2\cdot\ln(x^2+y^2)}$. Es reicht wegen der Stetigkeit der Exponentialfunktion, nur den Exponenten zu betrachten. Dieser ist nämlich wieder eine Funktion in x und y. Wenn diese stetig ist, ist auch die Gesamtfunktion als Verkettung zweier stetiger Funktionen stetig. Wir definieren daher:

$$h(x,y) := x^2y^2\cdot\ln(x^2+y^2).$$

Wir interessieren uns jetzt für den Grenzwert

$$\lim_{(x,y)\to(0,0)} h(x,y).$$

Sie könnten sich auch bei dieser Funktion auf die Suche nach zwei Folgenpaaren machen, entlang derer die entsprechenden Folgen der Funktionswerte unterschiedliche Grenzwerte liefern. Dieses Vorhaben wird aber scheitern (siehe Verständnisfrage 3).

Stattdessen hoffen wir, dass der Grenzwert existiert (also entlang jedes Folgenpaares denselben Wert annimmt). Das zeigen wir gleich auf zwei unterschiedliche Arten.

Was ist das Problem in Aufgabenteil b)? Der Ausdruck x^2y^2 konvergiert für $(x,y)\to(0,0)$ gegen Null, der Ausdruck $\ln(x^2+y^2)$ andererseits divergiert gegen $-\infty$. Es bleibt die Frage, welcher der beiden Ausdrücke sich beim Grenzübergang „durchsetzen“ kann.

1. Lösungsweg: Wir können eine Abschätzung machen, bei der wir versuchen, eine der beiden Variablen x oder y loszuwerden. Damit verfolgen wir nämlich wieder unsere Merkregel, den Grenzwert auf einen Grenzwert in einer Variablen zurückzuführen.

Z Nachdem (x, y) gegen $(0, 0)$ laufen soll, können wir uns in der folgenden Abschätzung auf $x^2 + y^2 < 1$ beschränken.

H Dann gilt nämlich $0 \geq \ln(x^2 + y^2) \geq \ln(x^2)$ und damit $|\ln(x^2 + y^2)| \leq |\ln(x^2)|$. Das liefert uns eine Abschätzung für $|h(x, y)|$:

A

$$0 \leq |x^2y^2 \ln(x^2 + y^2)| \overset{x^2+y^2<1}{\leq} |x^2y^2 \ln(x^2)| \overset{y^2<1}{\leq} |x^2 \ln(x^2)|$$

Es gilt: $\lim\limits_{x \to 0} x^2 \ln(x^2) = 0$ (leicht zu zeigen mit den Regeln von de l'Hospital). Also gilt nach dem Einschließungskriterium auch

H

$$\lim_{(x,y)\to(0,0)} x^2y^2 \ln(x^2 + y^2) = 0$$

und damit

$$\lim_{(x,y)\to(0,0)} g(x, y) = e^{\lim_{(x,y)\to(0,0)} h(x,y)} = e^0 = 1.$$

B Der Grenzwert existiert also tatsächlich. Wenn wir $g(0, 0) := 1$ setzen, dann ist g stetig.

Wieder hat es zum Ziel geführt, die Grenzwertberechnung auf eine gewöhnliche Grenzwertberechnung mit nur einer Variablen zurückzuführen (diesmal, indem wir die Variable y durch Abschätzungen losgeworden sind).

2. Lösungsweg: Es gibt einen kleinen Trick, den Sie sich unbedingt für Aufgaben dieser Art merken sollten.

T Der Grenzübergang $(x, y) \to (0, 0)$ ist schwer zu handhaben, weil „x und y sich auch unterschiedlich schnell dem Punkt $(0, 0)$ nähern können". Man rechnet daher in sogenannte *Polarkoordinaten* um. Dieser Trick eignet sich insbesondere dann, wenn der Ausdruck $x^2 + y^2$ vorkommt, wie Sie gleich merken werden.

Setze $x = r\cos(t)$, $y = r\sin(t)$ mit $r \geq 0$ und $t \in [0, 2\pi)$ (siehe Abbildung 7.3).

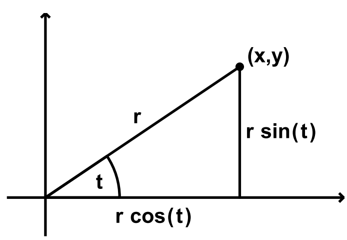

Abbildung 7.3: Darstellung des Punktes (x, y) in Polarkoordinaten

A Dann gilt nämlich: $(x, y) \to (0, 0) \iff r \searrow 0$.

P

Durch diesen Trick gewinnen wir den Vorteil, dass einer der Parameter alleine über die Konvergenz entscheidet! Wieder haben wir uns entsprechend unserer Merkregel verhalten.

H

Jetzt setzen wir das in die Funktion h ein:

$$\begin{aligned} h(x,y) &= r^2\cos^2(t)\cdot r^2\sin^2(t)\cdot\ln(r^2\cos^2(t)+r^2\sin^2(t)) \\ &= r^4\cos^2(t)\sin^2(t)\cdot\ln(r^2). \end{aligned}$$

A

Nachdem sin und cos betragsmäßig kleiner oder gleich 1 sind, gilt:

$$0 \leq |r^4\cos^2(t)\sin^2(t)\cdot\ln(r^2)| \leq r^4\cdot|\ln(r^2)| = r^4|2\ln(r)| = 2r^3|r\cdot\ln(r)|.$$

Nachdem die rechte Seite wegen

$$\lim_{r\searrow 0} r\ln(r) = 0$$

(nachzuweisen mit den Regeln von de l'Hospital) gegen Null geht, gilt nach dem Einschließungskriterium auch

H

$$\lim_{(x,y)\to(0,0)} h(x,y) = 0$$

und damit

$$\lim_{(x,y)\to(0,0)} g(x,y) = e^{\lim_{(x,y)\to(0,0)} h(x,y)} = e^0 = 1.$$

B

Der Grenzwert existiert also tatsächlich. Wenn wir $g(0,0) := 1$ setzen, dann ist g stetig.

Verständnisfragen 7.1

1. Was wäre gewesen, wenn in Teil a) der Grenzwert der Folge der Funktionswerte entlang des zweiten Folgenpaares $\{(u_n, v_n)\} = \{(\frac{1}{n}, \frac{1}{n})\}$ ebenfalls den Wert Null geliefert hätte?
2. Warum gilt in Teil b) $(x,y) \to (0,0) \Leftrightarrow r \searrow 0$? Wieso spielt hier die Variable t keine Rolle?
3. Setzen Sie zur Übung einige konkrete Folgenpaare $\{(x_n, y_n)\}$ (wobei x_n und y_n Nullfolgen sind) in die Funktion h aus Teil b) ein und berechnen Sie den Grenzwert $\lim\limits_{n\to\infty} h(x_n, y_n)$. Kommt dabei wirklich immer derselbe Wert heraus?

Arbeitsauftrag 7.1 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.3.

7.2 Integrationsmethoden

Aufgabe 7.2 Berechnen Sie die Integrale in a) und b) mit Hilfe von Substitution und die Integrale in c) und d) mittels partieller Integration!

a) $\int \frac{\ln x}{x} \, dx$

b) $\int_1^2 \frac{e^x}{e^{2x} + 2e^x + 1} \, dx$

c) $\int xe^x \, dx$

d) $\int \sqrt{x} \ln x \, dx$

Ausführliche Lösung 7.2 a) Integration mit Hilfe von Substitution funktioniert immer dann, wenn man ein Produkt aus zwei Funktionen integrieren soll, bei dem einer der beiden Faktoren „im Wesentlichen" die Ableitung des anderen darstellt[2].

Das Integral $\int \frac{\ln x}{x} \, dx = \int \ln x \cdot \frac{1}{x} \, dx$ ist ein prototypisches Beispiel für diesen Fall! Hier gilt nämlich sogar exakt, dass die Ableitung von $\ln x$ gleich dem anderen Faktor $\frac{1}{x}$ ist. Warum das wichtig ist, werden Sie gleich sehen:

Man substituiert nun also jenen Faktor, der abgeleitet den anderen ergibt, durch eine neue Variable u:

$$u = \ln x.$$

Wir leiten u mit Hilfe der Leibnizschen Schreibweise nach x ab:

$$\frac{du}{dx} = \frac{1}{x}$$

und formen nach dx um:

$$dx = x \cdot du.$$

Was bedeutet diese Substitution für das zu berechnende Integral? Wir setzen ein:

$$\int \ln x \cdot \frac{1}{x} \, dx = \int u \cdot \frac{1}{x} \cdot x \, du = \int u \, du.$$

Wir haben also das Integral nach x in ein (viel einfacheres) Integral nach u umgeformt. Entscheidend war dabei, dass wir die Variable x kürzen konnten. Und das erreicht man eben nur dann, wenn die Ableitung des substituierten Terms gerade (bis auf eine multiplikative Konstante) gleich dem anderen Term ist.

Wir können jetzt das Integral berechnen und am Ende wieder $u = \ln x$ rücksubstituieren. Schließlich wollen wir ja das *ursprüngliche* Integral berechnen:

$$\int \frac{\ln x}{x} \, dx = \int u \, du = \frac{u^2}{2} + C = \frac{(\ln x)^2}{2} + C,$$

[2] Dabei handelt es sich um eine Faustregel. Es gibt auch noch einige andere Fälle, in denen Substitution weiterhilft.

wobei $C \in \mathbb{R}$ die sogenannte Integrationskonstante ist. (Stammfunktionen sind eindeutig bis auf eine additive Konstante.)

b) Integration durch Substitution funktioniert sogar noch etwas allgemeiner als oben beschrieben. Es ist allerdings hilfreich, wenn sich die Ableitung des substituierten Terms mit dem nichtsubstituierten Teil kürzen lässt.

Zur Berechnung des Integrals

$$\int_1^2 \frac{e^x}{e^{2x}+2e^x+1}\,dx$$

kann man jedes e^x im Nenner durch u substituieren.

e^x ergibt abgeleitet wieder e^x, was sich dann mit dem Zähler kürzen lässt, wie wir weiter unten sehen werden.

$$\begin{aligned} u &= e^x \\ \frac{du}{dx} &= e^x \\ dx &= \frac{du}{e^x} \end{aligned}$$

Zusätzlich haben wir hier die Grenzen mitzusubstituieren. Wenn $1 \leq x \leq 2$ gilt, dann gilt $e^1 \leq e^x \leq e^2$, also $e \leq u \leq e^2$.

Wir erhalten insgesamt:

$$\begin{aligned} \int_1^2 \frac{e^x}{e^{2x}+2e^x+1}dx &= \int_e^{e^2} \frac{e^x}{(u^2+2u+1)}\frac{du}{e^x} = \int_e^{e^2} \frac{1}{(u+1)^2}du \\ &= \int_e^{e^2} (u+1)^{-2}du = [-(u+1)^{-1}]_e^{e^2} \\ &= -\frac{1}{e^2+1} + \frac{1}{e+1}. \end{aligned}$$

Durch das Mitsubstituieren der Grenzen ist uns übrigens auch gleich das Rücksubstituieren erspart geblieben. Alternativ könnte man auch zuerst das unbestimmte Integral berechnen (wie in Teil a) inkl. Rücksubstitution) und erst danach das bestimmte Integral durch Einsetzen der (ursprünglichen) Grenzen ermitteln.

c) Auch die partielle Integration dient zum Berechnen von Integralen über Produkte. Die zugehörige Formel lautet: $\int fg' = fg - \int f'g$. Hier ist ganz allgemein zu beachten, dass der Faktor, der als f gewählt wird (also als abzuleitende Funktion), beim Ableiten „einfacher" wird (also z. B. Polynome – beim Ableiten wird der Grad des Polynoms ja um 1 kleiner).

Wieder haben wir ein prototypisches Beispiel vor uns:

$$\int x \cdot e^x\,dx.$$

Wir wählen hier nach obiger Überlegung $f = x$ und $g' = e^x$. Daraus lassen sich $f' = 1$ und $g = e^x$ leicht berechnen[3]. Dann können wir schon in die Formel einsetzen:

$$\int x \cdot e^x \, dx = x \cdot e^x - \int 1 \cdot e^x \, dx = xe^x - e^x + C.$$

Das zu berechnende Integral wurde also durch Anwenden dieser Methode viel einfacher!

d) Ist einer der beiden Faktoren im Integranden $\ln x$, so lohnt es sich, (ausnahmsweise) diesen Faktor als abzuleitende Funktion f zu wählen. $\ln x$ liefert abgeleitet nämlich $\frac{1}{x}$, was sich dann eventuell gut mit dem anderen Faktor vereinfachen lässt (das heißt, selbst im Integral $\int x^2 \ln x \, dx$ sollten sie $f = \ln x$ und $g' = x^2$ wählen und nicht umgekehrt).

Zur Berechnung des Integrals $\int \sqrt{x} \ln x \, dx$ wählen wir also $f = \ln x$ und $g' = \sqrt{x}$. Daraus lassen sich leicht $f' = \frac{1}{x}$ und $g = \frac{x^{\frac{3}{2}}}{\frac{3}{2}} = \frac{2}{3}x^{\frac{3}{2}}$ berechnen.

Es ergibt sich durch Anwenden der partiellen Integration:

$$\begin{aligned}\int \sqrt{x} \ln x \, dx &= \frac{2}{3}x^{\frac{3}{2}} \cdot \ln x - \int \frac{2}{3}x^{\frac{3}{2}} \cdot \frac{1}{x} \, dx = \frac{2}{3}x^{\frac{3}{2}} \cdot \ln x - \frac{2}{3} \int x^{\frac{1}{2}} \, dx \\ &= \frac{2}{3}x^{\frac{3}{2}} \cdot \ln x - \frac{4}{9}x^{\frac{3}{2}} + C.\end{aligned}$$

Verständnisfragen 7.2

1. Wie kann man nach der Berechnung unbestimmter Integrale ganz einfach überprüfen, ob das Ergebnis richtig ist?
2. Wie kommt man in Teil b) darauf, gerade die beiden e^x im Nenner zu substituieren? Woher weiß man, dass man nicht auch e^x im Zähler substituieren muss?
3. Warum lohnt es sich in Teil d) $f = \ln x$ zu setzen, d. h. den Logarithmus als zu differenzierende Funktion zu wählen? Ginge es auch anders?

Übung 7.2 Ordnen Sie nun selbst den einzelnen Bearbeitungsschritten der Musterlösung die Phasen aus Kapitel 3 zu! Tragen Sie dazu die jeweiligen Buchstaben in die kleinen Quadrate am Seitenrand ein!

Arbeitsauftrag 7.2 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen, der Übung und dem Arbeitsauftrag finden Sie in Abschnitt 11.3.

[3] Eigentlich müssten wir schreiben: f sei definiert durch $f(x) := x$ und g' sei definiert durch $g'(x) := e^x$. Der Übersichtlichkeit halber verzichten wir hier bewusst auf das Anführen der Funktionsargumente.

7.3 Uneigentliche Integrale

Aufgabe 7.3 Entscheiden Sie, ob die folgenden uneigentlichen Integrale existieren!

a) $\displaystyle\int_0^\infty x^k e^{-x} dx$, wobei $k \in \mathbb{N}$

b) $\displaystyle\int_0^1 |\ln x| dx$

c) $\displaystyle\int_0^1 \frac{\sin x \cdot \ln x}{x} dx$

d) $\displaystyle\int_0^\infty \frac{(1+x^3)e^{-x}}{\sqrt{x}} dx$

P

Ausführliche Lösung 7.3 Manche uneigentlichen Integrale lassen sich nicht direkt berechnen. Die Schwierigkeit bestünde in den Aufgabenteilen c) und d) darin, eine Stammfunktion zu den gegebenen Funktionen anzugeben. Allerdings sollen wir ja nur entscheiden, ob die Integrale existieren und nicht berechnen, welchen Wert sie haben.

Dabei helfen Abschätzungen, wie Sie später sehen werden. Zunächst aber zu den Teilen a) und b). Hier werden wir die uneigentlichen Integrale sogar berechnen.

K

a) Die Form des Integrals deutet darauf hin, dass wir hier partiell integrieren müssen. Dabei können wir erreichen, dass der Exponent von x um 1 kleiner wird. Wenn wir diesen Schritt noch $(k-1)$-mal wiederholen, können wir erreichen, dass im Exponenten Null steht. Diese Idee verfolgen wir nun.

Z

Wir arbeiten dabei zunächst einmal mit dem Integral $\displaystyle\int_0^R x^k e^{-x} dx$ und führen den Grenzübergang $R \to +\infty$ erst am Ende durch. Uneigentliche Integrale sind ja – wie der Name schon andeutet – keine Integrale, sondern als Grenzwerte von Integralen definiert.

H

$$\int_0^R x^k e^{-x} dx = \left[-x^k e^{-x}\right]_0^R + k \cdot \int_0^R x^{k-1} e^{-x} dx = -\frac{R^k}{e^R} + k \cdot \int_0^R x^{k-1} e^{-x} dx$$

Den ersten Summanden können wir praktischerweise auswerten. Es gilt: $\displaystyle\lim_{R\to+\infty} \frac{R^k}{e^R} = 0$, was man z. B. durch k-maliges Anwenden der Regeln von de l'Hospital beweisen kann.

P

Bleibt also nur noch das Integral $\displaystyle\int_0^\infty x^{k-1} e^{-x} dx$, das große Ähnlichkeit zu unserem Ausgangsintegral hat. Es ist aber schon eine Spur einfacher, da der Exponent von x jetzt nur noch $k-1$ ist.

H

Das uneigentliche Integral $\displaystyle\int_0^\infty x^k e^{-x} dx$ existiert also genau dann, wenn das uneigentliche Integral $\displaystyle\int_0^\infty x^{k-1} e^{-x} dx$ existiert.

P Wenn wir den Schritt der partiellen Integration nun noch $(k-1)$-mal wiederholen, dann kommen wir zu der Erkenntnis, dass das uneigentliche Integral $\int_0^\infty x^k e^{-x} dx$ genau dann existiert, wenn das uneigentliche Integral $\int_0^\infty x^0 e^{-x} dx = \int_0^\infty e^{-x} dx$ existiert.

H Dieses können wir aber berechnen:

$$\int_0^\infty e^{-x} dx = \lim_{R\to+\infty} \left[-e^{-x}\right]_0^R = \lim_{R\to+\infty} \left(1-e^{-R}\right) = 1.$$

P Nachdem bei jeder partiellen Integration der Exponent von x im Integranden als Faktor vor das neue Integral kommt, also nacheinander die Zahlen k, $k-1$, $k-2$, …, 2, 1, ergibt sich:

$$\int_0^\infty x^k e^{-x} dx = k!\ .$$

B Insbesondere existiert das uneigentliche Integral.

P b) Hier muss man zunächst erkennen, dass das Problem in der linken Intervallgrenze liegt. $\ln x$ ist ja an der Stelle $x=0$ nicht definiert, da $\lim_{x\searrow 0} \ln x = -\infty$ gilt.

K D. h., wir untersuchen zunächst das Integral $\int_a^1 |\ln x| dx$ und führen den Grenzübergang $a \searrow 0$ wieder am Ende durch:

H

$$\begin{aligned}\int_a^1 |\ln x| dx &= -\int_a^1 \ln x dx = -\left[x\ln x - x\right]_a^1 \\ &= -(-1 - a\ln a + a) = 1 + a\ln a - a.\end{aligned}$$

Der Grenzübergang liefert

$$\lim_{a\searrow 0}(1 + a\ln a - a) = 1,$$

B da $\lim_{a\searrow 0} a\ln a = 0$ gilt, wie man mit den Regeln von de l'Hospital beweisen kann:

H

$$\lim_{a\searrow 0} a\ln a = \lim_{a\searrow 0} \frac{\ln a}{\frac{1}{a}} = \frac{-\infty}{\infty} = \lim_{a\searrow 0} \frac{\frac{1}{a}}{-\frac{1}{a^2}} = \lim_{a\searrow 0}(-a) = 0.$$

B Damit existiert auch dieses uneigentliche Integral, es hat den Wert 1.

P c) Dieser Aufgabenteil ist noch etwas leichter als Teil d), weil wir das Integral gleich über dem *gesamten* Integrationsbereich $(0,1]$ durch eine integrierbare Funktion abschätzen können[4]. Das Integral ist *deshalb* uneigentlich, weil der Integrand für $x \to 0$ (linke Integrationsgrenze) gegen $-\infty$ divergiert. Die rechte Integrationsgrenze 1 macht keine Schwierigkeiten.

[4] Das gilt aber nur, weil die Funktion $x \mapsto \frac{\sin x \cdot \ln x}{x}$ selbst auf jedem Intervall $[a,1]$ mit $a \in (0,1)$ Riemann-integrierbar ist. Ein Gegenbeispiel wäre die Funktion $f : [0,1] \to \mathbb{R}$ mit $f(x) = \begin{cases} 0 & x \in [0,1]\cap\mathbb{Q} \\ 1 & x \in [0,1]\setminus\mathbb{Q}\end{cases}$. Es gilt zwar $|f(x)| \le 1$ für alle $x \in [0,1]$ und $\int_0^1 1 dx = 1$, aber f selbst ist nicht Riemann-integrierbar.

Zur gegebenen Funktion $\frac{\sin x \cdot \ln x}{x}$ findet man keine Stammfunktion, man kann allerdings $\left|\frac{\sin x \cdot \ln x}{x}\right| \leq |\ln x|$ abschätzen, da $|\sin x| \leq |x|$ gilt. K

Nachdem sogar das Integral $\int_0^1 |\ln x| dx$ existiert (siehe Teil b)), existiert erst recht das Integral H

$$\int_0^1 \left|\frac{\sin x \cdot \ln x}{x}\right| dx$$

(da der Integrand ja über das gesamte Intervall $(0,1)$ kleiner als $|\ln x|$ ist).

Damit ist auch $\frac{\sin x \cdot \ln x}{x}$ auf $(0,1)$ integrierbar und wir sind fertig. B

d) Dieser Aufgabenteil ist etwas schwieriger, weil der Ausdruck $\frac{(1+x^3)e^{-x}}{\sqrt{x}}$ für $x \to 0$ (linke Integrationsgrenze) divergiert und auch die rechte Integrationsgrenze (∞) Probleme bereitet. Es machen also beide Integrationsgrenzen für sich das Integral zu einem uneigentlichen Integral – auf verschiedene Arten! P

In solchen Fällen „zerteilt“ man den Integrationsbereich geschickt. Wir schreiben also z. B. Z

$$\int_0^\infty \frac{(1+x^3)e^{-x}}{\sqrt{x}} dx = \int_0^1 \frac{(1+x^3)e^{-x}}{\sqrt{x}} dx + \int_1^\infty \frac{(1+x^3)e^{-x}}{\sqrt{x}} dx$$

und untersuchen die beiden Teilintegrale auf Integrierbarkeit.

Sind beide integrierbar, so existiert auch das ursprüngliche Integral. B

Betrachten wir zunächst das erste Teilintegral, also jenes über den Bereich $(0,1)$.

Nachdem dort $1 + x^3 \leq 2$ und $e^{-x} \leq 1$ gilt (skizzieren Sie die beiden Funktionen, falls Sie Schwierigkeiten haben, das zu sehen!) und alle Ausdrücke im Integranden positiv sind, können wir abschätzen: T

$$0 \leq \frac{(1+x^3)e^{-x}}{\sqrt{x}} \leq \frac{2 \cdot 1}{\sqrt{x}}.$$ H

Das uneigentliche Integral $\int_0^1 \frac{2}{\sqrt{x}} dx$ existiert (Sie können es leicht direkt berechnen) und damit auch das uneigentliche Integral $\int_0^1 \frac{(1+x^3)e^{-x}}{\sqrt{x}} dx$.

Jetzt fehlt nur noch das Intervall $(1,\infty)$. B

Nachdem dort $\sqrt{x} \geq 1$ gilt, können wir abschätzen: Z

$$0 \leq \frac{(1+x^3)e^{-x}}{\sqrt{x}} \leq (1+x^3)e^{-x}.$$ H

Der rechte Ausdruck ist integrierbar wegen Teil a) und damit erst recht der Ausdruck $\frac{(1+x^3)e^{-x}}{\sqrt{x}}$, womit wir fertig sind.

Das uneigentliche Integral $\int_0^\infty \frac{(1+x^3)e^{-x}}{\sqrt{x}} dx$ existiert also. B

Verständnisfragen 7.3

1. Man könnte die Lösung zu Teil a) noch sauberer als Induktionsbeweis aufschreiben. Welcher Teil des Beweises entspräche dann dem Induktionsanfang, welcher Teil dem Induktionsschritt?
2. Wieso gilt in Teil b) $\int_a^1 |\ln x| dx = -\int_a^1 \ln x dx$ und $\int_a^1 \ln x dx = [x \ln x - x]_a^1$?
3. Wieso haben wir das Integrationsintervall in Teil d) gerade bei 1 geteilt? Hätte man es auch an einer anderen Stelle teilen können?
4. Wieso haben wir die beiden Abschätzungen $1 + x^3 \leq 2$ und $e^{-x} \leq 1$, die wir in Teil d) für den Integrationsbereich $(0, 1)$ benutzt haben, nicht auch für den Integrationsbereich $(1, \infty)$ verwendet?

Arbeitsauftrag 7.3 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.3.

7.4 Differenzierbarkeit von Funktionen $f : \mathbb{R}^2 \to \mathbb{R}$

Aufgabe 7.4 Bestimmen Sie alle Punkte $(x, y) \in \mathbb{R}^2$, in denen die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x, y) = |xy|$ differenzierbar ist.

P

Ausführliche Lösung 7.4 Zunächst machen wir uns klar, dass diese Funktion keine Schwierigkeiten macht, solange $xy \neq 0$ ist, da ja auch die gewöhnliche Betragsfunktion $g : \mathbb{R} \to \mathbb{R}$ mit $g(t) := |t|$ in jedem $t \neq 0$ differenzierbar ist. Warum man so argumentieren kann, überlegen wir uns kurz:

H

Die Funktion h mit dem Funktionsterm $h(x, y) := xy$ ist differenzierbar in jedem Punkt (x, y), da die partiellen Ableitungen $\frac{\partial h}{\partial x} = y$ und $\frac{\partial h}{\partial y} = x$ überall stetig sind. Die Funktion g mit dem Funktionsterm $g(t) = |t|$ ist wie gesagt in jedem $t \neq 0$ differenzierbar. Schließlich ist dann auch $f(x, y) = g(h(x, y)) = g \circ h(x, y)$ als Komposition differenzierbarer Funktionen für $xy \neq 0$ differenzierbar.

B

Damit haben wir schon einen großen Teil abgehandelt, f ist also zumindest in allen Punkten (x, y) mit $xy \neq 0$ differenzierbar.

Exkurs 1

P

Es kann passieren, dass man den Funktionsterm zwar formal partiell ableiten kann, die partiellen Ableitungsfunktionen aber an manchen Stellen gar nicht definiert sind. Z. B.: $l(x, y) = \sqrt{x^2 + y^2}$.

H

Durch formales partielles Ableiten erhält man: $\frac{\partial l}{\partial x}(x, y) = \frac{x}{\sqrt{x^2+y^2}}$.

Hier müsste man also die partiellen Ableitungen im Punkt $(0,0)$ extra untersuchen.

Zurück zur Funktion f: Jetzt betrachten wir alle Punkte (x,y) mit $xy = 0$. Für solche Punkte muss gelten, dass entweder $x = 0$ oder $y = 0$ ist.

Nachdem die Funktion f symmetrisch in x und y ist, reicht es, einen der beiden Fälle zu betrachten. Der andere Fall läuft dann analog.

Sei jetzt o.B.d.A. $y = 0$.

Wir untersuchen also die Differenzierbarkeit der Funktion f in Punkten der Form $(x_0, 0)$ mit $x_0 \in \mathbb{R}$. Eine notwendige Bedingung für Differenzierbarkeit ist die Existenz der partiellen Ableitungen $\frac{\partial f}{\partial x}(x_0,0)$ und $\frac{\partial f}{\partial y}(x_0,0)$.

Diese müssen wir mit der Definition der partiellen Ableitung berechnen, da wir $f(x,y) = |xy|$ nicht formal ableiten können.

Es gilt: $\dfrac{\partial f}{\partial x}(x_0,0) = \lim\limits_{t\to 0} \dfrac{f(x_0+t,0) - f(x_0,0)}{t} = \lim\limits_{t\to 0} \dfrac{0-0}{t} = \lim\limits_{t\to 0} 0 = 0.$

Diese partielle Ableitung existiert also schon einmal für jedes x_0.

Es gilt: $\dfrac{\partial f}{\partial y}(x_0,0) = \lim\limits_{t\to 0} \dfrac{f(x_0,t) - f(x_0,0)}{t} = \lim\limits_{t\to 0} \dfrac{|x_0 t|}{t} = \begin{cases} |x_0| & t \searrow 0 \\ -|x_0| & t \nearrow 0 \end{cases}.$

Diese partielle Ableitung existiert hingegen nur, wenn $x_0 = 0$ ist, für jede andere Zahl gilt ja $|x_0| \neq -|x_0|$.

Damit ist jedenfalls schon einmal geklärt, dass f in *keinem* der Punkte $(x_0, 0)$ mit $x_0 \neq 0$ differenzierbar ist, da ja nicht einmal die partiellen Ableitungen existieren.

Bleibt noch ein einziger Punkt zu untersuchen, nämlich $a = (0,0)$. Von ihm wissen wir schon, dass in ihm beide partiellen Ableitungen existieren. Wir haben also eine Chance auf Differenzierbarkeit in $(0,0)$.

Das prüfen wir nun mit der Definition der Differenzierbarkeit nach.

Wir müssen dazu untersuchen, ob es eine lineare Abbildung T gibt mit

$$\lim_{0\neq h\to 0} \frac{f(a+h) - f(a) - Th}{||h||} = 0. \tag{7.1}$$

Falls es eine solche lineare Abbildung T gibt, dann muss sie – salopp gesprochen – die Jacobi-Matrix in a sein, also die Matrix der partiellen Ableitungen in $a = (0,0)$.

Wir nehmen nun an, dass f in $(0,0)$ differenzierbar ist, dass es also eine solche lineare Abbildung T gibt. (Diese Annahme wird sich als richtig bewähren, falls Gleichung (7.1) stimmt; sie wird verworfen werden müssen, wenn Gleichung (7.1) nicht stimmt.)

Es gilt dann in unserem Beispiel $T = \begin{pmatrix} 0 & 0 \end{pmatrix}$, weil wir ja oben gezeigt haben, dass $\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0$ ist.

Machen wir uns zunächst klar, wie die anderen beteiligten Objekte aussehen: f ist eine Funktion von $\mathbb{R}^2$ nach $\mathbb{R}$, demnach muss gelten: $h \in \mathbb{R}^2$. Schreiben wir also $h = (h_1, h_2)$.

H

Damit ist:

$$\frac{f(a+h)-f(a)-Th}{||h||} = \frac{f(h_1,h_2)-f(0,0)-\begin{pmatrix}0 & 0\end{pmatrix}\cdot\begin{pmatrix}h_1\\h_2\end{pmatrix}}{\sqrt{h_1^2+h_2^2}} = \frac{|h_1h_2|-0-0}{\sqrt{h_1^2+h_2^2}}.$$

Dieser Ausdruck ist jedenfalls immer größer gleich Null.

K

Um zeigen zu können, dass er gegen 0 konvergiert, werden wir ihn noch nach oben abschätzen:

H

Es gilt: $0 \leq \frac{|h_1h_2|}{\sqrt{h_1^2+h_2^2}} = \frac{|h_1|}{\sqrt{h_1^2+h_2^2}} \cdot |h_2| \leq 1 \cdot |h_2| \longrightarrow 0$ für $|h| \to 0$.

B

Unsere Annahme hat sich als richtig erwiesen, im Punkt $a = (0,0)$ ist die Funktion f also tatsächlich differenzierbar!

Exkurs 2

P

Angenommen, es hätte Folgendes gegolten: $\frac{f(a+h)-f(a)-Th}{||h||} = \ldots = \frac{|h_1|}{\sqrt{h_1^2+h_2^2}}$, wobei $T = \begin{pmatrix}\frac{\partial f}{\partial x}(a) & \frac{\partial f}{\partial y}(a)\end{pmatrix}$.

Dieser Ausdruck konvergiert nicht gegen 0, da z. B. für die Nullfolge $\{(h_1,h_2)\} = \{(\frac{1}{n},0)\}$ gilt:

$$\frac{|h_1|}{\sqrt{h_1^2+h_2^2}} = \frac{|\frac{1}{n}|}{\sqrt{\frac{1}{n^2}}} = \frac{\frac{1}{n}}{\frac{1}{n}} = 1 \neq 0.$$

In diesem Fall wäre also f in a *nicht* differenzierbar, obwohl beide partiellen Ableitungen in a existieren.

Verständnisfragen 7.4

1. Wo liegen die Punkte (x,y) mit $xy = 0$ in der x-y-Ebene? Diese Punkte waren bei der Differenzierbarkeitsuntersuchung ja gerade die interessanten.
2. An welcher Stelle der Lösung wird bewiesen, dass die partiellen Ableitungen in $(0,0)$ existieren?
3. Wieso dürfen wir einfach schreiben, dass $T = \begin{pmatrix}0 & 0\end{pmatrix}$ gilt? Wir wissen doch an dieser Stelle noch gar nicht, ob f an der Stelle $(0,0)$ differenzierbar ist.

Arbeitsauftrag 7.4 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.3.

7.5 Mehrdimensionale Kettenregel

Aufgabe 7.5 Sei $f : \mathbb{R}^2 \to \mathbb{R}$ eine zweimal differenzierbare Funktion mit

$$\frac{\partial f}{\partial x}(0,1) = \frac{\partial f}{\partial y}(0,1) = \frac{\partial^2 f}{\partial x^2}(0,1) = \frac{\partial^2 f}{\partial y^2}(0,1) = \frac{\partial^2 f}{\partial y \partial x}(0,1) = 1.$$

Sei $g(s,t) = f(s+t, e^{s+t} + t)$. Berechnen Sie $\frac{\partial^2 g}{\partial s \partial t}(0,0)$.

Ausführliche Lösung 7.5 Dies ist eine Aufgabe zur Kettenregel. Gegeben ist nämlich eine Funktion g mit $g(s,t) = f(s+t, e^{s+t} + t)$ und wir sollen g zweimal partiell ableiten.

Man kann also auch schreiben $g(s,t) = f(h(s,t))$, wobei

$$h(s,t) := \begin{pmatrix} x(s,t) \\ y(s,t) \end{pmatrix} := \begin{pmatrix} s+t \\ e^{s+t} + t \end{pmatrix}.$$

Damit wird noch deutlicher, dass g eine Verknüpfung von Funktionen ist, zu deren Ableitung man die Kettenregel verwenden muss. Zuerst sollen wir partiell nach s ableiten.

Wir beweisen dazu die „Formel“ $\frac{\partial g}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}$, die Sie sich auch für spätere Zwecke merken sollten.

Es gilt nach der Kettenregel: $g' = f' \cdot h'$, wobei $f' = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix}$ und $h' = \begin{pmatrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial t} \end{pmatrix}$.

Also: $g' = \begin{pmatrix} \frac{\partial g}{\partial s} & \frac{\partial g}{\partial t} \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial t} \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} & \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} \end{pmatrix}$.

Daraus liest man die obige Formel ab. Sie gilt übrigens auch allgemeiner für Funktionen $g : \mathbb{R}^n \to \mathbb{R}$, nämlich $\frac{\partial g}{\partial s} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial s} + \ldots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial s}$, wie man sich leicht überlegen kann.

Nun aber zurück zur Aufgabe:

$$\begin{aligned} \frac{\partial g}{\partial s}(s,t) &= \frac{\partial f}{\partial x}(x(s,t), y(s,t))\frac{\partial x}{\partial s}(s,t) + \frac{\partial f}{\partial y}(x(s,t), y(s,t))\frac{\partial y}{\partial s}(s,t) \qquad (7.2) \\ &= \frac{\partial f}{\partial x}(x(s,t), y(s,t)) \cdot 1 + \frac{\partial f}{\partial y}(x(s,t), y(s,t)) \cdot e^{s+t} \end{aligned}$$

Jetzt müssen wir dieses Ergebnis noch partiell nach t ableiten, um zu $\frac{\partial^2 g}{\partial t \partial s}$ zu kommen.

Dazu leiten wir die beiden Summanden separat ab, wobei $\frac{\partial f}{\partial x}$ und $\frac{\partial f}{\partial y}$ selbst wieder verknüpfte Funktionen von $\mathbb{R}^2$ nach $\mathbb{R}$ sind, für deren Ableitung also abermals die obige Formel angewandt werden kann. Beim zweiten Summanden ist außerdem die Produktregel anzuwenden:

$$\begin{aligned}\frac{\partial^2 g}{\partial t \partial s}(s,t) = &\frac{\partial^2 f}{\partial x^2}(x,y)\cdot 1 + \frac{\partial^2 f}{\partial y \partial x}(x,y)\cdot(e^{s+t}+1) \\ &+ \left(\frac{\partial^2 f}{\partial x \partial y}(x,y)\cdot 1 + \frac{\partial^2 f}{\partial y^2}(x,y)\cdot(e^{s+t}+1)\right)e^{s+t} + \frac{\partial f}{\partial y}(x,y)\cdot e^{s+t}.\end{aligned} \tag{7.3}$$

Bemerkung: Die ersten beiden Summanden rühren dabei vom ersten Summanden aus Gleichung (7.2) her, die letzten beiden Summanden vom zweiten Summanden aus Gleichung (7.2).

Jetzt müssen wir nur noch an der Stelle $(0,0)$ auswerten und ausnutzen, dass (zum Glück) $g(0,0) = f(0,1)$ gilt.
Wir erhalten:

$$\frac{\partial^2 g}{\partial t \partial s}(0,0) = 1\cdot 1 + 1\cdot 2 + 1\cdot 1 + 1\cdot 2 + 1\cdot 1 = 7.$$

Verständnisfragen 7.5

1. Was sind die Definitions- bzw. Wertebereiche der Funktionen x, y, h und g?
2. Wieso haben wir „Glück“, dass $g(0,0) = f(0,1)$ gilt?

Übung 7.5 Ordnen Sie nun selbst den einzelnen Bearbeitungsschritten der Musterlösung die Phasen aus Kapitel 3 zu! Tragen Sie dazu die jeweiligen Buchstaben in die kleinen Quadrate am Seitenrand ein!

Arbeitsauftrag 7.5 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen, der Übung und dem Arbeitsauftrag finden Sie in Abschnitt 11.3.

7.6 Jacobi- und Hesse-Matrix

Aufgabe 7.6 Bestimmen Sie die Hesse-Matrix der Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit

$$f(x, y) = 2x + y - xy + x^3,$$

ohne die partiellen Ableitungen zu berechnen! Hinweis: Verwenden Sie den nachstehenden Satz 7.1.

Satz 7.1 Sei $f : \mathbb{R}^n \to \mathbb{R}$ eine dreimal stetig differenzierbare Funktion, $A(x)$ eine $(1 \times n)$-Matrix und $B(x)$ eine symmetrische $(n \times n)$-Matrix. Angenommen, es gilt für $h \in \mathbb{R}^n$

$$f(x + h) = f(x) + A(x)h + \frac{1}{2}h^T B(x)h + r(h),$$

wobei $\lim\limits_{h \to 0} \frac{r(h)}{||h||^2} = 0$. Dann gilt: $A(x) = Df(x)$ und $B(x) = D^2 f(x)$.

Ausführliche Lösung 7.6 In dieser Aufgabe lernen Sie eine Methode kennen, wie Sie aus Polynomen in mehreren Veränderlichen relativ einfach die Jacobi- bzw. die Hesse-Matrix ablesen können. Wir sollen dazu den obigen Satz verwenden.

f ist als Polynom natürlich dreimal stetig differenzierbar. Diese Voraussetzung des Satzes ist also schon einmal erfüllt.

Um den Satz anwenden zu können, muss zuerst $f(x + h)$ (hier also $f(x + h_1, y + h_2)$) berechnet werden und anschließend sortiert man die Terme danach, ob sie konstant, linear oder quadratisch in h sind. Führen wir dies durch:

$$\begin{aligned}
f(x + h_1, y + h_2) &= 2(x + h_1) + y + h_2 - (x + h_1)(y + h_2) + (x + h_1)^3 \\
&= 2x + 2h_1 + y + h_2 - xy - xh_2 - yh_1 - h_1h_2 + x^3 + 3x^2h_1 + 3xh_1^2 + h_1^3 \\
&= 2x + y - xy + x^3 + (2 - y + 3x^2)h_1 + (1 - x)h_2 - h_1h_2 + 3xh_1^2 + h_1^3 \\
&= f(x, y) + (2 - y + 3x^2)h_1 + (1 - x)h_2 - h_1h_2 + 3xh_1^2 + h_1^3.
\end{aligned}$$

Als Nächstes können wir statt $(2 - y + 3x^2)h_1 + (1 - x)h_2$ schreiben:

$$\begin{pmatrix} 2 - y + 3x^2 & 1 - x \end{pmatrix} \cdot \begin{pmatrix} h_1 \\ h_2 \end{pmatrix},$$

also in der Form $A(x)h$ wie im Satz gefordert.

Jetzt müssen wir noch die quadratischen Terme $-h_1h_2 + 3xh_1^2$ in der Form

$$\frac{1}{2}h^T B(x)h$$

schreiben, wobei B eine symmetrische (2×2)-Matrix ist.

Dazu berechnen wir zunächst für eine allgemeine symmetrische Matrix $B(x)$:

$$\begin{aligned}\frac{1}{2}h^T B(x)h &= \frac{1}{2}\cdot\begin{pmatrix}h_1 & h_2\end{pmatrix}\cdot\begin{pmatrix}b_{11} & b_{12}\\ b_{12} & b_{22}\end{pmatrix}\cdot\begin{pmatrix}h_1\\ h_2\end{pmatrix}\\ &= \frac{1}{2}\cdot\begin{pmatrix}h_1 & h_2\end{pmatrix}\cdot\begin{pmatrix}b_{11}h_1+b_{12}h_2\\ b_{12}h_1+b_{22}h_2\end{pmatrix}\\ &= \frac{1}{2}\cdot(b_{11}h_1^2+b_{12}h_1h_2+b_{12}h_1h_2+b_{22}h_2^2)\\ &= \frac{b_{11}}{2}h_1^2+b_{12}h_1h_2+\frac{b_{22}}{2}h_2^2.\end{aligned}$$

Durch Koeffizientenvergleich mit $-h_1h_2+3xh_1^2$ erhalten wir: $b_{11}=6x$, $b_{12}=-1$ und $b_{22}=0$.

Es gilt also:

$$-h_1h_2+3xh_1^2=\frac{1}{2}\cdot\begin{pmatrix}h_1 & h_2\end{pmatrix}\cdot\begin{pmatrix}6x & -1\\ -1 & 0\end{pmatrix}\cdot\begin{pmatrix}h_1\\ h_2\end{pmatrix}.$$

Jetzt sind wir fast fertig. Wir müssen nur noch – wie im Satz gefordert – überprüfen, ob $\frac{r(h)}{||h||^2}\to 0$ für $h\to 0$.

Bei uns ist der Restterm $r(h)=h_1^3$.

Durch eine Abschätzung nach oben erhalten wir:

$$\left|\frac{h_1^3}{h_1^2+h_2^2}\right|=\frac{h_1^2}{h_1^2+h_2^2}\cdot|h_1|\le 1\cdot|h_1|\xrightarrow{h\to 0}0.$$

Damit ist also $Df(x,y)=\begin{pmatrix}2-y+3x^2 & 1-x\end{pmatrix}$ (was gar nicht gefragt war) und

$$D^2f(x,y)=\begin{pmatrix}6x & -1\\ -1 & 0\end{pmatrix}.$$

Verständnisfragen 7.6

1. Warum sortieren wir die Terme zu Beginn der Lösung gerade so, dass für $r(h)$ nur Terme der Ordnung ≥ 3 übrigbleiben? Könnte es nicht sein, dass auch in $r(h)$ konstante, lineare oder quadratische Terme vorkommen?
2. Sie können sich das formale Berechnen des Ausdrucks $\frac{1}{2}h^T B(x)h$ in zukünftigen Aufgaben ersparen. Was könnten Sie sich – nachdem Sie jetzt die Aufgabenlösung kennen – als Merkregel zum Ablesen der Hesse-Matrix aus $f(x+h)$ einprägen?
3. Welche Rolle spielt das Symbol x im Satz, welche Rolle in unserer Aufgabe?

Übung 7.6 Ordnen Sie nun selbst den einzelnen Bearbeitungsschritten der Musterlösung die Phasen aus Kapitel 3 zu! Tragen Sie dazu die jeweiligen Buchstaben in die kleinen Quadrate am Seitenrand ein!

Arbeitsauftrag 7.6 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen, der Übung und dem Arbeitsauftrag finden Sie in Abschnitt 11.3.

7.7 Lokale Extremstellen

Aufgabe 7.7 Bestimmen Sie alle Punkte, in denen die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x,y) = y^2 - yx^2$ ein lokales Extremum annimmt.

Ausführliche Lösung 7.7 Wir sollen in dieser Aufgabe alle Punkte finden, in denen die Funktion f mit $f(x,y) = y^2 - yx^2$ ein lokales Extremum annimmt. B

Dazu berechnen wir zunächst die Jacobi-Matrix und setzen sie gleich Null, um mögliche Kandidaten für Extremalstellen zu finden: K

$$Df(x,y) = \begin{pmatrix} -2xy & 2y - x^2 \end{pmatrix} \overset{!}{=} \begin{pmatrix} 0 & 0 \end{pmatrix}.$$ H

Aus der ersten Gleichung erhalten wir sofort, dass entweder $x = 0$ oder $y = 0$ gelten muss. Aus der zweiten Gleichung bekommen wir, dass dann auch die jeweils andere Variable den Wert Null annehmen muss.

Der einzige kritische Punkt von f ist also $(x_0, y_0) := (0,0)$. B

An dieser Stelle können wir versuchen, das folgende Eigenwertkriterium zu verwenden. K

Eigenwertkriterium: Seien $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ zweimal stetig differenzierbar und $x_0 \in U$ ein kritischer Punkt. Dann gilt:

i) Sind alle Eigenwerte der Hesse-Matrix $H = D^2 f(x_0)$ echt positiv, so besitzt f in x_0 ein lokales Minimum.

ii) Sind alle Eigenwerte von H echt negativ, so besitzt f in x_0 ein lokales Maximum.

iii) Hat H sowohl mindestens einen echt positiven als auch mindestens einen echt negativen Eigenwert, so hat f in x_0 einen Sattelpunkt.

Wir berechnen also in unserer Aufgabe die Hesse-Matrix $H = D^2 f(x,y)$ und setzen den kritischen Punkt $(0,0)$ ein: B

$$D^2 f(x,y) = \begin{pmatrix} -2y & -2x \\ -2x & 2 \end{pmatrix},$$ A

$$D^2 f(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}.$$

Man kann sofort ablesen (Diagonalmatrix), dass $D^2 f(0,0)$ die Eigenwerte 2 und 0 hat.

K Mit dem Eigenwertkriterium kann man also nichts über die Qualität des kritischen Punktes aussagen[5]. Wir werden daher die Funktion f selbst genauer unter die Lupe nehmen.

Entweder wir können beweisen, dass die Funktion f entlang *jeder beliebigen* Kurve ausgehend vom kritischen Punkt $(0,0)$ lokal wächst (bzw. fällt). Dann muss nämlich in $(0,0)$ ein lokales Minimum (bzw. Maximum[6]) vorliegen (es reicht nicht, nur alle „geradlinigen Richtungen" zu betrachten). Oder wir können nachweisen, dass es sowohl (mindestens) eine Kurve gibt, entlang derer f wächst *und* (mindestens) eine Kurve, entlang derer f fällt. Dann besitzt f im Punkt $(0,0)$ einen Sattelpunkt. Wenn man gar keine Idee hat, was im vorliegenden Beispiel eher zutrifft, dann sollte man eventuell zuerst versuchen zu prüfen, ob f in $(0,0)$ einen Sattelpunkt hat (das ist meistens leichter). Manchen Funktionen sieht man aber auch schon an, welche Art von kritischer Punkt vorliegt.

Wir werden nun versuchen, zwei solche Richtungen zu finden, eine entlang derer f wächst und eine entlang derer f fällt. Es ist zunächst ratsam, ganz einfache Richtungen zu untersuchen, z. B. entlang der beiden Achsen.

H

$$f(x,0) = 0$$

B Entlang der x-Achse ist die Funktion also konstant gleich Null. Das hilft uns nicht.

H

$$f(0,y) = y^2$$

P Entlang der y-Achse wächst die Funktion ausgehend von $(0,0)$ in beide Richtungen. Falls f in $(0,0)$ ein lokales Extremum hat, dann kann es also nur noch ein Minimum sein. Wir versuchen aber jetzt, vielleicht doch eine Richtung zu finden, entlang derer f fällt.

K Die nächstschwierigere Möglichkeit ist, sich entlang einer Geraden ausgehend vom kritischen Punkt $(0,0)$ zu bewegen, z. B. entlang der 1. Mediane $y = x$:

H

$$f(x,x) = x^2 - x^3.$$

B Diese Funktion wächst aber (leider) für kleine x. Das ist für unsere Zwecke unbrauchbar, denn so eine Kurve haben wir ja schon!

P Als Nächstes könnten Sie auch andere Geraden probieren, z. B. $y = 2x$ oder $y = \frac{1}{2}x$ oder allgemein $y = ax$. All das liefert Ihnen aber nicht das Gewünschte.

K Wir untersuchen daher f entlang der nächstschwierigeren Kurven: $y = ax^2$, also entlang von Parabeln ausgehend von $(0,0)$.

T Wie wir den Parameter a wählen, halten wir uns noch offen.

$$f(x,ax^2) = a^2x^4 - ax^2x^2 = (a^2 - a)x^4$$

[5] Übrigens wäre hier auch das Hauptminorenkriterium, das Sie vielleicht aus Ihrer Vorlesung kennen, nicht anwendbar.

[6] In unserem Fall ist $D^2f(0,0)$ positiv semidefinit, d. h. falls in $(0,0)$ überhaupt ein lokales Extremum vorliegt, so muss es sich um ein lokales Minimum handeln.

Diese Wahl war auch deshalb geschickt, weil sich die beiden Terme jetzt schön zusammenfassen ließen (auch darauf kann man achten, wenn man versucht, günstige Kurven zu finden). P

Jetzt sind wir aber schon fast fertig. B

Für $a \in (0,1)$ gilt nämlich $f(x, ax^2) = (a^2 - a)x^4 < 0$. D. h. f fällt ausgehend von $(0,0)$ entlang jeder Parabel $y = ax^2$ mit $a \in (0,1)$ (siehe Abbildung 7.4). H

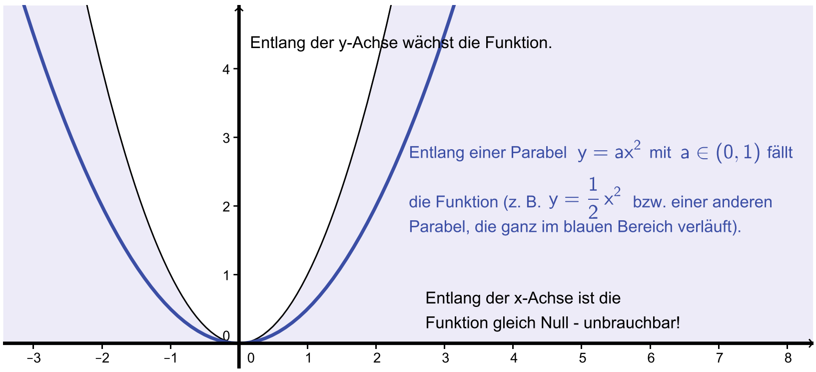

Abbildung 7.4: Kurven, entlang derer die Funktion f wächst bzw. fällt

Damit haben wir also wie gewünscht eine solche Richtung gefunden und sind fertig. An der Stelle $(0,0)$ besitzt die Funktion f also einen Sattelpunkt. B

Verständnisfragen 7.7

1. Wieso kann man mit dem Eigenwertkriterium in dieser Aufgabe keine Aussage über die Qualität des kritischen Punktes machen?
2. Es wird behauptet, dass die Funktion $f(x,x) = x^2 - x^3$ für kleine x wächst. Bis zu welchen Werten von x wächst sie – jeweils ausgehend von $x = 0$ – in beide Richtungen?
3. Wir haben festgestellt, dass $f(a, ax^2)$ für $a \in (0,1)$ ausgehend von $(x,y) = (0,0)$ fällt. Was passiert, wenn man $a > 1$ wählt? Wie könnte man dadurch die Lösung abkürzen?

Arbeitsauftrag 7.7 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.3.

7.8 Lokale Umkehrbarkeit

Aufgabe 7.8 Wir betrachten das Gleichungssystem

$$\begin{aligned} x^3 + y &= s \\ -y^3 + x &= t \end{aligned}$$

mit den Unbekannten $(x, y) \in \mathbb{R}^2$ und den gegebenen Parametern $(s, t) \in \mathbb{R}^2$.

a) Beweisen Sie, dass das Gleichungssystem für jedes Paar $(s, t) \in \mathbb{R}^2$ eine eindeutige Lösung $x := x(s, t)$, $y := y(s, t)$ besitzt!

b) Beweisen Sie, dass die Funktionen $x(s, t)$ und $y(s, t)$ stetig differenzierbar auf $\mathbb{R}^2$ sind und berechnen Sie $x'(0, 2)$ und $y'(0, 2)$.

P

Ausführliche Lösung 7.8 In dieser Aufgabe soll deutlich werden, dass man über Lösungen von Gleichungssystemen Aussagen machen kann, obwohl man diese Lösungen gar nicht explizit berechnen kann! In Teil a) machen wir eine Aussage über die Existenz und die Eindeutigkeit von Lösungen eines bestimmten Gleichungssystems, in Teil b) eine Aussage über die Ableitung der Lösungsfunktionen $x(s, t)$ und $y(s, t)$ an einer bestimmten Stelle.

K

a) Am einfachsten wäre diese Aufgabe zu lösen, wenn man aus dem Gleichungssystem

$$\begin{aligned} x^3 + y &= s \\ -y^3 + x &= t \end{aligned} \tag{7.4}$$

direkt die Lösung (in Abhängigkeit von s und t) angeben, also einfach „ausrechnen" könnte. Dann würden wir nämlich sehen, dass es für jede Wahl von s und t eine Lösung gibt und dass diese eindeutig ist. Warum das nicht geht, sehen wir gleich:

H

Drücken wir nämlich aus der ersten Gleichung y explizit aus

$$y = s - x^3 \tag{7.5}$$

und setzen das in die zweite Gleichung ein, so erhalten wir

$$-(s - x^3)^3 + x = t, \tag{7.6}$$

P

also eine Gleichung *neunten* Grades in x. Und solche Gleichungen kann man im Allgemeinen nicht lösen (für ganz bestimmte s und t kann man allerdings eine Lösung erraten, wie wir weiter unten sehen werden).

B

Man kann also keine Lösungsfunktionen $x(s, t)$ und $y(s, t)$ angeben, die uns für jedes Paar (s, t) die zugehörige Lösung (x, y) des Gleichungssystems liefern.

Aber wir können trotzdem eine Aussage über die Existenz und die Eindeutigkeit der Lösung machen.

Dazu definieren wir eine Funktion f durch $f(x) := -(s - x^3)^3 + x$ und zeigen, dass diese Funktion jeden Wert in $\mathbb{R}$ genau einmal annimmt, also auch den Wert t wie in Gleichung (7.6) gewünscht.

Differenzieren liefert $f'(x) = -3(s - x^3)^2 \cdot (-3x^2) + 1$. Es gilt $f'(x) > 0$, wie man leicht einsieht. Außerdem ist f eine Polynomfunktion ungeraden Grades und damit gilt: $\lim\limits_{x\to+\infty} f(x) = +\infty$ und $\lim\limits_{x\to-\infty} f(x) = -\infty$. Aus diesen Eigenschaften folgt, dass f für jeden Wert s bijektiv ist und insbesondere jeden Wert t genau einmal annimmt.
Es gibt also zu jedem Wertepaar (s, t) ein eindeutig bestimmtes x.

Nachdem $y = s - x^3$ gilt (siehe Gleichung (7.5)), ergibt sich daraus auch ein eindeutig bestimmtes y und wir sind fertig.

b) In diesem Teil ist die Rede von Funktionen $x(s, t)$ und $y(s, t)$. Das „riecht“ nach einer Umkehrfunktion der durch Gleichungssystem (7.4) gegebenen „Funktion“ (die aber noch gar nicht wirklich die Gestalt einer Funktion hat) sowie nach dem Satz über die lokale Umkehrbarkeit von Funktionen.

Satz 7.2 (lokale Umkehrbarkeit) Sei $U \subseteq \mathbb{R}^n$ offen und $a \in U$, $f : U \to \mathbb{R}^n$ sei stetig differenzierbar und es sei $\det Df(a) \neq 0$. Sei außerdem $b = f(a)$. Dann existieren offene Mengen $V, W \subseteq \mathbb{R}^n$ mit $a \in V$, $b \in W$, sodass $f : V \to W$ bijektiv und die zu f inverse Funktion $g : W \to V$ stetig differenzierbar ist. Es gilt $Dg(y) = [Df(g(y))]^{-1}$.

Um diesen Satz anwenden zu können, lohnt es sich zu definieren: $f : \mathbb{R}^2 \to \mathbb{R}^2$ mit

$$f(x, y) := \begin{pmatrix} x^3 + y \\ -y^3 + x \end{pmatrix}. \tag{7.7}$$

Die Funktion f ist stetig differenzierbar (da die partiellen Ableitungen existieren und stetig sind). Außerdem ist f bijektiv, wie wir aus Teil a) folgern können. Um Aussagen über die Ableitung der Umkehrfunktion f^{-1} machen zu können, berechnen wir zunächst die Jacobi-Matrix:

$$Df(x, y) = \begin{pmatrix} 3x^2 & 1 \\ 1 & -3y^2 \end{pmatrix}. \tag{7.8}$$

Es gilt: $\det Df(x, y) = -9x^2y^2 - 1 < 0$, also ist sie insbesondere verschieden von Null.

Damit ist nach dem Satz über die lokale Umkehrbarkeit f^{-1} mit $f^{-1}(s, t) = \begin{pmatrix} x(s,t) \\ y(s,t) \end{pmatrix}$ eine stetig differenzierbare Funktion und $Df^{-1}(s, t) = (Df(x(s, t), y(s, t)))^{-1}$.

Wir wollen die Ableitung von f^{-1} gerne konkret für den Punkt $(s, t) = (0, 2)$ berechnen.

Dazu müssen wir $Df(x, y)$ mit $x = x(0, 2)$ und $y = y(0, 2)$ kennen. Wir müssen also zunächst x und y für das Gleichungssystem (7.4) mit $s = 0$ und $t = 2$ lösen.

Das klappt in diesem einfachen Fall auch:

$$\begin{aligned} x^3 + y &= 0 \\ -y^3 + x &= 2. \end{aligned} \tag{7.9}$$

Formt man die erste Gleichung nach $y = -x^3$ um und setzt dies in die zweite ein, so erhält man: $x^9 + x = 2$. Daraus liest man leicht die (eindeutige) Lösung $x = 1$ ab. y ergibt sich zu $y = -1^3 = -1$.

Wir können jetzt $Df(1, -1)$ berechnen:

$$Df(1,-1) = \begin{pmatrix} 3 & 1 \\ 1 & -3 \end{pmatrix}.$$

Damit erhalten wir

$$Df^{-1}(0,2) = \begin{pmatrix} x'(0,2) \\ y'(0,2) \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 1 & -3 \end{pmatrix}^{-1} = -\frac{1}{10} \begin{pmatrix} -3 & -1 \\ -1 & 3 \end{pmatrix}$$

und sind fast fertig.

Man liest nämlich $x'(0,2) = \begin{pmatrix} \frac{3}{10} & \frac{1}{10} \end{pmatrix}$ und $y'(0,2) = \begin{pmatrix} \frac{1}{10} & -\frac{3}{10} \end{pmatrix}$ ab.

Verständnisfragen 7.8

1. Ist jede streng monoton wachsende Funktion $f : \mathbb{R} \to \mathbb{R}$ mit

   $$\lim_{x \to +\infty} f(x) = +\infty$$

   und

   $$\lim_{x \to -\infty} f(x) = -\infty$$

   automatisch bijektiv?

2. Wieso ist zwar im Satz über die lokale Umkehrbarkeit die Rede von Umgebungen V und W, jedoch nicht in unserer Anwendung des Satzes? Wieso dürfen wir sofort folgern, dass f^{-1} auf *ganz* $\mathbb{R}^2$ stetig differenzierbar ist?

3. Stellen Sie sich vor, *Sie* hätten diese Aufgabe gestellt. Welche Punkte hätte man statt des Punktes $(0, 2)$ z. B. noch angeben können, damit man die Aufgabe in ähnlicher Weise lösen kann? D. h. für welche Punkte (s, t) kann man eine Lösung des Gleichungssystems leicht erraten? Geben Sie drei Beispiele an!

Arbeitsauftrag 7.8 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.3.

7.9 Implizite Funktionen

Aufgabe 7.9 a) Beweisen Sie: Es gibt eine Zahl $\eta > 0$, sodass das Gleichungssystem

$$\begin{aligned} x + t\sin(x+y) &= 0 \\ y + t\cos(xy) &= 0 \end{aligned}$$

für jeden Parameter $t \in (-\eta, \eta)$ eine Lösung besitzt. Außerdem ist $(x(t), y(t))$ stetig differenzierbar auf $(-\eta, \eta)$.

b) Bestimmen Sie $x'(0)$ und $y'(0)$.

Ausführliche Lösung 7.9 Diese Aufgabe ist eine paradigmatische Aufgabe zum Satz über implizite Funktionen, den wir an dieser Stelle formulieren:

K

Satz 7.3 (implizite Funktionen) Sei $D \subseteq \mathbb{R}^{n+m}$ offen und $f : D \to \mathbb{R}^n$ sei stetig differenzierbar. Sei $(a,b) \in \mathbb{R}^n \times \mathbb{R}^m$ mit $f(a,b) = 0$ und mit

$$\det \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a,b) & \dots & \frac{\partial f_1}{\partial x_n}(a,b) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(a,b) & \dots & \frac{\partial f_n}{\partial x_n}(a,b) \end{pmatrix} \neq 0.$$

Dann gibt es eine Umgebung $U \subseteq \mathbb{R}^m$ von b und eine eindeutig bestimmte stetig differenzierbare Abbildung $\varphi: U \to \mathbb{R}^n$ mit $\varphi(b) = a$ und $f(\varphi(y), y) = 0$ für alle $y \in U$.

P

Zwar ist in unserer Aufgabe gar keine Funktion gegeben, sondern ein Gleichungssystem, diesen Trick kennen Sie jetzt aber schon. Wir können aus dem Gleichungssystem ja ganz leicht eine passende Funktion *machen*:

A

Wir definieren also:

$$f(x,y,t) := \begin{pmatrix} x + t\sin(x+y) \\ y + t\cos(xy) \end{pmatrix}.$$

P

Am Ende wollen wir gerne (siehe Aufgabenstellung) Funktionen $x(t)$ und $y(t)$ bekommen. D. h. von den insgesamt drei Variablen (x, y und t) soll am Ende nur noch t frei wählbar sein (in einer Umgebung von Null, also in $(-\eta, \eta)$), während x und y von t abhängig sein sollen.

Es ist wichtig, dass Sie sich zu Beginn immer überlegen, welche Variablen am Ende frei und welche abhängig sein sollen (in der Formulierung des Satzes sind übrigens die Variablen $y_1, y_2, \cdots, y_m$ frei und die Variablen $x_1, x_2, \cdots, x_n$ abhängig).

B

a) Um den Satz anwenden zu können, müssen wir zunächst einen Punkt (x,y,t) finden, der $f(x,y,t) = 0$ erfüllt.

A

Nachdem wir laut Aufgabenstellung schon wissen, dass wir t später aus einer Umgebung von Null wählen wollen, suchen wir also eigentlich einen Punkt $(x,y,0)$ mit

$f(x,y,0) = 0$. Das ist aber einfach – es funktioniert nur der Punkt $(x,y,t) = (0,0,0)$, wie man sieht, wenn man $t = 0$ in den Funktionsterm einsetzt.

B

Als Nächstes müssen wir prüfen, ob die Jacobi-Determinante – die entsteht, wenn man f ausschließlich nach den abhängigen Variablen x und y ableitet – in diesem Punkt ungleich Null ist:

H

$$D_{(x,y)}f(x,y,t) = \begin{pmatrix} 1 + t\cos(x+y) & t\cos(x+y) \\ -t\sin(xy)\cdot y & 1 - t\sin(xy)\cdot x \end{pmatrix},$$

A

$$\det D_{(x,y)}f(0,0,0) = \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \neq 0.$$

H

Damit können wir den Satz also anwenden: Es gibt eine Umgebung $U = (-\eta, \eta)$ von $t = 0$ und eine eindeutig bestimmte stetig differenzierbare Funktion $\varphi(t) = (x(t), y(t))$ mit $\varphi(0) = (x(0), y(0)) = (0,0)$ und $f(\varphi(t), t) = f(x(t), y(t), t) = 0$ auf U.

B

Das Gleichungssystem ist demnach für jeden Wert $t \in (-\eta, \eta)$ lösbar und die Lösungsfunktion $\varphi(t) = (x(t), y(t))$ ist wie gewünscht stetig differenzierbar.

B

b) Nach Teil a) gilt:

$$\begin{aligned} x(t) + t\sin(x(t) + y(t)) &= 0 \\ y(t) + t\cos(x(t)\cdot y(t)) &= 0. \end{aligned}$$

K

Wir differenzieren diese beiden Gleichungen implizit nach t und erhalten:

H

$$\begin{aligned} x'(t) + \sin(x(t) + y(t)) + t\cos(x(t) + y(t))\cdot(x'(t) + y'(t)) &= 0 \\ y'(t) + \cos(x(t)\cdot y(t)) - t\sin(x(t)\cdot y(t))\cdot(x(t)\cdot y(t))' &= 0. \end{aligned}$$

K

Jetzt brauchen wir nur noch $t = 0$ einzusetzen, alles auszuwerten und die Gleichungen nach $x'(0)$ bzw. $y'(0)$ umzustellen:

H

$$\begin{aligned} x'(0) + \sin(0+0) + 0 &= 0 \\ y'(0) + \cos(0\cdot 0) - 0 &= 0. \end{aligned}$$

Wir erhalten: $x'(0) = 0$ und $y'(0) = -1$ und sind fertig.

Verständnisfragen 7.9

1. Warum erfüllt in Teil a) nur der Punkt $(0,0,0)$ die Bedingung $f(x,y,0) = 0$? Und was wäre, wenn es noch weitere Punkte gegeben hätte, die diese Bedingung erfüllen?
2. Wieso haben wir beim impliziten Differenzieren der zweiten Zeile in Teil b) den Ausdruck $(x(t)y(t))'$ nicht weiter ausgerechnet?

Arbeitsauftrag 7.9 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.3.

8 Musterlösungen aus der Linearen Algebra 1

8.1 Unterraumkriterium

Aufgabe 8.1 Überprüfen Sie, ob die Mengen

$$U_1 := \{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \mid \sum_{i=1}^{n-1} x_i = x_n\} \text{ und } U_2 := \{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \mid \sum_{i=1}^{n} x_i = 1\}$$

Unterräume des $\mathbb{R}^n$ sind!

B

Ausführliche Lösung 8.1 Um die Frage beantworten zu können, müssen wir wissen, was ein Unterraum ist. Schauen wir uns dazu die Definition an:

Definition Sei V ein K-Vektorraum, wobei K ein Körper ist. Eine Teilmenge $U \subseteq V$ heißt genau dann Unterraum von V, wenn folgende drei Bedingungen erfüllt sind:

(i) $U \neq \emptyset$.

(ii) Für alle $u \in U$ und alle $a \in K$ gilt: $a \cdot u \in U$.

(iii) Für alle $u, u' \in U$ gilt: $u + u' \in U$.

Wir müssen also überprüfen, ob *die Bedingungen (i)-(iii)* in den gegebenen Beispielen erfüllt sind oder nicht. Aber reicht das? Nein, wir dürfen nicht vergessen zu überprüfen, ob auch die *Voraussetzungen der Definition* erfüllt sind!

K

Es ergeben sich folgende Strategien, um zu prüfen, ob eine Menge ein Unterraum ist oder nicht.

- U ist ein Unterraum, wenn wir beweisen können, dass sowohl *alle* Voraussetzungen als auch *alle* Bedingungen der Definition erfüllt sind.
- U ist kein Unterraum, wenn wir beweisen können, dass *mindestens ein* Teil der Definition (sprich eine Voraussetzung oder eine Bedingung) nicht erfüllt ist.

H

Die Voraussetzungen der Definition sind sowohl für U_1 als auch für U_2 erfüllt.

B

Wir beginnen mit U_1 und untersuchen, ob die Bedingungen (i)-(iii) erfüllt sind oder nicht.

C. Ableitinger, A. Herrmann, *Lernen aus Musterlösungen zur Analysis und Linearen Algebra*,
DOI 10.1007/978-3-658-02353-9_8, © Springer Fachmedien Wiesbaden 2013

Z Um ein Gespür für die Menge zu erhalten, ist es sinnvoll, sich zunächst ein paar Elemente zu überlegen, die in der Menge liegen. Die Menge U_1 besteht aus Elementen des $\mathbb{R}^n$, also aus Vektoren $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, die durch folgende Eigenschaft charakterisiert sind: Die Summe der ersten $n-1$ Komponenten muss gerade die n-te Komponente ergeben. Diese Eigenschaft erfüllen z. B. die folgenden Vektoren:

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ \vdots \\ 1 \\ n-1 \end{pmatrix}, \begin{pmatrix} -5 \\ \vdots \\ -5 \\ (n-1)\cdot(-5) \end{pmatrix}, \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

K Diese Beispiele liefern keine Gegenbeispiele zu den Bedingungen (i)-(iii), also liegt die Vermutung nahe, dass es sich tatsächlich um einen Unterraum handeln könnte.

B Um sicherzugehen, müssen wir das jetzt beweisen. Das heißt, wir beweisen nacheinander, dass die drei Bedingungen erfüllt sind:

H *Zu (i):* Die Bedingung ist natürlich erfüllt, da wir uns oben ja schon Elemente überlegt haben, die in U_1 liegen. Also ist U_1 nicht leer.

A *Zu (ii):* Wir „übersetzen" die allgemeine Definition für unseren speziellen Fall: Ist $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in U_1$, dann muss auch $a \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ für alle $a \in \mathbb{R}$ in U_1 liegen.

B Beweisen wir das nun:

Z Seien $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in U_1$ und $a \in \mathbb{R}$ beliebig.

Wir zeigen: $a \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in U_1$.

P Wir müssen also überprüfen, ob $a \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ die charakterisierende Eigenschaft von U_1 erfüllt.

A Da $a \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} ax_1 \\ \vdots \\ ax_n \end{pmatrix}$ ist, müssen wir also $\sum_{i=1}^{n-1} (ax_i) = ax_n$ beweisen.

Machen wir das: H

$$\sum_{i=1}^{n-1}(ax_i) \overset{\text{Distr.}}{=} a \cdot \left(\sum_{i=1}^{n-1} x_i\right) \overset{\text{Vor.}}{=} ax_n. \tag{8.1}$$

Wir haben also bekannte Rechenregeln (hier: Distributivgesetz) und die charakterisierende Eigenschaft von U_1 verwendet, um zu beweisen, dass auch $a \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ diese Eigenschaft besitzt. P

Zu (iii): Hier müssen wir zeigen: Sind $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ Elemente von U_1, so liegt auch $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ in U_1. K

Wir gehen wie oben vor: B

Seien $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in U_1$ beliebig. Z

Da $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$ ist, müssen wir $\sum_{i=1}^{n-1}(x_i + y_i) = x_n + y_n$ zeigen: A

$$\sum_{i=1}^{n-1}(x_i + y_i) \overset{\text{Komm. }+}{=} \sum_{i=1}^{n-1} x_i + \sum_{i=1}^{n-1} y_i \overset{\text{Vor.}}{=} x_n + y_n.$$ H

Zusammenfassung: Wir haben nun bewiesen, dass sowohl alle Voraussetzungen als auch alle Bedingungen der Definition „Unterraum“ erfüllt sind. Daraus können wir schließen, dass U_1 ein Unterraum des $\mathbb{R}^n$ ist. B

Kommen wir damit zur zweiten Menge U_2 und untersuchen, ob es sich dabei um einen Unterraum handelt. B

Auch hier versuchen wir zunächst eine Vorstellung von der Menge zu bekommen, indem wir nach Elementen suchen, die in U_2 enthalten sind. Das sind z. B. folgende: Z

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}; \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}; \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \\ \vdots \\ 0 \end{pmatrix}; \begin{pmatrix} \frac{1}{n} \\ \vdots \\ \frac{1}{n} \end{pmatrix}.$$

B Man hat jetzt drei Möglichkeiten um weiter vorzugehen:

K *1.:* Man sieht anhand der überlegten Beispiele schon, dass U_2 kein Unterraum ist.

H Denn z. B. liegt $\begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix} + \begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix} = \begin{pmatrix}1\\1\\0\\\vdots\\0\end{pmatrix}$ nicht in U_2, da $1+1+0+\ldots+0 = 2 \neq 1$. Oder auch $n \cdot \begin{pmatrix}\frac{1}{n}\\\vdots\\\frac{1}{n}\end{pmatrix} = \begin{pmatrix}1\\\vdots\\1\end{pmatrix} \notin U_2$, weil $1+\ldots+1 = n \neq 1$ (wenn $n > 1$ ist).

B Falls man das nicht sieht, geht man folgendermaßen vor:

K *2.:* Da jeder Unterraum den Nullvektor des Vektorraumes enthalten muss, testen wir, ob das im gegebenen Beispiel der Fall ist.

H $\begin{pmatrix}0\\\vdots\\0\end{pmatrix}$ liegt nicht in U_2, da $0+\ldots+0=0\neq 1$ ist. Daher kann U_2 kein Unterraum sein.

P *Bemerkung:* Es gibt aber auch Mengen, in denen der Nullvektor zwar liegt, die aber trotzdem keine Unterräume sind. In solchen Fällen hilft Strategie 1 oder 3.

K *3.:* Wir versuchen die drei Bedingungen nachzuweisen. Gelingt es, dann liegt ein Unterraum vor. Scheitert der Beweis an einer Stelle, können wir daraus ein Gegenbeispiel generieren.

B Wir werden diese dritte Strategie auf U_2 anwenden und tun jetzt einfach mal so, als wüssten wir noch nicht, dass $\begin{pmatrix}0\\\vdots\\0\end{pmatrix}$ nicht in U_2 liegt und damit U_2 kein Unterraum ist.

Wir beginnen mit Bedingung (ii) (genauso könnten wir mit Bedingung (iii) beginnen. Dass Bedingung (i) erfüllt ist, haben wir schon gezeigt). Das machen wir genauso, wie wir das auch für U_1 schon gemacht haben:

Z Seien $\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix} \in U_2$ und $a \in \mathbb{R}$ beliebig.

Wir zeigen: $a \cdot \begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix} \in U_2$.

Dazu müssen wir zeigen, dass $a \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ die charakterisierende Eigenschaft von U_2 besitzt. Da $a \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} ax_1 \\ \vdots \\ ax_n \end{pmatrix}$ ist, müssen wir also $\sum_{i=1}^{n}(ax_i) = 1$ beweisen.

A

Machen wir das:

H

$$
\begin{aligned}
\sum_{i=1}^{n}(ax_i) &\overset{\text{Distr.}}{=} a\Big(\sum_{i=1}^{n} x_i\Big) \\
&\overset{\text{Vor.}}{=} a \cdot 1 \\
&= a.
\end{aligned}
$$

Und jetzt sehen wir, dass für $a \neq 1$ das Element $a \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ nicht in U_2 liegt.

Daraus kann man sich nun ein Gegenbeispiel basteln. Das macht man, weil dadurch der Beweis kürzer und eleganter wird als die Überlegungen, die wir bis hierher angestellt haben. Das sieht dann z. B. so aus:

B

Behauptung: U_2 ist kein Unterraum des $\mathbb{R}^n$.

H

Beweis: $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in U_2$, da $1 + 0 + \ldots + 0 = 1$ ist. Und $2 \in \mathbb{R}$. Aber

$$
2 \cdot \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}
$$

liegt nicht in U_2, da $2 + 0 + \ldots + 0 = 2 \neq 1$ ist. Damit ist Bedingung (ii) nicht erfüllt und U_2 ist kein Unterraum.

Bemerkung: Um Bedingung (ii) zu widerlegen, reicht es also *ein* Element u aus U_2 und *ein* Element a aus $\mathbb{R}$ zu finden, für die $a \cdot u \in U_2$ *nicht* gilt.

P

Zusammenfassung der Strategien um zu testen, ob es sich bei einer gegebenen Menge $U \subseteq V$ um einen Unterraum handelt oder nicht:

1. Wir testen, ob der Nullvektor 0_V des Vektorraumes V auch in U liegt:

   - Wenn ja, dann ist $U \neq \emptyset$ und somit Bedingung (i) erfüllt.

   - Wenn nein, dann ist U kein Unterraum, denn jeder Unterraum muss 0_V enthalten.

2. Gilt $0_V \in U$ muss man weitermachen. Man überlegt sich einige Elemente, die in U liegen. Liefern diese ein Gegenbeispiel zu *mindestens einer* der Bedingungen (ii) oder (iii)?
3. Findet man kein Gegenbeispiel, dann versucht man die Bedingungen allgemein zu beweisen. Gelingt es, so ist U ein Unterraum. Scheitert der Beweis, so generiert man daraus ein Gegenbeispiel.

Bemerkung: Den zweiten Punkt kann man durchaus auch weglassen, aber es ist oft hilfreich, ein Gespür für die zu untersuchende Menge zu gewinnen.

Verständnisfragen 8.1

1. Was sind die Voraussetzungen der Definition „Unterraum“ und warum sind sie für U_1 und U_2 erfüllt?
2. Warum liefern die für U_1 gefundenen Beispielvektoren keine Gegenbeispiele zu den Bedingungen (i)-(iii)? Was muss man sich dafür überlegen?
3. Was ist die Voraussetzung, die in (8.1) (Beweis für Bedingung (ii) beim Unterraum U_1) eingeht?

Arbeitsauftrag 8.1 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.4.

8.2 Lineare Abhängigkeit und Unabhängigkeit

Aufgabe 8.2 Überprüfen Sie, ob die Teilmengen

$$M := \left\{ \begin{pmatrix} 3 \\ 5 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 6 \\ 2 \end{pmatrix} \right\} \text{ und } N := \left\{ \begin{pmatrix} 2 \\ -3 \\ -6 \end{pmatrix}, \begin{pmatrix} 1 \\ 5 \\ 7 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} \right\}$$

des $\mathbb{R}^3$ linear abhängig sind oder nicht. *Hinweis:* Das ist gleichbedeutend damit, die lineare (Un-)Abhängigkeit der Vektoren der Mengen zu überprüfen.

Ausführliche Lösung 8.2 Ist eine Menge linear abhängig, so existiert laut Definition eine nichttriviale Linearkombination aus den Vektoren der Menge, die den Nullvektor ergibt. Kann man andererseits den Nullvektor *nur* durch die triviale Linearkombination der Vektoren der Menge darstellen, so ist die Menge linear unabhängig.

Aus diesen Definitionen (oder Äquivalenzen zu den Definitionen)[1] der Begriffe *linear abhängig* und *linear unabhängig* können wir aber schon eine Strategie ableiten, wie wir vorzugehen haben, um die Mengen auf lineare Abhängigkeit zu überprüfen.

[1] Je nachdem wie Sie die Begriffe definiert haben, wird das eine oder das andere zutreffen.

Wir nehmen eine beliebige Linearkombination der Vektoren der Menge, die den Nullvektor ergeben soll. Danach versuchen wir durch Umformungen zu überprüfen, ob die gewählten Skalare alle Null sein müssen. Ist dies nicht der Fall, müssen wir solche Skalare finden, die nicht alle gleich Null sind.

Führen wir dies zunächst für die Menge M durch:

Seien $\alpha, \beta, \gamma \in \mathbb{R}$ beliebig mit

$$\alpha \cdot \begin{pmatrix} 3 \\ 5 \\ 2 \end{pmatrix} + \beta \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \gamma \cdot \begin{pmatrix} 3 \\ 6 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Es folgt durch Ausrechnen der linken Seite der Gleichung, dass

$$\begin{pmatrix} 3\alpha + \beta + 3\gamma \\ 5\alpha + \beta + 6\gamma \\ 2\alpha + \beta + 2\gamma \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

ist. Da zwei Vektoren aber nur dann gleich sind, wenn ihre einzelnen Einträge übereinstimmen, bekommt man drei Gleichungen:

$$3\alpha + \beta + 3\gamma = 0 \tag{8.2}$$
$$5\alpha + \beta + 6\gamma = 0 \tag{8.3}$$
$$2\alpha + \beta + 2\gamma = 0. \tag{8.4}$$

Nun müssen wir diese drei Gleichungen nach α, β und γ auflösen.

Das können wir machen, indem wir beispielsweise Gleichung (8.2) von Gleichung (8.3) und Gleichung (8.4) von Gleichung (8.2) abziehen:

$$(8.3) - (8.2) : 2\alpha + 3\gamma = 0 \tag{8.5}$$
$$(8.2) - (8.4) : \alpha + \gamma = 0. \tag{8.6}$$

Ziehen wir nun Gleichung (8.6) zweimal von Gleichung (8.5) ab, erhalten wir:

$$(8.5) - 2 \cdot (8.6) : \ \gamma = 0$$

Eingesetzt in (8.6) liefert das wiederum $\alpha = 0$, und mit Hilfe von (8.2),(8.3) oder (8.4) folgt auch $\beta = 0$.

Also müssen α, β und γ Null sein. Das bedeutet aber, dass die Menge M linear unabhängig ist.

Kommen wir damit zur Menge N.

Auch hier nehmen wir uns wieder beliebige $\alpha, \beta, \gamma \in \mathbb{R}$ mit

$$\alpha \cdot \begin{pmatrix} 2 \\ -3 \\ -6 \end{pmatrix} + \beta \cdot \begin{pmatrix} 1 \\ 5 \\ 7 \end{pmatrix} + \gamma \cdot \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

her und schauen, was dann für α, β und γ gelten muss.

Dafür formen wir erstmal wie oben um:

$$\begin{pmatrix} 2\alpha + \beta + 3\gamma \\ -3\alpha + 5\beta + 2\gamma \\ -6\alpha + 7\beta + \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Also haben wir die drei Gleichungen:

$$2\alpha + \beta + 3\gamma = 0 \tag{8.7}$$
$$-3\alpha + 5\beta + 2\gamma = 0 \tag{8.8}$$
$$-6\alpha + 7\beta + \gamma = 0. \tag{8.9}$$

Ziehen wir nun Gleichung (8.9) dreimal von Gleichung (8.7) und zweimal von Gleichung (8.8) ab, erhalten das neue Gleichungssystem:

$$20\alpha - 20\beta = 0 \tag{8.10}$$
$$9\alpha - 9\beta = 0 \tag{8.11}$$
$$-6\alpha + 7\beta + \gamma = 0. \tag{8.12}$$

Aus den ersten beiden Gleichungen (8.10) und (8.11) folgt jeweils $\alpha = \beta$. Setzen wir das in Gleichung (8.12) ein, so erhalten wir $\gamma = -\beta$. Für β erhalten wir keine Bedingung, wir können es frei wählen.

Wählen wir zum Beispiel $\beta = 1$, dann ist $\alpha = 1$ und $\gamma = -1$ und es gilt:

$$1 \cdot \begin{pmatrix} 2 \\ -3 \\ -6 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 5 \\ 7 \end{pmatrix} - 1 \cdot \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Wir haben also eine nichttriviale Linearkombination der Vektoren aus N gefunden, die den Nullvektor ergibt. Folglich ist die Menge N linear abhängig.

Verständnisfragen 8.2

1. Was ist die triviale Linearkombination von Vektoren, die den Nullvektor ergibt? Was ist eine nichttriviale Linearkombination, die den Nullvektor ergibt?
2. Bei der Menge N haben wir zum Schluss β frei gewählt. Hätten wir das auch für α oder γ machen können? Hätten wir auch α *und* β frei wählen können?
3. Angenommen, Sie hätten schon gleich gesehen, dass N linear abhängig ist, weil die Beziehung
$$1 \cdot \begin{pmatrix} 2 \\ -3 \\ -6 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 5 \\ 7 \end{pmatrix} - 1 \cdot \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$
zwischen den Vektoren aus N besteht. Hätten Sie dann trotzdem den Lösungsweg wie beschrieben gehen müssen?
4. Hätten wir im Fall N auch $\beta = 0$ wählen können statt $\beta = 1$?

Übung 8.2 Ordnen Sie nun selbst den einzelnen Bearbeitungsschritten der Musterlösung die Phasen aus Kapitel 3 zu! Tragen Sie dazu die jeweiligen Buchstaben in die kleinen Quadrate am Seitenrand ein!

Arbeitsauftrag 8.2 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen, der Übung und dem Arbeitsauftrag finden Sie in Abschnitt 11.4.

8.3 Bestimmung einer Basis eines Unterraumes

Aufgabe 8.3 Gegeben sei der Unterraum U des $\mathbb{R}$-Vektorraumes $\mathbb{R}^3$ durch:

$$U := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x + 3y + 2z = 0 \right\}.$$

Bestimmen Sie eine Basis von U!

Ausführliche Lösung 8.3 Bevor wir anfangen eine Basis zu bestimmen, schauen wir uns die Definition einer Basis an. Man kann eine Basis auf verschiedene Weise definieren. Wir wollen hier folgende Definition benutzen: Eine Basis ist ein linear unabhängiges Erzeugendensystem des Unterraumes. Was heißt das? Nun, einerseits muss jeder Vektor des Unterraumes als Linearkombination der Basisvektoren geschrieben werden können. Denn das heißt gerade, dass die Basis ein Erzeugendensystem des Unterraumes ist. Außerdem müssen die Basisvektoren auch linear unabhängig sein. B

Wir müssen also Vektoren aus U finden, die schon ganz U erzeugen (1.). Und dann müssen wir noch sicherstellen, dass die von uns gefundenen Vektoren auch tatsächlich linear unabhängig sind (2.).

Zu 1.: Wie findet man solche Vektoren? Man nimmt sich zunächst einen *beliebigen* Vektor aus U und versucht diesen als eine Linearkombination *ganz konkreter* Vektoren aus U zu schreiben. Dazu müssen wir die Rechengesetze der Vektorrechnung und die Eigenschaften, die ein Vektor des Unterraumes U besitzt, ausnutzen. K

Wir nehmen uns also einen beliebigen Vektor u aus dem Unterraum U. Nach Definition von U gibt es dann $x, y, z \in \mathbb{R}$ mit Z

$$u = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \text{ und } x + 3y + 2z = 0.$$

Der nächste Schritt ist, die Eigenschaft von u als Element des Unterraums „einzusetzen“. B

H Damit ist gemeint, dass wir z. B. $x + 3y + 2z = 0$ nach $x = -3y - 2z$ umformen können, um dieses dann für x in den Vektor u einzusetzen. So erhalten wir:

$$u = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -3y - 2z \\ y \\ z \end{pmatrix}.$$

B Wir wollen u als Linearkombination konkreter Vektoren aus U schreiben.

K Das schaffen wir, indem wir den Vektor mit Hilfe von Vektorrechnungen weiter umformen. Und zwar schreiben wir ihn gezielt als Addition von Vektoren, sodass in den Summanden (das sind Vektoren) nur noch je eine Variable (hier y bzw. z) auftaucht:

H

$$u = \begin{pmatrix} -3y - 2z \\ y \\ z \end{pmatrix} = \begin{pmatrix} -3y \\ y \\ 0 \end{pmatrix} + \begin{pmatrix} -2z \\ 0 \\ z \end{pmatrix}.$$

Und jetzt ziehen wir die Variablen noch als Skalare aus den Vektoren heraus:

$$u = y \cdot \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix} + z \cdot \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}.$$

Wunderbar, nun haben wir doch wirklich u als Linearkombination der beiden Vektoren $\begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$ aus U geschrieben.

A Da u aus U beliebig gewählt war, können wir also jedes Element von U als Linearkombination dieser beiden Vektoren schreiben. Das heißt, $\begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$ bilden ein Erzeugendensystem von U.

B *Zu 2.*: Wir müssen nun noch prüfen, ob $\begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$ linear unabhängig sind.

K Das machen wir, indem wir zeigen, dass aus

$$\alpha \cdot \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix} + \beta \cdot \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \text{mit } \alpha, \beta \in \mathbb{R}$$

stets $\alpha = \beta = 0$ folgt.

H Seien also $\alpha, \beta \in \mathbb{R}$ beliebig mit

$$\alpha \cdot \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix} + \beta \cdot \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

gegeben. Dann folgt:

$$\begin{pmatrix} -3\alpha - 2\beta \\ \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Und daraus ergibt sich:

$$-3\alpha - 2\beta = 0 \text{ und } \alpha = 0 \text{ und } \beta = 0.$$

Also folgt insbesondere $\alpha = \beta = 0$. Damit sind die beiden Vektoren linear unabhängig.

Zusammen mit 1. folgt dann auch, dass $\begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$ eine Basis von U bilden.

Verständnisfragen 8.3

1. Warum sind $\begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$ aus U?
2. Warum erhält man durch die getätigten Umformungen für u:

   $$u = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -3y - 2z \\ y \\ z \end{pmatrix} = y \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$

   automatisch Vektoren aus U?
3. Ist die von uns gefundene Basis die einzige Basis, die es zu U gibt?

Arbeitsauftrag 8.3 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.4.

8.4 Rechnen mit Matrizen

Aufgabe 8.4 Untersuchen Sie, ob es Matrizen $A, B \in M_{2\times 2}(K)$ mit

$$A^2 = B^2 = -E_2 \text{ und } A \cdot B = -B \cdot A$$

gibt, wenn $K = \mathbb{R}$ bzw. $K = \mathbb{C}$ ist. (Hinweis: E_2 bezeichnet die Einheitsmatrix)!

Ausführliche Lösung 8.4 Wir versuchen, soweit es geht, die beiden Fälle ($K = \mathbb{R}$ und $K = \mathbb{C}$) gemeinsam zu betrachten, um uns so Arbeit zu ersparen. Wir haben für A, B zwei Bedingungen gegeben. Wir sollten nacheinander überlegen, welche Informationen wir aus diesen Bedingungen über A und B erhalten. Überlegen wir zunächst, welche Matrizen $X \in M_{2\times 2}(K)$ die Bedingung $X^2 = -E_2$ erfüllen.

Wir schreiben $X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ mit $a, b, c, d \in K$, dann ist:

$$X^2 = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + bc & ab + bd \\ ca + dc & cb + d^2 \end{pmatrix} \overset{!}{=} \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Das heißt, es müssen die Gleichungen

$$a^2 + bc = -1, \tag{8.13}$$
$$ab + bd = 0, \tag{8.14}$$
$$ca + dc = 0 \qquad \text{und} \tag{8.15}$$
$$cb + d^2 = -1 \tag{8.16}$$

erfüllt sein.

K

Als Nächstes überlegen wir, welche Folgerungen wir aus diesen Gleichungen für a, b, c, d ziehen können. Ziel ist es also, Variablen aus den Gleichungen zu eliminieren und einige der Variablen durch möglichst wenige der anderen auszudrücken.

H

Wären nun b oder c gleich 0, dann würde aus Gleichung (8.13) $a^2 = -1$ folgen, was in $\mathbb{R}$ nicht möglich ist. Also gilt in $\mathbb{R}$: $b \neq 0$ und $c \neq 0$.

K

Im Fall $K = \mathbb{C}$ können wir dies natürlich nicht folgern, wir können aber trotzdem zunächst davon ausgehen, dass $b \neq 0$ und $c \neq 0$ gilt. Denn finden wir auf diesem Wege am Ende Matrizen A und B, die die Bedingungen erfüllen, sind wir ohnehin schon fertig. Nur falls wir auf diesem Wege zu einem Widerspruch gelangen, müssen wir uns daran erinnern, dass wir für b, c eine Einschränkung vorgenommen haben, die nicht unbedingt gilt, und dann diese Fälle noch gesondert betrachten.

B

Aber nun erst einmal weiter:

H

Sind b und c ungleich 0, können wir aus Gleichung (8.14) folgern:

$$0 = ab + bd = b(a + d) \overset{b \neq 0}{\Rightarrow} a + d = 0 \Rightarrow d = -a.$$

Und für c folgt aus Gleichung (8.13):

$$a^2 + bc = -1 \Rightarrow c = -\frac{1 + a^2}{b}.$$

Damit hat X also die Gestalt:

$$X = \begin{pmatrix} a & b \\ -\frac{1+a^2}{b} & -a \end{pmatrix}.$$

K

Mehr Informationen können wir aus $X^2 = -E_2$ nicht erhalten, deswegen schauen wir uns nun die zweite Bedingung an, die wir gegeben haben.

A

A und B sollen zwei Matrizen mit $A^2 = B^2 = -E_2$ sein. Das heißt:

$$A = \begin{pmatrix} a & r \\ -\dfrac{1 + a^2}{r} & -a \end{pmatrix} \text{ und } B = \begin{pmatrix} b & s \\ -\dfrac{1 + b^2}{s} & -b \end{pmatrix}, \text{ wobei } r, s \neq 0.$$

Außerdem soll noch $AB = -BA$ gelten. Also schauen wir uns die beiden Produkte an:

B

H

$$AB = \begin{pmatrix} ab - r\dfrac{1+b^2}{s} & as - rb \\ -b\dfrac{1+a^2}{r} + a\dfrac{1+b^2}{s} & -s\dfrac{1+a^2}{r} + ab \end{pmatrix}$$

und

$$-BA = -\begin{pmatrix} ba - s\dfrac{1+a^2}{r} & br - sa \\ -a\dfrac{1+b^2}{s} + b\dfrac{1+a^2}{r} & -r\dfrac{1+b^2}{s} + ba \end{pmatrix}.$$

Durch Vergleich der Matrizeneinträge erhalten wir:

$$ab - r\frac{1+b^2}{s} = -\left(ba - s\frac{1+a^2}{r}\right), \tag{8.17}$$

$$as - rb = -(br - sa), \tag{8.18}$$

$$-b\frac{1+a^2}{r} + a\frac{1+b^2}{s} = -\left(-a\frac{1+b^2}{s} + b\frac{1+a^2}{r}\right), \tag{8.19}$$

$$-s\frac{1+a^2}{r} + ab = -\left(-r\frac{1+b^2}{s} + ba\right). \tag{8.20}$$

K

Jetzt untersuchen wir anhand dieser Gleichungen die Abhängigkeiten der Variablen voneinander, um so Informationen über die Existenz der Matrizen zu erhalten.

H

In den Gleichungen (8.18) und (8.19) stehen wahre Aussagen (links und rechts steht das Gleiche), diese liefern uns also keine weiteren Bedingungen für die Matrizen A und B. Außerdem erkennt man, dass die Gleichungen (8.17) und (8.20) gleich sind.

B

[*Bemerkung:* Gleichung (8.17) kann man auch auf einem etwas anderen Wege erhalten:

T

$$(AB)^2 = A\underbrace{(BA)}_{=-AB}B = A(-AB)B = -A^2B^2 = -(-E_2)(-E_2) = -E_2,$$

A

und damit hat $(AB)^2$ die gleiche Form wie alle X mit $X^2 = -E_2$. Das heißt, es gilt: $ab - r\frac{1+b^2}{s} = -(-s\frac{1+a^2}{r} + ab)$.]

K Um weitere Bedingungen für die Matrizen zu erhalten, bleibt uns nichts anderes übrig als Gleichung (8.17) nach einer Variablen umzuformen – wir wollen das hier mal nach b auflösen:

H

$$\begin{aligned}
& ab - r\frac{1+b^2}{s} = -\left(ba - s\frac{1+a^2}{r}\right) && \mid \cdot rs \\
\Leftrightarrow\quad & abrs - r^2(1+b^2) = -abrs + s^2(1+a^2) && \mid +(abrs);\ \mid -(s^2(1+a^2)) \\
\Leftrightarrow\quad & 2abrs - r^2(1+b^2) - s^2(1+a^2) = 0 && \mid \cdot(-1) \\
\Leftrightarrow\quad & r^2(1+b^2) - 2abrs + s^2(1+a^2) = 0 && \\
\Leftrightarrow\quad & r^2 \cdot b^2 - 2ars \cdot b + \left(s^2(1+a^2) + r^2\right) = 0 && \mid : r^2 \\
\Leftrightarrow\quad & b^2 - \frac{2as}{r}b + \frac{s^2(1+a^2)+r^2}{r^2} = 0. &&
\end{aligned}$$

Die p-q-Formel liefert dann:

$$\begin{aligned}
b =& \frac{as}{r} \pm \sqrt{\frac{(as)^2}{r^2} - \frac{s^2(1+a^2)+r^2}{r^2}} \\
=& \frac{as}{r} \pm \frac{1}{|r|}\sqrt{(as)^2 - [s^2(1+a^2)+r^2]} \\
=& \frac{as}{r} \pm \frac{1}{|r|}\sqrt{-(s^2+r^2)}.
\end{aligned}$$

P Das heißt, b existiert in $\mathbb{R}$ nur dann, wenn der Ausdruck unter der Wurzel ≥ 0 ist.

H Da $s, r \neq 0$ ist, gilt aber:

$$-(s^2+r^2) < 0.$$

Also gibt es in $\mathbb{R}$ keine Lösung.

B Es existieren demnach keine Matrizen $A, B \in M_{2\times 2}(\mathbb{R})$ mit $A^2 = B^2 = -E_2$ und $AB = -BA$.

P Betrachten wir nun Matrizen in $M_{2\times 2}(\mathbb{C})$, dann dürfen wir auch negative Zahlen unter der Wurzel stehen haben.

Z Wenn wir also die Matrizen A, B wie eben wählen:

$$A = \begin{pmatrix} a & r \\ -\frac{1+a^2}{r} & -a \end{pmatrix} \text{ und } B = \begin{pmatrix} b & s \\ -\frac{1+b^2}{s} & -b \end{pmatrix} \quad \text{mit } r, s \neq 0,$$

H dann gibt es nach den obigen Umformungen zu vorgegebenen r, s und a eine Zahl b, sodass $A^2 = B^2 = -E_2$ und $AB = -BA$ erfüllt sind.

Z Wählen wir zum Beispiel $s = 1$, $r = 1$ und $a = 0$, so können wir zwei konkrete Matrizen hinschreiben, die die gewünschten Bedingungen erfüllen.

H

Dann ist

$$\begin{aligned} b &= \frac{as}{r} \pm \frac{1}{r}\sqrt{-(s^2+r^2)} \\ &= 0 \pm 1\sqrt{-(1+1)} \\ &= \pm\sqrt{-2} \\ &= \pm\sqrt{2}\cdot i. \end{aligned}$$

B

Also sind z. B. für $b = +\sqrt{2}\cdot i$:

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

und

$$B = \begin{pmatrix} \sqrt{2}\cdot i & 1 \\ 1 & -\sqrt{2}\cdot i \end{pmatrix}$$

zwei Matrizen, die die Bedingungen erfüllen (wie sich leicht auch noch einmal durch direktes Nachrechnen überprüfen lässt). In $M_{2\times 2}(\mathbb{C})$ existieren folglich Matrizen mit $A^2 = B^2 = -E_2$ und $AB = -BA$.

P

Wir sehen also, dass die Existenz von Matrizen mit bestimmten Eigenschaften stark davon abhängt, über welchem Körper man die Matrizen betrachtet (also aus welchem Körper die Einträge der Matrizen stammen).

Verständnisfragen 8.4

1. In der Lösung wird behauptet, dass wir aus

$$-E_2 = X^2 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^2$$

nur $b, c \neq 0$, $d = -a$ und $c = -\frac{1+a^2}{b}$ folgern können. Warum können wir nicht auch noch zum Beispiel b durch a ausdrücken und somit noch mehr Informationen erhalten?

2. Die Variablen a und b spielen verschiedene Rollen in X bzw. A und B. Vergleichen Sie die Matrizen A und B mit der Matrix X und bestimmen Sie, welchen Variablen in X die Variablen der Matrizen A und B jeweils entsprechen!

3. In der Lösung wurde die Gleichung $b^2 - \frac{2as}{r}b + \frac{s^2(1+a^2)+r^2}{r^2} = 0$ mit Hilfe der p-q-Formel gelöst. Was ist hier p und was ist q?

Arbeitsauftrag 8.4 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.4.

8.5 Basis des Kerns einer Matrix

Aufgabe 8.5 Bestimmen Sie eine Basis des Kerns der (3×7)-Matrix

$$A := \begin{pmatrix} 1 & 2 & 1 & 3 & 2 & 6 & 3 \\ 2 & 4 & 0 & 6 & 3 & 8 & 2 \\ 3 & 6 & 1 & 9 & 2 & 8 & 5 \end{pmatrix}.$$

Ausführliche Lösung 8.5 Der Kern einer Matrix $A \in M_{m \times n}(K)$ ist definiert als

$$\ker(A) := \{x \in K^n \mid Ax = 0\}.$$

Da Ax als lineares Gleichungssystem aufgefasst werden kann, ist der Kern einer Matrix nichts anderes als die *Lösungsmenge des zugehörigen homogenen linearen Gleichungssystems*. Insofern ist es also interessant, den Kern einer Matrix berechnen zu können.

Was müssen wir nun machen, um den Kern zu bestimmen? Wir könnten uns wie bei der Basisbestimmung für Unterräume einen beliebigen Vektor aus dem Kern wählen und diesen dann als Linearkombination konkreter Vektoren schreiben (vgl. Aufgabe 8.3).

Es gibt für den Kern von Matrizen aber noch eine andere Möglichkeit zur Bestimmung einer Basis. Bringt man die Matrix durch elementare Zeilenumformungen auf *Treppennormalform*, so ändert sich einerseits durch die gemachten Umformungen der Kern der Matrix nicht und andererseits kann man in der Treppennormalform leicht eine Basis ablesen.

Eine Matrix liegt in Treppennormalform vor, wenn sie folgende Form besitzt:

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * & 0 & * & \cdots \\ 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots \\ 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 & 1 & * & \cdots \\ & & & & & & & & & & & & \ddots & \end{pmatrix}.$$

Eine Stufe beginnt also immer mit einer 1 und in der Spalte, in der diese 1 steht, befinden sich sonst nur Nullen. Beachten Sie, dass die Stufen auch jeweils in aufeinanderfolgenden Spalten beginnen können und die erste Stufe auch in der ersten Spalte beginnen kann (im Gegensatz zur gezeichneten Matrix).

Wie wir aus so einer Treppennormalform eine Basis ablesen können, sehen wir gleich noch. Zunächst bestimmen wir erst einmal die Treppennormalform für unsere Matrix A.

Zur Bestimmung der Treppennormalform kann man den Gauß-Jordan-Algorithmus benutzen:

$$\begin{pmatrix} 1 & 2 & 1 & 3 & 2 & 6 & 3 \\ 2 & 4 & 0 & 6 & 3 & 8 & 2 \\ 3 & 6 & 1 & 9 & 2 & 8 & 5 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 3 & 2 & 6 & 3 \\ 0 & 0 & -2 & 0 & -1 & -4 & -4 \\ 0 & 0 & -2 & 0 & -4 & -10 & -4 \end{pmatrix}$$

$$\rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 3 & 2 & 6 & 3 \\ 0 & 0 & -2 & 0 & -1 & -4 & -4 \\ 0 & 0 & 0 & 0 & -3 & -6 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 3 & 2 & 6 & 3 \\ 0 & 0 & 1 & 0 & \frac{1}{2} & 2 & 2 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 \end{pmatrix}$$

$$\rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 3 & 0 & 2 & 3 \\ 0 & 0 & 1 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 0 & 3 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 \end{pmatrix}.$$

Wir haben nun die Treppennormalform von A bestimmt. Wie lesen wir daraus eine Basis ab? Im Endeffekt geht das ganz einfach.

Wir füllen unsere Matrix wie folgt auf: Beginnt in der i-ten Spalte keine Stufe[2], so schreiben wir in die i-te Zeile der Matrix eine neue Zeile (die alte rutscht dabei um eine Zeile nach unten). Diese zugefügte Zeile besteht bis auf eine -1 in der i-ten Spalte nur aus Nullen.

Für unsere Matrix müssen wir also Zeilen für die 2., 4., 6. und 7. Spalte einfügen und erhalten so die Matrix:

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

Die i-ten Spalten (das sind die Spalten, in denen in der Treppennormalform keine Stufe beginnt) bilden dann eine Basis von $\ker(A)$.

In unserem Fall können wir aus dieser Matrix also die Basis

$$b_1 := \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, b_2 := \begin{pmatrix} 3 \\ 0 \\ 0 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, b_3 := \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 2 \\ -1 \\ 0 \end{pmatrix}, \; b_4 := \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}$$

ablesen.

Und damit sind wir fertig!

[2] Man spricht bei den Stellen der Matrix, in denen eine Stufe beginnt, auch von *Pivotstellen*.

Verständnisfragen 8.5

1. Was sind elementare Zeilenumformungen?
2. Was bedeutet es, dass sich durch die elementaren Zeilenumformungen der Kern der Matrix nicht verändert? Die Matrix ändert sich doch durch die Umformungen!
3. Untersuchen Sie, welche Umformungsschritte beim Finden der Treppennormalform gemacht wurden! Überprüfen Sie ferner, ob diese tatsächlich dem Gauß-Jordan-Algorithmus entsprechen!
4. Wieso funktioniert das Ablesen der Basis wie in der Anleitung beschrieben? Überlegen Sie sich, welcher Zusammenhang zwischen den eingefügten (-1)en und der freien Wählbarkeit der Komponenten eines Vektors $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix}$ aus dem Kern der Matrix besteht!

Arbeitsauftrag 8.5 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.4.

8.6 Basisergänzung und Basis des Faktorraumes

Aufgabe 8.6 $U := \{\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} \in \mathbb{R}^4 \mid a_1 + 2a_2 = a_3 + 2a_4\}$ ist ein Unterraum von $V := \mathbb{R}^4$. Bestimmen Sie die Dimension von U, ergänzen Sie $M := \{\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}\}$ zu einer Basis von U und geben Sie eine Basis von V/U an.

K

Ausführliche Lösung 8.6 Die Dimension eines Unterraumes können wir ermitteln, indem wir eine Basis des Unterraumes bestimmen. Wie das geht, haben wir in Aufgabe 8.3 gesehen.

B

Wir möchten an dieser Stelle nicht den ganzen Weg vorführen, sondern uns lieber auf die neuen Inhalte konzentrieren. Deshalb geben wir Ihnen hier eine Basis vor, den Lösungsweg finden Sie am Ende dieser Aufgabe.

H

Die Vektoren $\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\2\\0\end{pmatrix}, \begin{pmatrix}0\\0\\-2\\1\end{pmatrix}$ bilden eine Basis von U. Da man also drei Basisvektoren hat, ist $\dim(U) = 3$.

B

Als Nächstes sollen wir M zu einer Basis von U ergänzen.

P

Wir wissen bereits, dass wir nur noch einen Vektor ergänzen müssen, da M zwei linear unabhängige Vektoren und eine Basis von U drei Vektoren enthält.

K

Man hat nun zwei Möglichkeiten:

1. Man findet direkt einen Vektor v, sodass $M \cup \{v\}$ eine Basis von U bildet.
2. Man tauscht die Vektoren von M in eine Basis von U hinein.

Wir wollen hier den zweiten Weg gehen, da er sich auch auf Beispiele übertragen lässt, in denen mehr als ein Vektor ergänzt werden muss. In diesen Fällen ist zwar der 1. Weg auch denkbar, aber er wird recht schnell kompliziert.

B

Der Austauschsatz von Steinitz besagt (vgl. Sie dazu auch mit Übung 10.14 und der entsprechenden Lösung in Abschnitt 12.4.2), dass wir linear unabhängige Vektoren $b'_1, \ldots, b'_r$ aus V gegen geeignete Vektoren aus einer Basis $B := \{b_1, \ldots, b_n\}$ von V austauschen dürfen.

A

Da die Vektoren in M linear unabhängig sind und auch in U liegen, können wir sie also in unsere gefundene Basis von U eintauschen.

B

Nur wissen wir noch nicht, gegen welche dieser drei Vektoren wir eintauschen dürfen. Dazu gleich mehr.

H

Wir sehen, dass wir den Vektor $\begin{pmatrix}1\\0\\1\\0\end{pmatrix}$ schon in der Basis haben. Deshalb müssen wir nur noch $\begin{pmatrix}0\\1\\0\\1\end{pmatrix}$ eintauschen.

Z

Nach dem Austauschlemma, auf dem der Beweis des Satzes von Steinitz beruht, darf man einen Vektor w gegen den Basisvektor b_i einer Basis $B := \{b_1, \ldots, b_n\}$ austauschen, falls $w = \sum_{j=1}^{n} \alpha_j b_i$ mit $\alpha_i \neq 0$ gilt.

B

Das heißt, da man w immer als Linearkombination der Basis schreiben kann, dürfen wir w immer gegen ein b_i eintauschen. Und zwar gerade gegen diejenigen b_i, die *wirklich* einen Beitrag zur Linearkombination leisten (bei denen also $\alpha_i \neq 0$ ist).

A Deshalb schreiben wir den Vektor, den wir in die Basis eintauschen wollen, als Linearkombination der Basisvektoren:

H
$$\begin{pmatrix}0\\1\\0\\1\end{pmatrix} = 0 \cdot \begin{pmatrix}1\\0\\1\\0\end{pmatrix} + 1 \cdot \begin{pmatrix}0\\1\\2\\0\end{pmatrix} + 1 \cdot \begin{pmatrix}0\\0\\-2\\1\end{pmatrix}.$$

Wir dürfen folglich $\begin{pmatrix}0\\1\\0\\1\end{pmatrix}$ entweder anstelle von $\begin{pmatrix}0\\1\\2\\0\end{pmatrix}$ oder anstelle von $\begin{pmatrix}0\\0\\-2\\1\end{pmatrix}$ in die Basis schreiben. Also ist zum Beispiel $\{\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\1\end{pmatrix}, \begin{pmatrix}0\\0\\-2\\1\end{pmatrix}\}$ eine neue Basis und wir haben somit M durch $\begin{pmatrix}0\\0\\-2\\1\end{pmatrix}$ zu einer Basis von U ergänzt.

B Es bleibt noch die Basis von V/U zu bestimmen.

Z Mit Hilfe der Dimensionsformel für Faktorräume[3] wissen wir aber, dass

$$\dim(V/U) = \dim(V) - \dim(U) = 4 - 3 = 1$$

ist. Damit bildet jeder Vektor aus V/U, der nicht der Nullvektor ist, eine Basis von V/U. Die Vektoren aus V/U haben die Form $v + U$, wobei $v \in V$ ist. Der Nullvektor ist $0_V + U = U$. Das heißt aber, dass wir v so wählen müssen, dass v nicht in U liegt, denn sonst würde $v + U = U$ gelten[4].

H Der Vektor $\begin{pmatrix}1\\0\\0\\0\end{pmatrix}$ liegt in V, aber nicht in U, da $1 + 2 \cdot 0 \neq 0 + 2 \cdot 0$ ist. Deshalb bildet $\{\begin{pmatrix}1\\0\\0\\0\end{pmatrix} + U\}$ eine Basis von V/U.

Anhang (Bestimmung einer Basis von U):

Wir nehmen uns einen beliebigen Vektor u aus U und schauen, durch welche Vektoren er erzeugt wird. Sei also $u \in U$ beliebig, dann gilt:

$$u = \begin{pmatrix}u_1\\u_2\\u_3\\u_4\end{pmatrix} \quad \text{mit } u_1 + 2u_2 = u_3 + 2u_4.$$

[3] Der Faktorraum wird oft auch als Quotient oder Quotientenraum bezeichnet.

[4] Zwei Elemente $v + U$, $v' + U$ aus V/U sind genau dann gleich, wenn $v - v' \in U$ ist.

Daraus folgt $u_3 = u_1 + 2u_2 - 2u_4$ und somit gilt:

$$u = \begin{pmatrix} u_1 \\ u_2 \\ u_1 + 2u_2 - 2u_4 \\ u_4 \end{pmatrix} = u_1 \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + u_2 \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix} + u_4 \begin{pmatrix} 0 \\ 0 \\ -2 \\ 1 \end{pmatrix}.$$

u wird also von den Vektoren $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 0 \\ -2 \\ 1 \end{pmatrix}$ erzeugt. Da u aber beliebig aus U war, werden alle Vektoren aus U von diesen drei Vektoren erzeugt. Das heißt, die drei Vektoren bilden ein Erzeugendensystem von U. Man sieht aber leicht, dass sie auch linear unabhängig sind, und damit ist $\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -2 \\ 1 \end{pmatrix} \right\}$ sogar eine Basis von U.

Verständnisfragen 8.6

1. Das Austauschlemma ist so formuliert worden, dass jedes w gegen einen der Basisvektoren ausgetauscht werden kann. Gilt das wirklich oder gibt es Ausnahmen?
2. Es wurde in der Lösung behauptet, dass die Vektoren aus M in U liegen und linear unabhängig sind. Warum gilt das wirklich?
3. Warum gilt in eindimensionalen Vektorräumen, dass jeder Vektor ungleich dem Nullvektor eine Basis bildet?

Arbeitsauftrag 8.6 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.4.

8.7 Homomorphismen / Lineare Abbildungen

Aufgabe 8.7 Seien V ein K-Vektorraum, f ein Homomorphismus (eine lineare Abbildung) von V nach V und U ein Unterraum von V mit $f(U) \subseteq U$. Des Weiteren sind die Abbildungen $f_U : U \to U$ mit $f_U(u) := f(u)$ für alle $u \in U$ und $\bar{f} : V/U \to V/U$ mit $\bar{f}(v + U) := f(v) + U$ für alle $v \in V$ Homomorphismen. Zeigen Sie, dass für die gegebenen Abbildungen $f, f_U, \bar{f}$ die Ungleichung

$$\dim(\ker(f)) \leq \dim(\ker(f_U)) + \dim(\ker(\bar{f}))$$

gilt.

K

Ausführliche Lösung 8.7 Um für diesen Beweis eine Idee zu gewinnen, lohnt es sich, einen Blick darauf zu werfen, welche Sätze und Definitionen beim Beweisen hilfreich sein könnten. Da wir mit Kern, Bild und Dimensionen arbeiten, sollten wir vor allem schauen, welche Aussagen und Formeln wir dazu zur Verfügung. Aber auch Aussagen über den Faktorraum V/U und die Abbildungen $f, f_U, \bar{f}$ interessieren uns. Sind V und W K-Vektorräume, $f : V \to W$ eine lineare Abbildung und U, X Unterräume von V, dann gilt:

$$
\begin{aligned}
&\ker(f) \subseteq V && \text{ist ein Unterraum von } V && (8.21)\\
&\operatorname{im}(f) \subseteq W && \text{ist ein Unterraum von } W && (8.22)\\
&V/\ker(f) \cong \operatorname{im}(f) && \textit{1. Homomorphiesatz} && (8.23)\\
&(X+U)/U \cong X/(X\cap U) && \textit{2. Homomorphiesatz} && (8.24)\\
&(V/U)/(X/U) \cong V/X && \textit{3. Homomorphiesatz} && (8.25)\\
&\dim(V/U) = \dim(V) - \dim(U) && && (8.26)\\
&\dim(V) = \dim(\ker(f)) + \operatorname{rg}(f) && && (8.27)\\
&U \cong X \Rightarrow \dim(U) = \dim(X) && && (8.28)\\
&\ker(f_U) = \ker(f) \cap U. && && (8.29)
\end{aligned}
$$

Wir haben also einige Aussagen (die evtl. auch in Ihrer Vorlesung bewiesen wurden), die uns weiterhelfen könnten. Vielleicht werden wir einige davon gar nicht brauchen. Nun fängt man an, Sachen auszuprobieren – auf die Sackgassen, die sich dabei ergeben, werden wir in dieser Lösung nicht weiter eingehen.

P

Aussage (8.29) sagt uns, dass $\ker(f_U) = \ker(f) \cap U$ gilt. Damit haben wir schon einmal eine Abbildung aus unserer Gleichung eliminiert. Dafür haben wir nun einen Durchschnitt zweier Unterräume zu betrachten.

Z

Nun sieht man aber, dass der 2. Homomorphiesatz (8.24) eine Aussage über den Schnitt zweier Unterräume trifft, vielleicht hilft uns das ja weiter!

A

Also übertragen wir ihn doch mal auf unseren Fall (hier ist $X = \ker(f)$ und $U = U$):

$$\ker(f)/(\ker(f) \cap U) \cong (\ker(f) + U)/U.$$

Hierfür brauchen wir auch gleich Aussage (8.21), welche besagt, dass der Kern von f ein Unterraum von V ist.

H

Daraus folgt mit Aussage (8.29):

$$\ker(f)/\ker(f_U) \cong (\ker(f) + U)/U.$$

B

Hat uns das jetzt unserem Ziel schon etwas näher gebracht, was meinen Sie?

K

Um das zu überprüfen, sehen wir uns die Dimensionen der involvierten Räume an, die wir da nun stehen haben. Dazu brauchen wir auch die Aussagen (8.26) und (8.28)!

H

$$
\begin{aligned}
\dim(\ker(f)) - \dim(\ker(f_U)) &\overset{(8.26)}{=} \dim(\ker(f)/\ker(f_U))\\
&\overset{(8.28)}{=} \dim((\ker(f) + U)/U)
\end{aligned}
$$

Wenn wir jetzt noch zeigen könnten, dass

$$\dim((\ker(f)+U)/U) \leq \dim(\ker(\bar{f}))$$

ist, dann wären wir nach einer kleinen Umformung der Ungleichung schon fertig.

Wir sind unserem Ziel also wirklich ein gutes Stück näher gekommen.

Die Ungleichung zeigen wir, indem wir beweisen, dass

$$(\ker(f)+U)/U \subseteq \ker(\bar{f})$$

gilt.

Das machen wir auf klassischem Wege: Wir nehmen uns ein beliebiges Element aus $(\ker(f)+U)/U$ und zeigen, dass es in $\ker(\bar{f})$ liegt.

Sei $v+U \in (\ker(f)+U)/U$ beliebig, wobei $v \in (\ker(f)+U)$ ist. Das heißt, v lässt sich als Summe von einem $v' \in \ker(f)$ und einem $u \in U$ schreiben: $v = v' + u$. Wir müssen nun zeigen, dass $\bar{f}(v+U) = U(= 0_{V/U})$ gilt:

$$\begin{aligned}
\bar{f}(v+U) &\overset{\text{Def.}\bar{f}}{=} f(v)+U \\
&\overset{v=v'+u}{=} f(v'+u)+U \\
&\overset{f\ \text{lin.}}{=} (f(v')+f(u))+U \\
&\overset{v'\in\ker(f)}{=} (0_V+f(u))+U = f(u)+U \\
&\overset{f(U)\subseteq U}{=} U.
\end{aligned}$$

Verständnisfragen 8.7

1. Wie folgt aus

   $$\dim(\ker(f)) - \dim(\ker(f_U)) = \dim((\ker(f)+U)/U)$$

   und

   $$\dim((\ker(f)+U)/U) \leq \dim(\ker(\bar{f}))$$

   die Behauptung

   $$\dim(\ker(f)) \leq \dim(\ker(f_U)) + \dim(\ker(\bar{f}))?$$

2. Warum kann grundsätzlich die Beziehung $(\ker(f)+U)/U \subseteq \ker(\bar{f})$ gelten? Nur dann nämlich ergibt der Lösungsansatz für diese Inklusion Sinn!
3. Warum folgt aus $(\ker(f)+U)/U \subseteq \ker(\bar{f})$, dass $\dim((\ker(f)+U)/U) \leq \dim(\ker(\bar{f}))$ gilt?

Arbeitsauftrag 8.7 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.4.

8.8 Zerlegung von Permutationen in Produkte aus Transpositionen

Aufgabe 8.8 Schreiben Sie die Permutationen

$$\sigma_1 := \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix},$$

$$\sigma_2 := \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix},$$

$$\sigma_3 := \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$$

als Produkte von Transpositionen!

Ausführliche Lösung 8.8 Transpositionen sind Permutationen, bei denen nur zwei Elemente i und j miteinander vertauscht werden:

$$\tau = \begin{pmatrix} 1 & 2 & \dots & i & \dots & j & \dots & n \\ 1 & 2 & \dots & j & \dots & i & \dots & n \end{pmatrix}.$$

Eine kürzere Schreibweise, die das Gleiche ausdrücken soll, ist: $\tau = (i\ j)$.

Um eine Permutation als ein Produkt von Transpositionen zu schreiben, lohnt es sich, zunächst die Permutation etwas genauer unter die Lupe zu nehmen. Hat man ein Gespür dafür, wie die Permutation abbildet, so findet man vielleicht schon geeignete Transpositionen.

σ_1 vertauscht 1 und 3 miteinander, 2 und 4 werden auf sich selbst abgebildet.

σ_1 ist demnach selbst eine Transposition. Somit ist $\sigma_1 = (1\ 3)$ schon als Produkt von Transpositionen dargestellt. (Das Produkt hat nur einen Faktor!)

σ_2 vertauscht 1 und 4 und es vertauscht 2 und 3 miteinander.

Deshalb können wir es recht leicht als Produkt der beiden folgenden Transpositionen (Vertauschung von 1 und 4; Vertauschung von 2 und 3) schreiben:

$$\sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix} = (2\ 3) \cdot (1\ 4).$$

Für die Pemutation σ_3 ist das nun nicht mehr so einfach, denn es werden nicht nur jeweils zwei Elemente vertauscht. Hier müssen wir uns eine andere Taktik überlegen. Wir können aber schrittweise vorgehen, indem wir erst einmal eine Transposition aufschreiben, die 1 auf das gewünschte Element abbildet (hier 2).

Dadurch ist die Transposition schon festgelegt. Denn soll 1 auf 2 abgebildet werden, muss ja 2 auf 1 abgebildet werden und 3 und 4 jeweils auf sich selbst:

$$\tau_1 := \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix} = (1\ 2).$$

Jetzt können wir uns den nächsten Schritt überlegen. Verketten wir diese Transposition mit einer anderen, so wollen wir ja, dass 1 insgesamt weiterhin auf 2 abgebildet wird.

Das heißt aber, dass 2 von der neuen Transposition auf sich selbst abgebildet werden muss:

$$\tau_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ & 2 & & \end{pmatrix}.$$

Nun werfen wir wieder einen Blick auf σ_3: Dort soll 2 auf 3 abgebildet werden. Wie bekommen wir das nun hin?

Von τ_1 wird 2 auf 1 abgebildet, also muss nun 1 auf 3 gehen, sodass bei der Verkettung insgesamt 2 auf 3 abgebildet wird. Damit erhalten wir τ_2:

$$\tau_2 := \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix} = (1\;3).$$

Gilt jetzt schon $\sigma_3 = \tau_2 \cdot \tau_1$?

Durch $\tau_2 \cdot \tau_1$ wird 1 auf 2 und 2 auf 3 abgebildet - das passt. 3 wird auf 1 abgebildet, sie soll aber auf 4 abgebildet werden. Also sind wir noch nicht fertig.

Die dritte Transposition muss demnach 1 auf 4 abbilden:

$$\tau_3 := \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix} = (1\;4).$$

Und nun gilt tatsächlich $\sigma_3 = \tau_3 \cdot \tau_2 \cdot \tau_1$.

Verständnisfragen 8.8

1. Was ist eigentlich das „Produkt“ zweier Transpositionen? Transpositionen sind doch Abbildungen, wie kann man die *multiplizieren*?
2. Müssen wir im Fall σ_3 unbedingt damit anfangen, eine Transposition zu finden, die 1 auf $\sigma_3(1)$ abbildet? Oder hätten wir auch 2, 3 oder 4 als „Startpunkte“ wählen können?
3. Überprüfen Sie, wohin 1, 2, 3 und 4 durch das Produkt der Transpositionen $\tau_3 \cdot \tau_2 \cdot \tau_1$ abgebildet werden. Kommt jeweils $\sigma_3(1)$, $\sigma_3(2)$, $\sigma_3(3)$ bzw. $\sigma_3(4)$ heraus?

Übung 8.8 Ordnen Sie nun selbst den einzelnen Bearbeitungsschritten der Musterlösung die Phasen aus Kapitel 3 zu! Tragen Sie dazu die jeweiligen Buchstaben in die kleinen Quadrate am Seitenrand ein!

Arbeitsauftrag 8.8 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen, der Übung und dem Arbeitsauftrag finden Sie in Abschnitt 11.4.

9 Musterlösungen aus der Linearen Algebra 2

9.1 Vandermondesche Determinante

Aufgabe 9.1 Seien $a_1, a_2, \ldots, a_n$ Elemente eines Körpers K, wobei $n \geq 2$ sei. Betrachte die $(n \times n)$-Matrix $V(a_1, \ldots, a_n) := (v_{ij})$ mit $v_{ij} := a_i^{(j-1)}$ für $1 \leq i, j \leq n$. Zeigen Sie: $\det(V(a_1, \ldots, a_n)) = \prod\limits_{1 \leq j < i \leq n} (a_i - a_j)$.

Ausführliche Lösung 9.1 In dieser Aufgabe soll die Formel für die Determinante der Vandermonde-Matrix $V(a_1, \ldots, a_n)$ bewiesen werden.

Zunächst sollte man sich zum Lösen dieser Aufgabe überlegen, wie die Vandermonde-Matrix aussieht. $V(a_1, \ldots, a_n) := (v_{ij})$ ist durch $v_{ij} = a_i^{(j-1)}$ gegeben. Das heißt, dass in der i-ten Zeile der Matrix Potenzen von a_i stehen, und zwar von $a_i^0 = 1$ bis a_i^{n-1}. Die Vandermonde-Matrix sieht demnach so aus:

$$V(a_1, \ldots, a_n) = \begin{pmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{pmatrix}.$$

Für den Beweis, dass

$$\det(V(a_1, \ldots, a_n)) = \prod_{1 \leq j < i \leq n} (a_i - a_j)$$

ist, bietet sich eine *Induktion* nach n an.

Der Induktionsanfang für $n = 2$ ist schnell gemacht:

$$\det(V(a_1, a_2)) = \begin{vmatrix} 1 & a_1 \\ 1 & a_2 \end{vmatrix} = 1 \cdot a_2 - 1 \cdot a_1 = \prod_{j=1,\ i=2} (a_i - a_j) = \prod_{1 \leq j < i \leq 2} (a_i - a_j).$$

Machen wir nun den Induktionsschritt von $n - 1$ nach n. Hat man hier erst einmal die richtige Idee zur Berechnung der Vandermondeschen Determinante, so ist der Beweis relativ zügig erbracht. Wir wollen ihn uns Schritt für Schritt überlegen.

C. Ableitinger, A. Herrmann, *Lernen aus Musterlösungen zur Analysis und Linearen Algebra*,
DOI 10.1007/978-3-658-02353-9_9, © Springer Fachmedien Wiesbaden 2013

Dazu sammeln wir zunächst alle uns zur Verfügung stehenden Möglichkeiten zur Berechnung einer Determinante und überlegen, welche uns davon für diesen speziellen Fall weiterhelfen. Dabei sollten wir immer im Blick haben, dass wir die Determinante für die $(n \times n)$-Matrix auf die uns bekannte Induktionsvoraussetzung (also die Determinante einer $(n-1 \times n-1)$-Vandermonde-Matrix) zurückführen wollen.

Machen wir uns zunächst eine Liste von Berechnungsverfahren:

- Formel für Dreiecksmatrizen,
- Kästchensatz: $\det \begin{pmatrix} A & * \\ 0 & B \end{pmatrix} = \det(A) \cdot \det(B)$,
- Laplacesche Entwicklung nach Zeilen oder Spalten,
- Elementarumformungen wie Vertauschen zweier Zeilen oder Spalten, Vervielfachen einer Zeile bzw. Spalte, Addition eines Vielfachen einer Zeile bzw. Spalte zu einer anderen (Achtung: Der Wert der Determinante ändert sich dabei!),
- Regel von Sarrus und zur Berechnung der Determinante von (2×2)-Matrizen.

Gut, welche davon können wir hier anwenden? Eine Dreiecksmatrix haben wir nicht gegeben und für den Kästchensatz müssten in der Matrix Kästchen auftauchen, in denen nur Nullen stehen. Diese beiden Formeln helfen uns also zunächst nicht weiter. Da wir eine allgemeine $(n \times n)$-Matrix vorliegen haben, können wir auch nicht die Regel von Sarrus oder die entsprechende Regel für (2×2)-Matrizen anwenden. Es bleibt also nur noch die Entwicklung nach Laplace oder die elementaren Umformungen. Eine Entwicklung nach Laplace empfiehlt sich allerdings auch nur, wenn möglichst viele Nullen in einer Zeile oder Spalte stehen. Denn andernfalls muss man viele neue Determinanten berechnen, da hat man nicht allzu viel gewonnen. Das heißt für unseren Fall aber nun, dass wir mit elementaren Umformungen versuchen, möglichst viele Nullen in einer Zeile oder Spalte zu bekommen. Dann können wir nach Laplace entwickeln und erhalten eine $(n-1 \times n-1)$-Matrix oder mehrere.

Dazu bieten sich nun mehrere Möglichkeiten an. Die erste, die einem da wohl ins Auge springt, ist, die erste Zeile von allen anderen abzuziehen, da dann in der ersten Spalte nur noch eine Eins und sonst Nullen stehen:

$$\begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix} = \begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 0 & a_2 - a_1 & a_2^2 - a_1^2 & \cdots & a_2^{n-1} - a_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_n - a_1 & a_n^2 - a_1^2 & \cdots & a_n^{n-1} - a_1^{n-1} \end{vmatrix}.$$

Jetzt könnte man mit Laplace nach der ersten Spalte entwickeln. Nur kann man auf diese Matrix nicht die Induktionsvoraussetzung anwenden, da es sich nicht um eine Vandermonde-Matrix handelt.

Man erhält so nämlich:

$$\ldots \overset{\text{Lapl.}}{=} \begin{vmatrix} a_2 - a_1 & a_2^2 - a_1^2 & \cdots & a_2^{n-1} - a_1^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_n - a_1 & a_n^2 - a_1^2 & \cdots & a_n^{n-1} - a_1^{n-1} \end{vmatrix}.$$

Mit einfachen Zeilen- oder Spaltenumformungen kommen wir hier nicht mehr weiter. Also überlegen wir uns jetzt, wie wir in der Ausgangsmatrix durch Spaltenumformungen möglichst viele Nullen in die erste Zeile bekommen können.

Man könnte das a_1^j-fache der ersten Spalte von der $(j+1)$-ten Spalte abziehen (für alle $j = 1, \dots, n-1$).

Dann erhalten wir allerdings eine ähnliche Matrix wie bei den Zeilenumformungen und auf die können wir wieder *nicht* die Induktionsvoraussetzung anwenden!

Welche andere Möglichkeit, um Nullen zu bekommen, gibt es nun noch? Wie wäre es damit:

Wir ziehen zunächst das a_1-fache der $(n-1)$-ten Spalte von der n-ten Spalte ab, dann das a_1-fache der $(n-2)$-ten Spalte von der $(n-1)$-ten Spalte und so weiter bis wir das a_1-fache der ersten Spalte von der zweiten abziehen.

Damit erhalten wir folgendes Ergebnis:

$$\begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & a_2 - a_1 & a_2^2 - a_1 \cdot a_2 & \cdots & a_2^{n-1} - a_1 \cdot a_2^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n - a_1 & a_n^2 - a_1 \cdot a_n & \cdots & a_n^{n-1} - a_1 \cdot a_n^{n-2} \end{vmatrix}.$$

Sieht das jetzt so viel besser aus? Da steht doch immer noch keine Vandermonde-Matrix! Stimmt, aber fast, wie man an dieser Umformung sieht:

$$\dots = \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & (a_2 - a_1) \cdot 1 & (a_2 - a_1) \cdot a_2 & \cdots & (a_2 - a_1) \cdot a_2^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & (a_n - a_1) \cdot 1 & (a_n - a_1) \cdot a_n & \cdots & (a_n - a_1) \cdot a_n^{n-2} \end{vmatrix}.$$

Jetzt entwickeln wir mit Laplace nach der ersten Zeile, es folgt:

$$\dots = \begin{vmatrix} (a_2 - a_1) & (a_2 - a_1) \cdot a_2 & \cdots & (a_2 - a_1) \cdot a_2^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ (a_n - a_1) & (a_n - a_1) \cdot a_n & \cdots & (a_n - a_1) \cdot a_n^{n-2} \end{vmatrix}.$$

Die Faktoren in den Zeilen darf man nun aber aus der Determinante herausziehen und erhält:

$$\dots = (a_2 - a_1) \cdot \ldots \cdot (a_n - a_1) \begin{vmatrix} 1 & a_2 & \cdots & a_2^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & \cdots & a_n^{n-2} \end{vmatrix}.$$

Nun haben wir tatsächlich wieder eine Vandermonde-Matrix erhalten, auf die wir die Induktionsvoraussetzung anwenden dürfen.

Man muss ein bisschen aufpassen, da die Indizes erst bei 2 beginnen. Für diese $(n-1 \times n-1)$-Matrix gilt nach Induktionsvoraussetzung:

$$\begin{vmatrix} 1 & a_2 & \cdots & a_2^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & \cdots & a_n^{n-2} \end{vmatrix} = \prod_{2 \le j < i \le n} (a_j - a_i).$$

Daraus folgt für die Vandermondesche Determinante:

$$\begin{aligned} V(a_1, \dots, a_n) = \dots &= (a_2 - a_1) \cdot \dots \cdot (a_n - a_1) \prod_{2 \le j < i \le n} (a_i - a_j) \\ &= \prod_{j=1,\ 1 < i \le n} (a_i - a_j) \prod_{2 \le j < i \le n} (a_i - a_j) \\ &= \prod_{1 \le j < i \le n} (a_i - a_j). \end{aligned}$$

Damit haben wir auch den Induktionsschritt zu Ende geführt.

Verständnisfragen 9.1

1. Der Beweis wurde mit Hilfe von Induktion durchgeführt. An sich kann man doch die Determinante auch direkt ausrechnen und somit zeigen, dass die Behauptung gilt. Warum macht man das nicht?
2. Welche Matrix erhält man durch die zweite vorgestellte Vorgehensweise (das a_1^j-fache der ersten Spalte der Ausgangsmatrix von der $j+1$-ten Spalte abziehen für $j = 1, \dots, n-1$)?
3. Wie ist $\prod_{1 \le j < i \le n} (a_i - a_j)$ zu verstehen? Schreiben Sie mit Hilfe der „Pünktchen-Schreibweise" das Produkt aus! Überlegen Sie sich dann, warum die im Beweis auftauchende Gleichheit

$$(a_2 - a_1) \cdot \dots \cdot (a_n - a_1) \prod_{2 \le j < i \le n} (a_i - a_j) = \prod_{j=1,\ 1 < i \le n} (a_i - a_j) \prod_{2 \le j < i \le n} (a_i - a_j)$$

richtig ist.

Übung 9.1 Ordnen Sie nun selbst den einzelnen Bearbeitungsschritten der Musterlösung die Phasen aus Kapitel 3 zu! Tragen Sie dazu die jeweiligen Buchstaben in die kleinen Quadrate am Seitenrand ein!

Arbeitsauftrag 9.1 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen, der Übung und dem Arbeitsauftrag finden Sie in Abschnitt 11.5.

9.2 Dualraum

Aufgabe 9.2 Sei $V = K^n$ der Vektorraum der Zeilenvektoren der Länge n über dem Körper K. Für $w \in V$ sei die Abbildung

$$f_w : V \to K \text{ durch } f_w(v) := w \cdot v^T$$

definiert.
Zeigen Sie:

(i) $f_w \in V^*$ für $w \in V$.

(ii) Ist $\mathcal{E} := \{e_1, \dots, e_n\}$ die Standardbasis von V, so ist $\mathcal{E}^* := \{f_{e_1}, \dots, f_{e_n}\}$ die duale Basis zu $\mathcal{E}$.

(iii) Für die zu id_V duale Abbildung $(\mathrm{id}_V)^*$ gilt: $(\mathrm{id}_V)^* = \mathrm{id}_{V^*}$.

Ausführliche Lösung 9.2 In dieser Aufgabe geht es um das Konzept des Dualraums. P

(i) Z. z.: $f_w \in V^*$ für alle $w \in V$. B

Als Erstes sollte man sich klar machen, was dafür zu zeigen ist.

Der Dualraum V^* ist der Vektorraum, der aus allen *linearen* Abbildungen $f : V \to K$ besteht (also aus linearen Abbildungen, die einem Vektor des Vektorraumes V ein Element aus dem Körper zuordnen). Z

Wir müssen deshalb einerseits zeigen, dass wir durch das Anwenden der Abbildung auf einen beliebgien Vektor v tatsächlich ein Element des Körpers erhalten. Andererseits müssen wir noch überprüfen, ob f_w auch linear ist.

Da f_w einen Vektor $v \in V$ auf $w \cdot v^T$ abbildet, müssen wir zeigen, dass $w \cdot v^T \in K$ ist. A

Bei v und w handelt es sich um Zeilenvektoren und wir können $v = (v_1, \dots, v_n)$ und $w = (w_1, \dots, w_n)$ für geeignete $v_i, w_i \in K$ schreiben. Z

Daraus folgt mit der Matrizenmultiplikation: H

$$w \cdot v^T = (w_1, \dots, w_n) \cdot (v_1, \dots, v_n)^T = (w_1, \dots, w_n) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \sum_{i=1}^{n} w_i v_i.$$

Da $w_i, v_i \in K$ sind, folgt aber direkt $w \cdot v^T = \sum_{i=1}^{n} w_i v_i \in K$. Also ist f_w wirklich eine Abbildung von V nach K.

Überprüfen wir nun die Linearität: Wir erinnern uns, dass dafür B

a) $\forall\, v, u \in V$: $f_w(v + u) = f_w(v) + f_w(u)$ und

b) $\forall\, v \in V$ und $\forall\, \lambda \in K$: $f_w(\lambda v) = \lambda f_w(v)$

zu zeigen ist.

Beweis von a): Seien $v, u \in V$ beliebig, dann gilt:

$$\begin{aligned}
f_w(v+u) &= w \cdot (v+u)^T && \text{Definition von } f_w \\
&= w \cdot (v^T + u^T) && \text{Rechenregel für Transponierte} \\
&= w \cdot v^T + w \cdot u^T && \text{Distributivgesetz} \\
&= f_w(v) + f_w(u). && \text{Definition von } f_w
\end{aligned}$$

Beweis von b): Sei $v \in V$ und $\lambda \in K$ beliebig, dann gilt:

$$\begin{aligned}
f(\lambda v) &= w \cdot (\lambda v)^T \\
&= w \cdot (\lambda v^T) \\
&= \lambda \cdot w \cdot v^T \\
&= \lambda \cdot f_w(v).
\end{aligned}$$

Damit haben wir gezeigt, dass $f_w : V \to K$ linear ist und somit in V^* liegt.

(ii) Z. z.: $\mathcal{E}^* = \{f_{e_1}, \ldots, f_{e_n}\}$ ist die duale Basis zur Standardbasis $\mathcal{E} = \{e_1, \ldots, e_n\}$.

Auch hier machen wir uns klar, was eigentlich zu zeigen ist. Dafür müssen wir wissen, wie die zu einer Basis von V duale Basis definiert ist.

Hat man eine Basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ von V gegeben, dann heißt die Basis

$$\mathcal{B}^* = \{f_1, \ldots, f_n\} \text{ mit } f_i(v_j) := \delta_{ij} \text{ für } i, j = 1, \ldots, n$$

die duale Basis zu $\mathcal{B}$. Hier müssen wir also beweisen, dass $f_{e_i}(e_j) = \delta_{ij}$ für $i, j = 1, \ldots, n$ gilt[1].

Das ist aber nicht allzu schwer, denn

$$\begin{aligned}
f_{e_i}(e_j) = e_i \cdot e_j^T = \underset{\substack{\uparrow \\ i\text{-te Spalte}}}{(0, \ldots, 0, 1, 0 \ldots, 0)} \cdot &\begin{array}{cc} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} & \begin{array}{l} \\ \\ \\ \leftarrow j\text{-te Zeile} \\ \\ \\ \\ \end{array} \end{array} \\
= \left\{ \begin{array}{ll} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j \end{array} \right\} = \delta_{ij}.
\end{aligned}$$

Also ist $\mathcal{E}^*$ tatsächlich die zur Standardbasis duale Basis.

(iii) Z. z.: $(\mathrm{id}_V)^* = \mathrm{id}_{V^*}$

Auch hier muss man sich zunächst klar machen, was die duale Abbildung zu einer vorgegebenen Abbildung ist.

[1] δ_{ij} ist das Kroneckersymbol, das heißt es gilt $\delta_{ij} = 1$ für $i = j$ und $\delta_{ij} = 0$ für $i \neq j$.

Allgemein gilt: Ist $A \in \mathrm{Hom}_K(V,W)$ eine lineare Abbildung von V nach W, dann heißt $A^* : W^* \to V^*$ mit $(A^*(f))(v) := f(A(v))$ die duale Abbildung zu A.

Übertragen auf $A := \mathrm{id}_V$ heißt das: $A \in \mathrm{Hom}_K(V,V)$, also ist $A^* \in \mathrm{Hom}_K(V^*,V^*)$. A^* ist also eine lineare Abbildung, die eine Linearform von V (also ein $f \in V^*$) wieder auf eine Linearform $g \in V^*$ abbildet und zwar wie folgt:

$$g = A^*(f) : V \mapsto K \text{ mit } (A^*(f))(v) := f(A(v)).$$

Machen Sie sich noch einmal genau klar, was hier eigentlich passiert! A^* ist eine Abbildung, die eine Abbildung $f \in V^*$ auf eine Abbildung $g = A^*(f) \in V^*$ abbildet. Und die Abbildung g ist dabei wie oben definiert (es handelt sich bei g um eine Abbildung von V nach K!):

$$g(v) := f(A(v)).$$

Alles verständlich soweit? Sich das klar zu machen, ist das Schwierige an der Aufgabe!

Nun sollen wir zeigen, dass $A^* = \mathrm{id}_{V^*}$ ist.

Das heißt, wir müssen zeigen, dass $A^*(f) = f$ ist für alle $f \in V^*$. Dafür nimmt man sich wie gewohnt ein beliebiges $f \in V^*$ und zeigt es für dieses f.

Gut, machen wir das! Sei $f \in V^*$ beliebig.

Wir müssen nun ja zeigen, dass $A^*(f) = f$ ist. Das heißt, wir müssen die Gleichheit von zwei Abbildungen zeigen. Und wie machen wir das? Richtig, indem wir zeigen, dass die beiden Abbildungen in jedem Punkt $v \in V$ (denn f und $A^*(f)$ bilden ja von V aus ab) übereinstimmen.

Also sei nun $v \in V$ beliebig gewählt. Wir zeigen $(A^*(f))(v) = f(v)$:

$$(A^*(f))(v) \overset{\text{Def. } A^*}{=} f(A(v)) \overset{A=\mathrm{id}_V}{=} f(\mathrm{id}_V(v)) = f(v).$$

Und fertig! Also gilt tatsächlich $(\mathrm{id}_V)^*(f) = A^*(f) = f$ für alle $f \in V^*$ und damit $(\mathrm{id}_V)^* = \mathrm{id}_{V^*}$.

Verständnisfragen 9.2

1. Schreiben Sie auch für den Beweis von b) „$\forall\ v \in V$ und $\forall\ \lambda \in K$: $f_w(\lambda v) = \lambda f_w(v)$" wie im Beweis von a) die Begründungen der einzelnen durchgeführten Schritte auf!

2. In der Lösung zu (iii) wird die duale Abbildung A^* für $A \in \mathrm{Hom}_K(V,V)$ beschrieben. Versuchen Sie für ein allgemeines $A \in \mathrm{Hom}_K(V,W)$ auch eine solche Beschreibung für die duale Abbildung zu formulieren! A^* ist eine lineare Abbildung, die eine Linearform …

3. In (iii) haben wir gezeigt, dass $A^* = \mathrm{id}_{V^*}$ gilt, indem wir für alle $f \in V^*$ die Gleichheit $A^*(f) = f$ bewiesen haben. Warum ist dieser Beweis ausreichend?

Arbeitsauftrag 9.2 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.5.

9.3 Jordansche Normalform ohne Basiswechselmatrix

Aufgabe 9.3 Bestimmen Sie die Jordansche Normalform der Matrix

$$A = \begin{pmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 7 & 1 & 2 & 1 \\ -17 & -6 & -1 & 0 \end{pmatrix} \in M_{4\times 4}(\mathbb{R}).$$

Ausführliche Lösung 9.3 Es gibt zwei Wege zur Bestimmung der Jordanschen Normalform[2] einer Abbildung bzw. Matrix A. Einmal kann man sich überlegen, wie die Jordansche Normalform aussehen muss, ohne dass man eine Transformationsmatrix[3] angibt. Zum anderen kann man sich eine Transformationsmatrix konstruieren und dann durch Matrizenmultiplikation die JNF berechnen. Letzteres ist weitaus aufwendiger, deswegen wählt man diesen Weg oft nur, wenn man die Transformationsmatrix für spätere Berechnungen braucht. Hier wollen wir zunächst nur die erste Methode durchführen[4].

Zunächst wollen wir uns überlegen, in welchen Schritten man dazu allgemein vorgehen kann:

1. Bestimme die Eigenwerte λ_i von A durch Berechnen der Nullstellen des charakteristischen Polynoms.
2. Die Anzahl aller Jordanblöcke zu einem festen Eigenwert λ_i ist gleich der Dimension des Eigenraumes $\dim(V(\lambda_i, A))$.
3. Die Größe m_i des größten Blocks zum Eigenwert λ_i ist gleich der Potenz, mit der der Linearfaktor $(x - \lambda_i)$ im Minimalpolynom vorkommt. m_i ist gleichzeitig auch die kleinste Zahl, für die $V^{m_i}(\lambda_i, A) = V^{m_i+1}(\lambda_i, A)$ gilt.

Diese Punkte reichen meist schon, um die Jordansche Normalform in eindeutiger Weise zu bestimmen (so auch im vorliegenden Fall). Machen wir dies nun konkret für die Matrix A:

1.

$$\mathrm{charpol}_A(x) = \begin{vmatrix} x-3 & -1 & 0 & 0 \\ 4 & x+1 & 0 & 0 \\ -7 & -1 & x-2 & -1 \\ 17 & 6 & 1 & x \end{vmatrix}$$

[2] Wird ab jetzt auch mit JNF abgekürzt.
[3] Häufig auch als Basiswechselmatrix bezeichnet.
[4] Die zweite Methode finden Sie in der ausführlichen Musterlösung zur nächsten Aufgabe.

$$\overset{\text{Lapl. 4. Sp.}}{=} \begin{vmatrix} x-3 & -1 & 0 \\ 4 & x+1 & 0 \\ 17 & 6 & 1 \end{vmatrix} + x \begin{vmatrix} x-3 & -1 & 0 \\ 4 & x+1 & 0 \\ -7 & -1 & x-2 \end{vmatrix}$$

$$\begin{aligned} &= (x-3)(x+1) + 4 + x[(x-3)(x+1)(x-2) + 4(x-2)] \\ &= (x^2 - 2x + 1) + x(x-2)[x^2 - 2x + 1] \\ &= (x - 2x + 1)[1 + x(x-2)] = (x-1)^2[x^2 - 2x + 1] \\ &= (x-1)^4. \end{aligned}$$

Die Matrix hat also nur den Eigenwert 1 und damit sind alle Diagonaleinträge der Jordanschen Normalform 1.

2. Berechnen wir die Anzahl aller Jordanmatrizen. Da wir uns dabei nur für die Dimension des Kerns interessieren, können wir diese auch mit Hilfe des Rangs bestimmen. Der Rang einer Matrix lässt sich häufig leichter berechnen, da man dazu nur die Anzahl der linear unabhängigen Zeilen bzw. Spalten bestimmen muss.

$$\dim(V(1, A)) = \dim(\ker(A - 1E_4)) = 4 - \operatorname{rg}(A - 1E_4)$$

$$= 4 - \operatorname{rg} \begin{pmatrix} 2 & 1 & 0 & 0 \\ -4 & -2 & 0 & 0 \\ 7 & 1 & 1 & 1 \\ -17 & -6 & -1 & -1 \end{pmatrix}$$

$$= 4 - \operatorname{rg} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 5 & 1 & 1 & 0 \\ -5 & -6 & -1 & 0 \end{pmatrix}$$

$$= 4 - \operatorname{rg} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -6 & -1 & 0 \end{pmatrix} = 4 - 2 = 2$$

Also hat die Matrix zwei Jordanblöcke.

Da bleiben aber nur die Möglichkeiten, dass beide die Größe 2 oder, dass einer Größe 3 und der andere Größe 1 hat. Denn insgesamt muss sich wieder eine Matrix der Größe 4 ergeben!

Wenn wir m (Exponent von $(x-1)$ im Minimalpolynom) kennen, wissen wir, was die maximale Größe der Jordanblöcke ist (es muss hier natürlich 2 oder 3 herauskommen), und damit kennen wir dann schon die Jordansche Normalform. Kommen wir deswegen zu

3. Bestimmen wir das Minimalpolynom von A: Wir wissen

$$\operatorname{minpol}_A(x) = (x-1)^m \text{ mit } 1 \leq m \leq 4.$$

Es gilt $A - 1E_4 \neq 0$, also berechnen wir als Nächstes $(A - 1E_4)^2$:

$$(A - 1E_4)^2 = \begin{pmatrix} 2 & 1 & 0 & 0 \\ -4 & -2 & 0 & 0 \\ 7 & 1 & 1 & 1 \\ -17 & -6 & -1 & -1 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Also ist $m = 2$, damit ist der größte Jordanblock eine (2×2)-Matrix.

Falls man das Minimalpolynom nicht kennt, muss man hier mit Hilfe des verallgemeinerten Eigenraumes argumentieren. Dazu bestimmen wir diesen. (Zur Erinnerung: Es gilt die Inklusionskette $V(\lambda_i, A) \subseteq V^2(\lambda_i, A) \subseteq \ldots$. Der Raum, ab dem die Kette zum ersten Mal konstant wird (also nicht nur „$\subseteq$" sondern „=" gilt), ist der verallgemeinerte Eigenraum.)

Wir berechnen also als Nächstes $V^2(1, A)$:

$$V^2(1, A) = \ker \begin{pmatrix} 2 & 1 & 0 & 0 \\ -4 & -2 & 0 & 0 \\ 7 & 1 & 1 & 1 \\ -17 & -6 & -1 & -1 \end{pmatrix}^2 = \ker \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \mathbb{R}^4.$$

Damit muss es sich bei $V^2(1, A)$ schon um den verallgemeinerten Eigenraum handeln. Es folgt also auch hier, dass $m = 2$ ist.

Da wir nun aber zwei Jordanblöcke haben, wobei der größte Block die Größe 2 hat und die Blöcke zusammen eine Matrix der Größe 4 ergeben müssen, gibt es demnach zwei Jordanblöcke der Größe 2.

Damit erhalten wir die Jordansche Normalform von A:

$$\operatorname{JNF}(A) = \left(\begin{array}{cc:cc} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hdashline 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right).$$

Verständnisfragen 9.3

1. Welche allgemeingültige Formel steckt hinter folgender, im zweiten Schritt verwendeten Gleichheit: $\dim(\ker(A - 1E_4)) = 4 - \operatorname{rg}(A - 1E_4)$?
2. Im dritten Schritt wurde behauptet, dass $A - 1E_4 \neq 0$ gilt. Woher wissen wir das zu dem Zeitpunkt bereits? Woran erkennt man dies auch ganz schnell an der Matrix A?
3. Warum folgt aus $A - 1E_4 \neq 0$ und $(A - 1E_4)^2 = 0$, dass $m = 2$ und somit $\operatorname{minpol}_A(x) = (x-1)^2$ ist?
4. Im zweiten Lösungsansatz für den dritten Schitt wurde aus $V^2(1, A) = \mathbb{R}^4$ gefolgert, dass $V^\infty(1, A) = V^2(1, A)$ ist. Warum folgt das wirklich?

Übung 9.3 Ordnen Sie nun selbst den einzelnen Bearbeitungsschritten der Musterlösung die Phasen aus Kapitel 3 zu! Tragen Sie dazu die jeweiligen Buchstaben in die kleinen Quadrate am Seitenrand ein!

Arbeitsauftrag 9.3 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen, der Übung und dem Arbeitsauftrag finden Sie in Abschnitt 11.5.

9.4 Jordansche Normalform mit Basiswechselmatrix

Aufgabe 9.4 Gegeben sei die Matrix $A = \begin{pmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 7 & 1 & 2 & 1 \\ -17 & -6 & -1 & 0 \end{pmatrix} \in M_{4\times 4}(\mathbb{R})$.

Bestimmen Sie eine Matrix S, sodass $S^{-1} \cdot A \cdot S$ in Jordanscher Normalform vorliegt.

Ausführliche Lösung 9.4 *Vorbemerkung:* Im folgenden wird ein Schema formuliert, um für einen Endomorphismus auf einem n-dimensionalen Vektorraum V bzw. für eine $(n \times n)$-Matrix eine Basis zu finden, bzgl. derer der Endomorphismus bzw. die Matrix in Jordanscher Normalform vorliegt. Das Schema ergibt sich aus einem konstruktiven Beweis zur Existenz der Jordanschen Normalform. (Informieren Sie sich gegebenenfalls in Lehrbüchern der Linearen Algebra darüber.) B

Zu der Matrix A wurde im letzten Abschnitt bereits eine Jordansche Normalform bestimmt, allerdings wurde keine Matrix S angegeben, sodass $S^{-1}AS$ die Jordansche Normalform ergibt. Das sollen wir hier nun tun.

Die gesuchte Matrix S ergibt sich, indem man die Vektoren von zyklischen Basen in die Spalten der Matrix S schreibt (um welche Basen es sich dabei genau handelt, sehen wir gleich). Ein Ergebnis der Linearen Algebra besagt, dass man verallgemeinerte Eigenräume in eine direkte Summe zyklischer Unterräume[5] zerlegen kann. Das heißt aber auch, dass wir den Vektorraum V (hier: $V = \mathbb{R}^4$) in zyklische Unterräume zerlegen können. (Denn den VR können wir bei trigonalisierbaren Endomorphismen/Matrizen als direkte Summe der verallgemeinerten Eigenräume schreiben. Wenn diese wiederum gleich direkten Summen aus zyklischen Unterräumen sind, folgt die Behauptung.) P

Auf folgende Weise findet man die Zerlegung eines verallgemeinerten Eigenraumes $\ker(A - \lambda E_n)^m$ in zyklische Unterräume (und insbesondere Basen zu diesen zyklischen Unterräumen, deren Vektoren die Spalten unserer Matrix S ergeben): K

[5] Die Unterräume sind dabei zyklisch bezüglich der Abbildung $A - \lambda E_n$, wobei λ der zum verallgemeinerten Eigenraum gehörige Eigenwert ist.

1. Wir ergänzen eine Basis von $\ker(A-\lambda E_n)^{m-1}$ durch $b_1^{(m)},\dots,b_{l_m}^{(m)}$ zu einer Basis von $\ker(A-\lambda E_n)^m$. Die Basis von $\ker(A-\lambda E_n)^{m-1}$ wird also durch l_m Vektoren zu einer Basis des verallgemeinerten Eigenraumes ergänzt. Die Ergänzung können wir vornehmen, da $\ker(A-\lambda E_n)^{m-1} \subsetneq \ker(A-\lambda E_n)^m$ ist.
2. Wir bilden
$$(A-\lambda E_n)(b_1^{(m)}),\dots,(A-\lambda E_n)(b_{l_m}^{(m)}).$$
Bei diesen Vektoren handelt es sich dann um Vektoren aus $\ker(A-\lambda E_n)^{m-1}$, die sogar linear unabhängig sind und nicht in $\ker(A-\lambda E_n)^{m-2}$ liegen.
3. Wir ergänzen eine Basis von $\ker(A-\lambda E_n)^{m-2}$ durch
$$(A-\lambda E_n)(b_1^{(m)}),\dots,(A-\lambda E_n)(b_{l_m}^{(m)})$$
und $b_1^{(m-1)},\dots,b_{l_{m-1}}^{(m-1)}$ zu einer Basis von $\ker(A-\lambda E_n)^{m-1}$.
4. Wir bilden
$$\begin{aligned}&(A-\lambda E_n)^2(b_1^{(m)}),\dots,(A-\lambda E_n)^2(b_{l_m}^{(m)}) \quad \text{und}\\ &(A-\lambda E_n)(b_1^{(m-1)}),\dots,(A-\lambda E_n)(b_{l_{m-1}}^{(m-1)}).\end{aligned}$$
Das sind linear unabhängige Vektoren aus $\ker(A-\lambda E_n)^{m-2}\setminus\ker(A-\lambda)^{m-3}$.
5. Wir ergänzen eine Basis von $\ker(A-\lambda E_n)^{m-3}$ durch die Vektoren aus 4. und $b_1^{(m-2)},\dots,b_{l_{m-2}}^{(m-2)}$ zu einer Basis von $\ker(A-\lambda E_n)^{m-2}$.
6. Wir bilden
$$\begin{aligned}&(A-\lambda E_n)^3(b_1^{(m)}),\dots,(A-\lambda E_n)^3(b_{l_m}^{(m)})\ ,\\ &(A-\lambda E_n)^2(b_1^{(m-1)}),\dots,(A-\lambda E_n)^2(b_{l_{m-1}}^{(m-1)}) \quad \text{und}\\ &(A-\lambda E_n)(b_1^{(m-2)}),\dots,(A-\lambda E_n)(b_{l_{m-2}}^{(m-2)}).\end{aligned}$$
Das sind linear unabhängige Vektoren aus $\ker(A-\lambda E_n)^{m-3}\setminus\ker(A-\lambda E_n)^{m-4}$.
7. Eine Basis von $\ker(A-\lambda E_n)^{m-4}$ ergänzen wir dann wiederum durch die Vektoren aus 6. und weitere Vektoren $(b_1^{(m-3)},\dots,b_{l_{m-3}}^{(m-3)})$ zu einer Basis von $\ker(A-\lambda E_n)^{m-3}$. Dann bilden wir wieder die Bilder *all* dieser Vektoren unter $(A-\lambda E_n)$. Und führen das Ganze so lange fort, bis wir eine Basis von $\ker(A-\lambda E_n)^0$ – also von $\{0_V\}$ – durch die vorhergehenden Vektoren und die zusätzlichen Vektoren $(b_1^{(1)},\dots,b_{l_1}^{(1)})$ zu einer von $\ker(A-\lambda E_n)^1$ ergänzen. (In dem Fall heißt das eigentlich, dass wir die bis dahin entwickelten Vektoren zu einer Basis des Eigenraumes ergänzen.)

8. Die zyklischen Basen zu den zyklischen Unterräumen sind dann:

$$\begin{aligned}
&(A-\lambda E_n)^{m-1}(b_1^{(m)}), (A-\lambda E_n)^{m-2}(b_1^{(m)}), \ldots, (A-\lambda E_n)(b_1^{(m)}), b_1^{(m)}\\
&\vdots\\
&(A-\lambda E_n)^{m-1}(b_{l_m}^{(m)}), (A-\lambda E_n)^{m-2}(b_{l_m}^{(m)}), \ldots, (A-\lambda E_n)(b_{l_m}^{(m)}), b_{l_m}^{(m)}\\
&(A-\lambda E_n)^{m-2}(b_1^{(m-1)}), \ldots, (A-\lambda E_n)(b_1^{(m-1)}), b_1^{(m-1)}\\
&\vdots\\
&(A-\lambda E_n)^{m-2}(b_{l_{m-1}}^{(m-1)}), \ldots, (A-\lambda E_n)(b_{l_{m-1}}^{(m-1)}), b_{l_{m-1}}^{(m-1)}\\
&\vdots\\
&\\
&\vdots\\
&b_1^{(1)}\\
&\vdots\\
&b_{l_1}^{(1)}
\end{aligned}$$

Es kann dabei durchaus vorkommen, dass in manchen Schritten keine neuen Vektoren $b_1^{(j)}, b_2^{(j)}, \ldots$ hinzukommen. Es ist nämlich so, dass zu jedem $b_l^{(j)}$, das wir in irgendeinem Schritt ergänzen müssen, ein Jordanblock der Größe j in der Jordanschen Normalform gehört. Dort müssen ja aber nicht alle Größen von m bis 1 auftauchen, deshalb muss auch nicht überall unbedingt etwas ergänzt werden.

In unserem Beispiel suchen wir die Zerlegung der verallgemeinerten Eigenräume in zyklische Unterräume. Dazu überlegen wir erst einmal, wie viele verallgemeinerte Eigenräume es eigentlich gibt, und suchen dann zu jedem die Zerlegung.

Zur Bestimmung der JNF hatten wir im Abschnitt 9.3 bereits das charakteristische Polynom von A berechnet.

Es war:

$$\mathrm{charpol}_A(x) = (x-1)^4.$$

Wir haben also nur den Eigenwert 1 und damit auch nur einen verallgemeinerten Eigenraum.

Um diesen zu zerlegen, müssem wir nach obigem Schema Basen der Kerne $\ker(A - 1E_4), \ker(A-1E_4)^2, \ldots, \ker(A-1E_4)^m$ bestimmen. Wir fangen also mit der Berechnung des Eigenraumes an. (Wir erinnern uns, dass $V(1,A) = \ker(A-1E_4) \subsetneq \ker(A-1E_4)^2 \subsetneq \ldots \subsetneq \ker(A-1E_4)^m = \ker(A-1E_4)^{m+1}$ für ein $m \in \mathbb{N}$ gilt, und dann $\ker(A-1E_4)^m$ gerade der verallgemeinerte Eigenraum ist.)

H

$$V(1,A) = \ker(A - 1E_4) = \ker \begin{pmatrix} 2 & 1 & 0 & 0 \\ -4 & -2 & 0 & 0 \\ 7 & 1 & 1 & 1 \\ -17 & -6 & -1 & -1 \end{pmatrix}$$

$$= \ker \begin{pmatrix} 1 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -\frac{5}{2} & 1 & 1 \\ 0 & \frac{5}{2} & -1 & -1 \end{pmatrix} = \ker \begin{pmatrix} 1 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & -\frac{2}{5} & -\frac{2}{5} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$= \ker \begin{pmatrix} 1 & 0 & \frac{1}{5} & \frac{1}{5} \\ 0 & 1 & -\frac{2}{5} & -\frac{2}{5} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ -2 \\ -5 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 0 \\ -5 \end{pmatrix} \right\rangle$$

$$= \langle e_1 - 2e_2 - 5e_3, e_1 - 2e_2 - 5e_4 \rangle$$

$$V^2(1,A) = \ker(A - 1E_4)^2 = \ker \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \mathbb{R}^4 = \langle e_1, e_2, e_3, e_4 \rangle$$

$e_1, \ldots, e_4$ sei dabei die Standardbasis des $\mathbb{R}^4$.

H

Wie man leicht einsieht, ist dann auch $\ker(A - 1E_4)^3 = \mathbb{R}^4$. Also ist $\ker(A - 1E_4)^2$ der verallgemeinerte Eigenraum zum Eigenwert 1.

B

Diesen wollen wir nun in zyklische Unterräume zerlegen und insbesondere zyklische Basen zu diesen Unterräumen bestimmen. Denn die Zusammensetzung dieser zyklischen Basen liefert uns gerade die Matrix S.

A

In unserem Fall besteht der oben beschriebene Vorgang nur aus 4 Schritten, da der verallgemeinerte Eigenraum $\ker(A - 1E_4)^2$ ist (also $m = 2$). Führen wir das Schema einmal durch:

K

1. Wir ergänzen eine Basis von $\ker(A - 1E_4)$ zu einer Basis von $\ker(A - 1E_4)^2$, indem wir die Basisvektoren des Eigenraumes in die Basis des verallgemeinerten Eigenraumes eintauschen (Austauschsatz von Steinitz!).

H

   Den Vektor $e_1 - 2e_2 - 5e_3$ dürfen wir gegen e_1, e_2 oder e_3 tauschen. Tauschen wir z. B. gegen e_3, so bekommen wir die neue Basis des verallgemeinerten Eigenraumes:

   $$\{e_1 - 2e_2 - 5e_3, e_1, e_2, e_4\}.$$

   Den Vektor $e_1 - 2e_2 - 5e_4$ dürfen wir für e_1, e_2 oder e_4 in diese neue Basis eintauschen. Machen wir das hier einmal für e_4, dann bekommen wir die neue Basis:

   $$\{e_1 - 2e_2 - 5e_3, e_1 - 2e_2 - 5e_4, e_1, e_2\}.$$

Demnach haben wir jetzt die Basis von $\ker(A - 1E_4)$ durch die Vektoren e_1 und e_2 zu einer Basis des verallgemeinerten Eigenraumes ergänzt.

2. Von den ergänzten Vektoren sollen wir nun die Bilder unter der Abbildung $A - 1E_4$ bilden:

$$(A - 1E_4)(e_1) = \begin{pmatrix} 2 & 1 & 0 & 0 \\ -4 & -2 & 0 & 0 \\ 7 & 1 & 1 & 1 \\ -17 & -6 & -1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ -4 \\ 7 \\ -17 \end{pmatrix}$$

$$(A - 1E_4)(e_2) = \begin{pmatrix} 2 & 1 & 0 & 0 \\ -4 & -2 & 0 & 0 \\ 7 & 1 & 1 & 1 \\ -17 & -6 & -1 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 1 \\ -6 \end{pmatrix}.$$

3. Nun müssen wir noch $(A - 1E_4)(e_1)$ und $(A - 1E_4)(e_2)$ zu einer Basis von $\ker(A - 1E_4)$ ergänzen.

   In dem Fall ist aber nicht viel zu tun, denn die beiden Vektoren sind aus dem Eigenraum und sie sind linear unabhängig. Außerdem wissen wir, dass der Eigenraum die Dimension 2 hat. Also müssen die beiden Vektoren schon eine Basis des Eigenraumes bilden. Wir brauchen also gar keine Vektoren mehr ergänzen.

4. Die gesuchte Basis ist dann:

$$\{(A - 1E_4)(e_1), e_1 (A - 1E_4)(e_2), e_2\}.$$

Jetzt müssen wir die Basisvektoren nur noch in der richtigen Reihenfolge in die Matrix S schreiben, sodass dann $S^{-1}AS$ eine Jordansche Normalform ergibt. Die richtige Reihenfolge ist oben schon formuliert.

Wichtig ist, dass man die einzelnen zyklischen Basen nacheinander als Spalten in die Matrix schreibt und zwar in der Reihenfolge vom höchsten Exponenten absteigend.

Für uns ergibt sich also: $S = \begin{pmatrix} 2 & 1 & 1 & 0 \\ -4 & 0 & -2 & 1 \\ 7 & 0 & 1 & 0 \\ -17 & 0 & -6 & 0 \end{pmatrix}$. Durch Berechnung der Inversen von S ergibt sich: $S^{-1} = \frac{1}{25} \begin{pmatrix} 0 & 0 & 6 & 1 \\ 25 & 0 & 5 & 5 \\ 0 & 0 & -17 & -7 \\ 0 & 25 & -10 & -10 \end{pmatrix}$. (Überprüfen Sie das Ergebnis!)

Und schließlich folgt:

$$S^{-1} \cdot A \cdot S = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

B

(Führen Sie diese Matrizenmultiplikation noch einmal selbst durch und überprüfen Sie so das Ergebnis!) Somit haben wir eine Transformationsmatrix gefunden, die A in eine Jordansche Normalform überführt.

Verständnisfragen 9.4

1. Zur Bestimmung der zyklischen Basen wird eine Basis von $\ker(A - \lambda E_n)^{m-1}$ durch die Vektoren $b_1^{(m)}, \ldots, b_{l_m}^{(m)}$ zu einer Basis von $\ker(A - \lambda E_n)^m$ ergänzt. Warum gilt für diese Vektoren, wie behauptet, dass

$$(A - \lambda E_n)\left(b_i^{(m)}\right), \quad i = 1, \ldots, l_m$$

   Element von $\ker(A - \lambda E_n)^{m-1}$ aber nicht von $\ker(A - \lambda E_n)^{(m-2)}$ ist? Und wozu ist es wichtig, dass dies gilt?

2. In der Lösung wurde nach der Berechnung von $\ker(A - 1E_4)^2$ gesagt: „Wie man leicht einsieht, ist dann auch $\ker(A - 1E_4)^3 = \mathbb{R}^4$“. Welche Begründung steckt hinter dieser Aussage?

3. Probieren Sie aus, wie sich die Matrix $S^{-1} \cdot A \cdot S$ verändert, wenn man die Basisvektoren in einer anderen Reihenfolge in die Matrix S schreibt:

   a) Vertauschte Reihenfolge (vom kleinsten Exponenten zum größten) innerhalb der Zyklen:

$$e_1, (A - 1E_4)(e_1), e_2, (A - 1E_4)(e_2).$$

   b) Zusammenfassung der Vektoren nach Größe des Exponenten:

$$e_1, e_2, (A - 1E_4)(e_1), (A - 1E_4)(e_2).$$

Arbeitsauftrag 9.4 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.5.

9.5 Invariante Unterräume

Aufgabe 9.5 Sei

$$A = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 1 & 3 \\ 2 & 0 & 3 \end{pmatrix} \in M_{3\times 3}(\mathbb{R}).$$

Bestimmen Sie alle A-invarianten Unterräume U von $V = \mathbb{R}^3$!

Ausführliche Lösung 9.5 Ein Unterraum U heißt A-invariant, wenn $A(U) \subseteq U$ gilt.

K

Da U ein Unterraum von V ist, muss $0 \leq \dim(U) \leq \dim(V) = 3$ gelten. Es ist daher sinnvoll, eine Fallunterscheidung nach den vier verschiedenen möglichen Dimensionen für A-invariante Unterräume zu machen.

H

Ist $\dim(U) = 0$, dann ist $U = \{0_V\}$ – also der Unterraum, der nur aus dem Nullvektor besteht, denn einen anderen Unterraum dieser Dimension gibt es nicht! Und U ist dann auch A-invariant, da $A(0_V) = 0_V$ gilt.

Der Fall $\dim(U) = 3$ ist ebenfalls schnell abgehandelt, denn dann folgt schon $U = V$, da $U \subseteq V$ und $\dim(V) = 3$ ist. Weil $A : V \to V$ eine Abbildung von V nach V beschreibt, gilt $A(V) \subseteq V$. Das heißt aber, dass V und damit natürlich auch U A-invariant ist.

K

Für die beiden anderen Fälle $\dim(U) = 1$ und $\dim(U) = 2$ werden uns folgende vier Aussagen helfen, um alle A-invarianten Unterräume zu bestimmen:

1. Das charakteristische Polynom von $A|_U$ teilt das charakteristische Polynom von A, wenn U A-invariant ist.
2. Zerfällt das charakteristische Polynom einer Matrix in lauter verschiedene Linearfaktoren, dann ist die Matrix diagonalisierbar und der zugehörige Vektorraum lässt sich als direkte Summe der Eigenräume schreiben.
3. Eine Matrix ist genau dann trigonalisierbar, wenn das charakteristische Polynom komplett in Linearfaktoren zerfällt. Der Vektorraum lässt sich dann als direkte Summe der verallgemeinerten Eigenräume schreiben.
4. Die Eigenräume $V(\lambda, A)$ und die verallgemeinerten Eigenräume $V^\infty(\lambda, A)$ einer Matrix A sind A-invariant.

Gut wäre es, gerade wegen des letzten Punktes, wenn wir die A-invarianten U in Abhängigkeit der Eigenräume und die verallgemeinerten Eigenräume bestimmen könnten.

Z

Es ist zunächst also sinnvoll, das charakteristische Polynom von A, die Eigenräume und verallgemeinerten Eigenräume zu bestimmen[6]:

H

$$\begin{aligned}
\mathrm{charpol}_A(x) = \det(xE_3 - A) = \det \begin{pmatrix} x-2 & 0 & -1 \\ 1 & x-1 & -3 \\ -2 & 0 & x-3 \end{pmatrix} \\
&= (x-2)(x-1)(x-3) - 2(x-1) \\
&= (x-1)[(x-2)(x-3) - 2] \\
&= (x-1)[x^2 - 5x + 4] \\
&= (x-1)^2(x-4).
\end{aligned}$$

[6] Wir verwenden gleich folgende Notation: $V^j(\lambda, A) = \ker(A - \lambda \mathrm{id}_V)^j$ für die Verallgemeinerung der Eigenräume.

Damit hat A die beiden Eigenwerte 1 und 4.

$$V(1,A) = \ker \begin{pmatrix} 1 & 0 & 1 \\ -1 & 0 & 3 \\ 2 & 0 & 2 \end{pmatrix} = \ker \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \langle e_2 \rangle$$

$$\begin{aligned} V^\infty(1,A) = V^2(1,A) &= \ker \begin{pmatrix} 1 & 0 & 1 \\ -1 & 0 & 3 \\ 2 & 0 & 2 \end{pmatrix}^2 \\ &= \ker \begin{pmatrix} 3 & 0 & 3 \\ 5 & 0 & 5 \\ 6 & 0 & 6 \end{pmatrix} = \ker \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \langle e_2, e_1 - e_3 \rangle \end{aligned}$$

$$\begin{aligned} V^\infty(4,A) = V(4,A) &= \ker \begin{pmatrix} -2 & 0 & 1 \\ -1 & -3 & 3 \\ 2 & 0 & -1 \end{pmatrix} \\ &= \ker \begin{pmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & -\frac{5}{6} \\ 0 & 0 & 0 \end{pmatrix} = \langle 3e_1 + 5e_2 + 6e_3 \rangle \end{aligned}$$

$\dim(U) = 1$:

Z Dann hat $\operatorname{charpol}_{A|_U}$ Grad 1 und teilt $\operatorname{charpol}_A$. Dafür gibt es zwei Möglichkeiten:

H a) $\operatorname{charpol}_{A|_U}(x) = (x-1)$. Damit ist $A\big|_U$ diagonalisierbar und es gilt

$$U = V(1, A|_U).$$

Nun gilt aber

$$\begin{aligned} U = &V(1, A|_U) = \ker(A\big|_U - 1 \cdot \mathrm{id}_U) = \ker((A - 1 \cdot \mathrm{id}_V)\big|_U) \\ &\subseteq \ker(A - 1 \cdot \mathrm{id}_V) = V(1,A) \end{aligned}$$

und $\dim(U) = 1 = \dim(V(1,A))$. Daraus folgt schließlich die Gleichheit.

B Wir haben also gezeigt: Wenn U ein eindimensionaler A-invarianter Unterraum mit $\operatorname{charpol}_{A|_U}(x) = (x-1)$ ist, dann ist $U = V(1,A)$. Da $V(1,A)$ eindeutig bestimmt ist, ist es also der einzige A-invariante Unterraum zu $\operatorname{charpol}_{A|_U}(x) = (x-1)$.

H b) $\operatorname{charpol}_{A_U}(x) = (x-4)$. Auch hier gilt also $U = V(4, A\big|_U)$. Und mit den gleichen Überlegungen wie oben folgt $U \subseteq V(4,A)$. Da die Dimensionen übereinstimmen, folgt $U = V(4,A)$. Also ist $V(4,A)$ der einzige A-invariante Unterraum mit $\operatorname{charpol}_{A|_U}(x) = (x-4)$.

$\dim(U) = 2$:

Z Dann hat $\operatorname{charpol}_{A|_U}$ Grad 2 und teilt $\operatorname{charpol}_A$. Dafür gibt es wieder zwei Möglichkeiten:

a) $\operatorname{charpol}_{A|_U}(x) = (x-1)^2$. Dann ist $A\big|_U$ trigonalisierbar und es gilt

$$U = V^\infty(1, A|_U).$$

Nun gibt es für $V^\infty(1, A|_U)$ zwei Möglichkeiten, entweder

$$V^\infty(1, A|_U) = V(1, A|_U) \text{ oder } V^\infty(1, A|_U) = V^2(1, A|_U).$$

Wir müssen die beiden Möglichkeiten separat untersuchen: K

Angenommen $U = V^\infty(1, A|_U) = V(1, A|_U)$: Da wir oben gesehen haben, dass H

$$V(1, A|_U) \subseteq V(1, A)$$

ist, würde dann folgen

$$2 = \dim(U) = \dim(V(1, A|_U)) \leq \dim(V(1, A)) = 1.$$

Deshalb kann $V^\infty(1, A|_U) = V(1, A|_U)$ nicht gelten.

Für $U = V^2(1, A|_U)$ folgt aber wieder:

$$\begin{aligned} U =& V^2(1, A|_U) = \ker((A|_U - \mathrm{id}_U)^2) = \ker((A - \mathrm{id}_V)^2|_U) \\ &\subseteq \ker((A - \mathrm{id}_V)^2) = V^2(1, A) = V^\infty(1, A). \end{aligned}$$

Wegen Dimensionsgleichheit folgt die Gleichheit der Unterräume und damit ist $V^\infty(1, A)$ der einzige zweidimensionale A-invariante Unterraum mit $\operatorname{charpol}_{A|_U}(x) = (x-1)^2$.

b) $\operatorname{charpol}_{A|_U}(x) = (x-1)(x-4)$. Dann ist $A|_U$ schon diagonalisierbar und $U = V(1, A|_U) \oplus V(4, A|_U)$. A

Mit den Ergebnissen für U eindimensional, folgt: H

$$U = V(1, A) \oplus V(4, A).$$

Damit ist U aber wieder eindeutig bestimmt.

Da die Eigenräume und verallgemeinerten Eigenräume A-invariant sind, sind die oben berechneten Unterräume U tatsächlich auch A-invariant. Als Ergebnis lässt sich festhalten, dass B

$\{0_V\}$
V
$\langle e_2 \rangle$
$\langle 3e_1 + 5e_2 + 6e_3 \rangle$
$\langle e_2, e_1 - e_3 \rangle$
$\langle e_2 \rangle \oplus \langle 3e_1 + 5e_2 + 6e_3 \rangle = \langle e_2, 3e_1 + 5e_2 + 6e_3 \rangle$

alle A-invarianten Unterräume sind.

Verständnisfragen 9.5

1. Was ist $A|_U$ und warum gilt im Fall $\dim(U) = 1$, dass der Grad von $\operatorname{charpol}_{A|_U}$ gleich 1 ist?
2. Warum gibt es für den Fall $\dim(U) = 1$ nur die beiden aufgeführten Möglichkeiten für das charakteristische Polynom von $A|_U$?
3. Warum kann im Fall $\operatorname{charpol}_{A|_U}(x) = (x-1)^2$ der verallgemeinerte Eigenraum von $A|_U$ zum Eigenwert 1 nur $V(1, A|_U)$ oder $V^2(1, A|_U)$ sein?

Arbeitsauftrag 9.5 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.5.

9.6 A-zyklisch, Primärkomponenten und rationale Jordannormalform

Aufgabe 9.6 Untersuchen Sie für die Matrix

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 0 & 1 & -2 \end{pmatrix} \in M_{6\times 6}(\mathbb{R}),$$

ob $V = \mathbb{R}^6$ A-zyklisch ist und geben Sie gegebenenfalls einen A-Erzeuger von V an! Bestimmen Sie ferner die Primärkomponenten von V bzgl. A und geben Sie die rationale Jordannormalform von A an!

K

Ausführliche Lösung 9.6 V ist genau dann A-zyklisch, wenn das charakteristische Polynom und das Minimalpolynom von A gleich sind. Deshalb bestimmen wir zunächst diese beiden Polynome.

Z

Schauen wir uns dazu die Matrix A genauer an. Sie zerfällt in zwei Blöcke, was uns die Berechnung des charakteristischen Polynoms und des Minimalpolynoms erleichtert.

$$A = \left(\begin{array}{ccc:ccc} 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 0 & 1 & -2 \end{array}\right)$$

Wir haben also die beiden Blöcke $A_1 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ und $A_2 = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & -2 \\ 0 & 1 & -2 \end{pmatrix}$.

Für das charakteristische Polynom und das Minimalpolynom von A gilt dann: **H**

$$\mathrm{charpol}_A = \mathrm{charpol}_{A_1} \cdot \mathrm{charpol}_{A_2} \quad \text{und} \quad \mathrm{minpol}_A = \mathrm{kgV}(\mathrm{minpol}_{A_1}, \mathrm{minpol}_{A_2}).$$

Wenn wir uns nun die beiden Matrizen A_1 und A_2 genauer anschauen, sehen wir, **Z**
dass sie die Form von Begleitmatrizen haben.

Und über diese Begleitmatrizen **K**

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix} \in M_{n\times n}(K)$$

gibt es einige Resultate, welche uns weiterhelfen können:

1. Der zugehörige Vektorraum ist zyklisch bzgl. der Matrix.
2. Charakteristisches Polynom und Minimalpolynom sind gleich.
3. $\mathrm{charpol}_A(x) = x^n + \sum_{i=0}^{n-1} a_i x^i$.

Das bedeutet für die Matrix A: V zerlegt sich schon wie folgt in zwei zyklische Unterräume W_1 und W_2 ($V = W_1 \oplus W_2$): Ist $B := \{e_1, e_2, e_3, e_4, e_5, e_6\}$ die kanonische Basis, dann ist **A**

$$W_1 = \langle e_1, e_2, e_3 \rangle \text{ und } W_2 = \langle e_4, e_5, e_6 \rangle$$

mit

$$A_1 := A|_{W_1} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \text{ und } A_2 := A|_{W_2} = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & -2 \\ 0 & 1 & -2 \end{pmatrix}.$$

Da zu W_1, W_2 Begleitmatrizen gehören, sind sie A_1- bzw. A_2-zyklisch. Damit ist

$$\mathrm{charpol}_{A_i} = \mathrm{minpol}_{A_i}$$

und wir können $\mathrm{charpol}_{A_1}, \mathrm{minpol}_{A_1}$ bzw. $\mathrm{charpol}_{A_2}, \mathrm{minpol}_{A_2}$ direkt aus den Matrizen ablesen:

$$\mathrm{charpol}_{A_1}(x) = \mathrm{minpol}_{A_1}(x) = x^3 - 1 \quad \text{und}$$
$$\mathrm{charpol}_{A_2}(x) = \mathrm{minpol}_{A_2}(x) = x^3 + 2x^2 + 2x + 1.$$ **H**

Jetzt zerlegen wir diese Polynome noch in irreduzible Faktoren (1 ist Nullstelle von $\mathrm{charpol}_{A_1}$ und -1 ist Nullstelle von $\mathrm{charpol}_{A_2}$!): **K**

H

$$\operatorname{charpol}_{A_1}(x) = \operatorname{minpol}_{A_1}(x) = (x-1)(x^2+x+1) \quad \text{und}$$
$$\operatorname{charpol}_{A_2}(x) = \operatorname{minpol}_{A_2}(x) = (x+1)(x^2+x+1).$$

Man überprüft leicht, dass x^2+x+1 in $\mathbb{R}$ keine Nullstelle hat und somit irreduzibel ist. Das charakteristische Polynom von A ist dann

$$\operatorname{charpol}_A(x) = \operatorname{charpol}_{A_1}(x) \cdot \operatorname{charpol}_{A_2}(x) = (x-1)(x+1)(x^2+x+1)^2$$

und das Minimalpolynom von A ist

$$\operatorname{minpol}_A(x) = \operatorname{kgV}(\operatorname{minpol}_{A_1}, \operatorname{minpol}_{A_2})(x) = (x-1)(x+1)(x^2+x+1).$$

A

Dann gilt aber $\operatorname{charpol}_A \neq \operatorname{minpol}_A$ und somit ist V nicht A-zyklisch.

P

Wie man sieht, kann man sich die Berechnung von $\operatorname{charpol}_A$ und minpol_A vereinfachen, indem man schaut, ob die Matrix nicht schon in Blöcke zerfällt. Dann braucht man nämlich nur noch von den kleineren Blöcken die charakteristischen Polynome und Minimalpolynome zu berechnen und bekommt dadurch schon die der ganzen Matrix. Noch besser ist es dann, wie in diesem Fall, wenn die Blöcke schon Begleitmatrizen sind. In einem solchen Fall kann man ja das charakteristische Polynom, welches dann gleich dem Minimalpolynom ist, direkt aus der Matrix ablesen.

B

Als Nächstes sollen wir noch die rationale Jordannormalform der Matrix bestimmen.

K

Dazu müssen wir als Erstes die Primärkomponenten berechnen und diese dann in zyklische Unterräume zerlegen.

Z

Wir haben $\operatorname{charpol}_A$ bereits in irreduzible, normierte Polynome zerlegt. Setze für $i = 1,2,3$: $g_i : \mathbb{R} \to \mathbb{R}$ mit

$$g_1(x) := x-1;\ g_2(x) := x+1 \text{ und } g_3(x) := x^2+x+1.$$

Dann ist

$$\operatorname{charpol}_A = g_1 \cdot g_2 \cdot g_3^2 \qquad \operatorname{minpol}_A = g_1 \cdot g_2 \cdot g_3.$$

A

Da die irreduziblen Faktoren jeweils mit Potenz 1 im Minimalpolynom auftauchen, ergeben sich die Primärkomponenten $V_i = \ker(g_i(A))$:

H

$$V_1 = \ker(g_1(A)) = \ker \begin{pmatrix} -1 & 0 & 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 1 & -3 \end{pmatrix}$$

$$= \ker \begin{pmatrix} 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\rangle = \langle e_1 + e_2 + e_3 \rangle,$$

$$V_2 = \ker(g_2(A)) = \ker \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}$$

$$= \ker \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \langle \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \rangle = \langle e_4 + e_5 + e_6 \rangle,$$

$$V_3 = \ker(g_3(A)) = \ker \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 & -1 & 1 \end{pmatrix}$$

$$= \ker \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$= \langle \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} \rangle = \langle e_1 - e_2, e_1 - e_3, e_4 + e_5, e_4 - e_6 \rangle.$$

Über das charakteristische Polynom und das Minimalpolynom der Primärkomponenten wissen wir Folgendes: A

$\mathrm{charpol}_{A|_{V_1}} = g_1$ und $\mathrm{minpol}_{A|_{V_1}} = g_1$, also ist V_1 schon A-zyklisch.

$\mathrm{charpol}_{A|_{V_2}} = g_2$ und $\mathrm{minpol}_{A|_{V_2}} = g_2$, also ist V_2 schon A-zyklisch.

$\mathrm{charpol}_{A|_{V_3}} = g_3^2$ und $\mathrm{minpol}_{A|_{V_3}} = g_3$, also ist V_3 nicht A-zyklisch.

Das heißt aber, dass wir nur mehr V_3 in zyklische Unterräume $V_{3,i}$ zu zerlegen brauchen. Das machen wir, indem wir ein Element aus $V_3 \backslash V_3^{(0)} = V_3 \backslash \{0\}$ nehmen (beachte: $V_3 = V_3^{(1)}$) und dazu den zyklischen Unterraum bilden[7]. K

[7] $V_i^{(k)} := \ker(g_i(A)^k)$

H

Wähle z. B. $w_1 := e_1 - e_2$, dann gilt:

$$A(e_1 - e_2) = e_2 - e_3 \qquad \text{und}$$
$$A^2(e_1 - e_2) = A(e_2 - e_3) = e_3 - e_1 = -1 \cdot (e_1 - e_2) - 1 \cdot (e_2 - e_3).$$

Also lässt sich $A^2(w_1)$ als Linearkombination von w_1 und $A(w_1)$ schreiben. Somit ist:[8]

$$V_{3,1} := U(w_1, A) = \langle e_1 - e_2, e_2 - e_3 \rangle = \langle e_1 - e_2, e_1 - e_3 \rangle \quad \text{und}$$
$$V_3 = U(w_1, A) \oplus \langle e_4 + e_5, e_4 - e_6 \rangle.$$

K

Es bleibt nun zu überprüfen, ob $\langle e_4 + e_5, e_4 - e_6 \rangle$ schon A-zyklisch ist.

H

Dazu betrachten wir $U(w_2, A)$ für $w_2 := e_4 + e_5$:

$$A(w_2) = e_5 + e_6$$
$$A^2(w_2) = A(e_5 + e_6) = -e_4 - 2e_5 - e_6 = -(e_4 + e_5) - (e_5 + e_6).$$

Also lässt sich $A^2(w_2)$ als Linearkombination von w_2 und $A(w_2)$ darstellen. Somit ist

$$V_{3,2} := U(w_2, A) = \langle e_4 + e_5, e_5 + e_6 \rangle = \langle e_4 + e_5, e_4 - e_6 \rangle.$$

B

Wir haben also die Zerlegung von V in zyklische Unterräume:

$$V = V_1 \oplus V_2 \oplus V_{3,1} \oplus V_{3,2}$$

bestimmt. Aus diesen Informationen können wir nun die rationale Jordannormalform von A ableiten.

K

Sie ist aus den einzelnen Blöcken zusammengesetzt, die zu den zyklischen Unterräumen gehören. Und die einzelnen Blöcke sehen wie folgt aus: Ist V A-zyklisch mit $\text{charpol}_A = g^e$ (g normiert und irreduzibel; $\text{grad}(g) = d$), dann gibt es eine Basis B bzgl. derer A die Form hat:

$$M_B^B(A) = \begin{pmatrix} P & & & \\ N & P & & \\ & \ddots & \ddots & \\ & & N & P \end{pmatrix} \in M_{de \times de}(K),$$

wobei $P \in M_{d \times d}(K)$ die Begleitmatrix zu g und $N = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix} \in K_{(d,d)}$ ist.

A

Das müssen wir nun auf $V_1, V_2, V_{3,1}$ und $V_{3,2}$ übertragen:

Für V_1 gilt: $\text{charpol}_{A|_{V_1}}(x) = g_1(x) = (x - 1)$. Also $e = 1$ und $d = 1$. Da e die Anzahl der Blöcke P angibt, haben wir hier nur einen Block (dieser hat die Länge d). Also gehört zu V_1 die Begleitmatrix zu g_1: $P_1 = (1)$.

Analog erhalten wir zu V_2 die Begleitmatrix zu $g_2(x) = x + 1$: $P_2 = (-1)$.

[8] Mit $U(w, A)$ wird der von w erzeugte A-zyklische Unterraum bezeichnet.

Für $V_{3,i}$ $(i = 1,2)$ ist $\operatorname{charpol}_{A|_{V_{3,i}}}(x) = g_3(x) = x^2 + x + 1$, also ist auch hier $e = 1$, aber $d = 2$. Also gehört zu $V_{3,i}$ die Begleitmatrix zu g_3: $P_3 = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$.

Insgesamt erhält man dann die rationale Jordannormalform von A [Beachten Sie: An den Positionen, an denen in der Matrix nichts steht, ist der Eintrag $= 0$]:

$$\operatorname{rJNF}(A) = \begin{pmatrix} P_1 & & & \\ & P_2 & & \\ & & P_3 & \\ & & & P_3 \end{pmatrix} = \begin{pmatrix} 1 & & & & & \\ & -1 & & & & \\ & & 0 & -1 & & \\ & & 1 & -1 & & \\ & & & & 0 & -1 \\ & & & & 1 & -1 \end{pmatrix}.$$

Verständnisfragen 9.6

1. In der Lösung wird behauptet, dass die Primärkomponenten $V_i = \ker(g_i(A))$ sind. Wie leitet sich dies aus dem allgemeinen Fall ab?
2. Beim Berechnen der Primärkomponenten mussten die Kerne von Matrizen bestimmt werden. Dabei wurde jeweils nur ein Umformungsschritt angegeben. Überlegen Sie, welche Zwischenschritte dort eigentlich noch stehen müssten!
3. Warum ist $\operatorname{kgV}(\operatorname{minpol}_{A_1}, \operatorname{minpol}_{A_2}) = g_1 g_2 g_3$?
4. Warum tauchen in der rationalen Jordannormalform keine Matrizen N mehr auf?

Arbeitsauftrag 9.6 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.5.

9.7 Vektorraum der selbstadjungierten Abbildungen

Aufgabe 9.7 Sei V ein n-dimensionaler K-Vektorraum mit Skalarprodukt über $K \in \{\mathbb{R}, \mathbb{C}\}$ und sei

$$\mathcal{H} := \{f \in \operatorname{Hom}_K(V,V) \mid f = f^{\mathrm{ad}}\}.$$

Zeigen Sie:

$$\dim(\mathcal{H}) = \begin{cases} n^2 & \text{für } K = \mathbb{C} \\ \frac{n^2+n}{2} & \text{für } K = \mathbb{R} \end{cases}.$$

Mit f^{ad} ist hier die zu f adjungierte Abbildung bezeichnet.

P **Ausführliche Lösung 9.7** In dieser Aufgabe ist also die Dimension des Vektorraumes $\mathcal{H}$ der *selbstadjungierten* Abbildungen zu bestimmen. Da es sich mit Abbildungen schwer arbeiten lässt (wir müssen schließlich eine Basis zu $\mathcal{H}$ finden), gehen wir zu Matrizen über. Das ist ein „Trick“, den man in der Linearen Algebra häufig anwendet.

Z Sei B eine Basis von V, dann gehört zu jedem Endomorphismus $f \in \operatorname{Hom}_K(V, V)$ eine eindeutig bestimmte (Darstellungs-)Matrix bzgl. dieser Basis: $M_B^B(f)$. Wir werden diese zugehörige Matrix mit A bezeichnen.

Die Darstellungsmatrizen sind aus dem Vektorraum der $(n \times n)$-Matrizen $M_{n\times n}(K)$. Zu diesem Vektorraum bilden die Elementarmatrizen E_{jk} $(1 \leq j, k \leq n)$ eine Basis.[9]

K Die Idee ist nun, aus dieser Basis eine Basis von $\mathcal{H}$ herzuleiten.

B Überlegen wir zunächst, was wir über selbstadjungierte Abbildungen wissen.

A Gegeben haben wir bereits, dass $f = f^{\mathrm{ad}}$ gelten muss (das ist ja gerade die Definition von selbstadjungiert). Für eine *hermitesche* Matrix gilt analog dazu $A = \bar{A}^T$. Denn betrachtet man die Darstellungsmatrizen der Abbildungen und der zu ihr adjungierten Abbildung bzgl. einer *Orthonormalbasis* B von V, so gilt die Beziehung:

$$M_B^B(f^{\mathrm{ad}}) = \overline{M_B^B(f)}^T.$$

Genauso gilt für hermitesche Matrizen, dass die zugehörige lineare Abbildung selbstadjungiert ist.

B Was wir uns bisher überlegt haben, ist, dass

$$\dim(\mathcal{H}) = \dim(\{A \in M_{n\times n}(K) \mid A = \bar{A}^T\})$$

ist und dass wir uns eine Basis des Vektorraumes der hermiteschen Matrizen (wir wollen ihn ab jetzt mit $\mathcal{H}'$ bezeichnen) mit Hilfe der Elementarmatrizen „basteln“ können.

Beginnen wir nun mit dem Fall $K = \mathbb{R}$:

H Dann wissen wir, dass das konjugiert Komplexe einer Zahl wieder die Zahl selbst ist, da bei reellen Zahlen kein Imaginärteil auftritt. Also gilt hier für eine hermitesche Matrix A:

$$A = \bar{A}^T = A^T.$$

Z Das heißt doch aber nichts anderes, als dass für die Matrizeneinträge $a_{jk} = a_{kj}$ für alle $k, j = 1, \ldots, n$ gilt.

H Und damit können wir die Matrix wie folgt darstellen:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix} = \sum_{l=1}^{n} a_{ll} E_{ll} + \sum_{j=1}^{n} \sum_{k=j+1}^{n} a_{jk}(E_{jk} + E_{kj}).$$

[9] Die Elementarmatrix E_{jk} ist die Matrix, bei der nur an der Position j, k (also j-te Zeile und k-te Spalte) eine Eins steht und sonst überall Nullen.

Daraus erkennen wir, dass $\{E_{ll}, E_{jk} + E_{kj} \mid 1 \leq l \leq n; 1 \leq j < k \leq n\}$ eine Basis von $\mathcal{H}'$ bildet.

Jetzt müssen wir nur noch die Anzahl der Elemente bestimmen und haben damit die Dimension von $\mathcal{H}'$ und somit natürlich auch von $\mathcal{H}$.

Von der Form E_{ll} gibt es n Matrizen und von $E_{jk} + E_{kj}$ gibt es $\frac{(n-1)\cdot n}{2}$ Matrizen, letzteres kann man sich auf zwei Arten überlegen.

1. Da $k > j$ gelten soll, haben wir

$$\begin{array}{lll} \text{für } j = 1 & n-1 & \text{Möglichkeiten für } k, \\ \text{für } j = 2 & n-2 & \text{Möglichkeiten für } k, \\ \quad\vdots & \quad\vdots & \\ \text{für } j = n-1 & 1 & \text{Möglichkeit für } k. \end{array}$$

   Das macht zusammen $(n-1) + (n-2) + \ldots + 1 = \sum_{i=1}^{n-1} i = \frac{(n-1)\cdot n}{2}$ Möglichkeiten.

2. Man kann sich das auch mit Hilfe von Kombinatorik überlegen.

   j und k geben uns die Position der Eins in der Elementarmatrix an. Beide können Werte zwischen 1 und n annehmen. Das heißt, wir ziehen zwei Kugeln aus einer Urne mit n Kugeln (die mit den Zahlen $1, \ldots, n$ beschriftet sind).

   Da $j \neq k$ ist, legen wir die Kugeln nicht zurück und da $E_{jk} + E_{kj} = E_{kj} + E_{jk}$ ist, ist auch die Reihenfolge der Ziehung egal. Damit können wir dafür die Anzahl der Möglichkeiten berechnen. Denn die Kombinatorik sagt, dass es für das Ziehen ohne Zurücklegen ohne Beachtung der Reihenfolge von 2 Kugeln aus n Kugeln gerade

$$\binom{n}{2} = \frac{n \cdot (n-1)}{2}$$

   Möglichkeiten gibt.

Insgesamt haben wir also $n + \frac{n\cdot(n-1)}{2} = \frac{2n+n^2-n}{2} = \frac{n^2+n}{2}$ Basiselemente, was ja zu zeigen war.

Kommen wir nun zu dem Fall $K = \mathbb{C}$:

Hier ist es sinnvoll, die Einträge der Matrix A in Real- und Imaginärteil zu zerlegen, also $a_{jk} = x_{jk} + iy_{jk}$ zu schreiben.

Dann erhalten wir aus der Bedingung $A = \bar{A}^T$:

$$a_{kj} = \bar{a}_{jk} = x_{jk} - iy_{jk}$$

für alle $k, j = 1, \ldots, n$. Insbesondere gilt dann $a_{jj} = \bar{a}_{jj}$, also sind die a_{jj} aus $\mathbb{R}$. Daraus erhalten wir:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ \bar{a}_{12} & a_{22} & \cdots & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ \bar{a}_{1n} & \cdots & \cdots & \bar{a}_{(n-1)n} & a_{nn} \end{pmatrix}$$

$$= \sum_{l=1}^{n} a_{ll} E_{ll} + \sum_{j=1}^{n} \sum_{k=j+1}^{n} x_{jk}(E_{jk} + E_{kj}) + \sum_{j=1}^{n} \sum_{k=j+1}^{n} y_{jk} \cdot i(E_{jk} - E_{kj}).$$

Also erzeugt die Menge $\{E_{ll}, E_{jk} + E_{kj}, i(E_{jk} - E_{kj}) \mid 1 \leq l \leq n; 1 \leq j < k \leq n\}$ schon ganz $\mathcal{H}'$. Außerdem sind diese Elemente auch linear unabhängig. Damit haben wir also auch in diesem Fall eine Basis gefunden.

B Und jetzt brauchen wir wieder nur zu zählen, die Hauptarbeit haben wir dafür schon im Fall $K = \mathbb{R}$ geleistet.

H Wir haben n Elemente der Form E_{ll}, $\frac{(n-1)\cdot n}{2}$ Elemente der Form $E_{jk} + E_{kj}$ und $\frac{(n-1)\cdot n}{2}$ Elemente der Form $i(E_{jk} - E_{kj})$. Also haben wir insgesamt $n + 2\frac{(n-1)\cdot n}{2} = n^2$ Basiselemente und damit hat $\mathcal{H}$ in diesem Fall die Dimension n^2.

Verständnisfragen 9.7

1. Zur Berechnung der Dimension von $\mathcal{H}$ im Fall $K = \mathbb{R}$ haben wir benutzt, dass $(n-1) + (n-2) + \ldots + 1 = \sum_{i=1}^{n-1} i = \dfrac{(n-1)\cdot n}{2}$ ist. Warum gilt diese Gleichheit? Welche allgemeine Formel steckt dahinter?

2. Warum folgt aus

$$A = \sum_{l=1}^{n} a_{ll} E_{ll} + \sum_{j=1}^{n} \sum_{k=j+1}^{n} a_{jk}(E_{jk} + E_{kj})$$

   schon, dass $\{E_{ll}, E_{jk} + E_{kj} \mid 1 \leq l \leq n; 1 \leq j < k \leq n\}$ eine Basis von $\mathcal{H}'$ ist?

3. In der Lösung wurde die Voraussetzung, dass V ein Vektorraum *mit Skalarprodukt* ist, nicht direkt benutzt. Warum kann man trotzdem nicht auf diese Voraussetzung verzichten?

Arbeitsauftrag 9.7 Verfassen Sie eine komprimierte Musterlösung zu der Aufgabe!

Lösungsvorschläge zu den Verständnisfragen und dem Arbeitsauftrag finden Sie in Abschnitt 11.5.

Teil III

Übungsteil

10 Verfassen ausführlicher Musterlösungen

Komprimierte Lösungen als Ausgangspunkt. In den Kapiteln 5 bis 9 sind viele ausführliche Musterlösungen dargestellt worden, die zum gründlichen Durcharbeiten gedacht sind. In diesem Kapitel sollen Sie nun selbst ausführliche Musterlösungen schreiben. Sie sollen sich dabei vorerst ausschließlich auf das Strukturieren, das treffende Formulieren und das Abwägen, welche Informationen in einer solchen ausführlichen Musterlösung vorkommen sollen, konzentrieren können. Zu diesem Zweck stellen wir Ihnen in den Abschnitten 10.1 bis 10.5 eine Reihe komprimierter Musterlösungen zur Verfügung. Diese sollen Sie zu ausführlichen Musterlösungen ausarbeiten. Sie werden dadurch noch vom kognitiv belastenden Aufgabenlösen an sich befreit, müssen also lediglich die vorliegende knapp gehaltene Lösung verstehen und sie detaillierter ausformulieren. Für den Beginn ist das aber sicher anspruchsvoll genug! In Ihrem Studium haben Sie ja ohnehin Gelegenheit, auch selbstständig Aufgaben zu lösen.

Wichtige Merkmale. Achten Sie bei Ihren ausführlichen Musterlösungen darauf, alle Argumentationsschritte detailliert aufzuführen, sodass sie für einen Leser gut nachvollziehbar wären! Strukturieren Sie Ihre Lösung, schreiben Sie Zwischenüberschriften und kennzeichnen Sie Nebenrechnungen als solche! Formulieren Sie immer präzise, was die Voraussetzungen sind, auf die Sie sich beziehen können, und was Sie in der vorliegenden Aufgabe noch zu zeigen haben! Kommentieren Sie einzelne Bearbeitungsschritte durch beschreibende und erklärende Texte! Achten Sie darauf, nicht ausufernd, sondern treffend zu formulieren! Fertigen Sie gegebenenfalls eine Skizze an, die es erlaubt, den vorliegenden Sachverhalt schnell zu erfassen!

Reihenfolge umstellen. In den komprimierten Musterlösungen kommt es manchmal vor, dass z. B. gleich zu Beginn eine Tätigkeit ausgeführt wird, auf die man nicht ohne Weiteres sofort gekommen wäre. Denken Sie etwa an die Fallunterscheidung in der komprimierten Lösung zu Aufgabe 1.1. Im eigentlichen Aufgabenlösungsprozess entwickelt sich die Idee erst viel später, gerade diese Fallunterscheidung zu machen.

Es ist aber ein Charakteristikum mathematischer Expertise, Aufgabenlösungen so prägnant und elegant wie möglich aufzuschreiben. Dabei interessiert dann nicht mehr die Entstehungsgeschichte einer Lösung, sondern nur mehr die klaren Fakten und die stringente Abfolge. Insofern müssen Sie bei der Ausarbeitung komprimierter Lösungen zu ausführlichen Lösungen darauf achten, diesen Prozess wieder umzukehren. Manche Bearbeitungsschritte müssen also in Ihrer Reihenfolge umgestellt werden, so wie es dem Aufgabenlösungsprozess entspricht. Dabei können Sie viel über das Lesen und Schreiben mathematischer Texte und über die Beziehung zwischen dem Lösungsprozess und seinem Endprodukt lernen.

C. Ableitinger, A. Herrmann, *Lernen aus Musterlösungen zur Analysis und Linearen Algebra*,
DOI 10.1007/978-3-658-02353-9_10, © Springer Fachmedien Wiesbaden 2013

Nichteindeutigkeit der Lösung. Sie können Ihre fertige ausführliche Musterlösung mit dem jeweiligen Lösungsvorschlag in Kapitel 12 vergleichen. Bedenken Sie dabei, dass es aber prinzipiell so viele Möglichkeiten gibt eine Lösung aufzuschreiben, wie es Aufgabenlöser gibt. Wundern Sie sich also nicht, wenn sich Ihre Lösung von der vorgeschlagenen in der Gestalt und teilweise auch in den Inhalten unterscheidet. Die Lösungsvorschläge dienen eher dazu, Ihnen weiterzuhelfen, wenn Sie eine Passage in der komprimierten Musterlösung nicht verstehen.

An die Teilprozesse denken. Die Auflistung der Teilprozesse aus Kapitel 3 kann Ihnen beim Verfassen der Musterlösung als Stütze und als Lösungsplan dienen. Machen Sie sich in jedem Schritt bewusst, welche Funktion und Bedeutung er im Aufgabenlösungsprozess einnimmt. Das erleichtert Ihnen das gut strukturierte Arbeiten. Erwähnen Sie in Ihrer Musterlösung ruhig explizit, wenn Sie einen Trick anwenden oder wenn ein Bearbeitungsschritt die Aufgabe öffnet, usw. Das ermöglicht es dem Leser, seinen Fokus auf die wichtigen Punkte in der Musterlösung zu richten.

10.1 Themen aus den mathematischen Grundlagen

10.1.1 Lösen von Ungleichungen

Übung 10.1 *Betrachten Sie die folgende Aufgabe:*

> Bestimmen Sie die Lösungsmengen der folgenden Gleichungen bzw. Ungleichungen in $\mathbb{R}$.
>
> a) $\sqrt{x+1} = x$
>
> b) $x + 1 > x^2$
>
> c) $|8x - 5| \leq |7x + 15|$
>
> d) $||x + 1| - 2| = 1$

Schreiben Sie zur folgenden komprimierten Musterlösung eine ausführliche Musterlösung! (Einen Lösungsvorschlag finden Sie in Abschnitt 12.1.1.)

Komprimierte Lösung 10.1 a) Kandidaten für Lösungen sind die Lösungen der quadratischen Gleichung $x^2 - x - 1 = 0$:

$$x_1 = \frac{1+\sqrt{5}}{2}, \quad x_2 = \frac{1-\sqrt{5}}{2}.$$

Nur x_1 ist Lösung von $\sqrt{x+1} = x$, da $x_2 = \dfrac{1-\sqrt{5}}{2} < 0$ gilt. Daher: $\mathbb{L} = \{\frac{1+\sqrt{5}}{2}\}$.

b) $x + 1 > x^2 \Leftrightarrow x^2 - x - 1 < 0$. Der Graph der Funktion $x \mapsto x^2 - x - 1$ ist eine nach oben geöffnete Parabel mit Nullstellen wie in a). Daher gilt $\mathbb{L} = (\frac{1-\sqrt{5}}{2}, \frac{1+\sqrt{5}}{2})$.

c) $8x-5=0 \Leftrightarrow x=\frac{5}{8}$ sowie $7x+15=0 \Leftrightarrow x=-\frac{15}{7}$.

1. Fall: $x \geq \frac{5}{8}$. Dann gilt:

$$\begin{aligned} & |8x-5| \leq |7x+15| \\ \Leftrightarrow \quad & 8x-5 \leq 7x+15 \\ \Leftrightarrow \quad & x \leq 20. \end{aligned}$$

Damit:

$$\mathbb{L}_1 = [\frac{5}{8}, 20].$$

2. Fall: $-\frac{15}{7} \leq x < \frac{5}{8}$. Dann gilt:

$$\begin{aligned} & |8x-5| \leq |7x+15| \\ \Leftrightarrow \quad & -(8x-5) \leq 7x+15 \\ \Leftrightarrow \quad & x \geq -\tfrac{2}{3}. \end{aligned}$$

Damit:

$$\mathbb{L}_2 = [-\frac{2}{3}, \frac{5}{8}).$$

3. Fall: $x < -\frac{15}{7}$. Dann gilt:

$$\begin{aligned} & |8x-5| \leq |7x+15| \\ \Leftrightarrow \quad & -(8x-5) \leq -(7x+15) \\ \Leftrightarrow \quad & x \geq 20. \end{aligned}$$

Damit:

$$\mathbb{L}_3 = \emptyset.$$

Insgesamt gilt: $\mathbb{L} = [-\frac{2}{3}, 20]$.

d) 1. Fall: $|x+1| = 1$. Dann gilt entweder $x+1=1$, also $x=0$. Oder es gilt $-(x+1) = 1$, also $x=-2$.

2. Fall: $|x+1| = 3$. Dann gilt entweder $x+1=3$, also $x=2$. Oder es gilt $-(x+1)=3$, also $x=-4$.

Insgesamt erhalten wir: $\mathbb{L} = \{-4, -2, 0, 2\}$.

10.1.2 Vollständige Induktion

Übung 10.2 *Betrachten Sie die folgende Aufgabe:*

> Beweisen Sie, dass die Gleichung
>
> $$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$
>
> für alle $n \in \mathbb{N}$ gilt.

Schreiben Sie zur folgenden komprimierten Musterlösung eine ausführliche Musterlösung! (Einen Lösungsvorschlag finden Sie in Abschnitt 12.1.2.)

Komprimierte Lösung 10.2 Beweis durch Induktion nach n:
Induktionsanfang für $n = 1$:

$$\sum_{k=0}^{1} \binom{1}{k} = \binom{1}{0} + \binom{1}{1} = 1 + 1 = 2^1$$

Damit ist die Behauptung für $n = 1$ richtig.

Induktionsvoraussetzung: Die Gleichung gilt für ein beliebiges, aber festes $n \in \mathbb{N}$.

Induktionsschritt von n nach $n + 1$:

$$\begin{aligned}
\sum_{k=0}^{n+1} \binom{n+1}{k} &= 1 + \sum_{k=1}^{n} \binom{n+1}{k} + 1 \\
&= 1 + \sum_{k=1}^{n} \left(\binom{n}{k} + \binom{n}{k-1} \right) + 1 = 1 + \sum_{k=1}^{n} \binom{n}{k} + \sum_{k=1}^{n} \binom{n}{k-1} + 1 \\
&= \sum_{k=0}^{n} \binom{n}{k} + \sum_{k=0}^{n} \binom{n}{k} = 2 \cdot \sum_{k=0}^{n} \binom{n}{k} \overset{\mathrm{IV}}{=} 2 \cdot 2^n = 2^{n+1}
\end{aligned}$$

Also gilt die Gleichheit auch für $n + 1$.

10.1.3 Mengenverknüpfungen

Übung 10.3 *Betrachten Sie die folgende Aufgabe:*

> Seien A, B und C bzw. A_i ($i \in \mathbb{N}$) beliebige Mengen. Beweisen Sie die folgenden Aussagen:
> a) $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$
> b) $\bigcap_{i \in \mathbb{N}} (A_i \cup C) = \left(\bigcap_{i \in \mathbb{N}} A_i \right) \cup C$

Schreiben Sie zur folgenden komprimierten Musterlösung eine ausführliche Musterlösung! (Einen Lösungsvorschlag finden Sie in Abschnitt 12.1.3.)

Komprimierte Lösung 10.3 a) Wir beweisen beide Richtungen gleichzeitig:

$$\begin{aligned}
& x \in A \setminus (B \cup C) \\
\Leftrightarrow\ & x \in A \wedge x \notin (B \cup C) \\
\Leftrightarrow\ & x \in A \wedge x \notin B \wedge x \notin C \\
\Leftrightarrow\ & (x \in A \wedge x \notin B) \wedge (x \in A \wedge x \notin C) \\
\Leftrightarrow\ & x \in A \setminus B \wedge x \in A \setminus C \\
\Leftrightarrow\ & x \in (A \setminus B) \cap (A \setminus C).
\end{aligned}$$

b) Wir beweisen wieder beide Richtungen gleichzeitig:

$$\begin{aligned}
& x \in \bigcap_{i \in \mathbb{N}} (A_i \cup C) \\
\Leftrightarrow \quad & \forall i \in \mathbb{N} : x \in A_i \cup C \\
\Leftrightarrow \quad & \forall i \in \mathbb{N} : x \in A_i \vee x \in C \\
\Leftrightarrow \quad & (\forall i \in \mathbb{N} : x \in A_i) \vee x \in C \\
\Leftrightarrow \quad & x \in \bigcap_{i \in \mathbb{N}} A_i \vee x \in C \\
\Leftrightarrow \quad & x \in \bigcap_{i \in \mathbb{N}} A_i \cup C.
\end{aligned}$$

10.1.4 Körperaxiome

Übung 10.4 *Betrachten Sie die folgende Aufgabe:*

> Sei $\text{Abb}(\mathbb{R}, \mathbb{R}) := \{f : \mathbb{R} \to \mathbb{R} | f \text{ ist Abbildung}\}$ die Menge aller reellwertigen Abbildungen. Darauf seien die beiden Verknüpfungen
>
> - $(f + g)(x) = f(x) + g(x)$ für alle $x \in \mathbb{R}$ und
> - $(f \cdot g)(x) := f(x) \cdot g(x)$ für alle $x \in \mathbb{R}$
>
> definiert. Bestimmen Sie das neutrale Element n bzgl. der Addition und das neutrale Element e bzgl. der Multiplikation. Handelt es sich bei der Menge mit den beiden Verknüpfungen um einen Körper? Begründen Sie Ihre Antwort.

Schreiben Sie zur folgenden komprimierten Musterlösung eine ausführliche Musterlösung! (Einen Lösungsvorschlag finden Sie in Abschnitt 12.1.4.)

Komprimierte Lösung 10.4 Setze $n : \mathbb{R} \to \mathbb{R}$ mit $n(x) = 0$ für alle $x \in \mathbb{R}$. Dann gilt für alle $f \in \text{Abb}(\mathbb{R}, \mathbb{R})$ und alle $x \in \mathbb{R}$:

$$(f + n)(x) = f(x) + n(x) = f(x) + 0 = f(x).$$

Also gilt $f + n = f$ für alle $f \in \text{Abb}(\mathbb{R}, \mathbb{R})$. Daraus folgt, dass n das neutrale Element bzgl. der Addition ist.

Setze $e : \mathbb{R} \to \mathbb{R}$ mit $e(x) = 1$ für alle $x \in \mathbb{R}$. Dann gilt für alle $f \in \text{Abb}(\mathbb{R}, \mathbb{R})$ und alle $x \in \mathbb{R}$:

$$(f \cdot e)(x) = f(x) \cdot e(x) = f(x) \cdot 1 = f(x).$$

Also gilt $f \cdot e = f$ für alle $f \in \text{Abb}(\mathbb{R}, \mathbb{R})$. Daraus folgt, dass e das neutrale Element bzgl. der Multiplikation ist.

Es handelt sich bei $\text{Abb}(\mathbb{R}, \mathbb{R})$ *nicht* um einen Körper, da das multiplikative Inverse nicht für alle Elemente aus $\text{Abb}(\mathbb{R}, \mathbb{R}) \setminus \{n\}$ existiert. Betrachte beispielsweise die Abbildung $f \in \text{Abb}(\mathbb{R}, \mathbb{R})$ mit $f(x) = x$.

Annahme: Es gibt ein $g \in \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ mit $f \cdot g = e$. Dann muss für alle $x \in \mathbb{R}$

$$1 = e(x) = (f \cdot g)(x) = f(x) \cdot g(x) = x \cdot g(x)$$

gelten. Insbesondere folgt dann: $1 = 0 \cdot g(0) = 0$. Dies ist aber ein Widerspruch.

10.2 Themen aus der Analysis 1

10.2.1 Stetigkeit mit Epsilon und Delta

Übung 10.5 *Betrachten Sie die folgende Aufgabe:*

> Beweisen Sie mit Hilfe der ε-δ-Definition, dass die Funktion $f : (0, \infty) \to \mathbb{R}$ mit $f(x) = \frac{1}{x}$ stetig ist!

Schreiben Sie zur folgenden komprimierten Musterlösung eine ausführliche Musterlösung! (Einen Lösungsvorschlag finden Sie in Abschnitt 12.2.1.)

Komprimierte Lösung 10.5 Sei $x_0 > 0$ beliebig. Wir zeigen, dass f in x_0 stetig ist. Sei $\varepsilon > 0$ beliebig und wähle $\delta := \min\{\frac{1}{2}x_0, \frac{1}{2}\varepsilon x_0^2\}$.

Dann gilt für alle x mit $|x - x_0| < \delta$ die Ungleichung: $|x - x_0| < \frac{1}{2}x_0$ und folglich $x > \dfrac{1}{2}x_0$. Somit:

$$\left|\frac{1}{x} - \frac{1}{x_0}\right| = \frac{|x - x_0|}{x x_0} < \frac{\delta}{\frac{1}{2}x_0^2} \leq \frac{\frac{1}{2}\varepsilon x_0^2}{\frac{1}{2}x_0^2} = \varepsilon.$$

10.2.2 Zwischenwertsatz

Übung 10.6 *Betrachten Sie die folgende Aufgabe:*

> Beweisen Sie, dass es eine reelle Zahl x gibt, die die folgende Gleichung erfüllt:
>
> $$x^5 + \frac{4}{1 + |x| + x^2} = 0.$$

Schreiben Sie zur folgenden komprimierten Musterlösung eine ausführliche Musterlösung! (Einen Lösungsvorschlag finden Sie in Abschnitt 12.2.2.)

Komprimierte Lösung 10.6 Sei $f(x) := x^5 + \frac{4}{1+|x|+x^2}$. Dann ist f auf $\mathbb{R}$ stetig. Außerdem ist $f(0) = 4 > 0$ und $f(-2) = (-2)^5 + \frac{4}{7} < 0$. Aus dem Zwischenwertsatz folgt: $\exists\, x_0 \in (-2, 0)$ mit $f(x_0) = 0$, d. h. $x_0^5 + \frac{4}{1+|x_0|+x_0^2} = 0$.

10.2.3 Funktionengrenzwerte ohne de l'Hospital

Übung 10.7 *Betrachten Sie die folgende Aufgabe:*

Entscheiden Sie, ob die folgenden Grenzwerte existieren! Falls ja, bestimmen Sie sie ohne die Regeln von de l'Hospital!

a) $\lim\limits_{y \to +\infty} \dfrac{y}{e^y}$

b) $\lim\limits_{t \searrow 0} t \ln t$

c) $\lim\limits_{x \searrow 0} \dfrac{x^x - 1}{x \ln x}$

Schreiben Sie zur folgenden komprimierten Musterlösung eine ausführliche Musterlösung! (Einen Lösungsvorschlag finden Sie in Abschnitt 12.2.3.)

Komprimierte Lösung 10.7 a) Da $e^y = 1 + y + \frac{y^2}{2!} + \frac{y^3}{3!} + \ldots$ ist, erhalten wir für $y > 0$:

$$e^y \geq 1 + y + \frac{y^2}{2}.$$

Damit ist:

$$0 \leq \frac{y}{e^y} \leq \frac{y}{1 + y + \frac{y^2}{2}} \overset{y \to +\infty}{\longrightarrow} 0.$$

Aus dem Einschließungskriterium folgt, dass der gefragte Grenzwert existiert.

b) Sei $y := -\ln t$. Dann ist $t = e^{-y}$. Es gilt $t \searrow 0$ genau dann, wenn $y \to +\infty$. Mit dem Ergebnis aus Teil a) erhalten wir:

$$\lim_{t \searrow 0} t \ln t = \lim_{y \to +\infty} e^{-y} \cdot (-y) = 0.$$

c) Sei $t := x \ln x$. Dann gilt wegen Teil b) $t \to 0$, wenn $x \searrow 0$. Also erhalten wir:

$$\frac{x^x - 1}{x \ln x} = \frac{e^{x \ln x} - 1}{x \ln x} = \frac{e^t - 1}{t} \overset{t \to 0}{\longrightarrow} e'(0) = 1.$$

10.2.4 Funktionenfolgen

Übung 10.8 *Betrachten Sie die folgende Aufgabe:*

Untersuchen Sie die folgenden Funktionenfolgen auf punktweise bzw. gleichmäßige Konvergenz!

a) $f_n(x) = x^n - x^{n+1}$ auf $[0, 1]$

b) $g_n(x) = \dfrac{x}{\sqrt{x^2 + \frac{1}{n}}}$ auf $(0, \infty)$

Schreiben Sie zur folgenden komprimierten Musterlösung eine ausführliche Musterlösung! (Einen Lösungsvorschlag finden Sie in Abschnitt 12.2.4.)

Komprimierte Lösung 10.8 a) Da $f_n(0)=f_n(1)=0$ für alle $n\in\mathbb{N}$ und $\lim_{n\to\infty} x^n = 0$ für $0 < x < 1$ gilt, konvergiert $f_n(x)$ für alle $x \in [0,1]$ punktweise gegen $f : [0,1] \to \mathbb{R}$ mit $f \equiv 0$. Für jedes $n \in \mathbb{N}$ suchen wir jetzt

$$a_n := \sup_{x\in[0,1]} |f_n(x) - f(x)| = \sup_{x\in[0,1]} (x^n - x^{n+1}).$$

Setze $h_n(x) := x^n - x^{n+1}$. Dann gilt

$$h_n'(x) = nx^{n-1} - (n+1)x^n \overset{!}{=} 0 \Leftrightarrow x = \frac{n}{n+1} \text{ oder } x = 0,$$
$$h_n(0) = h_n(1) = 0 \text{ und },$$
$$h_n(\frac{n}{n+1}) = (\frac{n}{n+1})^n \cdot (1 - \frac{n}{n+1}) > 0.$$

Also nimmt h_n auf $[0,1]$ das Maximum an der Stelle $x_0 := \frac{n}{n+1}$ an. Damit ist:

$$0 \le a_n = h_n(x_0) = \left(\frac{n}{n+1}\right)^n \cdot \left(1 - \frac{n}{n+1}\right) < 1 \cdot \left(1 - \frac{n}{n+1}\right) \overset{n\to\infty}{\longrightarrow} 0.$$

Daraus folgt, dass f_n auf $[0,1]$ gleichmäßig gegen $f : [0,1] \to \mathbb{R}$ mit $f \equiv 0$ konvergiert.

b) Für $x > 0$ gilt: $\lim_{n\to\infty} g_n(x) = 1$. Damit ist die punktweise Konvergenz gegen $g : (0,\infty) \to \mathbb{R}$ mit $g \equiv 1$ gezeigt.

Setze

$$a_n := \sup_{x>0} |g_n(x) - 1| = \sup_{x>0} \left(1 - \frac{x}{\sqrt{x^2 + \frac{1}{n}}}\right).$$

Dann ist

$$a_n \ge 1 - \frac{\frac{1}{\sqrt{n}}}{\sqrt{\frac{1}{n} + \frac{1}{n}}} = 1 - \frac{1}{\sqrt{2}}.$$

Daraus folgt: $\lim_{n\to\infty} a_n \neq 0$, also konvergiert g_n nicht gleichmäßig gegen g.

10.3 Themen aus der Analysis 2

10.3.1 Grenzwert einer Funktion $f : \mathbb{R}^2 \to \mathbb{R}$

Übung 10.9 *Betrachten Sie die folgende Aufgabe:*

Bestimmen Sie alle Punkte $(a,b) \in \mathbb{R}^2$, für die der folgende Grenzwert existiert:

$$\lim_{(x,y)\to(a,b)} \frac{xy}{x+y}.$$

Schreiben Sie zur folgenden komprimierten Musterlösung eine ausführliche Musterlösung! (Einen Lösungsvorschlag finden Sie in Abschnitt 12.3.1.)

Komprimierte Lösung 10.9 *1. Fall:* $a+b \neq 0$, dann ist

$$\lim_{(x,y)\to(a,b)} \frac{xy}{x+y} = \frac{ab}{a+b}.$$

2. Fall: $(a,b) = (a,-a)$, $a \neq 0$. Für $(x,y) \to (a,-a)$ folgt: $xy \to -a^2 \neq 0$ und $x+y \to 0$. Damit existiert der Grenzwert in diesem Fall nicht.

3. Fall: $(a,b) = (0,0)$. Wähle $\{(x_n,y_n)\} = \{(\frac{1}{n},0)\}$. Es gilt: $\lim\limits_{n\to\infty} \dfrac{\frac{1}{n}\cdot 0}{\frac{1}{n}+0} = 0.$

Wähle $\{(\tilde{x}_n,\tilde{y}_n)\} = \{(\frac{1}{n},\frac{1}{n^2}-\frac{1}{n})\}$. Es gilt: $\lim\limits_{n\to\infty} \dfrac{\frac{1}{n}\cdot\left(\frac{1}{n^2}-\frac{1}{n}\right)}{\frac{1}{n^2}} = -1.$

Der Grenzwert existiert in diesem Fall nicht.

10.3.2 Differenzierbarkeit einer Funktion $f : \mathbb{R}^3 \to \mathbb{R}^3$

Übung 10.10 *Betrachten Sie die folgende Aufgabe:*

> Benutzen Sie die Definition der Differenzierbarkeit um zu beweisen, dass die Funktion $f : \mathbb{R}^3 \to \mathbb{R}^3$ mit
>
> $$f(x,y,z) = \begin{pmatrix} xz - y \\ x^2 + z \\ x + 2y \end{pmatrix}$$
>
> im Punkt
>
> a) $a = (0,0,0)$
>
> b) $\tilde{a} = (1,1,0)$
>
> differenzierbar ist. Bestimmen Sie auch die Ableitung $Df(a)$, ohne partielle Ableitungen zu berechnen!

Schreiben Sie zur folgenden komprimierten Musterlösung eine ausführliche Musterlösung! (Einen Lösungsvorschlag finden Sie in Abschnitt 12.3.2.)

Komprimierte Lösung 10.10 a) $f(x,y,z) = f(0,0,0) + \begin{pmatrix} -y \\ z \\ x+2y \end{pmatrix} + \begin{pmatrix} xz \\ x^2 \\ 0 \end{pmatrix}.$

Die Abbildung $(x,y,z) \mapsto \begin{pmatrix} -y \\ z \\ x+2y \end{pmatrix}$ ist linear. Um zu zeigen, dass diese Abbildung gleich $Df(0,0,0)$ ist, müssen wir beweisen, dass

$$\lim_{(x,y,z)\to(0,0,0)} \frac{1}{\sqrt{x^2+y^2+z^2}} \cdot \begin{pmatrix} xz \\ x^2 \\ 0 \end{pmatrix} = 0$$

ist:

$$\left| \frac{1}{\sqrt{x^2+y^2+z^2}} \cdot \begin{pmatrix} xz \\ x^2 \\ 0 \end{pmatrix} \right| = \sqrt{\frac{x^2z^2 + x^4}{x^2+y^2+z^2}} = |x| \sqrt{\frac{x^2+z^2}{x^2+y^2+z^2}} \leq |x| \to 0$$

für $(x, y, z) \to (0, 0, 0)$.

b) $f(1+x, 1+y, z) = \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix} + \begin{pmatrix} z-y \\ 2x+z \\ x+2y \end{pmatrix} + \begin{pmatrix} xz \\ x^2 \\ 0 \end{pmatrix}.$

Da wir schon in a) gezeigt haben, dass $\lim\limits_{(x,y,z)\to(0,0,0)} \frac{1}{\sqrt{x^2+y^2+z^2}} \cdot \begin{pmatrix} xz \\ x^2 \\ 0 \end{pmatrix} = 0$ ist, ist

$$(x, y, z) \mapsto \begin{pmatrix} z-y \\ 2x+z \\ x+2y \end{pmatrix}$$

die Ableitung $Df(1, 1, 0)$.

10.3.3 Mehrdimensionale Kettenregel

Übung 10.11 *Betrachten Sie die folgende Aufgabe:*

> Gegeben sei eine differenzierbare Funktion $g : \mathbb{R}^3 \to \mathbb{R}$.
>
> a) Stellen Sie die partiellen Ableitungen der Funktion f mit
>
> $$f(x, y) = g(x^2y, x+y, 2xy^3)$$
>
> durch die partiellen Ableitungen der Funktion g dar!
>
> b) Sei $Dg(1, 0, -2) = (0, -1, 1)$. Berechnen Sie $Df(-1, 1)$.

Schreiben Sie zur folgenden komprimierten Musterlösung eine ausführliche Musterlösung! (Einen Lösungsvorschlag finden Sie in Abschnitt 12.3.3.)

Komprimierte Lösung 10.11 a) Seien $u := x^2y$, $v := x+y$, $w := 2xy^3$.

$$\frac{\partial f}{\partial x}(x, y) = \frac{\partial g}{\partial u}(u, v, w) \cdot 2xy + \frac{\partial g}{\partial v}(u, v, w) \cdot 1 + \frac{\partial g}{\partial w}(u, v, w) \cdot 2y^3$$

$$\frac{\partial f}{\partial y}(x, y) = \frac{\partial g}{\partial u}(u, v, w) \cdot x^2 + \frac{\partial g}{\partial v}(u, v, w) \cdot 1 + \frac{\partial g}{\partial w}(u, v, w) \cdot 6xy^2$$

b) $Df(-1, 1) = \begin{pmatrix} 0 \cdot (-2) + (-1) \cdot 1 + 1 \cdot 2 & 0 \cdot 1 + (-1) \cdot 1 + 1 \cdot (-6) \end{pmatrix} = \begin{pmatrix} 1 & -7 \end{pmatrix}$

10.3.4 Lokale Umkehrbarkeit

Übung 10.12 *Betrachten Sie die folgende Aufgabe:*

> Beweisen Sie: Es gibt eine Umgebung V des Punktes $(0,0)$, sodass das Gleichungssystem
>
> $$\begin{aligned} e^{2x+y} - \cos(xy) &= s \\ e^x - \cos(x+y) &= t \end{aligned}$$
>
> für jeden Punkt $(s,t) \in V$ eine Lösung $x := x(s,t)$, $y := y(s,t)$ besitzt.

Schreiben Sie zur folgenden komprimierten Musterlösung eine ausführliche Musterlösung! (Einen Lösungsvorschlag finden Sie in Abschnitt 12.3.4.)

Komprimierte Lösung 10.12 Sei $f(x,y) := \begin{pmatrix} e^{2x+y} - \cos(xy) \\ e^x - \cos(x+y) \end{pmatrix}$. Dann gilt:

$$Df(x,y) = \begin{pmatrix} e^{2x+y} \cdot 2 + \sin(xy) \cdot y & e^{2x+y} + \sin(xy) \cdot x \\ e^x + \sin(x+y) & \sin(x+y). \end{pmatrix}$$

Für $(s,t) = (0,0)$ hat das Gleichungssystem eine Lösung, denn es gilt $f(0,0) = (0,0)$.

$$\det Df(0,0) = \det \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} = -1 \neq 0.$$

Aus dem Satz über die lokale Umkehrbarkeit folgt: Es gibt Umgebungen U von $(0,0)$ und V von $f(0,0) = (0,0)$, sodass $f : U \to V$ bijektiv ist. Also hat das Gleichungssystem für $(s,t) \in V$ eine Lösung.

10.4 Themen aus der Linearen Algebra 1

10.4.1 Gauß-Jordan-Algorithmus

Übung 10.13 *Betrachten Sie die folgende Aufgabe:*

> Bestimmen Sie die Lösungsmenge des folgenden Gleichungssystems mit Koeffizienten aus $\mathbb{R}$ mit Hilfe des Gauß-Jordan-Algorithmus.
>
> $$\begin{array}{cccccccc} & 4x_2 & +3x_3 & +x_4 & +5x_5 & = & 1 \\ x_1 & +3x_2 & +2x_3 & +4x_4 & +2x_5 & = & 1 \\ 3x_1 & +x_2 & +2x_3 & +x_4 & & = & 1 \\ 2x_1 & +2x_2 & +3x_3 & -2x_4 & +3x_5 & = & 1 \end{array}$$

Schreiben Sie zur folgenden komprimierten Musterlösung eine ausführliche Musterlösung! (Einen Lösungsvorschlag finden Sie in Abschnitt 12.4.1.)

Komprimierte Lösung 10.13

$$\left(\begin{array}{ccccc|c} 0 & 4 & 3 & 1 & 5 & 1 \\ 1 & 3 & 2 & 4 & 2 & 1 \\ 3 & 1 & 2 & 1 & 0 & 1 \\ 2 & 2 & 3 & -2 & 3 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccccc|c} 1 & 3 & 2 & 4 & 2 & 1 \\ 0 & 4 & 3 & 1 & 5 & 1 \\ 3 & 1 & 2 & 1 & 0 & 1 \\ 2 & 2 & 3 & -2 & 3 & 1 \end{array}\right)$$

$$\rightsquigarrow \left(\begin{array}{ccccc|c} 1 & 3 & 2 & 4 & 2 & 1 \\ 0 & 4 & 3 & 1 & 5 & 1 \\ 0 & -8 & -4 & -11 & -6 & -2 \\ 0 & -4 & -1 & -10 & -1 & -1 \end{array}\right)$$

$$\rightsquigarrow \left(\begin{array}{ccccc|c} 1 & 3 & 2 & 4 & 2 & 1 \\ 0 & 1 & \frac{3}{4} & \frac{1}{4} & \frac{5}{4} & \frac{1}{4} \\ 0 & -8 & -4 & -11 & -6 & -2 \\ 0 & -4 & -1 & -10 & -1 & -1 \end{array}\right)$$

$$\rightsquigarrow \left(\begin{array}{ccccc|c} 1 & 0 & -\frac{1}{4} & \frac{13}{4} & -\frac{7}{4} & \frac{1}{4} \\ 0 & 1 & \frac{3}{4} & \frac{1}{4} & \frac{5}{4} & \frac{1}{4} \\ 0 & 0 & 2 & -9 & 4 & 0 \\ 0 & 0 & 2 & -9 & 4 & 0 \end{array}\right)$$

$$\rightsquigarrow \left(\begin{array}{ccccc|c} 1 & 0 & -\frac{1}{4} & \frac{13}{4} & -\frac{7}{4} & \frac{1}{4} \\ 0 & 1 & \frac{3}{4} & \frac{1}{4} & \frac{5}{4} & \frac{1}{4} \\ 0 & 0 & 1 & -\frac{9}{2} & 2 & 0 \\ 0 & 0 & 2 & -9 & 4 & 0 \end{array}\right)$$

$$\rightsquigarrow \left(\begin{array}{ccccc|c} 1 & 0 & 0 & \frac{17}{8} & -\frac{5}{4} & \frac{1}{4} \\ 0 & 1 & 0 & \frac{29}{8} & -\frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 1 & -\frac{9}{2} & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Aus der letzten Matrix können wir die Lösungsmenge ablesen:

$$\mathbb{L} = \{ \begin{pmatrix} \frac{1}{4} - \frac{17}{8}s + \frac{5}{4}t \\ \frac{1}{4} - \frac{29}{8}s + \frac{1}{4}t \\ \frac{9}{2}s - 2t \\ s \\ t \end{pmatrix} \mid s, t \in \mathbb{R}\}.$$

10.4.2 Basisergänzung

Übung 10.14 *Betrachten Sie die folgende Aufgabe:*

> Ergänzen Sie die Vektoren $w_1 := \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$, $w_2 := \begin{pmatrix} -2 \\ 0 \\ 5 \end{pmatrix}$ zu einer Basis des $\mathbb{Q}^3$.

Schreiben Sie zur folgenden komprimierten Musterlösung eine ausführliche Musterlösung! (Einen Lösungsvorschlag finden Sie in Abschnitt 12.4.2.)

Komprimierte Lösung 10.14 w_1, w_2 sind aus $\mathbb{Q}^3$ und sie sind linear unabhängig. Wegen des Austauschsatzes von Steinitz können wir sie deshalb gegen zwei Vektoren der Standardbasis $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ von $\mathbb{Q}^3$ tauschen.

Da $w_1 = 3 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ gilt, dürfen wir w_1 z. B. gegen $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ tauschen und erhalten die neue Basis von $\mathbb{Q}^3$: $\left\{ w_1, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$. Da nun aber $w_2 = -\frac{2}{3} \cdot w_1 + \frac{2}{3} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 5 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ ist, dürfen wir w_2 z. B. gegen $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ eintauschen und erhalten die neue Basis von $\mathbb{Q}^3$: $\left\{ w_1, w_2, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$. Damit haben wir w_1, w_2 durch den Vektor $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ zu einer Basis von $\mathbb{Q}^3$ ergänzt.

10.4.3 Invertierbarkeit von Matrizen

Übung 10.15 *Betrachten Sie die folgende Aufgabe:*

> Gegeben sei die Matrix $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & a \\ 1 & a & 1 \end{pmatrix}$ mit $a \in \mathbb{R}$. Für welche $a \in \mathbb{R}$ ist die Matrix invertierbar? Bestimmen Sie für diese Werte von a die inverse Matrix A^{-1}.

Schreiben Sie zur folgenden komprimierten Musterlösung eine ausführliche Musterlösung! (Einen Lösungsvorschlag finden Sie in Abschnitt 12.4.3.)

Komprimierte Lösung 10.15

$$\left.\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & a & 0 & 1 & 0 \\ 1 & a & 1 & 0 & 0 & 1 \end{array}\right. \quad \rightsquigarrow \quad \begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & a-1 & -1 & 1 & 0 \\ 0 & a-1 & 0 & -1 & 0 & 1 \end{array}$$

Für $a = 1$ ist dann A nicht invertierbar. Für $a \neq 1$ folgt durch weitere Zeilenumformungen:

$$\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -\frac{1}{a-1} & \frac{1}{a-1} & 0 \\ 0 & 1 & 0 & -\frac{1}{a-1} & 0 & \frac{1}{a-1} \end{array} \quad \rightsquigarrow \quad \begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -\frac{1}{a-1} & 0 & \frac{1}{a-1} \\ 0 & 0 & 1 & -\frac{1}{a-1} & \frac{1}{a-1} & 0 \end{array}$$

$$\rightsquigarrow \quad \begin{array}{ccc|ccc} 1 & 0 & 1 & 1+\frac{1}{a-1} & 0 & -\frac{1}{a-1} \\ 0 & 1 & 0 & -\frac{1}{a-1} & 0 & \frac{1}{a-1} \\ 0 & 0 & 1 & -\frac{1}{a-1} & \frac{1}{a-1} & 0 \end{array}$$

$$\rightsquigarrow \quad \begin{array}{ccc|ccc} 1 & 0 & 0 & 1+\frac{2}{a-1} & -\frac{1}{a-1} & -\frac{1}{a-1} \\ 0 & 1 & 0 & -\frac{1}{a-1} & 0 & \frac{1}{a-1} \\ 0 & 0 & 1 & -\frac{1}{a-1} & \frac{1}{a-1} & 0. \end{array}$$

Und damit ist

$$A^{-1} = \frac{1}{a-1} \begin{pmatrix} a+1 & -1 & -1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}$$

10.4.4 Kommutierende Matrizen

Übung 10.16 *Betrachten Sie die folgende Aufgabe:*

> Sei K ein Körper und $n \in \mathbb{N}$. Zeigen Sie, dass jede Matrix $A \in M_{n\times n}(K)$ mit $AX = XA$ für alle $X \in M_{n\times n}(K)$ die Gestalt $A = aE_n$ für ein $a \in K$ hat.

Schreiben Sie zur folgenden komprimierten Musterlösung eine ausführliche Musterlösung! (Einen Lösungsvorschlag finden Sie in Abschnitt 12.4.4.)

Komprimierte Lösung 10.16 Da $AX = XA$ für alle $X \in M_{n\times n}(K)$ ist, gilt insbesondere für die Elementarmatrizen E_{ij} (1 in Zeile i und Spalte j, alle anderen Einträge 0):

$$AE_{ij} = E_{ij}A.$$

Nun ist aber

$$AE_{ij} = \begin{array}{c} \begin{array}{ccccccc} & & & j\text{-te Spalte} & & & \\ & & & \downarrow & & & \end{array} \\ \begin{pmatrix} 0 & \cdots & 0 & a_{1i} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{ii} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{ni} & 0 & \cdots & 0 \end{pmatrix} \end{array}$$

$$\text{und} \quad E_{ij}A = \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \\ a_{j1} & \cdots & a_{jj} & \cdots & a_{jn} \\ 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} \begin{array}{l} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \\ \end{array}$$

Die Einträge a_{ii} bzw. a_{jj} tauchen dabei jeweils in der i-ten Zeile und j-ten Spalte auf. Da die beiden Matrizen gleich sein müssen, gilt $a_{ii} = a_{jj}$, $a_{li} = 0$ und $a_{jk} = 0$ für alle $l \neq i$ und $k \neq j$.

Setzt man also $a := a_{11}$, so ist $A = aE_n$.

10.5 Themen aus der Linearen Algebra 2

10.5.1 Eigenwerte, Eigenräume, charakteristisches Polynom und Minimalpolynom

Übung 10.17 *Betrachten Sie die folgende Aufgabe:*

> Sei $V := \{f \in \mathbb{R}[x] \mid \operatorname{grad}(f) \leq n\}$ der Vektorraum der reellen Polynome mit Grad kleiner oder gleich n und sei die lineare Abbildung $A : V \to V$ durch
>
> $$A(f) := f' \quad \text{für alle } f \in V$$
>
> gegeben. Bestimmen Sie die Eigenwerte, Eigenräume, charakteristisches Polynom und Minimalpolynom dieser Abbildung.

Schreiben Sie zur folgenden komprimierten Musterlösung eine ausführliche Musterlösung! (Einen Lösungsvorschlag finden Sie in Abschnitt 12.5.1.)

Komprimierte Lösung 10.17 Wir stellen die Matrix zur Abbildung bzgl. der Basis $B = \{1, x, x^2, x^3, \ldots, x^n\}$ von V auf:

$$M_B^B(A) = \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 2 & 0 & \cdots & 0 \\ \vdots & 0 & 0 & 3 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 0 & n \\ 0 & \cdots & \cdots & \cdots & 0 & 0 \end{pmatrix}.$$

Dann ist $\operatorname{charpol}_A(x) = x^{n+1}$ und 0 der einzige Eigenwert.

Da $\operatorname{rg}(A) = \operatorname{rg}(M_B^B(A)) = n$ und $A(1) = 0$ ist, bildet $x^0 = 1$ eine Basis von $\ker(A) = V(0, A)$.

Weiterhin gilt:

$$\begin{aligned} &A(x^n) = n \cdot x^{n-1} \neq 0 \\ &A^2(x^n) = n \cdot (n-1) \cdot x^{n-2} \neq 0 \\ &\vdots \\ &A^n(x^n) = n \cdot (n-1) \cdot \ldots 2 \cdot 1 \cdot x^0 \neq 0 \\ &A^{n+1}(x^n) = 0. \end{aligned}$$

Also ist $\operatorname{minpol}_A(x) = x^{n+1}$.

10.5.2 Diagonalisierbarkeit

Übung 10.18 *Betrachten Sie die folgende Aufgabe:*

Überprüfen Sie, ob die folgenden Matrizen über $\mathbb{Q}$ bzw. $\mathbb{C}$ diagonalisierbar sind:

$$A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -2 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Schreiben Sie zur folgenden komprimierten Musterlösung eine ausführliche Musterlösung! (Einen Lösungsvorschlag finden Sie in Abschnitt 12.5.2.)

Komprimierte Lösung 10.18 Zu A:

Die Matrix A besteht aus zwei Blöcken: $A = \left(\begin{array}{cc|cc} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 \\ 0 & 0 & -2 & 4 \end{array}\right)$. Sie ist genau dann diagonalisierbar, wenn die beiden Blöcke diagonalisierbar sind.

Setze $A_1 := \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$, dann ist $\text{charpol}_{A_1}(x) = (x-2)^2$.

$\text{minpol}_{A_1}(x) = (x-2)^2$, da $A - 2E_2 \neq 0$ ist. Da das Minimalpolynom nicht in lauter verschiedene Linearfaktoren zerfällt, ist A_1 nicht diagonalisierbar und damit ist auch A nicht diagonalisierbar (weder über $\mathbb{Q}$ noch über $\mathbb{C}$).

Alternativ (ohne Minimalpolynom):
Berechnung des Eigenraums:

$$V(2, A_1) = \ker \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rangle$$

Da $V(2, A_1) \neq K^2$ ist ($K \in \{\mathbb{Q}, \mathbb{C}\}$), ist A_1 und damit auch A nicht diagonalisierbar.

Zu B:

$$\begin{aligned} \text{charpol}_B(x) &= \begin{vmatrix} x & -1 & 0 & 0 \\ 0 & x & 0 & -1 \\ -1 & 0 & x & 0 \\ 0 & 0 & -1 & x \end{vmatrix} = x \begin{vmatrix} x & 0 & -1 \\ 0 & x & 0 \\ 0 & -1 & x \end{vmatrix} + 1 \begin{vmatrix} 0 & 0 & -1 \\ -1 & x & 0 \\ 0 & -1 & x \end{vmatrix} \\ &= x^4 - 1 = (x-1)(x+1)(x^2+1) \end{aligned}$$

$x^2 + 1$ hat über $\mathbb{Q}$ keine Nullstellen. Deshalb ist B über $\mathbb{Q}$ nicht diagonalisierbar.

Über $\mathbb{C}$ zerfällt das charakteristische Polynom in Linearfaktoren:

$$\text{charpol}_B(x) = (x-1)(x+1)(x-i)(x+i).$$

Damit ist B diagonalisierbar.

10.5.3 Linearformen

Übung 10.19 *Betrachten Sie die folgende Aufgabe:*

> Seien V ein K-Vektorraum (K Körper) und $f, g \in V^*$ Linearformen (d. h. $f, g : V \to K$ linear) mit folgender Eigenschaft: Ist $v \in V$ und $f(v) = 0$, so ist auch $g(v) = 0$. Zeigen Sie, dass es ein $\lambda \in K$ gibt mit $g = \lambda \cdot f$.

Schreiben Sie zur folgenden komprimierten Musterlösung eine ausführliche Musterlösung! (Einen Lösungsvorschlag finden Sie in Abschnitt 12.5.3.)

Komprimierte Lösung 10.19 Nach Voraussetzung ist $\ker(f) \subseteq \ker(g) \subseteq V$.

Ist $\ker(g) = V$, so ist g die Nullabbildung. Setze dann $\lambda := 0$.

Sei nun $\ker(g) \subsetneq V$.

Sei $v \in V \setminus \ker(g)$. Dann ist $g(v) \neq 0$ und deshalb $f(v) \neq 0$. Da $\dim(K) = 1$ ist, folgt $\text{im}(f) = \text{im}(g) = K$.

Da $V/\ker(f) \cong \mathrm{im}(f) = K = \mathrm{im}(g) \cong V/\ker(g)$ und $\ker(f) \subseteq \ker(g)$ ist, folgt nun $U := \ker(f) = \ker(g)$ und $V = U \oplus \langle v \rangle$.

Sei $f(v) = k_1 \neq 0$ und $g(v) = k_2 \neq 0$. Setze $\lambda := k_2 \cdot k_1^{-1} \in K$.

Für $u \in U$ ist $(\lambda f)(u) = \lambda \cdot f(u) = 0 = g(u)$. Für v gilt: $(\lambda f)(v) = \lambda \cdot f(v) = k_2 = g(v)$.

Wegen der Linearität von λf und g folgt jetzt $(\lambda f)(w) = g(w)$ für alle $w \in V$.

Somit gilt $g = \lambda f$.

10.5.4 Skalarprodukt

Übung 10.20 *Betrachten Sie die folgende Aufgabe:*

> Sei V ein K-Vektorraum ($K \in \{\mathbb{R}, \mathbb{C}\}$) mit Skalarprodukt (auch abgekürzt mit SKP) f und sei $A : V \to V$ eine lineare Abbildung mit $f(A(v), v) = 0$ für alle $v \in V$.
>
> a) Zeigen Sie für $K = \mathbb{C}$, dass $A = 0$ gilt.
>
> b) Finden Sie zu $K = \mathbb{R}$ ein Beispiel für A und f, sodass $A \neq 0$ ist, aber dennoch alle obigen Forderungen erfüllt sind.

Schreiben Sie zur folgenden komprimierten Musterlösung eine ausführliche Musterlösung! (Einen Lösungsvorschlag finden Sie in Abschnitt 12.5.4.)

Komprimierte Lösung 10.20 Zu (a): Für alle $v, w \in V$ hat man:

$$\begin{aligned} 0 &= f(A(v+w), v+w) = f(A(v) + A(w), v+w) \\ &= f(A(v), v) + f(A(v), w) + f(A(w), v) + f(A(w), w) \\ &= f(A(v), w) + f(A(w), v) \end{aligned}$$

und

$$\begin{aligned} 0 &= f(A(v+iw), v+iw) = f(A(v) + iA(w), v+iw) \\ &= f(A(v), v) + \bar{i} f(A(v), w) + i f(A(w), v) + i\bar{i} f(A(w), w) \\ &= \bar{i} f(A(v), w) + i f(A(w), v) \\ &= i(f(A(w), v) - f(A(v), w)). \end{aligned}$$

Damit hat man

$$f(A(v), w) + f(A(w), v) = 0 \quad \text{und} \quad f(A(v), w) - f(A(w), v) = 0.$$

Es folgt $f(A(v), w) = 0$ für alle $v, w \in V$. Insbesondere ist dann für $w = A(v)$: $f(A(v), A(v)) = 0$ für alle $v \in V$ und somit ist $A(v) = 0$ für alle $v \in V$.

Zu (b): Betrachte $V := \mathbb{R}^2$ mit dem Standardskalarprodukt f und dem Endomorphismus A, der durch $A(e_1) = e_2, A(e_2) = -e_1$ gegeben ist (e_1, e_2 sei die Standardbasis des $\mathbb{R}^2$). Für alle $v \in \mathbb{R}^2$ gibt es $\lambda, \mu \in \mathbb{R}$ mit $v = \lambda e_1 + \mu e_2$. Somit folgt:

$$\begin{aligned}
f(A(v), v) &= f(\lambda A(e_1) + \mu A(e_2), \lambda e_1 + \mu e_2) \\
&= \lambda^2 f(e_2, e_1) + \lambda\mu f(e_2, e_2) + \lambda\mu f(-e_1, e_1) + \mu^2 f(-e_1, e_2) \\
&= 0 + \lambda\mu - \lambda\mu + 0 \\
&= 0.
\end{aligned}$$

Teil IV

Lösungsvorschläge

11 Lösungsvorschläge zu Teil II

11.1 Lösungen zu Kapitel 5

Antworten zu den Verständnisfragen 5.1 – Summen- und Produktzeichen

1. Es gilt definitionsgemäß $m^0 = 1$. Diese Definition legt man deshalb fest, damit z. B. die Potenzrechenregel $\frac{a^n}{a^m} = a^{n-m}$ auch für $\frac{a^3}{a^3}$ gilt. Den Faktor 1 braucht man dann natürlich nicht weiter mitschleppen.

2. Es ist möglich. In gewisser Weise ist der Ausdruck $4 + 3n$ sogar naheliegender als der Term $1 + 3n$. Immerhin ist der Nenner des ersten Faktors ja 4 und ab dann wird in jedem Schritt die Zahl 3 addiert. Der Index n muss dann von 0 bis 4 laufen. Der Zähler wird durch diese Laufindexveränderung allerdings etwas komplizierter: $(n+1)^2 + 1$. Man könnte also schreiben:

$$\frac{2}{4} \cdot \frac{5}{7} \cdot \frac{10}{10} \cdot \frac{17}{13} \cdot \frac{26}{16} = \prod_{n=0}^{4} \frac{(n+1)^2 + 1}{4 + 3n}.$$

3. Man hätte $3(k+1) + 5$, also $3k + 8$ schreiben müssen.

Komprimierte Lösung zu Aufgabe 5.1 – Summen- und Produktzeichen

a) $\displaystyle\sum_{k=2}^{5} \left((-1)^k \cdot \frac{k-3}{k+7} \right)$

$$= (-1)^2 \cdot \frac{2-3}{2+7} + (-1)^3 \cdot \frac{3-3}{3+7} + (-1)^4 \cdot \frac{4-3}{4+7} + (-1)^5 \cdot \frac{5-3}{5+7}$$
$$= -\frac{1}{9} - 0 + \frac{1}{11} - \frac{1}{6} = -\frac{37}{198}$$

b) $\displaystyle\prod_{j=-3}^{5} m^j = m^{-3} \cdot m^{-2} \cdot m^{-1} \cdot m^0 \cdot m^1 \cdot m^2 \cdot m^3 \cdot m^4 \cdot m^5 = m^9$

c) $\displaystyle\frac{2}{4} \cdot \frac{5}{7} \cdot \frac{10}{10} \cdot \frac{17}{13} \cdot \frac{26}{16} = \prod_{n=1}^{5} \frac{n^2 + 1}{1 + 3n}$

d) $\displaystyle\sum_{k=1}^{n} \frac{(k-2)^2}{3^{k+1}} = \sum_{k=0}^{n-1} \frac{(k-1)^2}{3^{k+2}}$

e) $\displaystyle\prod_{l=3}^{6} \frac{l+3}{l^2} = \prod_{l=-2}^{1} \frac{l+8}{(l+5)^2}$

C. Ableitinger, A. Herrmann, *Lernen aus Musterlösungen zur Analysis und Linearen Algebra*,
DOI 10.1007/978-3-658-02353-9_11, © Springer Fachmedien Wiesbaden 2013

Antworten zu den Verständnisfragen 5.2 – Vollständige Induktion

1. Da die Aussage für alle $n \in \mathbb{N}$ mit $n \geq 2$ gezeigt werden soll, muss der Induktionsanfang als $n = 2$ gewählt werden. (Die Ungleichung stimmt für $n = 1$ im Übrigen auch nicht.)

2. Entscheidend für den Beweis ist, dass wir aus einer wahren Aussage, die zu zeigende Aussage folgern! Da wir im Beweis damit beginnen, die zu zeigende Aussage auf eine wahre zurückzuführen, sind vor allem die Rückrichtungen wichtig. Statt

$$\begin{aligned} & (3n+2) \cdot 2 \geq 3(n+1) + 2 \\ \Leftrightarrow\ & 6n + 4 \geq 3n + 5 && |(-3n);\ |(-4) \\ \Leftrightarrow\ & 3n \geq 1 \end{aligned}$$

könnten wir also auch

$$\begin{aligned} & 3n \geq 1 \\ \Rightarrow\ & 6n + 4 \geq 3n + 5 \\ \Rightarrow\ & (3n+2) \cdot 2 \geq 3(n+1) + 2 \end{aligned}$$

schreiben und hätten immer noch einen vollständigen Beweis formuliert.

Komprimierte Lösung zu Aufgabe 5.2 – Vollständige Induktion

Induktionsanfang: Für $n = 2$ ist

$$2^{n+1} = 2^{2+1} = 2^3 = 8 \geq 8 = 3 \cdot 2 + 2 = 3n + 2,$$

damit ist der Induktionsanfang gezeigt.

Induktionsschritt von n nach $n + 1$: Es gelte $2^{n+1} \geq 3n + 2$ für ein beliebiges, aber festes $n \geq 2$.

Es folgt:

$$2^{(n+1)+1} = 2^{n+1} \cdot 2 \overset{\text{IV}}{\geq} (3n + 2) \cdot 2.$$

Es bleibt also noch zu zeigen: $(3n + 2) \cdot 2 \geq 3(n + 1) + 2$.

Beweis:

$$\begin{aligned} & (3n+2) \cdot 2 && \overset{?}{\geq} && 3(n+1) + 2 \\ \Leftrightarrow\ & 6n + 4 && \overset{?}{\geq} && 3n + 5 && |(-3n);\ |(-4) \\ \Leftrightarrow\ & 3n && \overset{?}{\geq} && 1 \end{aligned}$$

Die letzte Ungleichung gilt, da $n \geq 2$ ist. Somit ist die Aussage auch für $n + 1$ bewiesen.

Antworten zu den Verständnisfragen 5.3 – Bild und Urbild

1. Wie bei Teil a) auch wurden zur Begründung der Implikationen nur Definitionen herangezogen und Definitionen sind immer „Wenn-Dann-Aussagen".

2. Ist $x \in f^{-1}(Y)$, so bedeutet das nach der Definition, dass x einerseits ein Element von X ist und dass weiterhin $f(x) \in Y$ ist. Also liegt x insbesondere in X. Die Inklusion $f^{-1}(Y) \subseteq X$ ist somit gezeigt. Liegt x dagegen in X, dann ist aber $f(x) \in Y$, weil f nun mal nach Y abbildet. Demnach ist $x \in X$ mit der Eigenschaft $f(x) \in Y$, wonach es wiederum aufgrund der Definition des Urbildes in $f^{-1}(Y)$ liegt.
3. Auf $f^{-1}(C) \subseteq X$.
4. f surjektiv $\Leftrightarrow\ f(X) = Y \ \Leftrightarrow\ \forall y \in Y\ \exists x \in X$ mit $y = f(x)$.
5. Nein, die Äquivalenz gilt dann nicht mehr. Es folgt zwar weiterhin aus der Surjektivität die Mengengleichheit, aber die andere Richtung ist nur für den Spezialfall $C = Y$ gegeben. Betrachte zum Beispiel die Abbildung $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^2$. Diese ist nicht surjektiv, da es zum Beispiel für $y = -1$ kein $x \in \mathbb{R}$ mit $f(x) = -1$ gibt. Trotzdem gilt für $C := \mathbb{R}^+ = \{x \in \mathbb{R} | x \geq 0\}$:
$$f(f^{-1}(C)) = f(\mathbb{R}) = \mathbb{R}^+ = C.$$

Komprimierte Lösung zu 5.3 – Bild und Urbild

a)

$$\begin{aligned} y \in f(A \cup B) &\Leftrightarrow \exists x \in A \cup B \text{ mit } y = f(x) \\ &\Leftrightarrow \exists x \in A \text{ oder } \exists x \in B \text{ mit } y = f(x) \\ &\Leftrightarrow y \in f(A) \text{ oder } y \in f(B) \\ &\Leftrightarrow y \in f(A) \cup f(B) \end{aligned}$$

b)

$$\begin{aligned} x \in f^{-1}(C \cap D) &\Leftrightarrow x \in X \text{ und } f(x) \in C \cap D \\ &\Leftrightarrow x \in X \text{ und } f(x) \in C \text{ und } f(x) \in D \\ &\Leftrightarrow x \in f^{-1}(C) \text{ und } x \in f^{-1}(D) \\ &\Leftrightarrow x \in f^{-1}(C) \cap f^{-1}(D) \end{aligned}$$

c)

$$\begin{aligned} y \in f(f^{-1}(C)) &\Rightarrow \exists x \in f^{-1}(C) \text{ mit } f(x) = y \\ &\Rightarrow \exists x \in X \text{ mit } f(x) = y \text{ und } f(x) \in C \\ &\Rightarrow y = f(x) \in C \end{aligned}$$

d) „f surjektiv $\Rightarrow$ Gleichheit“: Es bleibt zu zeigen: $C \subseteq f(f^{-1}(C))$: Sei $y \in C$, dann existiert (da f surjektiv ist) ein $x \in X$ mit $f(x) = y$. Außerdem ist $f(x) \in C$ und x somit Element von $f^{-1}(C)$. Insgesamt gilt: Es existiert ein $x \in f^{-1}(C)$ mit $y = f(x)$. Dann gilt also: $y \in f(f^{-1}(C))$.

„Gleichheit $\Rightarrow f$ surjektiv“: Wenn die Gleichheit für alle Teilmengen von Y gelten soll, dann gilt sie insbesondere auch für Y (als Teilmenge seiner selbst): $f(f^{-1}(Y)) = Y$. Da $f^{-1}(Y) = X$ ist, folgt $f(X) = Y$ und damit auch, dass f surjektiv ist.

Antworten zu den Verständnisfragen 5.4 – Injektivität und Surjektivität

1. Für f wurde benutzt, dass es in die Menge Y abbildet, also die Elemente $f(x)$ mit $x \in X$ aus Y stammen.

   Für g wurde die Wohldefiniertheit einer Abbildung (jedem Element aus dem Definitionsbereich wird *genau* ein Element aus dem Bildbereich zugeordnet) ausgenutzt.

2. Als Erstes nimmt man sich ein beliebiges, aber festes n aus N. Zu diesem n muss man dann ein m aus M finden, für das $f(m) = n$ gilt. Das m findet man, indem man die Informationen, die man über f hat, ausnutzt. (In unserem Beisiel war das die Eigenschaft der Verknüfung von g mit f. In anderen Fällen kann man oftmals die Abbildungsvorschrift $n \overset{!}{=} f(m) = \ldots$ nach m umstellen.) Zum Schluss muss man noch sicherstellen, dass das so gefundene m auch in M liegt.

   Achtung: In der Beschreibung wurde hier die Surjektivität einer Abbildung von M nach N beschrieben; in der Aufgabe hatten wir die Surjektivität der Abbildung g von Y nach X überprüft. Achten Sie auf die nötigen Notationsanpassungen!

Lösungsvorschlag zu Übung 5.4 – Injektivität und Surjektivität

Von oben nach unten: B K B Z H B K B Z K Z H A

Komprimierte Lösung zu Aufgabe 5.4 – Injektivität und Surjektivität

Zu zeigen: f injektiv.

Seien $x_1, x_2 \in X$ mit $f(x_1) = f(x_2)$ beliebig. Da g eine Abbildung ist, muss auch $g(f(x_1)) = g(f(x_2))$ sein. Aus $(g \circ f)(x) = x$ für alle $x \in X$ folgt dann schon, dass $x_1 = x_2$ gilt. Damit haben wir gezeigt, dass f injektiv ist.

Zu zeigen: g surjektiv.

Sei $x \in X$ beliebig. Setze $y := f(x)$. Dann ist $y \in Y$, weil f von X nach Y abbildet, und $g(y) = g(f(x)) = x$ nach der Voraussetzung für $g \circ f$. Somit ist g surjektiv.

Antworten zu den Verständnisfragen 5.5 – Äquivalenzrelation und -klassen

1. Für $\sim$ reflexiv: $a = (x, y)$. Für $\sim$ symmetrisch: $a = (x_1, y_1)$, $b = (x_2, y_2)$. Für $\sim$ transitiv: $a = (x_1, y_1)$, $b = (x_2, y_2)$, $c = (x_3, y_3)$.

2. Im Fall $x_2 = 0$ wird von den Gleichungen $x_1 \cdot y_2 = x_2 \cdot y_1$ und $x_2 \cdot y_3 = x_3 \cdot y_2$ ausgegangen. Da nun $x_2 = 0$ ist, erhalten wir durch Einsetzen in die Gleichungen: $x_1 \cdot y_2 = 0$ und $0 = x_3 \cdot y_2$. Da für den Fall $x_2 = 0$ nun aber $y_2 \neq 0$ ist, können wir daraus $x_1 = 0$ bzw. $x_3 = 0$ folgern. Im Fall $y_2 = 0$ wird ebenfalls von den Gleichungen $x_1 \cdot y_2 = x_2 \cdot y_1$ und $x_2 \cdot y_3 = x_3 \cdot y_2$ ausgegangen. Da nun $y_2 = 0$ ist, erhalten wir durch Einsetzen in die Gleichungen: $0 = x_2 \cdot y_1$ und $x_2 \cdot y_3 = 0$. Da für den Fall $y_2 = 0$ nun aber $x_2 \neq 0$ ist, können wir daraus $y_1 = 0$ bzw. $y_3 = 0$ folgern, was wiederum $x_1 \cdot y_3 = x_3 \cdot y_1$ (auf beiden Seiten der Gleichung steht Null) und somit $(x_1, y_1) \sim (x_3, y_3)$ impliziert.

3. Nein, dann wäre es keine Äquivalenzrelation mehr. Für die Reflexivität und die Symmetrie haben wir nicht ausgenutzt, dass die Elemente ungleich $(0,0)$ sein müssen. Im Beweis für die Transitivität ist dies aber eingeflossen. D. h., dass es nur dort zu Problemen führen kann. Und dies macht es auch tatsächlich. Zum Beispiel gilt: $(1,2) \sim (0,0)$ und $(0,0) \sim (3,4)$, aber nicht $(1,2) \sim (3,4)$, wie man schnell nachprüfen kann. Damit wäre die Relation nicht mehr transitiv.

4. Um ein Repräsentantensystem zu bestimmen, muss man zunächst wissen, wie die Äquivalenzklassen aussehen. Das haben wir in der Aufgabe gemacht. Ein Repräsentantensystem ist eine Teilmenge R der zugrunde liegenden Menge M (hier also von $(\mathbb{R} \times \mathbb{R}) \setminus \{(0,0)\}$), und zwar so, dass jedes Element der Menge M in *genau einer* der Äquivalenzklassen der Elemente von R liegt. Man muss also die ganze Menge $(\mathbb{R} \times \mathbb{R}) \setminus \{(0,0)\}$ mit den Äquivalenzklassen der Repräsentanten abdecken, darf aber kein Element der Menge doppelt abdecken[1]. In unserem Fall müssen wir also die ganze Ebene (bis auf den Koordinatenursprung) abdecken. Das bekommen wir hin, indem wir alle möglichen Steigungen $\frac{b}{a}$ für die Geraden in Betracht ziehen. Damit wir dabei nichts doppelt verwerten, können wir z. B. $a = 1$ setzen und b alle reellen Zahlen durchwandern lassen. Damit haben wir alle reellwertigen Steigungen abgedeckt. Die y-Achse fehlt uns aber noch – den Fall $a = 0$ haben wir so nämlich außer Acht gelassen. Wir müssen einen Repräsentanten der Form $(0, b)$ noch hinzunehmen. Ein Repräsentantensystem wäre dann beispielsweise die Menge $\{(1,b) | b \in \mathbb{R}\} \cup \{(0,1)\}$. Es besteht demnach aus unendlich vielen Elementen. Den exakten Beweis, dass es sich dabei tatsächlich um ein Repräsentantensystem handelt, werden wir an dieser Stelle nicht führen. Noch ein kleiner Hinweis: Die Wahl des Repräsentantensystems ist nicht eindeutig – es gibt unendlich viele andere Möglichkeiten, ein Repräsentantensystem anzugeben!

Komprimierte Lösung zu Aufgabe 5.5 – Äquivalenzrelation und -klassen

1. $\sim$ ist reflexiv: Da $x_1 \cdot y_1 = x_1 \cdot y_1$ folgt: $(x_1, y_1) \sim (x_1, y_1)$ für alle $(x_1, y_1) \in (\mathbb{R} \times \mathbb{R}) \setminus \{(0,0)\}$.

2. $\sim$ ist symmetrisch:

$$\begin{aligned}(x_1, y_1) \sim (x_2, y_2) &\Rightarrow x_1 \cdot y_2 = x_2 \cdot y_1 \\ &\Rightarrow x_2 \cdot y_1 = x_1 \cdot y_2 \\ &\Rightarrow (x_2, y_2) \sim (x_1, y_1)\end{aligned}$$

3. $\sim$ ist transitiv:

$$\begin{aligned}&(x_1, y_1) \sim (x_2, y_2) \text{ und } (x_2, y_2) \sim (x_3, y_3) \\ &\qquad \Rightarrow x_1 \cdot y_2 = x_2 \cdot y_1 \text{ und } x_2 \cdot y_3 = x_3 \cdot y_2 \\ &\qquad \Rightarrow x_1 \cdot y_2 \cdot x_2 \cdot y_3 = x_2 \cdot y_1 \cdot x_3 \cdot y_2 \\ &\qquad \overset{x_2 \cdot y_2 \neq 0}{\Rightarrow} x_1 \cdot y_3 = x_3 \cdot y_1 \\ &\qquad \Rightarrow (x_1, y_1) \sim (x_3, y_3)\end{aligned}$$

[1] Dass kein Element doppelt abgedeckt wird, ist im Übrigen gleichbedeutend damit, dass keine Äquivalenzklasse doppelt verwendet wird.

Falls $x_2 = 0$ und $y_2 \neq 0$ ist, sind $x_1 = 0$ und $x_3 = 0$. Daraus folgt: $(x_1, y_1) \sim (x_3, y_3)$. Analog geht man für den Fall $x_2 \neq 0$ und $y_2 = 0$ vor.

Aus 1.-3. folgt die Behauptung, dass $\sim$ eine Äquivalenzrelation ist.

Wir bestimmen nun noch die Äquivalenzklasse $[(a, b)]$:

1. Fall: $a \neq 0$:

$$\begin{aligned}[(a,b)] &= \{(x,y) \in (\mathbb{R} \times \mathbb{R}) \backslash \{(0,0)\} | (a,b) \sim (x,y)\} \\ &= \{(x,y) \in (\mathbb{R} \times \mathbb{R}) \backslash \{(0,0)\} | a \cdot y = x \cdot b\} \\ &= \{(x,y) \in (\mathbb{R} \times \mathbb{R}) \backslash \{(0,0)\} | y = \frac{b}{a} \cdot x\}\end{aligned}$$

Die Äquivalenzklasse von (a, b) ist also die Menge der Punkte auf der Geraden $y = \frac{b}{a} \cdot x$ ohne den Ursprung.

2. Fall: $a = 0$:

$$\begin{aligned}[(0,b)] &= \{(x,y) \in (\mathbb{R} \times \mathbb{R}) \backslash \{(0,0)\} | (0,b) \sim (x,y)\} \\ &= \{(x,y) \in (\mathbb{R} \times \mathbb{R}) \backslash \{(0,0)\} | 0 \cdot y = x \cdot b\} \\ &= \{(x,y) \in (\mathbb{R} \times \mathbb{R}) \backslash \{(0,0)\} | x = 0\}\end{aligned}$$

Die Äquivalenzklasse von $(0, b)$ entspricht also den Punkten auf der y-Achse ohne den Ursprung.

11.2 Lösungen zu Kapitel 6

Antworten zu den Verständnisfragen 6.1 – Supremum und Infimum

1. Nein. Das Berechnen der beiden Grenzwerte kann zwar in manchen Fällen eine Idee liefern, welche Zahlen als Infimum bzw. Supremum in Frage kommen, doch ein Beweis ist das nicht. Es kann nämlich eine Situation wie in Abbildung 11.1 eintreten.

   Dort findet man die graphische Darstellung der Menge

   $$M = \left\{(x^3 - 3x^2 + 4) \cdot e^{-x} \mid x \in \mathbb{R}\right\}.$$

   Man erkennt an dieser Abbildung, dass die Menge kein Infimum besitzt und dass das Supremum bei einem Wert zwischen 4 und 6 liegen wird. Berechnet man lediglich die beiden Grenzwerte

   $$\lim_{x \to -\infty} (x^3 - 3x^2 + 4) \cdot e^{-x} = -\infty \text{ und } \lim_{x \to +\infty} (x^3 - 3x^2 + 4) \cdot e^{-x} = 0,$$

   so erkennt man das nicht.

2. Für $\varepsilon_0 = 3$ liefert der Ausdruck $\frac{6-2\varepsilon_0}{\varepsilon_0} = \frac{6-3\cdot 2}{3} = 0$. Die beiden Ungleichungen $x_0 > 0$ bzw. $x_0 > \frac{6-2\varepsilon_0}{\varepsilon_0}$ stellen dann also dieselbe Bedingung an x_0. Für $\varepsilon_0 > 3$

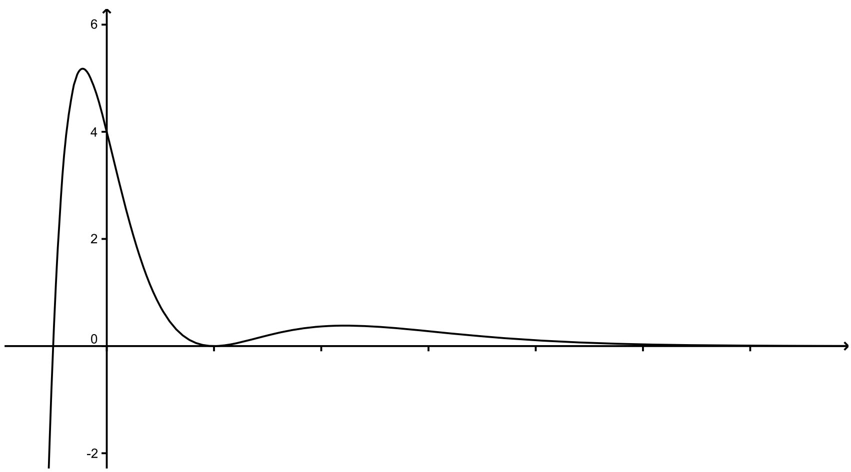

Abbildung 11.1: Graphische Darstellung der Menge M (Aufgabe 6.1)

bzw. $\varepsilon_0 < 3$ ist jeweils eine Bedingung *stärker* als die andere. Die Fallunterscheidung dient dazu herauszufinden, welche Bedingung durch die jeweils andere schon automatisch miterfüllt ist und daher vernachlässigt werden kann.

3. Wir möchten eine Zahl x_0 nennen, die die Bedingung $x_0 < \frac{2\varepsilon_0}{3-\varepsilon_0}$ erfüllt, egal welchen Wert ε_0 im Intervall $(0,3)$ annimmt. Eine konkrete Zahl können wir für x_0 leider nicht angeben, weil der Ausdruck $\frac{2\varepsilon_0}{3-\varepsilon_0}$ beliebig nahe an 0 herankommt. Auch können wir nicht einfach $x_0 = \frac{2\varepsilon_0}{3-\varepsilon_0} - 1$ nehmen, da dieser Ausdruck für bestimmte Werte von ε_0 negativ würde, wir aber ein positives x_0 angeben müssen. Wir brauchen also eine positive Zahl, die in Abhängigkeit von ε_0 kleiner als die Zahl $\frac{2\varepsilon_0}{3-\varepsilon_0}$ ist, z. B. einfach die Hälfte dieser Zahl.

Komprimierte Lösung zu Aufgabe 6.1 – Supremum und Infimum

Supremum: Da $\frac{3x}{x+2} < 3$ für alle $x > 0$ gilt, ist 3 eine obere Schranke von M. Um zu zeigen, dass $\sup M = 3$ gilt, müssen wir beweisen:

Zu jedem $\varepsilon > 0$ gibt es ein $x > 0$ mit $\frac{3x}{x+2} > 3 - \varepsilon$.

Wenn $\varepsilon \geq 3$ ist, dann gilt die Ungleichung für jedes $x > 0$.

Sei also $0 < \varepsilon < 3$. Dann ist

$$\frac{3x}{x+2} > 3 - \varepsilon \Leftrightarrow x > \frac{6-2\varepsilon}{\varepsilon}.$$

Wir können dann $x := \frac{6-2\varepsilon}{\varepsilon} + 1$ wählen.

Infimum: Da $\frac{3x}{x+2} > 0$ für alle $x > 0$ gilt, ist 0 eine untere Schranke von M. Um zu zeigen, dass $\inf M = 0$ gilt, müssen wir beweisen:

Zu jedem $\varepsilon > 0$ gibt es ein $x > 0$ mit $\frac{3x}{x+2} < \varepsilon$.

Wenn $\varepsilon \geq 3$ ist, dann gilt die Ungleichung für jedes $x > 0$.

Sei also $0 < \varepsilon < 3$. Dann ist

$$\frac{3x}{x+2} < \varepsilon \Leftrightarrow x < \tfrac{2\varepsilon}{3-\varepsilon}.$$

Da $\frac{2\varepsilon}{3-\varepsilon} > 0$ ist, können wir $x := \frac{\varepsilon}{3-\varepsilon}$ wählen.

Antworten zu den Verständnisfragen 6.2 – Konvergenz von Folgen

1. Es macht die Aufgabe überhaupt erst zugänglich. Das Problem in der ursprünglichen Darstellung war ja gerade, dass mit jedem Folgenglied die Anzahl der Summanden wächst. Durch die Summenbildung hat man nur mehr einen einzigen Ausdruck zu untersuchen, was viel einfacher ist.
2. In diesem Fall würde die Folge gegen $+\infty$ oder $-\infty$ divergieren.
3. Es reicht nicht, den Ausdruck $\frac{2+\frac{1}{\sqrt{n}}}{\sqrt{n}+5^{-n}}$ nur nach oben gegen eine Nullfolge abzuschätzen, wenn man das Einschließungskriterium verwenden möchte. Man muss ihn auch gegen eine Nullfolge *nach unten* abschätzen. Nachdem der Ausdruck für alle $n \in \mathbb{N}$ positiv ist, gelingt das mit der primitiven Nullfolge, deren Folgenglieder alle gleich 0 sind. Die Zahl 0 ganz links in der Ungleichungskette kann man daher als konstante Nullfolge interpretieren.

Komprimierte Lösung zu Aufgabe 6.2 – Konvergenz von Folgen

a) $a_n = \sum_{k=1}^{n} \left(-\frac{1}{2}\right)^k = \frac{1-\left(-\frac{1}{2}\right)^{n+1}}{1-\left(-\frac{1}{2}\right)} - 1 = \frac{1-\left(-\frac{1}{2}\right)^{n+1}}{\frac{3}{2}} - 1 \overset{n\to\infty}{\longrightarrow} \frac{2}{3} - 1 = -\frac{1}{3}$, nach der geometrischen Summe.

b) $b_n = \frac{r_0 + r_1 n + \ldots + r_k n^k}{s_0 + s_1 n + \ldots + s_k n^k} = \frac{\frac{r_0}{n^k} + \frac{r_1}{n^{k-1}} + \cdots + \frac{r_{k-1}}{n} + r_k}{\frac{s_0}{n^k} + \frac{s_1}{n^{k-1}} + \cdots + \frac{s_{k-1}}{n} + s_k} \overset{n\to\infty}{\longrightarrow} \frac{r_k}{s_k}$

c) $0 \leq c_n = \frac{2+\frac{1}{\sqrt{n}}}{\sqrt{n}+5^{-n}} = \frac{2+\frac{1}{\sqrt{n}}}{\sqrt{n}+\frac{1}{5^n}} \leq \frac{2+\frac{1}{\sqrt{n}}}{\sqrt{n}} = \frac{2\sqrt{n}+1}{n} = \frac{2}{\sqrt{n}} + \frac{1}{n} \overset{n\to\infty}{\longrightarrow} 0$. Aus dem Einschließungskriterium folgt $\lim_{n\to\infty} c_n = 0$.

Antworten zu den Verständnisfragen 6.3 – Cauchyfolgen

1. Das Minus zwischen den Brüchen wird einfach in den Zähler des zweiten Bruches hineingezogen. Aus $(-1)^m$ wird dann $-(-1)^m = (-1)^{m+1}$. Wir möchten gerne die Dreiecksungleichung verwenden, die in ihrer üblichen Form ein Plus zwischen den beiden Termen stehen hat:

$$|x+y| \leq |x| + |y|.$$

2. Weil das Ersetzen von m^2 durch n^2 im Nenner zu einer *Verkleinerung* des Ausdrucks führen würde – wir würden uns dadurch also unsere gewünschte Abschätzung nach oben zerstören.

3. Ja, eigentlich schon. n_0 selbst muss ja nicht größer als $\sqrt{\frac{2}{\varepsilon}}$ sein. Nachdem n und m aber echt größer als n_0 sind, wären sie bei Wahl von $n_0 := \left[\sqrt{\frac{2}{\varepsilon}}\right]$ mindestens $\left[\sqrt{\frac{2}{\varepsilon}}\right] + 1$ (n und m sind ja natürliche Zahlen) und damit insbesondere größer als $\sqrt{\frac{2}{\varepsilon}}$. Und genau das ist ja die Bedingung an n und m.

Lösungsvorschlag zu Übung 6.3 – Cauchyfolgen

Von oben nach unten: B P H K H P K H P Z H P H T B

Komprimierte Lösung zu Aufgabe 6.3 – Cauchyfolgen

Beweis: Sei $\varepsilon > 0$ und sei o.B.d.A. $n \geq m$. Wähle $n_0 := \left[\sqrt{\frac{2}{\varepsilon}}\right] + 1$. Für $n, m > n_0$ gilt dann:

$$\begin{aligned} |x_n - x_m| &= \left|\frac{(-1)^n}{n^2} - \frac{(-1)^m}{m^2}\right| = \left|\frac{(-1)^n}{n^2} + \frac{(-1)^{m+1}}{m^2}\right| \leq \left|\frac{(-1)^n}{n^2}\right| + \left|\frac{(-1)^{m+1}}{m^2}\right| \\ &= \frac{1}{n^2} + \frac{1}{m^2} = \frac{m^2+n^2}{n^2m^2} \leq \frac{2n^2}{n^2m^2} = \frac{2}{m^2} < \varepsilon. \end{aligned}$$

Antworten zu den Verständnisfragen 6.4 – Konvergenz von Reihen

1. Selbstverständlich nicht. Es reicht, eines der Kriterien zu verwenden, sofern es zu einem Ergebnis führt. Wir haben in dieser Aufgabe nur deswegen mehrere Kriterien angewandt, um Ihnen deren Gebrauch zu verdeutlichen.

2. Vorsicht: Man darf hier nicht einfach argumentieren, dass $\frac{2n^2+3n+1}{2n^2+3n}$ gegen 1 konvergiert und damit auch der Ausdruck $\left(\frac{2n^2+3n+1}{2n^2+3n}\right)^n$ ebenfalls gegen 1. Es gilt nämlich z. B. $\lim\limits_{n\to\infty}\left(\frac{n+1}{n}\right)^n = e$, obwohl $\lim\limits_{n\to\infty}\frac{n+1}{n} = 1$.

   Der Grenzwert von $\left(\frac{2n^2+3n+1}{2n^2+3n}\right)^n \cdot \frac{n+1}{2n+3}$ mit $n \to \infty$ ist übrigens $\frac{1}{2}$, was aber eben nicht einfach nachzuweisen ist.

3. Nein, z. B. ist die Folge $a_n = 2 + \frac{1}{n}$ monoton fallend, ihr Grenzwert ist aber 2. Man muss beim Leibniz-Kriterium also wirklich alle drei Voraussetzungen (Nullfolge, alternierend und monoton fallend) prüfen, keine folgt aus einer der beiden anderen!

Komprimierte Lösung zu Aufgabe 6.4 – Konvergenz von Reihen

a) $\sqrt[n]{\left|\left(\frac{n}{2n+1}\right)^n\right|} = \frac{n}{2n+1} = \frac{1}{2+\frac{1}{n}}$ und $\lim\limits_{n\to\infty}\frac{1}{2+\frac{1}{n}} = \frac{1}{2}$. Nach dem Wurzelkriterium konvergiert die Reihe absolut. Insbesondere konvergiert sie auch.

b) Es sei $a_n := \frac{(-1)^n}{n-\sqrt{n}}$. Klar: a_n ist eine alternierende Folge.

Wir zeigen: $|a_{n+1}| \leq |a_n|$ für alle $n \geq 2$.

$$\begin{array}{rccc} & |a_{n+1}| & \leq & |a_n| \\ \Leftrightarrow & \frac{1}{n+1-\sqrt{n+1}} & \leq & \frac{1}{n-\sqrt{n}} \\ \Leftrightarrow & n-\sqrt{n} & \leq & n+1-\sqrt{n+1} \\ \Leftrightarrow & -\sqrt{n} & \leq & 1-\sqrt{n+1} \\ \Leftrightarrow & \sqrt{n+1} & \leq & 1+\sqrt{n} \\ \overset{n\in\mathbb{N}}{\Leftrightarrow} & n+1 & \leq & n+1+2\sqrt{n} \\ \Leftrightarrow & 0 & \leq & \sqrt{n} \end{array}$$

$|a_n|$ ist also monoton fallend.

Außerdem gilt: $\lim\limits_{n\to\infty} |a_n| = \lim\limits_{n\to\infty} \dfrac{1}{n-\sqrt{n}} = \lim\limits_{n\to\infty} \dfrac{\frac{1}{n}}{1-\frac{1}{\sqrt{n}}} = 0$, $|a_n|$ ist demnach eine Nullfolge. Nach dem Leibniz-Kriterium konvergiert die Reihe.

Es gilt: $|a_n| = \frac{1}{n-\sqrt{n}} \geq \frac{1}{n}$. Die harmonische Reihe ist also eine divergierende Minorante für die Reihe $\sum\limits_{n=2}^{\infty} \left| \dfrac{(-1)^n}{n-\sqrt{n}} \right|$. D. h., $\sum\limits_{n=2}^{\infty} \dfrac{(-1)^n}{n-\sqrt{n}}$ konvergiert nicht absolut.

Antworten zu den Verständnisfragen 6.5 – Folgenstetigkeit

1. Nachdem die Funktion f auf ganz $\mathbb{R}$ definiert ist, ist es nicht selbstverständlich, dass die einzige ganze Zahl x_0, in der die Funktion f stetig ist, gerade in dem gezeichneten Bildausschnitt liegt.
2. Sie wird in der Gleichungskette
$$\begin{aligned} \lim_{n\to\infty} f(x_0 - \frac{1}{n}) &= \lim_{n\to\infty} \left[x_0 - \frac{1}{n}\right] (1-(x_0-\frac{1}{n})) \\ &= \lim_{n\to\infty} (x_0-1)(1-x_0+\frac{1}{n}) \\ &= (x_0-1)(1-x_0) \end{aligned}$$
im zweiten Gleichheitszeichen verwendet. Sie gilt, weil die Zahl $\frac{1}{n}$ für $n \in \mathbb{N}$ immer im Intervall $(0,1]$ liegt. Damit liegt $x_0 - \frac{1}{n}$ im Intervall $[x_0-1, x_0)$. Da x_0 eine ganze Zahl ist, haben alle Zahlen in diesem Intervall als nächstkleinere ganze Zahl die Zahl $x_0 - 1$. Es gilt also $\left[x_0 - \frac{1}{n}\right] = x_0 - 1$.
3. Aus $[x_n] \in \{0,1\}$ folgt, dass $[x_n] \leq 1$ gilt. Diese Ungleichung verwenden wir beim Übergang vom dritten zum vierten Ausdruck in der darauffolgenden Ungleichungskette.

Komprimierte Lösung zu Aufgabe 6.5 – Folgenstetigkeit

Sei $x_0 \notin \mathbb{Z}$. Da sowohl $x \mapsto [x]$ als auch $x \mapsto (1-x)$ in x_0 stetig ist, ist auch f als Produkt stetiger Funktionen stetig in x_0.

Sei $x_0 \in \mathbb{Z}$. Betrachten wir die Folge $x_n := x_0 - \frac{1}{n}$, dann ist:

$$\begin{aligned}\lim_{n\to\infty} f(x_0 - \frac{1}{n}) &= \lim_{n\to\infty} \left[x_0 - \frac{1}{n}\right](1-(x_0 - \frac{1}{n})) = \lim_{n\to\infty}(x_0-1)(1-x_0+\frac{1}{n}) \\ &= (x_0-1)(1-x_0).\end{aligned}$$

Außerdem gilt:

$$f(\lim_{n\to\infty}(x_0 - \frac{1}{n})) = f(x_0) = [x_0](1-x_0) = x_0(1-x_0).$$

Es gilt:

$$(x_0-1)(1-x_0) = x_0(1-x_0) \Leftrightarrow x_0 = 1.$$

f ist also jedenfalls nicht stetig in $x_0 \in \mathbb{Z}$, sofern $x_0 \neq 1$.

Sei nun $x_0 = 1$. Sei x_n eine beliebige Folge, die gegen 1 konvergiert. Dann gibt es ein n_0, sodass für alle $n \geq n_0$ gilt:

$$|[x_n](1-x_n)| = |[x_n]||1-x_n| \leq |1-x_n| \overset{n\to\infty}{\longrightarrow} 0.$$

Also ist $\lim\limits_{n\to\infty}[x_n](1-x_n) = 0 = [1]\cdot(1-1) = f(1)$ und daher ist f stetig in $x_0 = 1$.

Antworten zu den Verständnisfragen 6.6 – Stetigkeit mit Epsilon und Delta

1. Wir möchten x loswerden, indem wir es gegen eine Konstante nach oben abschätzen. x_0 und δ sind beide konstant! Außerdem kommen alle drei genannten Variablen in der Ungleichung $|x - x_0| < \delta$ vor. Einen anderen Zusammenhang zwischen den drei Variablen haben wir ohnehin nicht zur Verfügung, deshalb liegt es nahe, mit dieser Ungleichung zu starten.

2. Wir starten mit der Ungleichung $|x - x_0| < \delta$. Wollten wir jetzt die Dreiecksungleichung anwenden, dann würden wir schreiben: $|x - x_0| = |x + (-x_0)| \leq |x| + |-x_0| = |x| + |x_0|$. Jetzt wissen wir, dass $|x - x_0|$ sowohl kleiner als δ als auch kleiner gleich $|x| + |x_0|$ sein muss. Leider können wir mit diesen Informationen aber keinen Vergleich zwischen δ und $|x| + |x_0|$ anstellen. Einen solchen streben wir allerdings an. Das ist der Grund, warum hier die Dreiecksungleichung keinen Nutzen bringt.

3. Bei Stetigkeitsbeweisen geht es gerade darum, zu vorgegebenem $\varepsilon > 0$ ein passendes $\delta > 0$ zu finden. Wir dürfen also selbst bestimmen, wie wir δ in Abhängigkeit von x_0 und ε wählen möchten (δ muss lediglich positiv sein). Dabei darf man also auch zwischendurch vorläufige Festlegungen machen. Wenn man also – wie in unserem Fall – benötigt, dass δ nicht größer als 1 sein soll, so kann man das schon einmal als *eine* Bedingung an δ festhalten. Es könnte auch passieren, dass man am Weg zum Endergebnis noch mehrere Bedingungen an δ stellen muss, z. B. $\delta \leq \frac{1}{2}$ oder $\delta \leq \frac{\varepsilon}{2|x_0|}$. Man darf dann nur am Ende nicht vergessen, *alle* diese Bedingungen bei der endgültigen Wahl von δ zu berücksichtigen! Das macht man meist dadurch, dass man δ als das Minimum aller oberen Schranken für δ wählt, so wie wir das in unserer Musterlösung auch gemacht haben.

4. Wir haben im Laufe des Beweises zwei Bedingungen an δ geknüpft. Zunächst haben wir festgelegt, dass $\delta \leq 1$ sein soll, und dann haben wir noch $\delta \leq \frac{\varepsilon}{\frac{1+2|x_0|}{x_0^2+1}}$ gefordert. Jetzt weiß man zunächst nicht, welche der beiden Bedingungen stärker ist (d. h., welche der beiden Zahlen 1 oder $\frac{\varepsilon}{\frac{1+2|x_0|}{x_0^2+1}}$ kleiner ist).

   Damit δ beide Bedingungen gleichzeitig erfüllt, wählt man das Minimum der beiden Zahlen. Ist dann z. B. 1 kleiner oder gleich $\frac{\varepsilon}{\frac{1+2|x_0|}{x_0^2+1}}$, so gilt $\delta \leq 1$ und erst recht $\delta \leq \frac{\varepsilon}{\frac{1+2|x_0|}{x_0^2+1}}$. Umgekehrt natürlich analog.

Lösungsvorschlag zu Übung 6.6 – Stetigkeit mit Epsilon und Delta

Von oben nach unten: P B H K   A P K H   B H K Z   H B A H

Komprimierte Lösung zu Aufgabe 6.6 – Stetigkeit mit Epsilon und Delta

Seien $x_0 \in \mathbb{R}$ und $\varepsilon > 0$. Wähle $0 < \delta < \min\{1, \varepsilon \cdot \frac{x_0^2+1}{1+2|x_0|}\}$. Dann ist für alle x mit $|x - x_0| < \delta$ auch $|f(x) - f(x_0)| < \varepsilon$, denn

$$\left| \frac{1}{x^2+1} - \frac{1}{x_0^2+1} \right| = \left| \frac{x_0^2+1-(x^2+1)}{(x^2+1)(x_0^2+1)} \right| \leq \frac{|x_0 - x| \cdot |x_0 + x|}{x_0^2+1} = ...$$

Da $|x - x_0| < \delta$, ist

$$... \leq \frac{|x - x_0| \cdot (|x_0| + |x|)}{x_0^2+1} \leq \frac{\delta \cdot (\delta + 2|x_0|)}{x_0^2+1} = ...$$

Nun ist zugleich $\delta < 1$, also $\delta^2 < \delta$, und daher

$$... = \frac{\delta^2 + 2\delta|x_0|}{x_0^2+1} < \frac{\delta(1+2|x_0|)}{x_0^2+1} = ...$$

Nach Wahl von δ gilt schließlich

$$... < \varepsilon.$$

Antworten zu den Verständnisfragen 6.7 – Gleichmäßig und Lipschitz-stetig

1. Die heikelste Stelle für die gleichmäßige Stetigkeit ist die Stelle $x = 0$. Dort ist die Funktion f nämlich am steilsten (siehe Abbildung 11.2).

2. Zunächst wird $|\sqrt{x} - \sqrt{y}|$ als Bruch mit Nenner 1 aufgefasst und mit dem Ausdruck $|\sqrt{x} + \sqrt{y}|$ erweitert:

$$|\sqrt{x} - \sqrt{y}| = \frac{|\sqrt{x} - \sqrt{y}|}{1} = \frac{|\sqrt{x} - \sqrt{y}| \cdot |\sqrt{x} + \sqrt{y}|}{|\sqrt{x} + \sqrt{y}|}.$$

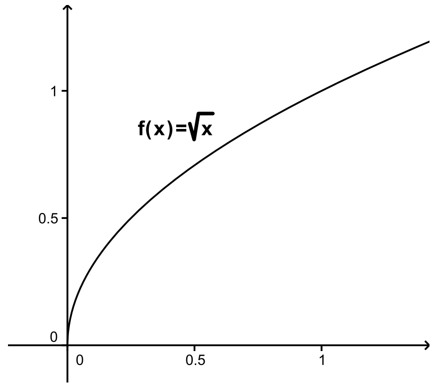

Abbildung 11.2: Graphische Darstellung der Funktion f

Im Zähler können wir die Rechenregel $|a| \cdot |b| = |ab|$ anwenden. Im Nenner stellen wir fest, dass $\sqrt{x} + \sqrt{y}$ positiv ist. Wir können also die Betragsstriche weglassen:

$$\frac{|\sqrt{x} - \sqrt{y}| \cdot |\sqrt{x} + \sqrt{y}|}{|\sqrt{x} + \sqrt{y}|} = \frac{|(\sqrt{x} - \sqrt{y})(\sqrt{x} + \sqrt{y})|}{\sqrt{x} + \sqrt{y}}.$$

Im Zähler wenden wir jetzt die dritte binomische Formel $(a+b)(a-b) = a^2 - b^2$ an und erhalten:

$$\frac{|(\sqrt{x} - \sqrt{y})(\sqrt{x} + \sqrt{y})|}{\sqrt{x} + \sqrt{y}} = \frac{|x-y|}{\sqrt{x} + \sqrt{y}}.$$

3. Nein. Es reicht, für jede vorgegebene Zahl L *ein* Wertepaar (x, y) zu finden, für das $\frac{1}{\sqrt{x}+\sqrt{y}} > L$ gilt. Und das kann eben auch ein Wertepaar der Form $(x, y) = (x, 0)$ sein.

Komprimierte Lösung zu Aufgabe 6.7 – Gleichmäßig und Lipschitz-stetig

a) Da f sogar auf $[0, 1]$ definiert und stetig ist, ist f auf $[0, 1]$ gleichmäßig stetig. Stetige Funktionen sind nämlich auf kompakten Intervallen gleichmäßig stetig. Insbesondere ist f auf $[0, 1)$ gleichmäßig stetig.

b) $|\sqrt{x} - \sqrt{y}| = \left|\frac{x-y}{\sqrt{x}+\sqrt{y}}\right| = |x-y| \cdot \frac{1}{|\sqrt{x}+\sqrt{y}|}$ für alle x, $y \in [0, 1)$ mit $(x, y) \neq (0, 0)$. Wählen wir $x \in (0, 1)$ beliebig und $y = 0$, so ist $|\sqrt{x}| = |x| \cdot \frac{1}{\sqrt{x}}$. Es gilt $\lim\limits_{x \to 0} \frac{1}{\sqrt{x}} = +\infty$. Somit existiert kein $L > 0$ mit $|\sqrt{x}| \leq |x| \cdot L$ für alle $x \in (0, 1)$ und f ist nicht Lipschitz-stetig.

Antworten zu den Verständnisfragen 6.8 – Differenzierbarkeit

1. In der Version
$$\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}$$
des Differentialquotienten ist x eine variable Stelle, während x_0 die feste Stelle ist, an der wir die Differenzierbarkeit nachweisen möchten. Man kann jetzt alternativ die Stelle x auch durch ihren Abstand zur festen Stelle x_0 angeben. Nennen wir diesen Abstand h, so gilt $x = x_0 + h$. (Dabei kann h natürlich auch negativ sein.) Wir haben sozusagen die Variabilität von x einfach auf eine neue Variable h übertragen, die den Abstand zwischen x und x_0 misst. Anstatt x gegen x_0 laufen zu lassen, lassen wir jetzt einfach den Abstand der beiden, also h, gegen Null laufen.

2. Würde man den Grenzwert gleich zu Beginn beim Ausdruck
$$\frac{(x_0+h)^2-x_0^2}{h}$$
auswerten, würde man sowohl im Zähler als auch im Nenner Null erhalten. Die Regeln von de l'Hospital können wir hier nicht anwenden, weil wir dazu wissen müssten, was die Ableitung von $(x_0+h)^2$ ist – dazu bräuchten wir aber schon das Ergebnis aus Aufgabenteil a), an dem wir aber gerade arbeiten. Durch Termumformungen lässt sich allerdings der Nenner wegkürzen und wir können das Problem umgehen.

3. Je nachdem, ob h positiv oder negativ ist, müssen wir den Wert von $g(0+h)$ mit Hilfe der oberen bzw. der unteren Zeile der Definition der Funktion g berechnen. Daraus ergibt sich die Fallunterscheidung. Falls $h > 0$ gilt, liefert $g(0+h) = g(h) = h$ und damit $\frac{g(0+h)-g(0)}{h} = \frac{h-0}{h} = 1$. Ist $h \leq 0$, so gilt $g(0+h) = g(h) = h^2$ und $\frac{g(0+h)-g(0)}{h} = \frac{h^2-0}{h} = h$.

4. $\lim\limits_{h\to 0}\dfrac{g(h)-g(0)}{h}$ ist ein Funktionengrenzwert (die Variable h nimmt ihre Werte kontinuierlich in $\mathbb{R}$ an), die beiden Grenzwerte
$$\lim_{n\to\infty}\frac{g(h_n)-g(0)}{h_n}$$
bzw.
$$\lim_{n\to\infty}\frac{g(\tilde{h}_n)-g(0)}{\tilde{h}_n}$$
sind Folgengrenzwerte (die Variable n nimmt Werte in $\mathbb{N}$ an).
Wie in der Lösung zu Teil a) dargestellt, existiert ein Funktionengrenzwert
$$\lim_{x\to x_0} g(x) = a$$
nur dann, wenn
$$\lim_{n\to\infty} g(x_n) = a$$

für alle Folgen $\{x_n\}_{n\in\mathbb{N}}$ mit $\lim\limits_{n\to\infty} x_n = x_0$ gilt. Die Definition des Funktionengrenzwertes greift also auf Folgengrenzwerte zurück. Das kann man nun ausnutzen, wenn man zeigen will, dass der Funktionengrenzwert nicht existiert. Dazu reicht es zu zeigen, dass die Grenzwerte der Folgen der Funktionswerte entlang zweier unterschiedlicher Folgen von Argumenten auch unterschiedliche Werte annehmen.

5. Nachdem die Funktion $x \mapsto |x|$ selbst nicht differenzierbar an der Stelle $x_0 = 0$ ist, liegt die Vermutung nahe, dass gerade diese Stelle Schwierigkeiten bereiten könnte. Im vorliegenden Fall bestätigt sich diese Vermutung allerdings nicht.
6. Eigentlich denkt man auch hier wieder in Folgengrenzwerten. Allerdings ist offensichtlich, dass $|h|$ entlang jeder Nullfolge h_n gegen Null läuft, sodass wir es nicht extra erwähnt haben.
7. Selbstverständlich nicht! Es reicht, wenn der Differentialquotient überhaupt existiert. Welchen Wert er annimmt, ist für die Differenzierbarkeit egal!

Komprimierte Lösung zu Aufgabe 6.8 – Differenzierbarkeit

a) $\lim\limits_{h\to 0} \dfrac{(x_0+h)^2 - x_0^2}{h} = \lim\limits_{h\to 0} \dfrac{2x_0h + h^2}{h} = 2x_0$. Daraus folgt, dass f differenzierbar in jedem $x_0 \in \mathbb{R}$ ist.

b) Für alle $x_0 \neq 0$ ist g offensichtlich differenzierbar. Sei nun $x_0 = 0$.

$$\frac{g(h) - g(0)}{h} = \begin{cases} 1 & h > 0 \\ h & h < 0 \end{cases}$$

Also ist

$$\lim_{h\searrow 0} \frac{g(h) - g(0)}{h} = 1 \neq 0 = \lim_{h\nearrow 0} \frac{g(h) - g(0)}{h}$$

und damit g nicht differenzierbar in $x_0 = 0$.

c) In $x_0 \neq 0$ ist l offensichtlich differenzierbar. Sei nun $x_0 = 0$.

$$\frac{l(h) - l(0)}{h} = \frac{h \cdot |h|}{h} = |h| \xrightarrow{h\to 0} 0$$

l ist also auch in $x_0 = 0$ differenzierbar.

Antworten zu den Verständnisfragen 6.9 – Taylorpolynom

1. Nachdem wir das Taylorpolynom zur Funktion $f(x) = \sqrt[4]{1+x}$ berechnet haben und $f(\frac{1}{2}) = \sqrt[4]{1.5}$ ist, muss man auch das zugehörige Taylorpolynom an der Stelle $\frac{1}{2}$ auswerten.
2. Er verwendet ein Taylorpolynom höheren Grades als 2, dadurch wird im Allgemeinen (allerdings nicht immer) auch der Fehler kleiner.
3. Das Restglied lässt sich leider nicht konkret berechnen, weil man vom Parameter θ lediglich weiß, dass er zwischen 0 und 1 liegt. Man muss also versuchen,

den Ausdruck mit θ irgendwie nach oben abzuschätzen. Das haben auch wir so gemacht. Dementsprechend erhält man für den Fehler meist nur eine obere Schranke und keinen genauen Wert.

Komprimierte Lösung zu Aufgabe 6.9 – Taylorpolynom

$$\begin{aligned} f(x) &= (1+x)^{\frac{1}{4}} \text{ und damit } f(0)=1. \\ f'(x) &= \frac{1}{4}\cdot(1+x)^{-\frac{3}{4}} \text{ und damit } f'(0)=\frac{1}{4}. \\ f''(x) &= -\frac{3}{16}\cdot(1+x)^{-\frac{7}{4}} \text{ und damit } f''(0)=-\frac{3}{16}. \end{aligned}$$

Damit:

$$T_{2,f,0}(h) = f(0) + \frac{f'(0)}{1!}h + \frac{f''(0)}{2!}h^2 = 1 + \frac{1}{4}h - \frac{3}{32}h^2.$$

$\sqrt[4]{1.5} \approx T_{2,f,0}\left(\frac{1}{2}\right) = \frac{141}{128}$. Laut Taylorformel gibt es eine Zahl $0<\theta<1$ mit

$$\sqrt[4]{1.5} = \frac{141}{128} + \frac{f^{(3)}(0+\frac{1}{2}\theta)}{3!}\left(\frac{1}{2}\right)^3.$$

Damit ist

$$\left|\sqrt[4]{1.5} - \frac{141}{128}\right| = \left|\frac{f^{(3)}(\frac{1}{2}\theta)}{3!}\left(\frac{1}{2}\right)^3\right| = \frac{7}{1024}\cdot\left(1+\frac{\theta}{2}\right)^{-\frac{11}{4}} < \frac{7}{1024},$$

da $(1+\frac{\theta}{2})^{-\frac{11}{4}} < 1$ gilt.

Antworten zu den Verständnisfragen 6.10 – Funktionenreihen

1. Wir zeigen zunächst $\dfrac{|x|}{x^2+1} < 1$ für $x \in \mathbb{R}$. Sei zunächst $x \geq 0$. Zu zeigen ist dann: $\frac{x}{x^2+1} < 1$ bzw. nach Umformung $x^2 - x + 1 > 0$. Es gilt

   $$x^2 - x + 1 \geq x^2 - 2x + 1 = (x-1)^2 > 0$$

   für $x \neq 1$. Für $x = 1$ gilt aber die ursprüngliche Ungleichung, da $\frac{|1|}{1^2+1} = \frac{1}{2} < 1$ ist.

   Sei nun $x < 0$. Dann ist zu zeigen: $\frac{-x}{x^2+1} < 1$ oder nach Umformung $x^2 + x + 1 > 0$. Es gilt $x^2 + x + 1 \overset{x<0}{>} x^2 + 2x + 1 = (x+1)^2 \geq 0$. Damit ist die erste Ungleichung gezeigt.

   Die zweite Ungleichung lässt sich sehr einfach beweisen: $\frac{n}{n^3+1} < \frac{1}{n^2} \Leftrightarrow n^3 < n^3 + 1$, was trivialerweise erfüllt ist.

2. Nachdem die ursprüngliche Variable x Werte in ganz $\mathbb{R}$ annehmen durfte, nimmt $|x|$ Werte in $[0,\infty)$ an. Weil wir bei der Definition von h indirekt auch die Variable $t := |x|$ definiert haben, gilt dieses Intervall also auch für t.

3. Es kann passieren, dass eine Funktion am „Rand ihres Definitionsbereiches“ noch größere oder kleinere Werte annimmt, als an den relativen Extremstellen – auch wenn Sie das vielleicht aus üblichen Schulbuchaufgaben nicht gewöhnt sind. Zum Beispiel nimmt die Funktion $p\colon [0,10] \to \mathbb{R}$ mit $p(x) = 2x^3 - 15x^2 + 36x$ ihr Maximum an der Stelle $x = 10$ und ihr Minimum an der Stelle $x = 0$ an, obwohl sie an der Stelle $x = 2$ ein relatives Maximum und an der Stelle $x = 3$ ein relatives Minimum besitzt, wie Sie leicht nachrechnen können.

4. Weil der Definitionsbereich nach rechts eigentlich unbeschränkt ist. Während man den linken Rand $t = 0$ einfach durch Einsetzen in $h(t)$ untersuchen kann, klappt das am rechten „Rand“ nicht. Hier muss man stattdessen den Funktionengrenzwert für $t \to \infty$ berechnen.

Lösungsvorschlag zu Übung 6.10 – Funktionenreihen

Von oben nach unten: K B A P K H P K H B H A P Z K H P H B H

Komprimierte Lösung zu Aufgabe 6.10 – Funktionenreihen

a) Sei $f_n(x) := \dfrac{\ln(1 + |nx|)}{(n^3+1)(x^2+1)}$. Da $\ln(1+t) \leq t$ für alle $t \geq 0$ gilt, erhalten wir:

$$|f_n(x)| \leq \frac{|nx|}{(n^3+1)(x^2+1)} \leq \frac{n}{n^3+1} \leq \frac{1}{n^2}.$$

Also ist: $||f_n||_\infty = \sup\limits_{x \in \mathbb{R}} |f_n(x)| \leq \dfrac{1}{n^2}$. Da die Reihe $\sum\limits_{n=1}^{\infty} \dfrac{1}{n^2}$ konvergiert, konvergiert nach dem Weierstraß-Kriterium die Reihe $\sum\limits_{n=1}^{\infty} f_n(x)$ gleichmäßig auf $\mathbb{R}$.

b) Sei $g_n(x) := \dfrac{x}{e^{|x|n} \cdot n}$. Wir berechnen $||g_n||_\infty = \sup\limits_{x \in \mathbb{R}} |g_n(x)|$.

$|g_n(x)| = \dfrac{1}{n}|x| \cdot e^{-|x|n}$, also müssen wir das Supremum der Funktion $h(t) := \dfrac{1}{n} \cdot t \cdot e^{-tn}$ auf $[0, +\infty)$ finden.

$$h(0) = 0,$$

$$\lim_{t \to +\infty} h(t) = 0,$$

$$h'(t) = \frac{1}{n} \cdot (t \cdot e^{-tn} \cdot (-n) + e^{-tn}).$$

$h'(t) = 0 \Leftrightarrow -tn + 1 = 0 \Leftrightarrow t = \frac{1}{n}$. Damit ist

$$||g_n||_\infty = \sup_{t \in [0,\infty)} \frac{1}{n} \cdot te^{-tn} = \frac{1}{n^2} \cdot e^{-\frac{1}{n} \cdot n} = \frac{1}{en^2}.$$

Da $\sum\limits_{n=1}^{\infty} \dfrac{1}{en^2}$ konvergiert, konvergiert die Reihe $\sum\limits_{n=1}^{\infty} g_n(x)$ nach dem Weierstraß-Kriterium gleichmäßig auf $\mathbb{R}$.

11.3 Lösungen zu Kapitel 7

Antworten zu den Verständnisfragen 7.1 – Funktionengrenzwerte

1. Dann hätte man Pech gehabt. Es ist in diesem Fall nämlich noch keine Aussage darüber möglich, ob der Funktionengrenzwert existiert. Man müsste dann weiter nach einem Folgenpaar suchen, entlang dem der Grenzwert der entsprechenden Folge der Funktionswerte einen anderen Wert als Null liefert. Ist das nicht möglich, so muss man beweisen, dass der Grenzwert der Folge der Funktionswerte entlang jedes Nullfolgenpaares $\{(x_n, y_n)\}$ den Wert Null liefert (siehe Teil b)).

2. Bei der Substitution $x = r\cos t$ bzw. $y = r\sin t$ gibt die Variable r den Abstand des Punktes (x, y) vom Koordinatenursprung an, da

$$|(x,y)| = \sqrt{r^2\cos^2 t + r^2\sin^2 t} = r$$

ist. Die Variable t steht für den Winkel, den man zwischen x-Achse und dem Punkt (x, y) gegen den Uhrzeigersinn misst. Wenn man nun möchte, dass sich der Punkt (x, y) dem Ursprung $(0, 0)$ nähert, so kommt es dabei nicht auf den Winkel, sondern bloß auf die Verkleinerung des Abstandes r an.

3. Ja, es kommt immer das Ergebnis Null heraus. Hier zwei Beispiele:

$$\lim_{n\to\infty} h(\frac{1}{n}, 0) = \lim_{n\to\infty} \frac{1}{n^2}\cdot 0^2 \cdot \ln(\frac{1}{n^2} + 0^2) = \lim_{n\to\infty} 0 = 0,$$

$$\begin{aligned} \lim_{n\to\infty} h(\frac{1}{n}, \frac{1}{n}) &= \lim_{n\to\infty} \frac{1}{n^2}\cdot\frac{1}{n^2}\cdot\ln(\frac{1}{n^2}+\frac{1}{n^2}) = \lim_{n\to\infty}\frac{1}{n^4}\cdot\ln(\frac{2}{n^2}) \\ &= \frac{-\infty}{\infty} \overset{\text{de l'H.}}{=} \lim_{n\to\infty} \frac{\frac{n^2}{2}\cdot\frac{-4}{n^3}}{4n^3} = \lim_{n\to\infty}\left(-\frac{1}{2n^4}\right) = 0. \end{aligned}$$

Komprimierte Lösung zu Aufgabe 7.1 – Funktionengrenzwerte

a) Falls f stetig in $(0, 0)$ wäre, dann wäre

$$f(0,0) = \lim_{n\to\infty} f(0, \frac{1}{n}) = \lim_{n\to\infty} 0 = 0$$

und

$$f(0,0) = \lim_{n\to\infty} f(\frac{1}{n}, \frac{1}{n}) = \lim_{n\to\infty} \frac{\frac{1}{n^2}}{\frac{1}{n^2}+\frac{1}{n^2}} = \frac{1}{2}.$$

Widerspruch.

b) $g(x,y) = (x^2+y^2)^{x^2y^2} = e^{x^2y^2\ln(x^2+y^2)}$. Setze $h(x,y) := x^2y^2\ln(x^2+y^2)$. Wir beweisen, dass $\lim\limits_{(x,y)\to(0,0)} h(x,y) = 0$ ist.

Seien $x = r\cos t$ und $y = r\sin t$. Dann gilt $(x,y)\to(0,0)$ genau dann, wenn $r \searrow 0$ gilt.

$$|h(x,y)| = |r^4\sin^2 t\cos^2 t\ln(r^2)| \le r^4|\ln(r^2)| = 2r^3|r\ln r| \xrightarrow{r\searrow 0} 0.$$

Setzen wir $g(0,0) := e^0 = 1$, so setzt dies die Funktion g stetig auf $\mathbb{R}^2$ fort.

Antworten zu den Verständnisfragen 7.2 – Integrationsmethoden

1. Man differenziert einfach das Ergebnis – dabei muss als Resultat wieder der Integrand herauskommen!
2. Ganz einfach: Weil es so funktioniert. Nachdem wir bei der Substitution den Ausdruck $\frac{du}{e^x}$ bekommen haben, wollen wir den Ausdruck e^x natürlich durch Kürzen loswerden. Nachdem er bei $\frac{du}{e^x}$ im Nenner steht, können wir ihn nur mit einem e^x im Zähler kürzen. Deswegen sollten wir also den Ausdruck e^x im Zähler vorerst stehenlassen.
3. Würde man $g' = \ln x$ setzen, so müssten wir erstens eine Stammfunktion von $\ln x$ kennen, um weitermachen zu können. ($x \ln x - x$ ist eine Stammfunktion, wie man durch einen kleinen Trick und partielle Integration feststellen kann.) Zweitens ist das dann verbleibende Integral viel komplizierter als das ursprüngliche Integral.

Lösungsvorschlag zu Übung 7.2 – Integrationsmethoden

Von oben nach unten: P B Z A H P K H P T P A H P P B K H P K A H

Komprimierte Lösung zu Aufgabe 7.2 – Integrationsmethoden

a) $\displaystyle\int \frac{\ln x}{x} dx = \left\{ \begin{matrix} u = \ln x \\ du = \frac{1}{x} dx \end{matrix} \right\} = \int u\, du = \frac{u^2}{2} + C = \frac{(\ln x)^2}{2} + C$, wobei $C \in \mathbb{R}$.

b) $$\int_1^2 \frac{e^x}{e^{2x} + 2e^x + 1} dx = \left\{ \begin{matrix} u = e^x \\ du = e^x dx \end{matrix} \right\} = \int_e^{e^2} \frac{du}{(u+1)^2} = \int_e^{e^2} (u+1)^{-2} du$$
$$= \left[-(u+1)^{-1} \right]_e^{e^2} = -\frac{1}{e^2+1} + \frac{1}{e+1}$$

c) $\displaystyle\int xe^x dx = \left\{ \begin{matrix} f = x & g = e^x \\ f' = 1 & g' = e^x \end{matrix} \right\} = xe^x - \int 1 \cdot e^x dx = xe^x - e^x + C$, wobei $C \in \mathbb{R}$.

d) $$\int \sqrt{x} \ln x = \left\{ \begin{matrix} f = \ln x & g = \frac{2}{3} x^{\frac{3}{2}} \\ f' = \frac{1}{x} & g' = \sqrt{x} \end{matrix} \right\} = \frac{2}{3} x^{\frac{3}{2}} \ln x - \int \frac{2}{3} x^{\frac{3}{2}} \frac{1}{x} dx$$
$$= \frac{2}{3} x^{\frac{3}{2}} \ln x - \frac{4}{9} x^{\frac{3}{2}} + C, \text{ wobei } C \in \mathbb{R}.$$

Antworten zu den Verständnisfragen 7.3 – Uneigentliche Integrale

1. Das Berechnen des Integrals
$$\int_0^\infty e^{-x} dx$$
ist der Induktionsanfang für $k = 0$. Wenn man nun als Induktionsvoraussetzung annimmt, dass
$$\int_0^\infty x^{k-1} e^{-x} dx = (k-1)!$$

gilt, dann entspricht dem Induktionsschritt die Herleitung von

$$\int_0^\infty x^k e^{-x} dx = k \cdot \int_0^\infty x^{k-1} e^{-x} dx,$$

die wir in der Lösung durchgeführt haben.

2. Die erste Gleichung gilt, weil der Logarithmus auf dem Integrationsintervall $(0, 1)$ nur negative Werte annimmt und das Auflösen des Betrags demnach ein Minus liefert.

   Das Integral $\int \ln x dx$ kann man durch geschickte partielle Integration lösen:

$$\begin{aligned}\int \ln x dx &= \int 1 \cdot \ln x dx = x \cdot \ln x - \int x \cdot \frac{1}{x} dx = x \ln x - \int 1 dx \\ &= x \ln x - x + C.\end{aligned}$$

3. Das Teilen bei 1 hat den Vorteil, dass man die involvierten Ausdrücke $(1 + x^3)$, e^{-x} sowie $\sqrt{x}$ an der Stelle 1 gut auswerten kann. Das wird bei den nachfolgenden Abschätzungen wichtig. Man hätte den Integrationsbereich allerdings auch an irgendeiner anderen Stelle $a \in (0, \infty)$ teilen können. Auch das hätte (mit etwas weniger eleganten Abschätzungen) zum Ziel geführt.
4. Die Abschätzung $1 + x^3 \leq 2$ haben wir deshalb nicht verwenden können, weil sie auf dem Intervall $(1, \infty)$ gar nicht stimmt! Die Abschätzung $e^{-x} \leq 1$ stimmt zwar auch auf $(0, \infty)$ (machen Sie eine Skizze, um das zu sehen!), jedoch hätte sie uns die Möglichkeit zerstört, Aufgabenteil a) zu verwenden. Außerdem wäre diese Abschätzung zu grob, da das uneigentliche Integral $\int_1^\infty \frac{1 + x^3}{\sqrt{x}} dx$ nicht existiert und wir daher keine Aussage über die Existenz des uneigentlichen Integrals $\int_1^\infty \frac{(1 + x^3) e^{-x}}{\sqrt{x}} dx$ machen könnten.

Komprimierte Lösung zu Aufgabe 7.3 – Uneigentliche Integrale

a) $\int_0^R x^k e^{-x} dx = \left[-x^k e^{-x}\right]_0^R + k \cdot \int_0^R x^{k-1} e^{-x} dx = -\frac{R^k}{e^R} + k \cdot \int_0^R x^{k-1} e^{-x} dx.$
Da $\lim_{R \to +\infty} \frac{R^k}{e^R} = 0$ gilt, existiert $\int_0^\infty x^k e^{-x} dx$ genau dann, wenn $\int_0^\infty x^{k-1} e^{-x} dx$ existiert. Wiederholung des Verfahrens liefert: $\int_0^\infty x^k e^{-x} dx$ existiert genau dann, wenn $\int_0^\infty x^0 e^{-x} dx = \int_0^\infty e^{-x} dx$ existiert. Es ist aber

$$\int_0^\infty e^{-x} dx = \lim_{R \to +\infty} \left[-e^{-x}\right]_0^R = \lim_{R \to +\infty} (1 - e^{-R}) = 1.$$

Daraus folgt, dass $\int_0^\infty x^k e^{-x} dx$ existiert und $\int_0^\infty x^k e^{-x} dx = k!$ gilt.

b) $\int_a^1 |\ln x| dx = -\int_a^1 \ln x dx = 1 + a \ln a - a \xrightarrow{a \searrow 0} 1$. Insbesondere existiert das uneigentliche Integral.

c) Da $|\sin x| \leq |x|$, erhalten wir $\left|\frac{\sin x \cdot \ln x}{x}\right| \leq |\ln x|$. Da $|\ln x|$ nach Teil b) auf $(0,1]$ integrierbar ist, ist auch $\left|\dfrac{\sin x \cdot \ln x}{x}\right|$ und damit auch $\dfrac{\sin x \cdot \ln x}{x} = -\left|\dfrac{\sin x \cdot \ln x}{x}\right|$ auf $(0,1]$ integrierbar.

d) $\displaystyle\int_0^\infty \frac{(1+x^3)e^{-x}}{\sqrt{x}}dx = \int_0^1 \frac{(1+x^3)e^{-x}}{\sqrt{x}}dx + \int_1^\infty \frac{(1+x^3)e^{-x}}{\sqrt{x}}dx$

$0 \leq \dfrac{(1+x^3)e^{-x}}{\sqrt{x}} \leq \dfrac{2 \cdot 1}{\sqrt{x}}$ für $x \in (0,1)$ und $\dfrac{1}{\sqrt{x}}$ ist auf $(0,1)$ integrierbar.

$0 \leq \dfrac{(1+x^3)e^{-x}}{\sqrt{x}} \leq (1+x^3)e^{-x}$ für $x \in (1,\infty)$ und $(1+x^3)e^{-x}$ ist auf $(0,\infty)$ integrierbar nach Teil a), also insbesondere auf $(1,\infty)$. Das uneigentliche Integral existiert demnach.

Antworten zu den Verständnisfragen 7.4 – Differenzierbarkeit von $f : \mathbb{R}^2 \to \mathbb{R}$

1. Nachdem für diese Punkte entweder $x = 0$ oder $y = 0$ gilt, liegen sie gerade auf den beiden Koordinatenachsen.
2. Der Punkt $(0,0)$ ist ein Spezialfall von Punkten der Form $(x_0,0)$. Für diese Punkte wird in der Lösung versucht, die partiellen Ableitungen auszurechnen. Die partielle Ableitung nach x ist sowieso für keinen der Punkte $(x_0,0)$ ein Problem, also insbesondere für $(0,0)$ nicht. Die partielle Ableitung nach y existiert nur – so wird festgestellt – falls $x_0 = 0$ ist, also wenn es sich um den Punkt $(0,0)$ handelt.
3. Wichtig ist in diesem Zusammenhang der Satz: „Wir *nehmen* nun *an*, dass f differenzierbar ist." *Wenn* wir nämlich annehmen, dass f differenzierbar ist, dann existiert eine entsprechende lineare Abbildung T. Diese setzt sich – falls die Annahme gerechtfertigt war – aus den partiellen Ableitungen von f im Punkt $(0,0)$ zusammen. Es hätte uns aber auch passieren können, dass diese Annahme zu einem Widerspruch führt.

Komprimierte Lösung zu Aufgabe 7.4 – Differenzierbarkeit von $f : \mathbb{R}^2 \to \mathbb{R}$

Da die Funktion g mit $g(t) := |t|$ in jedem Punkt $t \neq 0$ differenzierbar ist, ist f differenzierbar für alle (x,y) mit $xy \neq 0$.

Sei also $xy = 0$. Daraus folgt $x = 0$ oder $y = 0$. Sei nun o.B.d.A. $y = 0$.

$$\frac{\partial f}{\partial y}(x,0) = \lim_{t\to 0} \frac{f(x,t) - f(x,0)}{t} = \lim_{t\to 0} \frac{|xt|}{t} = \begin{cases} |x| & t \searrow 0 \\ -|x| & t \nearrow 0 \end{cases}$$

Also existiert $\frac{\partial f}{\partial y}(x,0)$ nur für $x = 0$. Daraus folgt, dass f in allen Punkten $(x,0)$ und $(0,x)$ mit $x \neq 0$ nicht differenzierbar ist.

$$a = (0,0) \quad h = (h_1,h_2) \quad \frac{\partial f}{\partial x}(0,0) = 0 = \frac{\partial f}{\partial y}(0,0)$$

$$\left|\frac{f(a+h)-f(a)-0\cdot h}{||h||}\right| = \frac{|h_1 h_2|}{\sqrt{h_1^2+h_2^2}} = |h_2|\cdot\frac{|h_1|}{\sqrt{h_1^2+h_2^2}} \leq |h_2| \overset{|h|\to 0}{\longrightarrow} 0$$

Daraus folgt, dass f in $(0,0)$ differenzierbar ist.

Antworten zu den Verständnisfragen 7.5 – Mehrdimensionale Kettenregel

1. $x:\mathbb{R}^2\to\mathbb{R}$, $y:\mathbb{R}^2\to\mathbb{R}$, $h:\mathbb{R}^2\to\mathbb{R}^2$ und $g:\mathbb{R}^2\to\mathbb{R}$.
2. Wir möchten $\frac{\partial^2 g}{\partial t\partial s}$ an der Stelle $(0,0)$ auswerten. Dazu müssen wir laut Gleichung (7.3) diverse partielle Ableitungen von f an der Stelle $(x(s,t),y(s,t))$ kennen. Es gilt aber $x(0,0)=0$ und $y(0,0)=1$, also können wir die in der Aufgabenstellung angegebenen Informationen über die partiellen Ableitungen von f an der Stelle $(0,1)$ nutzen.

Lösungsvorschlag zu Übung 7.5 – Mehrdimensionale Kettenregel

Von oben nach unten: P A B K H B H B K H B A H

Komprimierte Lösung zu Aufgabe 7.5 – Mehrdimensionale Kettenregel

$$\frac{\partial g}{\partial s}(s,t) = \frac{\partial f}{\partial x}(x,y)\cdot 1 + \frac{\partial f}{\partial y}(x,y)\cdot e^{s+t},$$

wobei $x=s+t$ und $y=e^{s+t}+t$ gilt. Daraus folgt:

$$\begin{aligned}\frac{\partial^2 g}{\partial t\partial s}(s,t) &= \frac{\partial^2 f}{\partial x^2}(x,y)\cdot 1 + \frac{\partial^2 f}{\partial y\partial x}(x,y)\cdot(e^{s+t}+1)\\ &+ \left(\frac{\partial^2 f}{\partial x\partial y}(x,y)\cdot 1 + \frac{\partial^2 f}{\partial y^2}(x,y)\cdot(e^{s+t}+1)\right)e^{s+t} + \frac{\partial f}{\partial y}(x,y)\cdot e^{s+t}.\end{aligned}$$

Somit:

$$\frac{\partial^2 g}{\partial t\partial s}(0,0) = 1\cdot 1 + 1\cdot 2 + 1\cdot 1 + 1\cdot 2 + 1\cdot 1 = 7.$$

Antworten zu den Verständnisfragen 7.6 – Jacobi- und Hesse-Matrix

1. Nein, das kann nicht sein, denn sonst würde man dadurch die Bedingung

   $$\lim_{h\to 0}\frac{r(h)}{||h||^2} = 0$$

   verletzen.
2. Der Eintrag b_{ii} in der Hesse-Matrix ist gerade das Doppelte des Koeffizienten von h_i^2. Der Eintrag $b_{ij}=b_{ji}$ in der Hesse-Matrix ist gleich dem Koeffizienten von $h_i h_j$.

3. Im Satz ist x ein Vektor mit n Komponenten, in der Aufgabe ist x die erste Komponente eines Vektors mit zwei Komponenten, also eine reelle Zahl. Es passiert sehr häufig, dass die Notationen von Aufgaben und Sätzen nicht zusammenpassen. Sie müssen in solchen Fällen erst einmal für sich klären, mit welchen Objekten Sie es jeweils zu tun haben und wie Sie diese Objekte miteinander in Verbindung bringen können.

Lösungsvorschlag zu Übung 7.6 – Jacobi- und Hesse-Matrix

Von oben nach unten: P A K A T H A B A H

Komprimierte Lösung zu Aufgabe 7.6 – Jacobi- und Hesse-Matrix

$$\begin{aligned} f(x+h_1, y+h_2) &= 2(x+h_1) + y + h_2 - (x+h_1)(y+h_2) + (x+h_1)^3 \\ &= f(x,y) + (2-y+3x^2)h_1 + (1-x)h_2 - h_1h_2 + 3xh_1^2 + h_1^3 \\ &= f(x,y) + \begin{pmatrix} 2-y+3x^2 & 1-x \end{pmatrix} \cdot \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \\ &+ \tfrac{1}{2} \cdot \begin{pmatrix} h_1 & h_2 \end{pmatrix} \cdot \begin{pmatrix} 6x & -1 \\ -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} + h_1^3 \end{aligned}$$

Außerdem ist: $\left| \dfrac{h_1^3}{h_1^2 + h_2^2} \right| \leq |h_1| \xrightarrow{h \to 0} 0$. Folglich: $D^2 f(x,y) = \begin{pmatrix} 6x & -1 \\ -1 & 0 \end{pmatrix}$.

Antworten zu den Verständnisfragen 7.7 – Lokale Extremstellen

1. Es liegt ein echt positiver Eigenwert und ein Eigenwert 0 vor. Dieses Szenario wird in keinem der Fälle i) bis iii) des Eigenwertkriteriums beschrieben. Deshalb ist hier keine Aussage möglich!

2. Definiere $g(x) := x^2 - x^3$. Es gilt $g'(x) = 2x - 3x^2 = 0 \Leftrightarrow x = 0$ oder $x = \frac{2}{3}$. Die Funktion g ist auf dem Intervall $[0, \frac{2}{3}]$ monoton wachsend, überall sonst monoton fallend. Ausgehend von $x = 0$ wächst sie also – wenn man nach rechts wandert – bis zum Wert $x = \frac{2}{3}$. Wandert man von $x = 0$ ausgehend nach links, so wächst sie sogar „dauerhaft“.

3. Für $a > 1$ ist $a^2 - a > 0$. D. h. $f(x, ax^2) = cx^4$ mit $c > 0$. Wir hätten dadurch Parabeln gefunden, entlang derer die Funktion f ausgehend von $(0,0)$ wächst. Wenn man also sofort den richtigen Blick für den Term $y^2 - yx^2$ hat und sieht, dass man am besten (x, ax^2) einsetzt, findet man in einem Schritt Kurven entlang derer f fällt bzw. wächst. Man hätte sich dann die Untersuchung der Funktion entlang der beiden Achsen, beliebiger Geraden, usw. ersparen können.

Komprimierte Lösung zu Aufgabe 7.7 – Lokale Extremstellen

$$\begin{aligned} Df(x,y) &= \begin{pmatrix} -2xy & 2y - x^2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix} \Leftrightarrow (x,y) = (0,0) \\ f(x, ax^2) &= a^2x^4 - ax^2 \cdot x^2 = (a^2 - a)x^4 \begin{cases} \leq 0 & a \in (0,1) \\ \geq 0 & a > 1 \end{cases} \end{aligned}$$

Gleichheit gilt jeweils nur für $x = 0$. Es folgt, dass f an der einzigen kritischen Stelle $(0,0)$ einen Sattelpunkt besitzt.

Antworten zu den Verständnisfragen 7.8 – Lokale Umkehrbarkeit

1. Nein. Wenn man allerdings zusätzlich die Stetigkeit von f fordert, stimmt die Aussage. Nachdem wir es in unserem Fall mit einer Polynomfunktion zu tun haben (Polynomfunktionen sind immer stetig), können wir die Bijektivität von f wie beschrieben folgern.
2. Der Satz über die *lokale* Umkehrbarkeit macht nur eine Aussage über die Invertierbarkeit von f in einer Umgebung von a. Nachdem wir aber in a) schon gezeigt haben, dass f auf ganz $\mathbb{R}^2$ invertierbar ist, können wir dies in Teil b) natürlich ausnutzen. f^{-1} ist also nach Teil a) auf ganz $\mathbb{R}^2$ definiert.

   Dass nun diese Funktion f^{-1} auch stetig differenzierbar ist, folgt aus dem Satz über die lokale Umkehrbarkeit. Nachdem nämlich die Determinante der Jacobi-Matrix $(Df(x,y))$ in *jedem* Punkt (x,y) ungleich Null ist, können wir den Satz auf jeden einzelnen Punkt anwenden.
3. Das klappt z. B. für die Punkte $(2,0)$, $(0,0)$, $(1,-1)$ und $(1,1)$ ebenso leicht.

Komprimierte Lösung zu Aufgabe 7.8 – Lokale Umkehrbarkeit

a) $y = s - x^3 \Rightarrow -(s-x^3)^3 + x = t$. Sei $f(x) := -(s-x^3)^3 + x$. Dann ist $f'(x) = -3(s-x^3)^2 \cdot (-3x^2) + 1 > 0$ für alle $x \in \mathbb{R}$. Daraus folgt, dass f eine streng monoton wachsende Funktion ist.

Da f eine Polynomfunktion ungeraden Grades ist, gilt:

$$\lim_{x\to+\infty} f(x) = +\infty \quad \text{und} \quad \lim_{x\to-\infty} f(x) = -\infty.$$

Also hat die Gleichung $f(x) = t$ eine eindeutige Lösung $x \in \mathbb{R}$. Aus $y = s - x^3$ erhalten wir den eindeutigen Wert für y.

b) Sei $f(x,y) := (x^3 + y, -y^3 + x)$. Dann ist f auf $\mathbb{R}^2$ stetig differenzierbar. Aus Teil a) folgt, dass $f : \mathbb{R}^2 \to \mathbb{R}^2$ eine Bijektion ist. Außerdem gilt:

$$\det Df(x,y) = \det \begin{pmatrix} 3x^2 & 1 \\ 1 & -3y^2 \end{pmatrix} = -9x^2y^2 - 1 < 0.$$

Also ist f^{-1} eine stetig differenzierbare Funktion und $f^{-1}(s,t) = (Df(x,y))^{-1}$, wobei $x = x(s,t)$ und $y = y(s,t)$ ist.

Für $(s,t) = (0,2)$ bekommen wir $(x,y) = (1,-1)$ und

$$\begin{pmatrix} x'(0,2) \\ y'(0,2) \end{pmatrix} = Df^{-1}(0,2) = (Df(1,-1))^{-1} = \begin{pmatrix} 3 & 1 \\ 1 & -3 \end{pmatrix}^{-1} = -\frac{1}{10} \begin{pmatrix} -3 & -1 \\ -1 & 3 \end{pmatrix}.$$

Antworten zu den Verständnisfragen 7.9 – Implizite Funktionen

1. Wenn wir den Funktionsterm für $(x,y,0)$ auswerten, erhalten wir:

$$f(x,y,0) := \begin{pmatrix} x + 0 \cdot \sin(x+y) \\ y + 0 \cdot \cos(xy) \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ gilt dann und nur dann, wenn sowohl $x = 0$ als auch $y = 0$ gilt.

Hätte es weitere Punkte $(x, y, 0)$ gegeben, die $f(x, y, 0) = 0$ erfüllen, so hätte man einen von ihnen wählen können, mit dem man weiterarbeitet.

2. Weil wir schon gewusst haben, dass wir für t den Wert Null einsetzen werden. Da der Ausdruck $(x(t)y(t))'$ mit t multipliziert wird, fällt er also ohnehin weg.

Komprimierte Lösung zu Aufgabe 7.9 – Implizite Funktionen

a) Für $t = 0$ erhalten wir $x = y = 0$. Sei $f(x, y, t) := \big(x + t\sin(x+y) \quad y + t\cos(xy)\big)$.

$$D_{(x,y)} f(x,y,t) = \begin{pmatrix} 1 + t\cos(x+y) & t\cos(x+y) \\ -t\sin(xy)\cdot y & 1 - t\sin(xy)\cdot x \end{pmatrix}$$

$$\det D_{(x,y)} f(0,0,0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \neq 0$$

Nach dem Satz über implizite Funktionen gibt es eine Umgebung $(-\eta, \eta)$ von $t = 0$ in $\mathbb{R}$, in der das Gleichungssystem eine Lösung besitzt. Außerdem folgt, dass $(x(t), y(t))$ stetig differenzierbar auf $(-\eta, \eta)$ ist.

b) $x(0) = 0$, $y(0) = 0$. Es gilt:

$$\begin{aligned} x'(t) + \sin(x(t) + y(t)) + t\cos(x(t) + y(t)) \cdot (x'(t) + y'(t)) &= 0 \\ y'(t) + \cos(x(t)y(t)) - t\sin(x(t)y(t)) \cdot (x(t)y(t))' &= 0. \end{aligned}$$

Wir setzen $t = 0$ ein und erhalten:

$$\begin{aligned} x'(0) + \sin(0+0) + 0 &= 0 \\ y'(0) + \cos(0\cdot 0) - 0 &= 0. \end{aligned}$$

Es gilt also $x'(0) = 0$ und $y'(0) = -1$.

11.4 Lösungen zu Kapitel 8

Antworten zu den Verständnisfragen 8.1 – Unterraumkriterium

1. Die Voraussetzungen sind: K ist ein Körper, V ist ein K-Vektorraum und $U \subseteq V$. Für U_1 und U_2 sind die Voraussetzungen erfüllt, da es sich bei $\mathbb{R}$ um einen Körper, bei $\mathbb{R}^n$ um einen $\mathbb{R}$-Vektorraum handelt und wie U_1 und U_2 aus Elementen (also Vektoren) des $\mathbb{R}^n$ bestehen (siehe Definition von U_1 und U_2).

2. Die Bedingung (i) ist erfüllt, da wir Vektoren gefunden haben, die in U_1 liegen. Hier liefern die Beispiele also schon einen *Beweis*, dass die Bedingung (i) erfüllt ist. Um zu zeigen, dass die Beispielvektoren kein Gegenbeispiel zu den Bedingungen (ii) und (iii) liefern, müsste man überprüfen, ob die Summe von je zwei der Vektoren sowie alle skalaren Vielfachen der Vektoren wieder in U_1 liegen.

3. Die Voraussetzung, die hier eingeht, ist, dass $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in U_1$ ist und damit $\sum_{i=1}^{n-1} x_i = x_n$ gilt.

Komprimierte Lösung zu Aufgabe 8.1 – Unterraumkriterium

Beh. 1: U_1 ist Unterraum von $\mathbb{R}^n$.

Beweis:

(i) $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in U_1$, da $\sum_{i=1}^{n-1} 0 = 0$ gilt. Daraus folgt $U_1 \neq \emptyset$.

(ii) Seien $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in U_1$ und $a \in K$ beliebig. Für $a \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} ax_1 \\ \vdots \\ ax_n \end{pmatrix}$ folgt dann:

$$\sum_{i=1}^{n-1} (ax_i) = a(\sum_{i=1}^{n-1} x_i) \overset{\text{Vor.}}{=} ax_n.$$

Also ist $a \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in U_1$.

(iii) Seien $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in U_1$ beliebig.

Für $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$ folgt:

$$\sum_{i=1}^{n-1} (x_i + y_i) = \sum_{i=1}^{n-1} x_i + \sum_{i=1}^{n-1} y_i \overset{\text{Vor.}}{=} x_n + y_n.$$

Damit gilt: $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in U_1$.

Beh. 2: U_2 ist kein Unterraum.

Beweis: $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \notin U_2$, da $\sum_{i=1}^{n} 0 = 0 \neq 1$ ist. Damit kann U_2 kein Unterraum sein.

Antworten zu den Verständnisfragen 8.2 – Lineare (Un-)Abhängigkeit

1. Den Nullvektor 0_V kann man immer als Linearkombination beliebiger anderer Vektoren $v_1, \ldots, v_n$ schreiben:

   $$0_V = 0 \cdot v_1 + \ldots + 0 \cdot v_n.$$

   Sind also die Skalare dieser Linearkombination alle gleich 0, spricht man von einer *trivialen* Linearkombination. Gilt dagegen

   $$0_V = \alpha_1 \cdot v_1 + \ldots + \alpha_n \cdot v_n,$$

   wobei mindestens ein α_i ungleich 0 ist, so spricht man von einer *nichttrivialen* Linearkombination.

2. Ja, wir hätten auch α oder γ frei wählen können. Die beiden Gleichungen, die wir erhalten haben ($\alpha = \beta$ und $\gamma = -\beta$), sagen uns, dass egal welche der Variablen wir festlegen, die anderen damit automatisch auch festgelegt sind. Wir dürfen deshalb auch nicht α *und* β frei wählen.

3. Nein, denn es kommt beim Nachweis der linearen Abhängigkeit nur darauf an, dass wir eine nichttriviale Linearkombination finden. Wie wir das tun, spielt keine Rolle. Wir müssen nur beweisen, dass die von uns gefundene Linearkombination tatsächlich den Nullvektor ergibt und dass mindestens einer der Skalare ungleich 0 ist.

4. Nein, denn so hätten wir die triviale Linearkombination erhalten und die wollten wir ja gerade nicht!

Lösungsvorschlag zu Übung 8.2 – Lineare (Un-)Abhängigkeit

Von oben nach unten: B K B Z H K H B Z H A

Komprimierte Lösung zu Aufgabe 8.2 – Lineare (Un-)Abhängigkeit

Zu M: Seien $\alpha, \beta, \gamma \in \mathbb{R}$ beliebig mit

$$\alpha \cdot \begin{pmatrix} 3 \\ 5 \\ 2 \end{pmatrix} + \beta \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \gamma \cdot \begin{pmatrix} 3 \\ 6 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Es folgt

$$\left\{\begin{array}{lcl} 3\alpha + \beta + 3\gamma & = & 0 \\ 5\alpha + \beta + 6\gamma & = & 0 \\ 2\alpha + \beta + 2\gamma & = & 0 \end{array}\right\} \Rightarrow \left\{\begin{array}{ccc} 2\alpha + 3\gamma & = & 0 \\ \alpha + \gamma & = & 0 \\ \alpha + \beta + 2\gamma & = & 0 \end{array}\right\} \Rightarrow \left\{\begin{array}{lcc} \gamma & = & 0 \\ \alpha & = & -\gamma \\ \beta & = & -2\alpha - 2\gamma \end{array}\right\}.$$

Damit folgt: $\alpha = \beta = \gamma = 0$. Der Nullvektor lässt sich also nur als triviale Linearkombination der Vektoren aus M schreiben. Folglich ist M linear unabhängig.

Zu N:

Da

$$1 \cdot \begin{pmatrix} 2 \\ -3 \\ -6 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 5 \\ 7 \end{pmatrix} - 1 \cdot \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

gilt, wie man durch Nachrechnen leicht überprüfen kann, haben wir eine nichttriviale Linearkombination gefunden und damit ist N linear abhängig.

Antworten zu den Verständnisfragen 8.3 – Bestimmung einer Basis

1. Damit die Vektoren in U liegen, muss es sich einerseits um Vektoren aus $\mathbb{R}^3$ handeln und andererseits muss $x + 3y + 2z = 0$ sein (d. h. in Worten: Die Summe aus einmal dem ersten Eintrag, dreimal dem zweiten Eintrag und zweimal dem dritten Eintrag des Vektors muss 0 ergeben). Ersteres ist sicherlich erfüllt und letzteres auch, da $1 \cdot (-3) + 3 \cdot 1 + 2 \cdot 0 = 0$ bzw. $1 \cdot (-2) + 3 \cdot 0 + 2 \cdot 1 = 0$ ist.
2. Zunächst hat man sich überlegt, dass sich jeder Vektor aus U in der Form $u = \begin{pmatrix} -3y - 2z \\ y \\ z \end{pmatrix}$ schreiben lässt. Das Umformen entspricht gewissermaßen dem Setzen von einmal $y = 1$, $z = 0$ (dadurch erhält man den ersten Summanden – den Vektor $\begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix}$, der dann natürlich immer noch in U liegt) und einmal $y = 0$, $z = 1$ (ergibt den Vektor $\begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$).
3. Nein, z. B. bilden auch die Vektoren $\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ -3 \end{pmatrix}$ eine Basis von U.

Komprimierte Lösung zu Aufgabe 8.3 – Bestimmung einer Basis

Sei $u \in U$ beliebig. Dann gibt es $x, y, z \in \mathbb{R}$ mit $u = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ und $x + 3y + 2z = 0$.

Es folgt:

$$u = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -3y - 2z \\ y \\ z \end{pmatrix} = y \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}.$$

Damit bilden $\begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$ ein Erzeugendensystem von U. Außerdem folgt für alle $\alpha, \beta \in \mathbb{R}$ mit $\alpha \cdot \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix} + \beta \cdot \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$:

$$\begin{pmatrix} -3\alpha - 2\beta \\ \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \Rightarrow \quad \alpha = \beta = 0.$$

Also sind $\begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$ linear unabhängig. Insgesamt folgt dann, dass die beiden Vektoren eine Basis von U bilden.

Antworten zu den Verständnisfragen 8.4 – Rechnen mit Matrizen

1. Aus Gleichung (8.13) hatten wir gefolgert, dass b und c ungleich 0 sein müssen (zumindest in $\mathbb{R}$). Aus Gleichung (8.14) folgte damit $d = -a$. Aus Gleichung (8.15) folgt das selbe! Aus dieser Gleichung bekommen wir also keine zusätzlichen Informationen. Setzen wir außerdem in Gleichung (8.16) die gefolgerte Beziehung $d = -a$ ein, so erhalten wir wieder Gleichung (8.13) – also erneut keine weiteren Informationen. (Gleichung (8.13) hatten wir dann noch nach c umgestellt.)
2. In A übernimmt a die Rolle von a und r die Rolle von b in X. In B übernimmt b die Rolle von a und s die Rolle von b in X.
3. $p = -\frac{2as}{r}$ und $q = \frac{s^2(1+a^2)+r^2}{r^2}$.

Komprimierte Lösung zu Aufgabe 8.4 – Rechnen mit Matrizen

Für $X = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_{2\times 2}(K)$ mit $X^2 = -E_2$ können wir – da

$$X^2 = \begin{pmatrix} a^2 + bc & ab + bd \\ ca + dc & cb + d^2 \end{pmatrix}$$

gilt – nachstehende Gleichheiten folgern:

$$\begin{aligned} a^2 + bc &= -1, \\ ab + bd &= 0, \\ ca + dc &= 0, \\ cb + d^2 &= -1. \end{aligned}$$

In $\mathbb{R}$ können dann b, c nicht 0 sein. Damit folgt aus den Gleichungen schon: $d = -a$ und $c = -\frac{1+a^2}{b}$. A und B sind damit Matrizen der Gestalt:

$$A = \begin{pmatrix} a & r \\ -\frac{1+a^2}{r} & -a \end{pmatrix} \text{ und } B = \begin{pmatrix} b & s \\ -\frac{1+b^2}{s} & -b \end{pmatrix}, \text{ wobei } r, s \neq 0.$$

Aus der Bedingung $AB = -BA$ können wir folgern: $ab - r\frac{1+b^2}{s} = -(ba - s\frac{1+a^2}{r})$.

Durch Umformungen nach b erhalten wir: $b = \frac{as}{r} \pm \frac{1}{|r|}\sqrt{-(s^2 + r^2)}$.

In $\mathbb{R}$ existiert b nicht, weil $-(s^2 + r^2) < 0$ ist. Also gibt es auch keine Matrizen $A, B \in M_{2\times 2}(\mathbb{R})$ mit $A^2 = B^2 = -E_2$ und $AB = -BA$.

In $\mathbb{C}$ ist auch die Wurzel aus negativen Zahlen definiert. Wir dürfen deshalb z. B. $s = 1, r = 1, a = 0$ wählen und erhalten so $b = \sqrt{2} \cdot i$.

Die Matrizen

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

und

$$B = \begin{pmatrix} \sqrt{2} \cdot i & 1 \\ 1 & -\sqrt{2} \cdot i \end{pmatrix}$$

erfüllen dann gerade die beiden Bedingungen. Damit existieren in $M_{2\times 2}(\mathbb{C})$ Matrizen A, B mit $A^2 = B^2 = -E_2$ und $AB = -BA$.

Antworten zu den Verständnisfragen 8.5 – Basis des Kerns einer Matrix

1.
   - Die Multiplikation einer Zeile mit einem Skalar ungleich 0
   - Die Addition des skalaren Vielfachen einer Zeile zu einer anderen
   - Vertauschen zweier Zeilen
2. Das bedeutet lediglich, dass der Kern der Ausgangsmatrix und der Kern der veränderten Matrix gleich sind.
3. *Schritt 1*: Wir ziehen das Zweifache der 1. Zeile von der 2. Zeile und das Dreifache der 1. Zeile von der 3. Zeile ab.

   Schritt 2: Wir ziehen die 2. Zeile von der 3. Zeile ab.

   Schritt 3: Wir teilen die 2. Zeile durch -2 und die 3. Zeile durch -3.

   Schritt 4: Wir ziehen die 3. Zeile $\frac{1}{2}$-mal von der 2. Zeile und 2-mal von der 1. Zeile ab.

   Schritt 5: Wir ziehen die 2. Zeile von der 1. Zeile ab.

   In Schritt 2 wurde direkt die 2. Zeile von der 3. Zeile abgezogen. Nach dem Gauß-Jordan-Algorithmus müsste man zunächst die 2. Zeile durch -2 teilen und dann das 2-fache dieser Zeile von der 3. Zeile abziehen. Manchmal lohnt es sich also, vom Gauß-Jordan-Algorithmus abzuweichen, da sich so der Lösungsweg verkürzt.

4. Schauen wir uns das im Beispiel genauer an: Ist der Vektor $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix}$ aus $\ker(A)$, so haben wir festgestellt, dass er auch im Kern der Matrix in Treppennormalform liegt. Demnach gilt:

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 & 3 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 + 3x_4 + x_6 + x_7 \\ x_3 + x_6 + 2x_7 \\ x_5 + 2x_6 \end{pmatrix}.$$

Damit hängen x_1, x_3 und x_5 jeweils von x_2, x_4, x_6 und x_7 ab (diese vier kann man also frei wählen). So erhält man auch auf diesem Wege die gleichen Basisvektoren (evtl. bis auf ein Vorzeichen). Probieren Sie es aus!

Komprimierte Lösung zu Aufgabe 8.5 – Basis des Kerns einer Matrix

Wir bringen die Matrix A auf Treppennormalform:

$$\begin{pmatrix} 1 & 2 & 1 & 3 & 2 & 6 & 3 \\ 2 & 4 & 0 & 6 & 3 & 8 & 2 \\ 3 & 6 & 1 & 9 & 2 & 8 & 5 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 3 & 2 & 6 & 3 \\ 0 & 0 & -2 & 0 & -1 & -4 & -4 \\ 0 & 0 & -2 & 0 & -4 & -10 & -4 \end{pmatrix}$$

$$\rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 3 & 2 & 6 & 3 \\ 0 & 0 & -2 & 0 & -1 & -4 & -4 \\ 0 & 0 & 0 & 0 & -3 & -6 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 3 & 2 & 6 & 3 \\ 0 & 0 & 1 & 0 & \frac{1}{2} & 2 & 2 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 \end{pmatrix}$$

$$\rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & 3 & 0 & 2 & 3 \\ 0 & 0 & 1 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 0 & 3 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 \end{pmatrix}.$$

Aus der Treppennormalform können wir eine Basis des Kerns nach dem Einfügen von (-1)en an den richtigen Stellen ablesen:

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

Die 2., 4., 6. und 7. Spalte dieser Matrix bilden eine Basis des Kerns von A.

Antworten zu den Verständnisfragen 8.6 – Basisergänzung und Faktorraum

1. Es gibt genau eine Ausnahme! Der Nullvektor darf *nicht* gegen einen Basisvektor getauscht werden, denn jede Menge, die den Nullvektor enthält, ist automatisch linear abhängig und damit *keine* Basis mehr. Der Nullvektor erfüllt allerdings auch nicht die Voraussetzung des Austauschlemmas, da keiner der Koeffizienten α_i in der Linearkombination $0 = \sum_{i=1}^{n} \alpha_i b_i$ ungleich Null ist.

2. $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \in U$, da $1 + 0 = 1 + 0$ ist.

   $\begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \in U$, da $0 + 2 = 0 + 2$ ist.

   Da für $\alpha, \beta \in \mathbb{R}$ aus $\alpha \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ stets $\alpha + 0 = 0$ und $0 + \beta = 0$, also insbesondere $\alpha = \beta = 0$ folgt, sind die beiden Vektoren linear unabhängig.

3. Jeder Vektor v, der nicht der Nullvektor ist, ist linear unabhängig, da dann $\lambda v = 0_V$ nur für $\lambda = 0$ gelten kann. Nun wissen wir aber auch, dass jede Basis des Vektorraums nur aus einem Vektor besteht (die Dimension ist 1). Und somit muss es sich bei *einem* linear unabhängigen Vektor schon um eine Basis handeln.

Komprimierte Lösung zu Aufgabe 8.6 – Basisergänzung und Faktorraum

Wir bestimmen eine Basis von U.

Sei $u \in U$ beliebig, dann gilt:

$$u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} \text{ mit } u_1 + 2u_2 = u_3 + 2u_4.$$

Daraus folgt $u_3 = u_1 + 2u_2 - 2u_4$ und somit gilt:

$$u = \begin{pmatrix} u_1 \\ u_2 \\ u_1 + 2u_2 - 2u_4 \\ u_4 \end{pmatrix} = u_1 \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + u_2 \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix} + u_4 \begin{pmatrix} 0 \\ 0 \\ -2 \\ 1 \end{pmatrix}.$$

Damit bilden $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 0 \\ -2 \\ 1 \end{pmatrix}$ ein Erzeugendensystem von U. Man sieht leicht, dass sie auch linear unabhängig sind und demnach eine Basis von U bilden.

Wir ergänzen M mit Hilfe des Austauschlemmas zu einer Basis von U. Da

$$\begin{pmatrix}0\\1\\0\\1\end{pmatrix} = 0 \cdot \begin{pmatrix}1\\0\\1\\0\end{pmatrix} + 1 \cdot \begin{pmatrix}0\\1\\2\\0\end{pmatrix} + 1 \cdot \begin{pmatrix}0\\0\\-2\\1\end{pmatrix}$$

ist, dürfen wir z. B. $\begin{pmatrix}0\\1\\0\\1\end{pmatrix}$ für $\begin{pmatrix}0\\1\\2\\0\end{pmatrix}$ in die Basis schreiben. Also können wir M durch $\begin{pmatrix}0\\0\\-2\\1\end{pmatrix}$ zu einer Basis von U ergänzen.

Wir bestimmen eine Basis von V/U:

Da $\dim(V/U) = \dim(V) - \dim(U) = 4 - 3 = 1$ ist und $\begin{pmatrix}1\\0\\0\\0\end{pmatrix} + U \neq U = 0_{V/U}$ gilt (denn $\begin{pmatrix}1\\0\\0\\0\end{pmatrix} \notin U$), ist $\left\{\begin{pmatrix}1\\0\\0\\0\end{pmatrix} + U\right\}$ eine Basis von V/U.

Antworten zu den Verständnisfragen 8.7 – Lineare Abbildungen

1. Nun, wenn beides gilt, so folgt $\dim(\ker(f)) - \dim(\ker(f_U)) \leq \dim(\ker(\bar{f}))$. Addieren wir bei dieser Ungleichung auf beiden Seiten $\dim(\ker(f_U))$, so erhalten wir gerade

   $$\dim(\ker(f)) \leq \dim(\ker(f_U)) + \dim(\ker(\bar{f})),$$

   was wir beweisen sollten.

2. Damit die Inklusion gelten kann, müssen die beiden Mengen im gleichen Vektorraum enthalten sein. Ihre Elemente müssen die gleiche Form haben. Da $\bar{f}$ eine Abbildung von V/U nach V/U ist, liegt ihr Kern in V/U. Aber auch $(\ker(f) + U)/U$ ist ein Unterraum von V/U, da $(\ker(f) + U)$ ein Unterraum von V ist.

3. Hier geht ein, dass wir über zwei ineinander enthaltene (Unter-)Vektorräume W_1, W_2 mit $W_1 \subseteq W_2$ schon Folgendes über ihre Dimensionen wissen:

   $$\dim(W_1) \leq \dim(W_2).$$

   In unserem Fall ist $W_1 = (\ker(f) + U)/U$ und $W_2 = \ker(\bar{f})$.

Komprimierte Lösung zu Aufgabe 8.7 – Lineare Abbildungen

Wenden wir den 2. Homomorphiesatz auf $X = \ker(f)$ und U an, erhalten wir:

$$\ker(f)/(\ker(f) \cap U) \cong (\ker(f) + U)/U.$$

Da aber $\ker(f) \cap U = \ker(f_U)$ ist, gilt:

$$\ker(f)/\ker(f_U) \cong (\ker(f) + U)/U.$$

Für die Dimensionen folgt dann:

$$\dim(\ker(f)) - \dim(\ker(f_U)) = \dim((\ker(f) + U)/U).$$

Es bleibt noch zu zeigen:

$$\dim((\ker(f) + U)/U) \leq \dim(\ker(\bar{f})).$$

Behauptung: $(\ker(f) + U)/U \subseteq \ker(\bar{f})$. (Daraus folgt, was noch zu zeigen ist!)

Beweis der Behauptung: Sei $v \in \ker(f) + U$ beliebig. Dann gibt es $v' \in \ker(f)$ und $u \in U$ mit $v = v' + u$.

Es folgt, da $f(v') = 0_V$ und $f(u) \in U$:

$$\bar{f}(v + U) = f(v) + U = f(v') + f(u) + U = U = 0_{V/U}.$$

Also ist $v + U \in \ker(\bar{f})$, was zu zeigen war.

Antworten zu den Verständnisfragen 8.8 – Zerlegung in Transpositionen

1. Mit dem Produkt ist hier die Hintereinanderausführung (Verkettung) der Transpositionen gemeint, wie man sie von Abbildungen her kennt!
2. Prinzipiell ist es egal, welches Element wir als „Startpunkt" wählen. Die Zerlegung in Transpositionen wird sich dadurch aber im Allgemeinen ändern. Die Zerlegung in Transpositionen ist also nicht eindeutig.
3. Es gilt:

$$\begin{aligned}
(1\,4) \cdot (1\,3) \cdot (1\,2)(1): &\quad 1 \overset{\tau_1}{\mapsto} 2 \overset{\tau_2}{\mapsto} 2 \overset{\tau_3}{\mapsto} 2 = \sigma_3(1)\\
(1\,4) \cdot (1\,3) \cdot (1\,2)(2): &\quad 2 \overset{\tau_1}{\mapsto} 1 \overset{\tau_2}{\mapsto} 3 \overset{\tau_3}{\mapsto} 3 = \sigma_3(2)\\
(1\,4) \cdot (1\,3) \cdot (1\,2)(3): &\quad 3 \overset{\tau_1}{\mapsto} 3 \overset{\tau_2}{\mapsto} 1 \overset{\tau_3}{\mapsto} 4 = \sigma_3(3)\\
(1\,4) \cdot (1\,3) \cdot (1\,2)(4): &\quad 4 \overset{\tau_1}{\mapsto} 4 \overset{\tau_2}{\mapsto} 4 \overset{\tau_3}{\mapsto} 1 = \sigma_3(4).
\end{aligned}$$

Lösungsvorschlag zu Übung 8.8 – Zerlegung in Transpositionen

Von oben nach unten: K Z H Z H K H K H Z H B A H

Komprimierte Lösung zu Aufgabe 8.8 – Zerlegung in Transpositionen

σ_1 ist selbst eine Transposition, also ist die Zerlegung in Transpositionen gegeben durch $\sigma_1 = (1\,3)$.

$\sigma_2 = (2\,3) \cdot (1\,4)$, denn σ_2 vertauscht 2 mit 3 und 1 mit 4.

$\sigma_3 = (1\,4) \cdot (1\,3) \cdot (1\,2)$, denn für dieses Produkt von Transpositionen gilt:

$$\begin{aligned}
(1\,4) \cdot (1\,3) \cdot (1\,2)(1): &\quad 1 \overset{\tau_1}{\mapsto} 2 \overset{\tau_2}{\mapsto} 2 \overset{\tau_3}{\mapsto} 2 = \sigma_3(1),\\
(1\,4) \cdot (1\,3) \cdot (1\,2)(2): &\quad 2 \overset{\tau_1}{\mapsto} 1 \overset{\tau_2}{\mapsto} 3 \overset{\tau_3}{\mapsto} 3 = \sigma_3(2),\\
(1\,4) \cdot (1\,3) \cdot (1\,2)(3): &\quad 3 \overset{\tau_1}{\mapsto} 3 \overset{\tau_2}{\mapsto} 1 \overset{\tau_3}{\mapsto} 4 = \sigma_3(3),\\
(1\,4) \cdot (1\,3) \cdot (1\,2)(4): &\quad 4 \overset{\tau_1}{\mapsto} 4 \overset{\tau_2}{\mapsto} 4 \overset{\tau_3}{\mapsto} 1 = \sigma_3(4).
\end{aligned}$$

11.5 Lösungen zu Kapitel 9

Antworten zu den Verständnisfragen 9.1 – Vandermondesche Determinante

1. Der Induktionsbeweis ist eleganter und auch kürzer. Die Schritte, die zur Anwendung der Induktionsvoraussetzung geführt haben, kann man durchaus wiederholt für die so entstehenden, kleineren Matrizen anwenden. So kommt man mit einer „Pünktchen-Schreibweise" im Beweis auch zum Ziel. Diese „Pünktchen" werden aber von vielen Mathematikern nicht gerne gesehen, sie bevorzugen den Induktionsbeweis. Da kann nämlich zwischendurch *wirklich* nichts schief gehen.

2. Man erhält:

$$\begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & a_2 - a_1 & a_2^2 - a_1^2 & \cdots & a_2^{n-1} - a_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n - a_1 & a_n^2 - a_1^2 & \cdots & a_n^{n-1} - a_1^{n-1} \end{vmatrix}$$

$$= \begin{vmatrix} a_2 - a_1 & a_2^2 - a_1^2 & \cdots & a_2^{n-1} - a_1^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_n - a_1 & a_n^2 - a_1^2 & \cdots & a_n^{n-1} - a_1^{n-1} \end{vmatrix}.$$

   Das ist sogar die gleiche Matrix wie bei den Zeilenumformungen!

3. Das Produkt ist ein Doppelprodukt:

$$\begin{aligned}
\prod_{1 \le j < i \le n} (a_i - a_j) = &\prod_{j=1}^{n-1} \prod_{i=j+1}^{n} (a_i - a_j) \\
= &(a_2 - a_1) \cdot (a_3 - a_1) \cdot \ldots \cdot (a_n - a_1) \\
&\cdot (a_3 - a_2) \cdot (a_4 - a_2) \cdot \ldots \cdot (a_n - a_2) \\
&\cdot \ldots \\
&\cdot (a_{n-1} - a_{n-2}) \cdot (a_n - a_{n-2}) \\
&\cdot (a_n - a_{n-1}).
\end{aligned}$$

   Die Gleichheit gilt ja, wenn $\prod\limits_{j=1,\ 1<i\le n} (a_i - a_j) = (a_2 - a_1) \cdot \ldots \cdot (a_n - a_1)$ ist. Wenn wir das Produkt wie oben ausschreiben, steht aber genau die rechte Seite da.

Lösungsvorschlag zu Übung 9.1 – Vandermondesche Determinante

Von oben nach unten: B Z K H B K Z H K H K Z K B Z H B H K A H B

Komprimierte Lösung zu Aufgabe 9.1 – Vandermondesche Determinante

Wir beweisen die Behauptung durch Induktion nach n.

Induktionsanfang $n = 2$:

Da

$$\det(V(a_1, a_2)) = \begin{vmatrix} 1 & a_1 \\ 1 & a_2 \end{vmatrix} = 1 \cdot a_2 - 1 \cdot a_1 = \prod_{j=1,\ i=2} (a_i - a_j) = \prod_{1 \le j < i \le 2} (a_i - a_j)$$

ist, gilt die Behauptung für $n = 2$.

Induktionsschritt von $n-1$ nach n:

$$\begin{aligned}
\det(V(a_1, \dots, a_n)) &= \begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix} \\
&= \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & a_2 - a_1 & a_2^2 - a_1 \cdot a_2 & \cdots & a_2^{n-1} - a_1 \cdot a_2^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n - a_1 & a_n^2 - a_1 \cdot a_n & \cdots & a_n^{n-1} - a_1 \cdot a_n^{n-2} \end{vmatrix} \\
&\overset{\text{Lapl.}}{=} \begin{vmatrix} a_2 - a_1 & a_2^2 - a_1 \cdot a_2 & \cdots & a_2^{n-1} - a_1 \cdot a_2^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_n - a_1 & a_n^2 - a_1 \cdot a_n & \cdots & a_n^{n-1} - a_1 \cdot a_n^{n-2} \end{vmatrix} \\
&= \begin{vmatrix} (a_2 - a_1) & (a_2 - a_1) \cdot a_2 & \cdots & (a_2 - a_1) \cdot a_2^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ (a_n - a_1) & (a_n - a_1) \cdot a_n & \cdots & (a_n - a_1) \cdot a_n^{n-2} \end{vmatrix} \\
&= (a_2 - a_1) \cdot \ldots \cdot (a_n - a_1) \begin{vmatrix} 1 & a_2 & \cdots & a_2^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & \cdots & a_n^{n-2} \end{vmatrix} \\
&\overset{\text{IV}}{=} (a_2 - a_1) \cdot \ldots \cdot (a_n - a_1) \prod_{2 \le j < i \le n} (a_i - a_j) \\
&= \prod_{j=1,\ 1 < i \le n} (a_i - a_j) \prod_{2 \le j < i \le n} (a_i - a_j) \\
&= \prod_{1 \le j < i \le n} (a_i - a_j)
\end{aligned}$$

Antworten zu den Verständnisfragen 9.2 – Dualraum

1. Es gilt:

$$\begin{aligned} f(\lambda v) &= w \cdot (\lambda v)^T && \text{Definition von } f_w \\ &= w \cdot (\lambda v^T) && \text{Rechenregel für Transponierte} \\ &= \lambda \cdot w \cdot v^T && \text{Rechenregeln für Matrizen (insb. Multipl. mit Skalaren)} \\ &= \lambda \cdot f_w(v). && \text{Definition von } f_w \end{aligned}$$

2. A^* ist eine lineare Abbildung, die eine Linearform f aus W^* auf eine Linearform g aus V^* abbildet. Und zwar so, dass für $g = A^*(f) : V \to K$ Folgendes gilt:

$$g(v) = f(A(v)) \quad \text{für alle } v \in V.$$

3. Bei A^* und id_{V^*} handelt es sich jeweils um Abbildungen von V^* nach V^*. Die Gleichheit zweier Abbildungen beweist man aber gerade dadurch, dass jedes Element der Grundmenge (hier V^*) durch beide Abbildungen auf das gleiche Element der Bildmenge (hier ebenfalls V^*) abgebildet wird. Man muss also zeigen, dass für alle $f \in V^*$ schon $A^*(f) = id_{V^*}(f)$ gilt. Da $id_{V^*}(f) = f$ ist, müssen wir also gerade $A^*(f) = f$ für alle $f \in V^*$ beweisen.

Komprimierte Lösung zu Aufgabe 9.2 – Dualraum

Zu (i): Für $v = (v_1, \dots, v_n)$ und $w = (w_1, \dots, w_n)$ aus V ist

$$f_w(v) = w \cdot v^T = \sum_{i=1}^{n} w_i v_i \in K.$$

Also ist tatsächlich $f_w : V \to K$.

Wegen

$$f_w(v+u) = w \cdot (v+u)^T = w \cdot (v^T + u^T) = w \cdot v^T + w \cdot u^T = f_w(v) + f_w(u)$$

und

$$f_w(\lambda v) = w \cdot (\lambda v)^T = w(\lambda \cdot v^T) = \lambda \cdot w \cdot v^T = \lambda \cdot f_w(v)$$

für alle $u, v \in V$ und $\lambda \in K$ ist dann $f_w \in V^*$.

Zu (ii): Man hat $e_i = (0, \dots, 0, 1, 0, \dots, 0) \in V$ mit der Eins an der i-ten Stelle. Für $i, j \in \{1, \dots, n\}$ gilt dann:

$$f_{e_i}(e_j) = e_i \cdot e_j^T = \delta_{ij}.$$

Damit ist $\mathcal{E}^* = \{f_{e_i} \mid 1 \le i \le n\}$ die zu $\mathcal{E}$ duale Basis.

Zu (iii): Für $v \in V$ und $f \in V^*$ gilt:

$$((id_V)^*(f))(v) \overset{\text{Def. duale Abb.}}{=} f(id_V(v)) = f(v),$$

d. h. $(id_V)^*(f) = f$. Also ist $(id_V)^* = id_{V^*}$.

Antworten zu den Verständnisfragen 9.3 – JNF ohne Basiswechselmatrix

1. Die Dimensionsformel $\dim(V) = \dim(\ker f) + \operatorname{rg}(f)$ für lineare Abbildungen $f : V \to W$ wurde hier auf die Matrix $A - 1E_4$ angewandt (d. h. $f = A - 1E_4$) und nach $\dim(\ker(A - 1E_4))$ umgestellt. Außerdem wurde noch verwendet, dass $V = \mathbb{R}^4$ also $\dim V = 4$ ist.

2. Zur Berechnung der Dimension von $\ker(A - 1E_4)$ haben wir im zweiten Schritt die Matrix $A - 1E_4$ berechnet. Die Matrix $A - 1E_4$ kann außerdem nur dann die Nullmatrix sein, wenn $A = E_4$ ist. Das ist aber nicht der Fall.

3. Das Minimalpolynom von A ist das Polynom, das A als Nullstelle besitzt und alle anderen Polynome schon teilt, die A als Nullstelle haben. Außerdem weiß man, dass das Minimalpolynom von der Form $(x-1)^m$ ist. Es reicht also das kleinste m zu bestimmen, sodass $(x-1)^m$ die Matrix A als Nullstelle hat. Wäre $\operatorname{minpol}_A(x) = x - 1$, müsste demnach $A - 1E_4 = \operatorname{minpol}_A(A) = 0$ sein. Das geht also nicht. Da $(A - 1E_4)^2 = 0$ ist, ist A Nullstelle des Polynoms $(x-1)^2$. Dieses muss dann wegen der vorherigen Überlegungen bereits das Minimalpolynom sein.

4. Da die Räume $V^k(\lambda_i, A)$ Unterräume des Vektorraums V (hier: $\mathbb{R}^4$) sind, können sie auch maximal so groß werden. Das heißt aber auch, wenn ein $V^k(\lambda_i, A) = V$ ist, dass danach nur noch Gleichheitszeichen in der Inklusionskette auftreten. Da im vorliegenden Fall $V(1, A) \subsetneq V^2(1, A) = \mathbb{R}^4$ gilt, muss $V^2(1, A)$ der verallgemeinerte Eigenraum sein.

Lösungsvorschlag zu Übung 9.3 – JNF ohne Basiswechselmatrix

Von oben nach unten: P K B H K H B H K B H B H

Komprimierte Lösung zu Aufgabe 9.3 – JNF ohne Basiswechselmatrix

$$\begin{aligned}
\operatorname{charpol}_A(x) &= \begin{vmatrix} x-3 & -1 & 0 & 0 \\ 4 & x+1 & 0 & 0 \\ -7 & -1 & x-2 & -1 \\ 17 & 6 & 1 & x \end{vmatrix} \\
&= \begin{vmatrix} x-3 & -1 & 0 \\ 4 & x+1 & 0 \\ 17 & 6 & 1 \end{vmatrix} + x \begin{vmatrix} x-3 & -1 & 0 \\ 4 & x+1 & 0 \\ -7 & -1 & x-2 \end{vmatrix} \\
&= (x-3)(x+1) + 4 + x[(x-3)(x+1)(x-2) + 4(x-2)] \\
&= (x^2 - 2x + 1) + x(x-2)[x^2 - 2x + 1] \\
&= (x - 2x + 1)[1 + x(x-2)] = (x-1)^2[x^2 - 2x + 1] \\
&= (x-1)^4
\end{aligned}$$

A hat also nur den Eigenwert 1.

Da

$$\dim(V(1,A)) = 4 - \mathrm{rg}(A - 1E_4) = 4 - \mathrm{rg}\begin{pmatrix} 2 & 1 & 0 & 0 \\ -4 & -2 & 0 & 0 \\ 7 & 1 & 1 & 1 \\ -17 & -6 & -1 & -1 \end{pmatrix}$$

$$= 4 - \mathrm{rg}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 5 & 1 & 1 & 0 \\ -5 & -6 & -1 & 0 \end{pmatrix}$$

$$= 4 - \mathrm{rg}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -6 & -1 & 0 \end{pmatrix} = 4 - 2 = 2$$

ist, hat die Jordansche Normalform zwei Blöcke.

Da $(A - 1E_4)^2 = 0$ und $A - 1E_4 \neq 0$ ist, ist $\mathrm{minpol}_A(x) = (x-1)^2$ bzw. $V^\infty(1,A) = V^2(1,A)$. Daraus folgt aber, dass der größte Block die Größe 2 hat.

Also gibt es zwei Blöcke der Größe 2 und die Jordansche Normalform sieht so aus:

$$\mathrm{JNF}(A) = \left(\begin{array}{cc:cc} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hdashline 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right).$$

Antworten zu den Verständnisfragen 9.4 – JNF mit Basiswechselmatrix

1. Wenn man eine Basis von $\ker(A - \lambda E_n)^{m-1}$ durch die Vektoren $b_1^{(m)}, \ldots, b_{l_m}^{(m)}$ zu einer Basis von $\ker(A - \lambda E_n)^m$ ergänzt, dann liegen diese Vektoren zwar in $\ker(A - \lambda E_n)^m$, aber nicht im $\ker(A - \lambda E_n)^{m-1}$ (denn dann wären sie nicht linear unabhängig von den Basisvektoren von $\ker(A - \lambda E_n)^{m-1}$).

   $(A - \lambda E_n)(b_i^{(m)})$ ist dann aus $\ker(A - \lambda E_n)^{m-1}$, denn

   $$(A - \lambda E_n)^{m-1}((A - \lambda E_n)(b_i^{(m)})) = (A - \lambda E_n)^m(b_i^{(m)}) \overset{b_i^{(m)} \in \ker(A-\lambda E_n)^m}{=} 0.$$

   $(A - \lambda E_n)(b_i^{(m)})$ ist nicht aus $\ker(A - \lambda E_n)^{m-2}$, denn

   $$(A - \lambda E_n)^{m-2}((A - \lambda E_n)(b_i^{(m)})) = (A - \lambda E_n)^{m-1}(b_i^{(m)}) \overset{b_i^{(m)} \notin \ker(A-\lambda E_n)^{m-1}}{\neq} 0.$$

Wichtig ist dies für den nächsten Schritt, denn dann lässt sich wirklich eine Basis von $\ker(A - \lambda E_n)^{m-2}$ durch die Vektoren $(A - \lambda E_n)(b_i^{(m)})$ und weitere Vektoren zu einer Basis von $\ker(A - \lambda E_n)^{m-1}$ ergänzen. (*Hinweis:* Eigentlich muss man sich dazu auch noch überlegen, dass diese Vektoren auch linear unabhängig sind!)

2. Bei der Berechnung von $\ker(A - 1E_4)^2$ hat man die Matrix $(A - 1E_4)^2$ berechnet und festgestellt, dass diese die Nullmatrix ist. Dann muss aber auch die Matrix $(A - 1E_4)^3 = (A - 1E_4)^2 \cdot (A - 1E_4)$ die Nullmatrix sein, und damit ihr Kern der ganze Raum ($\mathbb{R}^4$).

3. Aus der ursprünglichen Reihenfolge wissen wir, dass

$$\begin{aligned}
A[(A-1E_4)(e_1)] &= 1 \cdot (A-1E_4)(e_1) + 0 \cdot e_1 + 0 \cdot (A-1E_4)(e_2) + 0 \cdot e_2 \\
A[e_1] &= 1 \cdot (A-1E_4)(e_1) + 1 \cdot e_1 + 0 \cdot (A-1E_4)e_2 + 0 \cdot (e_2) \\
A[(A-1E_4)(e_2)] &= 0 \cdot (A-1E_4)(e_1) + 0 \cdot e_1 + 1 \cdot (A-1E_4)(e_2) + 0 \cdot e_2 \\
A[e_2] &= 0 \cdot (A-1E_4)(e_1) + 0 \cdot e_1 + 1 \cdot (A-1E_4)(e_2) + 1 \cdot e_2
\end{aligned}$$

gilt. Durch Vertauschen der Reihenfolge ändert sich die Reihenfolge der Koeffizienten in der Matrix. (Die Koeffizienten zu dem Bild eines Basisvektors werden in die entsprechende Spalte der Matrix geschrieben. Beachten Sie, dass sich die Summationsreihenfolge ändert, wenn sich die Reihenfolge der Basiselemente ändert!)

a) $S^{-1}AS = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$

b) $S^{-1}AS = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$

Komprimierte Lösung zu Aufgabe 9.4 – JNF mit Basiswechselmatrix

Zuerst berechnen wir das charakteristische Polynom (aus Abschnitt 9.3 kennen wir dies bereits): $\mathrm{charpol}_A(x) = (x-1)^4$.

Wir bestimmen nun den Eigenraum bis hin zum verallgemeinerten Eigenraum unter Angabe von Basen:

$$V(1, A) = \ker(A - 1E_4) = \ker \begin{pmatrix} 2 & 1 & 0 & 0 \\ -4 & -2 & 0 & 0 \\ 7 & 1 & 1 & 1 \\ -17 & -6 & -1 & -1 \end{pmatrix} \tag{11.1}$$

$$= \ker \begin{pmatrix} 1 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -\frac{5}{2} & 1 & 1 \\ 0 & \frac{5}{2} & -1 & -1 \end{pmatrix} = \ker \begin{pmatrix} 1 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & -\frac{2}{5} & -\frac{2}{5} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$= \ker \begin{pmatrix} 1 & 0 & \frac{1}{5} & \frac{1}{5} \\ 0 & 1 & -\frac{2}{5} & -\frac{2}{5} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ -2 \\ -5 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 0 \\ -5 \end{pmatrix} \right\rangle$$

$$= \langle e_1 - 2e_2 - 5e_3, e_1 - 2e_2 - 5e_4 \rangle,$$

$$V^2(1, A) = \ker(A - 1E_4)^2 = \ker \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \mathbb{R}^4 = \langle e_1, e_2, e_3, e_4 \rangle.$$

Damit ist $\ker(A - 1E_4)^2$ der verallgemeinerte Eigenraum.

Wir ergänzen nun die Basis von $\ker(A - 1E_4)$ zu einer Basis von $\ker(A - 1E_4)^2$:

Dazu tauschen wir den Vektor $e_1 - 2e_2 - 5e_3$ gegen e_3 in der Basis des verallgemeinerten Eigenraumes und bekommen so die neue Basis des verallgemeinerten Eigenraumes:

$$\{e_1 - 2e_2 - 5e_3, e_1, e_2, e_4\}.$$

$e_1 - 2e_2 - 5e_4$ dürfen wir z. B. gegen e_4 eintauschen. Wir erhalten die neue Basis:

$$\{e_1 - 2e_2 - 5e_3, e_1 - 2e_2 - 5e_4, e_1, e_2\}.$$

e_1, e_2 ergänzt damit die Basis von $\ker(A - 1e_4)$ zu einer Basis von $\ker(A - 1e_4)^2$.

Wir berechnen $(A - 1E_4)e_1$ und $(A - 1E_4)(e_2)$:

$$(A - 1E_4)e_1 = \begin{pmatrix} 2 & 1 & 0 & 0 \\ -4 & -2 & 0 & 0 \\ 7 & 1 & 1 & 1 \\ -17 & -6 & -1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ -4 \\ 7 \\ -17 \end{pmatrix},$$

$$(A - 1E_4)e_2 = \begin{pmatrix} 2 & 1 & 0 & 0 \\ -4 & -2 & 0 & 0 \\ 7 & 1 & 1 & 1 \\ -17 & -6 & -1 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 1 \\ -6 \end{pmatrix}.$$

Die Vektoren $(A - 1E_4)(e_1)$ und $(A - 1E_4)(e_2)$ bilden dann schon eine Basis von $\ker(A - 1E_4)$. Und somit ist die gesuchte Basis:

$$\{(A - 1E_4)(e_1), e_1, (A - 1E_4)(e_2), e_2\}$$

und die zugehörige Transformationsmatrix S:

$$S = \begin{pmatrix} 2 & 1 & 1 & 0 \\ -4 & 0 & -2 & 1 \\ 7 & 0 & 1 & 0 \\ -17 & 0 & -6 & 0 \end{pmatrix}.$$

Antworten zu den Verständnisfragen 9.5 – Invariante Unterräume

1. $A|_U$ ist die Einschränkung von A auf den Unterraum U. Das heißt, $A|_U$ bildet von U nach U ab mit $A|_U(u) := A(u)$ für alle $u \in U$. (Weil U A-invariant ist, landen wir mit der Abbildung wirklich in U.)

   Ist allgemein ein Endomorphismus $f : V \to V$ gegeben, so gilt für den Grad des charakteristischen Polynoms stets:

   $$\operatorname{grad}(\operatorname{charpol}_f) = \dim V$$

   In diesem Fall ist $V = U$, $f = A|_U$ und $\dim U = 1$, woraus schließlich

   $$\operatorname{grad}(\operatorname{charpol}_{A|_U}) = 1$$

   folgt.

2. Wir wissen, dass der Grad von $\operatorname{charpol}_{A|_U}$ gleich 1 ist und dass $\operatorname{charpol}_{A|_U}$ das charakteristische Polynom von A teilt. Dieses haben wir bereits in irreduzible Faktoren zerlegt. Da nun aber die irreduziblen Polynome im Ring der Polynome auch prim sind, wissen wir, dass ein Teiler eines Polynoms stets aus den gleichen irreduziblen Faktoren aufgebaut ist wie das Polynom, das es teilt. Oder anders formuliert: Teilt ein Polynom f ein anderes Polynom g, so ist jeder irreduzible Faktor von f stets auch ein irreduzibler Faktor von g.

   Da $\operatorname{charpol}_A(x) = (x-1)^2(x-4)$ ist, kann $\operatorname{charpol}_{A|_U}$ nur aus den Faktoren $(x-1)$ (eventuell zum Quadrat) oder $(x-4)$ bestehen. Da es Grad 1 besitzt, kann es aber auch nur aus einem dieser irreduziblen Faktoren bestehen, was uns gerade die beiden genannten Möglichkeiten liefert.

3. Das folgt aus Dimensionsgründen: Wir wissen, dass $\dim(V^\infty(1, A|_U)) = 2$ sein muss, da der zugehörige Linearfaktor $x - 1$ im charakteristischen Polynom mit Potenz 2 vorkommt.

   Außerdem gilt $\{0_V\} \subsetneq V(1, A|_U) \subseteq V^2(1, A|_U) \subseteq V^\infty(1, A|_U)$. Also hat der Eigenraum $V(1, A|_U)$ mindestens Dimension 1 und maximal Dimension 2. Für Dimension 2 muss er dann schon gleich dem verallgemeinerten Eigenraum sein. Für Dimension 1 ist die Inklusion $V(1, A|_U) \subseteq V^2(1, A|_U)$ eine echte Inklusion und damit hat $V^2(1, A|_U)$ mindestens Dimension 2 und ist der verallgemeinerte Eigenraum.

Komprimierte Lösung zu Aufgabe 9.5 – Invariante Unterräume

Ist U A-invarianter Unterraum von V, so gilt $\text{charpol}_{A|_U} | \text{charpol}_A$.

Berechne das charakteristische Polynom:

$$\begin{aligned}
\text{charpol}_A(x) = \det(xI_3 - A) &= \det \begin{pmatrix} x-2 & 0 & -1 \\ 1 & x-1 & -3 \\ -2 & 0 & x-3 \end{pmatrix} \\
&= (x-2)(x-1)(x-3) - 2(x-1) \\
&= (x-1)[(x-2)(x-3) - 2] \\
&= (x-1)[x^2 - 5x + 4] \\
&= (x-1)^2(x-4).
\end{aligned}$$

A hat also die Eigenwerte 1 und 4.

Berechne nun die Eigenräume und die verallgemeinerten Eigenräume:

$$V(1,A) = \ker \begin{pmatrix} 1 & 0 & 1 \\ -1 & 0 & 3 \\ 2 & 0 & 2 \end{pmatrix} = \ker \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \langle e_2 \rangle,$$

$$\begin{aligned}
V^\infty(1,A) = & V^2(1,A) = \ker \begin{pmatrix} 1 & 0 & 1 \\ -1 & 0 & 3 \\ 2 & 0 & 2 \end{pmatrix}^2 \\
= & \ker \begin{pmatrix} 3 & 0 & 3 \\ 5 & 0 & 5 \\ 6 & 0 & 6 \end{pmatrix} = \ker \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \langle e_2, e_1 - e_3 \rangle,
\end{aligned}$$

$$\begin{aligned}
V^\infty(4,A) = & V(4,A) = \ker \begin{pmatrix} 2 & 0 & -1 \\ 1 & 3 & -3 \\ -2 & 0 & 1 \end{pmatrix} \\
= & \ker \begin{pmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & -\frac{5}{6} \\ 0 & 0 & 0 \end{pmatrix} = \langle 3e_1 + 5e_2 + 6e_3 \rangle.
\end{aligned}$$

Nun zu den A-invarianten Unterräumen U:

1. $\dim(U) = 0$, dann ist $U = \{0_V\}$. Damit ist U A-invariant.
2. $\dim(U) = 1$. Dann gibt es folgende Möglichkeiten:
   a) $\text{charpol}_{A|_U}(x) = (x-1)$ und $U = V(1,A) = \langle e_2 \rangle$
   b) $\text{charpol}_{A|_U}(x) = (x-4)$ und $U = V(4,A) = \langle 3e_1 + 5e_2 + 6e_3 \rangle$
3. $\dim(U) = 2$. Dann gibt es folgende Möglichkeiten:
   a) $\text{charpol}_{A|_U}(x) = (x-1)^2$ und $U = V^\infty(1,A) = \langle e_2, e_1 - e_3 \rangle$

b) $\mathrm{charpol}_{A|_U}(x) = (x-1)(x-4)$ und $U = V(1,A) \oplus V(4,A) = \langle e_2, \begin{pmatrix} 3 \\ 5 \\ 6 \end{pmatrix} \rangle$

4. $\dim(U) = 3$. Dann ist $U = V$ und damit A-invariant.

Sämtliche A-invarianten Unterräume sind also:

$$\{0_V\},\ \langle e_2 \rangle,\ \langle 3e_1 + 5e_2 + 6e_3 \rangle,\ \langle e_2, e_1 - e_3 \rangle,\ \langle e_2, 3e_1 + 5e_2 + 6e_3 \rangle,\ V.$$

Antworten zu den Verständnisfragen 9.6 – Rationale Jordannormalform

1. Ist das Minimalpolynom einer Matrix A durch $\mathrm{minpol}_A = g_1^{k_1} \cdot \ldots \cdot g_l^{k_l}$ mit irreduziblen g_i gegeben, so sind die Primärkomponenten definiert als $V_i := \ker(g_i(A)^{k_i})$.

   Da in unserem Fall die $k_i = 1$ sind, folgt $V_i = \ker(g_i(A)^1) = \ker(g_i(A))$.

2. V_1:

   1. Zeile mal (-1) nehmen, dann von der 2. abziehen.

   Diese neue 2. Zeile mal (-1) nehmen und von der 3. Zeile abziehen.

   4. Zeile mal (-1) nehmen und von der 5. Zeile abziehen.

   Neue 5. Zeile mal (-1) nehmen und von der 6. Zeile abziehen.

   Neue 6. Zeile durch (-6) teilen und dann von der 4. Zeile abziehen

   … und 3-mal zur 5. Zeile addieren.

   V_2:

   1. Zeile von der 2. abziehen.

   Die neue 2. Zeile von der 3. Zeile abziehen.

   Die neue 3. Zeile durch 2 teilen und dann zur 2. addieren

   … und von der 1. subtrahieren.

   4. Zeile von der 5. Zeile abziehen.

   Diese neue 5. Zeile von der 6. Zeile abziehen.

   V_3:

   1. Zeile von 2. und 3. Zeile abziehen.

   4. Zeile von 5. und 6. Zeile abziehen.

3. Da wir die beiden Minimalpolynome bereits in irreduzible Faktoren zerlegt haben, können wir daraus einfach das kleinste gemeinsame Vielfache ablesen. Man muss sich nur die einzelnen irreduziblen Faktoren in den Polynomen anschauen. Die höchste Potenz, mit der sie in einem der beiden Polynome vorkommen, ist dann gleich der Potenz, mit der sie auch im kgV vorkommen müssen. Da g_1, g_2, g_3 aber in den Minimalpolynomen jeweils nur mit Potenz 1 auftauchen, ergibt sich daraus $\mathrm{kgV}(\mathrm{minpol}_{A_1}, \mathrm{minpol}_{A_2}) = g_1 g_2 g_3$.

4. Für alle vorkommenden zyklischen Unterräume ist $e = 1$. Für $e = 1$ ist die zugehörige Matrix $M_B^B(A) = (P) \in M_{d\times d}(K)$ und somit taucht N dort nicht mehr auf.

Komprimierte Lösung zu Aufgabe 9.6 – Rationale Jordannormalform

Setze $W_1 := \langle e_1, e_2, e_3\rangle$ und $W_2 := \langle e_4, e_5, e_6\rangle$, sowie

$$A_1 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \text{ und } A_2 = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & -2 \\ 0 & 1 & -2 \end{pmatrix}.$$

Dann ist $V = W_1 \oplus W_2$ und $A_i = A\big|_{W_i}$ für $i = 1, 2$.

Nun liest man ab, da A_1, A_2 Begleitmatrizen sind:

$$\begin{aligned} \operatorname{charpol}_{A_1}(x) =& \operatorname{minpol}_{A_1}(x) = x^3 - 1 = (x-1)(x^2+x+1) \\ \operatorname{charpol}_{A_2}(x) =& \operatorname{minpol}_{A_2}(x) = x^3 + 2x^2 + 2x + 1 = (x+1)(x^2+x+1). \end{aligned}$$

Setze

$$g_1(x) := x - 1, \quad g_2(x) := x + 1, \quad g_3(x) := x^2 + x + 1,$$

dann sind g_1, g_2, g_3 über $\mathbb{R}$ irreduzibel.

Damit ist ferner:

$$\begin{aligned} \operatorname{charpol}_A =& \operatorname{charpol}_{A_1} \cdot \operatorname{charpol}_{A_2} = g_1 \cdot g_2 \cdot g_3^2, \\ \operatorname{minpol}_A =& \operatorname{kgV}(\operatorname{minpol}_{A_1}, \operatorname{minpol}_{A_2}) = \operatorname{kgV}(g_1 g_3, g_2 g_3) = g_1 g_2 g_3. \end{aligned}$$

Insbesondere ist $\operatorname{charpol}_A \neq \operatorname{minpol}_A$ und V ist damit *nicht* A-zyklisch.

Als Primärkomponenten erhält man:

$$\begin{aligned} V_1 &= \ker(g_1(A)) = V(A, 1) = \langle e_1 + e_2 + e_3\rangle, \\ V_2 &= \ker(g_2(A)) = V(A, -1) = \langle e_4 + e_5 + e_6\rangle \text{ und} \\ V_3 &= \ker(g_3(A)) = V_{3,1} \oplus V_{3,2}, \text{ wobei} \\ V_{3,1} &= \langle e_1 - e_2, e_2 - e_3\rangle \text{ und} \\ V_{3,2} &= \langle e_4 + e_5, e_5 + e_6\rangle. \end{aligned}$$

$V_{3,1}$ und $V_{3,2}$ sind A-zyklisch. Beachte

$$g_3(A) = A^2 + A + I = \left(\begin{array}{ccc:ccc} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 & -1 & 1 \end{array}\right).$$

Setze $P := \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$ (Begleitmatrix zu g_3). Da g_3 zur Potenz 1 im Minimalpolynom auftaucht, erhalten wir als rationale Jordannormalform:

$$\mathrm{rJNF}(A) = \left(\begin{array}{cc|cc} 1 & & & \\ & -1 & & \\ \hline & & P & \\ & & & P \end{array}\right).$$

Antworten zu den Verständnisfragen 9.7 – Selbstadjungierte Abbildungen

1. Da steht die Gaußsche Summe: $(n-1)+(n-2)+\ldots+1 = \sum_{i=1}^{n-1} i$. Und für diese haben Sie sicher schon einmal bewiesen, dass $\sum_{i=1}^{n-1} i = \frac{(n-1)\cdot n}{2}$ gilt.

2. Nun, zunächst folgt nur, da wir A in $\mathcal{H}'$ beliebig gewählt hatten, dass die Menge ein Erzeugendensystem bildet. Die lineare Unabhängigkeit der angegebenen Menge muss man sich eigentlich noch überlegen:

   Stellen wir die Nullmatrix als Linearkombination der Elemente der Menge dar:

   $$(0) = \sum_{l=1}^{n} a_{ll}E_{ll} + \sum_{j=1}^{n}\sum_{k=j+1}^{n} a_{jk}(E_{jk}+E_{kj}).$$

   Wenn wir die Matrizen noch einmal ausschreiben:

   $$\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix},$$

   dann folgt durch Vergleich der Einträge, dass alle $a_{ll} = 0$ und alle $a_{jk} = 0$ sind. Also kann man die Nullmatrix nur als triviale Linearkombination der Elemente der Menge schreiben. Das heißt aber gerade, dass die Menge linear unabhängig ist.

   Da die Menge dann sowohl linear unabhängig als auch ein Erzeugendensystem von $\mathcal{H}'$ ist, handelt es sich wirklich um eine Basis von $\mathcal{H}'$.

3. Man braucht die Existenz des Skalarproduktes, da mit dessen Hilfe die adjungierte Abbildung und die Orthonormalbasis (deren Existenz wir im Beweis benutzt haben) *definiert* sind. Die Voraussetzung ist also nötig, um mit den verwendeten Begriffen arbeiten zu dürfen.

Komprimierte Lösung zu Aufgabe 9.7 – Selbstadjungierte Abbildungen

Schreibe die Abbildungen aus $\mathcal{H}$ als $(n \times n)$-Matrizen über K bzgl. einer Orthonormalbasis B von V. Dann ist

$$\mathcal{H} \cong \{A \in M_{n\times n}(K) \mid A = \bar{A}^T\}.$$

1. Fall: $K = \mathbb{R}$

$\{A \in M_{n\times n}(K) \mid A = A^T\}$ hat die Basis

$$\{E_{jk} + E_{kj}, E_{ll} \mid 1 \leq l \leq n,\ 1 \leq j < k \leq n\}.$$

Dabei ist E_{jk} die Matrix mit 1 an Position (j,k) und Eintrag 0 an allen anderen Positionen.

Es gilt dann: $\dim(\mathcal{H}) = \binom{n}{2} + n = \frac{n(n-1)}{2} + n = \frac{n^2+n}{2}$.

2. Fall: $K = \mathbb{C}$

$\{A \in M_{n\times n}(K) \mid A = \bar{A}^T\}$ hat die Basis

$$\{E_{jk} + E_{kj}, i(E_{jk} - E_{kj}), E_{ll} \mid 1 \leq l \leq n,\ 1 \leq j < k \leq n\}.$$

Damit ist $\dim(\mathcal{H}) = 2\binom{n}{2} + n = n(n-1) + n = n^2$.

12 Ausführliche Musterlösungen zu Teil III

12.1 Themen aus den mathematischen Grundlagen

12.1.1 Lösen von Ungleichungen

In dieser Aufgabe kann man lernen, mit wichtigen, manchmal „unbeliebten" mathematischen Operationen und Symbolen umzugehen: *Wurzel*, *Betrag* und *Ungleichungen*. Sie sind vermutlich deshalb unbeliebt, weil sich bei der Arbeit mit ihnen leicht kleine Fehler und Ungenauigkeiten einschleichen können.

a) Wir sollen die Lösungsmenge der Gleichung $\sqrt{x+1} = x$ in $\mathbb{R}$ bestimmen. Gefährlich ist es hier, einfach auf beiden Seiten zu quadrieren. Dadurch würde man nämlich die Lösungsmenge verändern – Quadrieren ist keine Äquivalenzumformung!

Wir müssen uns daher zunächst die Gleichung genauer ansehen. Auf der linken Seite steht jedenfalls eine nichtnegative[1] Zahl, da die Wurzel aus einer Zahl immer nichtnegativ ist. (Auch das ist sehr wichtig zu wissen! Z. B. ist $\sqrt{4} = 2$ und *nicht* $\sqrt{4} = \pm 2$.) Damit steht aber auch auf der rechten Seite der Gleichung eine nichtnegative Zahl. Sonst würde ja keine Gleichheit herrschen. Wir wissen also schon: $x \geq 0$.

Außerdem dürfen wir (weil wir Lösungen in $\mathbb{R}$ suchen) die Wurzel nur aus nichtnegativen Zahlen ziehen. Es muss also $x + 1 \geq 0$ gelten. Das ist gleichbedeutend mit $x \geq -1$. Diese Bedingung ist durch die obige Forderung $x \geq 0$ aber automatisch abgesichert. Wir beachten nach dem Quadrieren, dass $x \geq 0$ gilt und beginnen zu rechnen:

$$\begin{aligned} & \sqrt{x+1} = x \\ \Leftrightarrow \quad & x + 1 = x^2 \quad \wedge \quad x \geq 0 \\ \Leftrightarrow \quad & x^2 - x - 1 = 0 \quad \wedge \quad x \geq 0. \end{aligned}$$

Lösen der quadratischen Gleichung liefert $x_{1,2} = \frac{1}{2} \pm \sqrt{\frac{1}{4} + 1} = \frac{1}{2} \pm \frac{\sqrt{5}}{2}$. Davon ist nur die erste Lösung nichtnegativ, also $x = \frac{1}{2} + \frac{\sqrt{5}}{2}$. Damit sind wir fertig: $\mathbb{L} = \{\frac{1}{2} + \frac{\sqrt{5}}{2}\}$.

[1] *Nichtnegativ* ist nicht dasselbe wie *positiv*! Die Zahl 0 ist nämlich *nichtnegativ*, sie ist aber nicht *positiv*. Nichtnegativ heißt also: ≥ 0.

C. Ableitinger, A. Herrmann, *Lernen aus Musterlösungen zur Analysis und Linearen Algebra*,
DOI 10.1007/978-3-658-02353-9_12,

b) Wir sollen die Lösungsmenge der Ungleichung $x+1 > x^2$ bestimmen. Diese Aufgabe kann man mit Hilfe einer kleinen Skizze ganz einfach lösen. Zunächst bringen wir alles auf eine Seite:

$$\begin{array}{rrcl} & x+1 & > & x^2 \\ \Leftrightarrow & -x^2+x+1 & > & 0 \\ \Leftrightarrow & x^2-x-1 & < & 0. \end{array}$$

Die quadratische *Gleichung* $x^2-x-1=0$ haben wir schon in a) gelöst: $x_{1,2} = \frac{1}{2} \pm \frac{\sqrt{5}}{2}$. In Abbildung 12.1 finden Sie eine graphische Darstellung des Polynoms x^2-x-1 mit seinen beiden Nullstellen (der Koeffizient von x^2 ist positiv, also ist die Parabel nach oben geöffnet). Wir fragen uns nun, für welche Werte von x das Polynom negative Werte annimmt! Das ist natürlich genau dann der Fall, wenn x zwischen den beiden Nullstellen liegt. Die Lösungsmenge der Ungleichung ist also das offene Intervall $\mathbb{L} = \left(\frac{1}{2} - \frac{\sqrt{5}}{2}, \frac{1}{2} + \frac{\sqrt{5}}{2}\right)$.

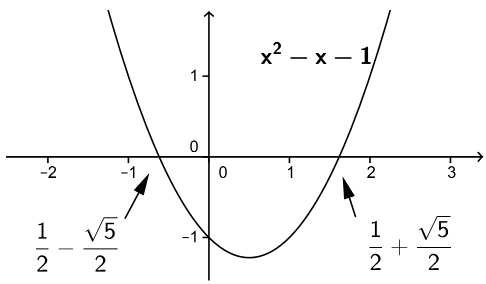

Abbildung 12.1: Graphische Darstellung des Polynoms x^2-x-1

c) Wir sollen die Lösungsmenge der Ungleichung $|8x-5| \leq |7x+15|$ bestimmen. Das ist die einzige der Aufgaben, die wirklich „Knochenarbeit" bedeutet! Wir müssen hier Fallunterscheidungen machen, je nachdem ob das, was zwischen den Betragsstrichen steht, negativ oder positiv ist. *Entweder* gilt:

1. Fall: $8x-5 \geq 0$ und $7x+15 \geq 0$. Das ist gleichbedeutend mit $x \geq \frac{5}{8}$ bzw. $x \geq -\frac{15}{7}$. In diesem Fall können wir einfach die Betragsstriche weglassen:

$$\begin{array}{rccl} & 8x-5 & \leq & 7x+15 \\ \Leftrightarrow & x & \leq & 20. \end{array}$$

Insgesamt erhalten wir $\mathbb{L}_1 = \left[\frac{5}{8}, 20\right]$. Damit sind nämlich alle drei Bedingungen erfüllt, wie man sich anhand einer Skizze deutlich machen kann.

Oder es gilt:

2. Fall: $8x-5 < 0$ und $7x+15 \geq 0$. Das ist gleichbedeutend mit $x < \frac{5}{8}$ bzw. $x \geq -\frac{15}{7}$. In diesem Fall müssen wir beim Wegnehmen der Betragsstriche der linken Seite ein negatives Vorzeichen verpassen:

$$\begin{array}{rccl} & -(8x-5) & \leq & 7x+15 \\ \Leftrightarrow & -10 & \leq & 15x \\ \Leftrightarrow & x & \geq & -\frac{2}{3}. \end{array}$$

Insgesamt erhalten wir $\mathbb{L}_2 = \left[-\frac{2}{3}, \frac{5}{8}\right)$.

Oder es gilt:

3. Fall: $8x - 5 < 0$ und $7x + 15 < 0$. Das ist gleichbedeutend mit $x < \frac{5}{8}$ bzw. $x < -\frac{15}{7}$. In diesem Fall bekommen beide Seiten beim Wegnehmen der Betragsstriche ein negatives Vorzeichen:

$$\begin{array}{rccc} & -(8x-5) & \leq & -(7x+15) \\ \Leftrightarrow & 5-8x & \leq & -7x-15 \\ \Leftrightarrow & x & \geq & 20. \end{array}$$

Eine Skizze macht deutlich, dass diese drei Bedingungen nicht gleichzeitig erfüllt sein können. Deshalb: $\mathbb{L}_3 = \emptyset$.

Oder es gilt:

4. Fall: $8x - 5 \geq 0$ und $7x + 15 < 0$. Hier brauchen wir gar nicht weiterzurechnen, da schon diese beiden Bedingungen nicht gleichzeitig erfüllt sein können. Also: $\mathbb{L}_4 = \emptyset$.

Oder-Verbindungen äußern sich bei der Gesamtlösungsmenge dadurch, dass man die Teillösungsmengen *vereinigt*. Es gilt also $\mathbb{L} = \mathbb{L}_1 \cup \mathbb{L}_2 \cup \mathbb{L}_3 \cup \mathbb{L}_4 = \left[-\frac{2}{3}, 20\right]$.

d) Wir sollen die Lösungsmenge der Gleichung $||x+1|-2| = 1$ bestimmen. Auch hier könnte man gleich eine Fallunterscheidung danach machen, ob der Term zwischen den Betragsstrichen positiv oder negativ ist. Es lohnt sich aber, diese Gleichung von einem höheren Standpunkt aus zu *interpretieren*. Dazu ein kleiner Exkurs:

Exkurs: Sehen wir uns die folgende einfache Aufgabe an: $|x - 5| = 2$. Wir suchen eine Zahl x, deren Abstand von der Zahl 5 gleich 2 ist (nichts anderes meint der Betrag: $|x - y|$ ist der Abstand zwischen x und y). Das ist aber leicht – welche Zahlen haben denn den Abstand 2 von der Zahl 5? Natürlich die Zahlen 3 und 7 (siehe Abbildung 12.2).

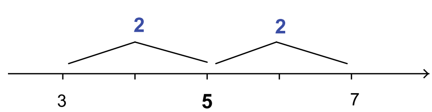

Abbildung 12.2: Suche nach Zahlen, die von der Zahl 5 den Abstand 2 haben

Damit zurück zu unserer Aufgabe. Wir setzen zunächst die gröbste Brille auf und interpretieren die äußeren Betragsstriche. Wir suchen also Zahlen $|x + 1|$, deren Abstand von der Zahl 2 gleich 1 ist. Welche Zahlen haben den Abstand 1 zur Zahl 2? Natürlich die Zahlen 1 und 3. Diese beiden Fälle betrachten wir getrennt:

1. Fall: $|x + 1| = 1$. Jetzt brauchen wir einen kleinen Trick, damit wir bei unserer „Abstandsinterpretation“ bleiben können. Wir schreiben $|x + 1|$ als $|x - (-1)|$ und suchen daher jetzt Zahlen x, die zur Zahl -1 einen Abstand von 1 haben. Das sind die Zahlen -2 und 0.

2. Fall: $|x+1| = 3$. Mit dem gleichen Trick wie oben können wir diese Gleichung schreiben als: $|x-(-1)| = 3$. Wir suchen also Zahlen x, deren Abstand von der Zahl -1 gleich 3 ist. Das sind die Zahlen -4 und 2.

Insgesamt erhalten wir also vier Zahlen, die diese Gleichung erfüllen:

$$\mathbb{L} = \{-4, -2, 0, 2\}.$$

12.1.2 Vollständige Induktion

Nachdem wir eine Aussage für natürliche Zahlen beweisen sollen, eignet sich ein Beweis mittels vollständiger Induktion[2].

Induktionsanfang: $n = 1$:

$$\sum_{k=0}^{1} \binom{1}{k} = \binom{1}{0} + \binom{1}{1} = 1 + 1 = 2 = 2^1$$

Damit ist der Induktionsanfang gezeigt.

Induktionsvoraussetzung: Wir nehmen nun an,

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

sei richtig für ein beliebiges, aber festes $n \in \mathbb{N}$.

(Mit der Formulierung der Induktionsvoraussetzung muss man immer vorsichtig sein! Wir setzen dabei voraus, dass die Formel für *ein bestimmtes* $n \in \mathbb{N}$ gelten soll – das ist nicht dasselbe wie die *allgemeine* Gültigkeit der Formel; es reicht also nicht, einfach die Aufgabenstellung noch einmal abzuschreiben und das als Induktionsvoraussetzung zu bezeichnen!)

Induktionsschritt: Wir möchten nun zeigen:

$$\sum_{k=0}^{n+1} \binom{n+1}{k} = 2^{n+1}.$$

Damit wir die Induktionsvoraussetzung anwenden können, müssen wir den Binomialkoeffizienten $\binom{n+1}{k}$ so umformen, dass oben statt $n+1$ nur noch n steht. An dieser Stelle müssen wir uns folgende Eigenschaft des Binomialkoeffizienten zu Nutze machen:

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

[2] Einfacher und eleganter kann man diese Gleichung aber auch mit Hilfe des Binomischen Lehrsatzes $(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}$ beweisen, indem man dort $a = b = 1$ setzt.

für alle $n, k \in \mathbb{N}$ mit $1 \leq k \leq n$.[3] Jetzt aber wenden wir diese Gleichung auf unsere Situation an:

$$\sum_{k=0}^{n+1} \binom{n+1}{k} \stackrel{?}{=} \sum_{k=0}^{n+1} \binom{n}{k} + \sum_{k=0}^{n+1} \binom{n}{k-1}.$$

Das Fragezeichen soll andeuten, dass die Gültigkeit des Gleichheitszeichens nicht unbedingt klar ist. Sehen wir uns die beiden entstandenen Summen einmal genauer an: Bei der ersten Summe kommt der Summand $\binom{n}{n+1}$ (für $k = n+1$) und bei der zweiten Summe der Summand $\binom{n}{-1}$ (für $k = 0$) vor. Beides wollen wir vermeiden, da solche Binomialkoeffizienten im Allgemeinen gar nicht definiert sind! Wir ziehen daher zunächst einmal aus der Ausgangssumme den ersten und den letzten Summanden heraus (diese beiden haben ja die Schwierigkeiten verursacht):

$$\sum_{k=0}^{n+1} \binom{n+1}{k} = \binom{n+1}{0} + \sum_{k=1}^{n} \binom{n+1}{k} + \binom{n+1}{n+1} = 1 + \sum_{k=1}^{n} \binom{n+1}{k} + 1.$$

Erst jetzt ziehen wir den Binomialkoeffizienten auseinander:

$$\cdots = 1 + \sum_{k=1}^{n} (\binom{n}{k} + \binom{n}{k-1}) + 1 = 1 + \sum_{k=1}^{n} \binom{n}{k} + \sum_{k=1}^{n} \binom{n}{k-1} + 1.$$

Als nächsten Schritt schreiben wir statt der beiden Einsen wieder geschickt Binomialkoeffizienten. In der ersten Summe „fehlt" nämlich der Binomialkoeffizient $\binom{n}{0}$, in der zweiten „fehlt" $\binom{n}{n}$. Definitionsgemäß sind diese beiden Binomialkoeffizienten jeweils gleich 1, genau wie gewünscht! In der zweiten Summe führen wir schließlich noch eine Indexverschiebung von k nach $k-1$ durch, sodass wir Folgendes erhalten:

$$\begin{aligned} \cdots &= \binom{n}{0} + \sum_{k=1}^{n} \binom{n}{k} + \sum_{k=0}^{n-1} \binom{n}{k} + \binom{n}{n} = \sum_{k=0}^{n} \binom{n}{k} + \sum_{k=0}^{n} \binom{n}{k} \\ &= 2 \cdot \sum_{k=0}^{n} \binom{n}{k} \stackrel{\text{IV}}{=} 2 \cdot 2^n = 2^{n+1}. \end{aligned}$$

Damit sind wir fertig.

Exkurs:

$$\begin{aligned} \binom{n}{k} + \binom{n}{k-1} &= \frac{n!}{k!(n-k)!} + \frac{n!}{(k-1)!(n-k+1)!} = \frac{n!(n-k+1) + kn!}{k!(n+1-k)!} \\ &= \frac{n!(n-k+1+k)}{k!(n+1-k)!} = \frac{(n+1)!}{k!(n+1-k)!} = \binom{n+1}{k} \end{aligned}$$

[3] Den Beweis dieser Aussage finden Sie am Ende der Lösung als Exkurs.

12.1.3 Mengenverknüpfungen

a) Eine intuitiv naheliegende und unter Studierenden meist beliebte, da anschauliche Strategie ist es, eine visuelle Darstellung zu solchen Aufgaben anzufertigen (siehe das Venn-Diagramm in Abbildung 12.3). Solche Visualisierungen können zwar hilfreich sein, um sich Zusammenhänge besser vorstellen zu können, einen allgemeingültigen Beweis ersetzen sie jedoch nicht. Es gibt genügend Beispiele, die zeigen, dass die Anschauung auch trügen kann. Wir werden daher in der folgenden Lösung auf anschauliche Argumente verzichten und lieber auf die formalen Definitionen der verwendeten Symbole zurückgreifen.

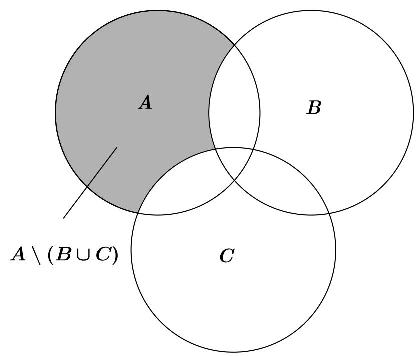

Abbildung 12.3: Venn-Diagramm zur Visualisierung der Menge $A \setminus (B \cup C)$ bzw. der Menge $(A \setminus B) \cap (A \setminus C)$

Den Beweis der Gleichheit zweier Mengen kann man zunächst einmal in zwei Teile gliedern. Zwei Mengen M_1 und M_2 gelten nämlich definitionsgemäß genau dann als gleich, wenn sowohl M_1 eine Teilmenge von M_2 als auch M_2 eine Teilmenge von M_1 ist.

Für unsere Aufgabe heißt das, dass wir zunächst $A \setminus (B \cup C) \subseteq (A \setminus B) \cap (A \setminus C)$ zeigen werden.

Aber wie zeigt man, dass eine Menge Teilmenge einer anderen Menge ist? Auch dazu muss man wieder bei einer Definition nachschlagen, nämlich bei jener der Teilmenge. Daraus gewinnt man folgende Strategie: Man wählt ein beliebiges Element aus der Menge M_1 und zeigt, dass dieses Element auch in M_2 liegt. Da das Element aus M_1 *beliebig* gewählt war, gilt der Beweis stellvertretend für alle Elemente aus M_1. Und wenn jedes Element aus M_1 auch in M_2 liegt, dann gilt $M_1 \subseteq M_2$.

Wir brechen mit dieser Strategie den Beweis letztlich auf die Ebene der Elemente herunter. Es ist üblicherweise leichter, Aussagen für einzelne Elemente zu formulieren, als für die gesamte Menge. Es macht uns den Beweis auf ganz natürliche Weise zugänglich.

1. Teil: Wir haben zu zeigen: $x \in A \setminus (B \cup C) \Rightarrow x \in (A \setminus B) \cap (A \setminus C)$.

Beweis: Wir starten dazu folgendermaßen: Sei x ein beliebiges Element aus $A \setminus (B \cup C)$. Mittels geschickter Umformungen müssen wir nun zeigen, dass x dann auch ein Element von $(A \setminus B) \cap (A \setminus C)$ ist.

Und schon wieder können wir direkt auf eine Definition zurückgreifen. Wir müssen nämlich übersetzen, was $x \in A \setminus (B \cup C)$ eigentlich bedeutet. Wir arbeiten uns von außen nach innen, d. h., wir übersetzen zunächst das Zeichen $\setminus$. $A \setminus (B \cup C)$ steht für die Differenzmenge aus A und $B \cup C$, das ist die Menge aller Elemente, die zwar in A, nicht aber in $B \cup C$ liegen. Also gelten für unser gewähltes x die Aussagen $x \in A$ und $x \notin (B \cup C)$.

Statt des Wortes „und" wird häufig auch das Symbol $\wedge$ verwendet, das anzeigt, dass sowohl die Aussage $x \in A$ als auch die Aussage $x \notin (B \cup C)$ wahr ist: $x \in A \wedge x \notin (B \cup C)$.

Als Nächstes benötigen wir die Definition des Zeichens $\cup$. $B \cup C$ steht für die Vereinigung der Mengen B und C. In dieser Menge liegen also Elemente, die entweder in B oder in C oder in beiden Mengen enthalten sind. Was soll dann aber $x \notin (B \cup C)$ bedeuten? x darf nun weder in B noch in C enthalten sein, sonst wäre es ja auch in $B \cup C$ enthalten. Wir dürfen demnach schreiben: $x \in A \wedge (x \notin B \wedge x \notin C)$.

Die drei Aussagen $x \in A$, $x \notin B$ und $x \notin C$ sind wahr. Wir dürfen sie jetzt auch in einer anderen Reihenfolge schreiben, da sie alle durch ein „und" verbunden sind. Das machen wir deshalb, um uns der gewünschten Aussage weiter zu nähern. Wir können schreiben: $x \in A \wedge x \notin B$ und $x \in A \wedge x \notin C$. Darin steckt letztlich keine neue Information – es lässt sich daran aber schon die Struktur der gewünschten Aussage erkennen. Jetzt ist es nur noch ein kleiner Schritt.

Wir wenden nochmal die Definition der Differenzmenge an, diesmal in die andere Richtung: $x \in A \wedge x \notin B$ kann man demnach schreiben als $x \in A \setminus B$ und $x \in A \wedge x \notin C$ als $x \in A \setminus C$. Nachdem beide Aussagen wahr sind, schreiben wir $x \in A \setminus B \wedge x \in A \setminus C$.

Als letzten Schritt arbeiten wir mit der Definition der Durchschnittsmenge zweier Mengen. Diese besteht nämlich aus allen Elementen, die sowohl in der einen als auch in der anderen Menge liegen. Genau dieser Fall liegt aber bei unserem Element x vor. Es liegt ja sowohl in $A \setminus B$ als auch in $A \setminus C$. Damit gilt insgesamt $x \in (A \setminus B) \cap (A \setminus C)$ und wir sind mit der einen Richtung fertig.

2. Teil: Wir haben zu zeigen, dass auch $(A \setminus B) \cap (A \setminus C)$ eine Teilmenge von $A \setminus (B \cup C)$ ist. Der Beweis verläuft ganz ähnlich wie jener der ersten Richtung. Wir gehen daher etwas flotter voran.

Auf die Ebene der Elemente herunter gebrochen lautet die zu beweisende Aussage: $x \in (A \setminus B) \cap (A \setminus C) \Rightarrow x \in A \setminus (B \cup C)$.

Beweis: Wir starten mit einem beliebigen Element x aus der Menge $(A \setminus B) \cap (A \setminus C)$ und arbeiten uns zunächst wieder von außen nach innen.

Die Definition der Durchschnittsmenge zweier Mengen erlaubt uns zu schreiben: $x \in A \setminus B \wedge x \in A \setminus C$.

Die Definition der Differenzmenge liefert uns: $(x \in A \land x \notin B) \land (x \in A \land x \notin C)$.

Jetzt können wir die Klammern weglassen, die Reihenfolge der einzelnen Aussagen vertauschen und die doppelt vorkommende Aussage $x \in A$ einmal weglassen. Der Informationsgehalt der Aussage ändert sich dadurch nicht: $x \in A \land x \notin B \land x \notin C$.

$x \notin B \land x \notin C$ bedeutet, dass x weder in B noch in C liegt und deshalb auch nicht in der Vereinigung der beiden Mengen. Wir erhalten: $x \in A \land x \notin (B \cup C)$.

Die Definition der Differenzmenge bringt uns nun zum Ziel: $x \in A \setminus (B \cup C)$.

b) Auch bei dieser Aufgabe handelt es sich um den Beweis einer Mengengleichheit. Wir teilen ihn wieder in zwei Teile auf.

1. Teil: Wir beweisen zunächst $\bigcap\limits_{i \in \mathbb{N}} (A_i \cup C) \subseteq \left(\bigcap\limits_{i \in \mathbb{N}} A_i\right) \cup C$ und formulieren die zu beweisende Aussage wieder auf der Ebene der Elemente: $x \in \bigcap\limits_{i \in \mathbb{N}} (A_i \cup C) \Rightarrow x \in \left(\bigcap\limits_{i \in \mathbb{N}} A_i\right) \cup C$.

Beweis: Wir starten bei der linken Aussage und wählen ein beliebiges Element $x \in \bigcap\limits_{i \in \mathbb{N}} (A_i \cup C)$. Die Definition des großen Durchschnittszeichens sagt aus, dass x in jeder einzelnen der Mengen $(A_i \cup C)$ liegt, ganz egal, welchen Index $i \in \mathbb{N}$ wir betrachten. Das können wir wie folgt ausdrücken: Für alle Indizes $i \in \mathbb{N}$ gilt $x \in (A_i \cup C)$.

Jetzt benutzen wir die Definition der Vereinigung zweier Mengen. $x \in (A_i \cup C)$ bedeutet demnach $x \in A_i \lor x \in C$. Unsere Aussage lautet dann: Für alle Indizes $i \in \mathbb{N}$ gilt $x \in A_i \lor x \in C$.

Wenn nun x nicht in C läge, müsste x in jeder der Mengen A_i liegen, damit die Aussage wahr bleibt. Insbesondere können wir schreiben: Für alle $i \in \mathbb{N}$ gilt $x \in A_i$ oder es gilt $x \in C$ (oder beide Aussagen sind wahr).

Die erste der beiden Aussagen können wir mit Hilfe einer Durchschnittsmenge schreiben und erhalten: $x \in \left(\bigcap\limits_{i \in \mathbb{N}} A_i\right)$ oder $x \in C$.

Jetzt müssen wir nur noch die durch das Wörtchen „oder“ verbundenen Aussagen mit Hilfe einer Vereinigungsmenge zusammenführen und erhalten $x \in \left(\bigcap\limits_{i \in \mathbb{N}} A_i\right) \cup C$.

2. Teil: Es bleibt jetzt noch zu zeigen: $\left(\bigcap\limits_{i \in \mathbb{N}} A_i\right) \cup C \subseteq \bigcap\limits_{i \in \mathbb{N}} (A_i \cup C)$.

Wir wählen zunächst wieder ein beliebiges Element der linken Menge: $x \in \left(\bigcap\limits_{i \in \mathbb{N}} A_i\right) \cup C$.

Die Definition der Vereinigungsmenge liefert uns $x \in \left(\bigcap\limits_{i \in \mathbb{N}} A_i\right) \lor x \in C$.

Nun schlüsseln wir auf, was das große Durchschnittszeichen hier bedeutet: Für jedes $i \in \mathbb{N}$ gilt: $x \in A_i$. Insgesamt haben wir damit abgeleitet, dass x in jeder der Mengen A_i liegt oder in der Menge C (dieses „oder“ ist hier nicht ausschließend gemeint).

Wenn nun aber x in jeder der Mengen A_i oder in C liegt, dann kann man auch sagen: Für jedes $i \in \mathbb{N}$ gilt: $x \in A_i \vee x \in C$.

Diese letzte Aussage lässt sich mit Hilfe einer Vereinigungsmenge kürzer schreiben: Für jedes $i \in \mathbb{N}$ gilt: $x \in A_i \cup C$.

x liegt also in jeder der Vereinigungsmengen $A_i \cup C$ und damit auch in deren Durchschnitt. Wir erhalten: $x \in \bigcap\limits_{i \in \mathbb{N}} (A_i \cup C)$ und sind fertig.

Man hätte sowohl in a) als auch in b) die beiden Richtungen des Beweises auch in einem gemeinsamen Schritt durchführen können. Für jede der getätigten Schlussfolgerungen gilt nämlich auch die Umkehrung, wie wir gesehen haben. Ein Beweis, der beide Richtungen gleichzeitig abhandelt, würde allerdings schnell unübersichtlich, wenn man ihn – so wie wir das gerade gemacht haben – in einer Mischung aus Worten und Symbolen aufschriebe. Sie sollten also unbedingt lernen, Beweise dieser Art auch rein formal zu notieren – es bleibt dann weniger Raum für Missverständnisse und der Beweis wird viel prägnanter.

12.1.4 Körperaxiome

Wir sollen als Erstes ein neutrales Element bzgl. der Addition bestimmen. Das neutrale Element ist dasjenige Element der Menge, das bei der Verknüpfung (hier: Addition) alle Elemente unverändert lässt. Das heißt, wir suchen ein Element $n \in \mathrm{Abb}(\mathbb{R}, \mathbb{R})$, für das $f + n = f$ gilt für *alle* Abbildungen $f \in \mathrm{Abb}(\mathbb{R}, \mathbb{R})$. Wie findet man dieses Element? Wir müssen erst einmal ein Gefühl für die Verknüpfung gewinnen. Dazu schauen wir uns diese genauer an. Wie ist die Addition zweier Abbildungen überhaupt definiert? Laut Aufgabenstellung werden die Abbildungen punktweise addiert. Ein Element x wird demnach durch die Summe der Abbildungen gerade auf die Summe der einzelnen Abbildungswerte abgebildet. Schauen wir uns dazu am besten ein Beispiel an: Nehmen wir mal die beiden Abbildungen f und g mit $f(x) = x^2$ und $g(x) = 5x - 3$, dann ist $f + g$ gegeben durch $(f + g)(x) = f(x) + g(x) = x^2 + (5x - 3)$. Für $x = 2$ gilt dann beispielsweise: $f(2) = 4$, $g(2) = 7$ und $(f + g)(2) = f(2) + g(2) = 4 + 7 = 11$. Es fällt auf, dass das Zeichen $+$ hier zwei verschiedene Rollen einnimmt. Es bezeichnet sowohl die bereits bekannte Addition auf $\mathbb{R}$ als auch die „neue“ Addition auf $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$. Um die beiden Additionen auseinander halten zu können, ist es zu Beginn evtl. sinnvoll, sie auch mit unterschiedlichen Zeichen zu versehen. Da meistens klar ist, wo welche Addition vorliegt, verzichten die Mathematiker auf eine solche Unterscheidung. Wir wollen sie hier dennoch machen. „+“ soll die Addition auf $\mathbb{R}$ bezeichnen, wie sie bereits aus der Schule mit all den geltenden Regeln bekannt ist. Die „neue“ Addition auf $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$ wollen wir hier mit $\oplus$ bezeichen. Das heißt, $f \oplus g$ ist wie folgt definiert:

$$(f \oplus g)(x) = f(x) + g(x) \quad \text{für alle } x \in \mathbb{R}.$$

Analog bezeichnen wir die Multiplikation auf $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$ mit $\odot$.

Kehren wir zurück zur eigentlichen Aufgabe, das neutrale Element n der Addition zu finden. Für dieses Element muss also nach unseren Vorüberlegungen

$$f(x) + n(x) = (f \oplus n)(x) = f(x)$$

für alle $x \in \mathbb{R}$ und für alle $f \in \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ gelten. Sehen Sie, wie n zu wählen ist, damit das funktioniert? Wir haben das Problem zunächst auf eine uns bekannte Situation zurückgeführt, denn $f(x)$, $n(x)$ und $f(x) + n(x)$ sind Elemente aus $\mathbb{R}$. Und dort wissen wir, dass $a + b = a$ nur für $b = 0$ gilt. Demnach muss $n(x) = 0$ sein und zwar für jedes $x \in \mathbb{R}$. Genau so können (und müssen) wir n dann auch definieren! Jetzt haben wir uns überlegt, wie wir n wählen sollten, und eigentlich ist auch schon klar, dass es mit diesem n (wir sprechen hier von der Nullabbildung) tatsächlich funktioniert. Aufschreiben würde ein Mathematiker dies nun aber genau umgekehrt. In etwa wie folgt:

Setze $n : \mathbb{R} \to \mathbb{R}$ mit $n(x) = 0$ für alle $x \in \mathbb{R}$. Dann gilt für alle $f \in \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ und alle $x \in \mathbb{R}$:

$$(f \oplus n)(x) = f(x) + n(x) = f(x) + 0 = f(x).$$

Also gilt $f \oplus n = f$ für alle $f \in \mathrm{Abb}(\mathbb{R}, \mathbb{R})$, weil zwei Abbildungen (hier: $f \oplus n$ und f) genau dann gleich sind, wenn sie an jeder Stelle übereinstimmen. Daraus folgt, dass n das neutrale Element bzgl. der Addition ist.

Ähnliche Überlegungen können wir nun für das neutrale Element der Multiplikation anstellen. Wir suchen eine Abbildung $e \in \mathrm{Abb}(\mathbb{R}, \mathbb{R})$, für die

$$f(x) \cdot e(x) = (f \odot e)(x) = f(x)$$

für alle $f \in \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ und alle $x \in \mathbb{R}$ gilt. Wie oben können wir mit Hilfe unserer Kenntnisse über reelle Zahlen nun folgern, dass $e(x) = 1$ für alle $x \in \mathbb{R}$ sein muss. Also werden wir e genau so definieren: Sei $e \in \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ mit $e(x) = 1$ für alle $x \in \mathbb{R}$. Dann gilt

$$(f \odot e)(x) = f(x) \cdot e(x) = f(x) \cdot 1 = f(x)$$

für alle $x \in \mathbb{R}$, woraus $f \odot e = f$ folgt, und zwar für alle $f \in \mathrm{Abb}(\mathbb{R}, \mathbb{R})$. Damit ist e wirklich das neutrale Element bzgl. der Multiplikation.

Nun sollen wir noch entscheiden, ob die Menge zusammen mit den beiden Verknüpfungen einen Körper bildet. Dazu müssen wir überprüfen, ob die Menge die Körperaxiome erfüllt. Werden sie *alle* erfüllt, so handelt es sich um einen Körper. Wird auch nur *eines nicht* erfüllt, so liegt kein Körper vor. Wir müssen also testen, ob alle Körperaxiome durch die vorliege Menge mit ihren Verknüpfungen erfüllt werden. Übertragen wir die Körperaxiome auf die Menge $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$ mit den Verknüpfungen $\oplus$ und $\odot$, so müssen wir folgende Punkte überprüfen:

1. $\oplus$ ist assoziativ, d. h. $(f \oplus g) \oplus h = f \oplus (g \oplus h)$ für alle $f, g, h \in \mathrm{Abb}(\mathbb{R}, \mathbb{R})$,
2. $\oplus$ ist kommutativ, d. h. $f \oplus g = g \oplus f$ für alle $f, g \in \mathrm{Abb}(\mathbb{R}, \mathbb{R})$,
3. Existenz eines neutralen Elements bzgl. $\oplus$, d. h. $\exists n \in \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ mit $f \oplus n = f$ für alle $f \in \mathrm{Abb}(\mathbb{R}, \mathbb{R})$.

4. Existenz des inversen Elements bzgl. $\oplus$, d. h. für alle $f \in \text{Abb}(\mathbb{R}, \mathbb{R})$ existiert ein $g \in \text{Abb}(\mathbb{R}, \mathbb{R})$ mit $f \oplus g = n$.
5. $\odot$ ist assoziativ, d. h. $(f \odot g) \odot h = f \odot (g \odot h)$ für alle $f, g, h \in \text{Abb}(\mathbb{R}, \mathbb{R})$,
6. $\odot$ ist kommutativ, d. h. $f \odot g = g \odot f$ für alle $f, g \in \text{Abb}(\mathbb{R}, \mathbb{R})$,
7. Existenz eines neutralen Elements bzgl. $\odot$, d. h. $\exists e \in \text{Abb}(\mathbb{R}, \mathbb{R}) \backslash \{n\}$ mit $f \odot e = f$ für alle $f \in \text{Abb}(\mathbb{R}, \mathbb{R})$.
8. Existenz des inversen Elements bzgl. $\odot$, d. h. für alle $f \in \text{Abb}(\mathbb{R}, \mathbb{R}) \backslash \{n\}$ existiert ein $g \in \text{Abb}(\mathbb{R}, \mathbb{R})$ mit $f \odot g = e$.
9. Distributivgesetz: Es gilt $f \odot (g \oplus h) = f \odot g \oplus f \odot h$ für alle $f, g, h \in \text{Abb}(\mathbb{R}, \mathbb{R})$.

Das 3. und das 7. Axiom wird durch die Menge erfüllt, wie wir bereits bewiesen haben. Die Gültigkeit der meisten anderen Axiome kann durch Zurückführen auf die entsprechenden Axiome in $\mathbb{R}$ nachgeprüft werden. Überlegen wir uns das beispielhaft an der Assoziativität der Addition:

Seien $f, g, h \in \text{Abb}(\mathbb{R}, \mathbb{R})$ beliebig. Es gilt dann für alle $x \in \mathbb{R}$:

$$\begin{aligned}
[(f \oplus g) \oplus h](x) &= (f \oplus g)(x) + h(x) \\
&= (f(x) + g(x)) + h(x) \\
&\stackrel{*}{=} f(x) + (g(x) + h(x)) \\
&= f(x) + (g \oplus h)(x) \\
&= [f \oplus (g \oplus h)](x).
\end{aligned}$$

Daraus folgt aber, dass $(f \oplus g) \oplus h = f \oplus (g \oplus h)$ ist. In der Gleichheitskette geht lediglich bei $*$ das Assoziativgesetz der Addition auf $\mathbb{R}$ ein, alle anderen Gleichheitszeichen beruhen auf der Definition von $\oplus$.

In ähnlicher Weise funktioniert dies auch für die Axiome 2, 5, 6 und 9. Lediglich bei den Axiomen, die eine Existenz fordern, kann es zu Problemen kommen. Als kritische Axiome kommen somit nur noch die Existenz des inversen Elements bzgl. der Addition bzw. der Multiplikation in Frage.

Für die Addition müssten wir zu jedem $f \in \text{Abb}(\mathbb{R}, \mathbb{R})$ ein $g \in \text{Abb}(\mathbb{R}, \mathbb{R})$ finden, für das $(f \oplus g)(x) = n(x) = 0$, also $f(x) + g(x) = 0$, für alle $x \in \mathbb{R}$ gilt. Durch Umformen der Gleichung nach g erhalten wir: $g(x) = -f(x)$ für alle $x \in \mathbb{R}$. Da g durch $g(x) := -f(x)$ wohldefiniert ist, finden wir also ein additiv inverses Element zu jedem f.

Für die Multiplikation können wir analoge Überlegungen anstellen. Für jede Abbildung $f \in \text{Abb}(\mathbb{R}, \mathbb{R}) \backslash \{n\}$ müssen wir demnach ein $g \in \text{Abb}(\mathbb{R}, \mathbb{R})$ finden, für das

$$f(x) \cdot g(x) = (f \odot g)(x) = e(x) = 1$$

für alle $x \in \mathbb{R}$ gilt. Genau wie oben können wir nun diese Gleichung nach g umstellen und erhalten:

$$g(x) = \frac{1}{f(x)}$$

für alle $x \in \mathbb{R}$. Ist das auch wohldefiniert? Nein, denn $f(x)$ kann durchaus auch Null sein. Für f ist zwar die Nullabbildung n ausgeschlossen, das heißt aber noch lange nicht, dass f nicht an bestimmten Stellen Null werden darf. Die Nullstellen einer Abbildung f machen uns also die Existenz des multiplikativen Inversen kaputt. Um dies nun noch stichhaltig zu formulieren, müssen wir eine Abbildung angeben, die eine Nullstelle besitzt und dann begründen, dass es zu dieser Abbildung kein Inverses geben kann. Dies funktioniert mit ziemlich vielen Abbildungen. Nehmen wir z. B. f mit $f(x) = x$ für alle $x \in \mathbb{R}$. Dann ist $f \neq n$ und es gilt $f(0) = 0$. f besitzt also eine Nullstelle.

Nun führen wir einen sauberen Widerspruchsbeweis. Angenommen, es gibt ein inverses Element g zur Abbildung f mit $f(x) = x$. Dann folgt aber

$$1 = e(x) = (f \odot g)(x) = f(x) \cdot g(x) = x \cdot g(x)$$

für alle $x \in \mathbb{R}$. Insbesondere gilt dann aber auch für $x = 0$:

$$1 = 0 \cdot g(0) = 0.$$

Und das ist ein Widerspruch, weil in der Menge der reellen Zahlen $1 \neq 0$ ist.

12.2 Themen aus der Analysis 1

12.2.1 Stetigkeit mit Epsilon und Delta

Zu untersuchen ist die Stetigkeit der Funktion f mit $f(x) = \frac{1}{x}$ auf dem Intervall $(0, \infty)$. Und zwar direkt mit der ε-δ-Definition.

Bemerkung: Zu Beginn sei erwähnt, dass man die Stetigkeit nur bei sehr einfachen Funktionen mit Hilfe der ε-δ-Definition zeigen kann. Für schwierigere (zusammengesetzte) Funktionen muss man zur Argumentation meist die Sätze über die Summe, das Produkt und/oder den Quotienten (kurz: die Komposition) stetiger Funktionen bzw. die zur ε-δ-Definition äquivalente Formulierung der Stetigkeit mit Hilfe von *Folgen* verwenden.

Nun aber zur Aufgabe: Um zu zeigen, dass eine Funktion auf einem ganzen Intervall stetig ist, muss man zeigen, dass sie in jedem einzelnen Punkt des Intervalls stetig ist. Wir wählen also irgendein beliebiges $x_0 \in (0, \infty)$ und zeigen, dass f in x_0 stetig ist.

Zur Erinnerung noch einmal die Definition der Stetigkeit, formuliert für unsere Situation: f heißt stetig in x_0 genau dann, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, sodass für alle $x \in (0, \infty)$ mit $|x - x_0| < \delta$ gilt: $|f(x) - f(x_0)| < \varepsilon$.

Wir müssen also zu jedem noch so kleinen $\varepsilon > 0$ ein passendes $\delta > 0$ angeben können, sodass aus der Voraussetzung $|x - x_0| < \delta$ folgt: $|f(x) - f(x_0)| < \varepsilon$.

Wir starten zunächst mit dem Ausdruck $|f(x) - f(x_0)|$. Schließlich möchten wir ja zeigen, dass dieser Ausdruck kleiner als das vorgegebene ε ist. Wir setzen ein und formen geschickt so um, dass der Ausdruck $|x - x_0|$ vorkommt:

$$|f(x) - f(x_0)| = \left|\frac{1}{x} - \frac{1}{x_0}\right| = \left|\frac{x_0 - x}{xx_0}\right| = \frac{|x - x_0|}{|xx_0|}.$$

Jetzt können wir nämlich ausnutzen, dass laut Voraussetzung $|x - x_0| < \delta$ gilt und danach noch den Betrag im Nenner weglassen, weil x und x_0 beide größer als Null sind:

$$\frac{|x - x_0|}{|xx_0|} < \frac{\delta}{|xx_0|} = \frac{\delta}{xx_0}.$$

Jetzt macht uns eigentlich nur noch eine Sache Probleme: Wir möchten ja, dass der Ausdruck $\frac{\delta}{xx_0}$ kleiner als ε ist. Wenn nun aber x sehr klein ist (über x wissen wir ja bis jetzt nichts weiter, als dass es in der δ-Umgebung des festen Punktes x_0 liegt), dann wird der Bruch sehr groß. Das müssen wir unbedingt verhindern! Aber wie?

Wir haben ja noch die Möglichkeit, δ frei zu wählen. Wir stehen dabei vor der Situation aus Abbildung 12.4.

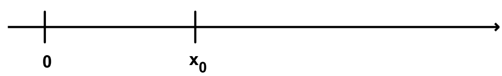

Abbildung 12.4: Ausgangslage vor der Wahl von δ

Um zu verhindern, dass x der Zahl 0 zu nahe kommt, können wir δ vorerst einmal kleiner oder gleich $\frac{x_0}{2}$ wählen (siehe Abbildung 12.5).

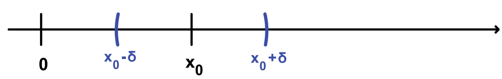

Abbildung 12.5: Situation nach der vorläufigen Wahl von δ

Damit haben wir erreicht, dass $x_0 - \frac{x_0}{2} < x < x_0 + \frac{x_0}{2}$ gilt. x kann also nicht mehr kleiner als $x_0 - \frac{x_0}{2} = \frac{x_0}{2}$ gewählt werden und damit kann auch der obige Bruch $\frac{\delta}{xx_0}$ nicht mehr beliebig groß werden. Geschafft! Mit $x > \frac{x_0}{2}$ gilt also:

$$\frac{\delta}{xx_0} < \frac{\delta}{\frac{x_0}{2}x_0} = \frac{\delta}{\frac{1}{2}x_0^2}.$$

Jetzt sind wir fast am Ziel. Wenn wir jetzt noch festsetzen, dass $\delta \leq \frac{1}{2}x_0^2\varepsilon$ gilt (und das dürfen wir – wir haben ja bei δ freie Wahl), dann gilt:

$$\frac{\delta}{\frac{1}{2}x_0^2} \leq \frac{\frac{1}{2}x_0^2\varepsilon}{\frac{1}{2}x_0^2} = \varepsilon.$$

Was haben wir also insgesamt über δ ausgesagt: Zum einen muss $\delta \leq \frac{x_0}{2}$ sein, zum anderen muss $\delta \leq \frac{1}{2}x_0^2\varepsilon$ gelten. Beide Bedingungen sind erfüllt, wenn wir $\delta := \min\{\frac{x_0}{2}, \frac{1}{2}x_0^2\varepsilon\}$ festsetzen.

Und genau dieses $\delta > 0$ ist auch in der komprimierten Musterlösung gewählt worden. Wenn man das macht, dann lassen sich nämlich alle obigen Umformungen durchführen und man erhält

$$|f(x) - f(x_0)| = \qquad \cdots \qquad < \qquad \cdots \qquad = \varepsilon,$$

was zu zeigen war!

12.2.2 Zwischenwertsatz

Wir sollen beweisen, dass es eine reelle Zahl x_0 gibt, die die Gleichung

$$x^5 + \frac{4}{1 + |x| + x^2} = 0$$

erfüllt.

Erstmal Vorsicht: Wir sollen die Gleichung *nicht lösen*, sondern nur beweisen, dass sie prinzipiell durch eine reelle Zahl x_0 *gelöst werden kann*!

Dazu eignet sich der Zwischenwertsatz. Wir definieren zu diesem Zweck die linke Seite der Gleichung als Funktion von x und zeigen, dass sie eine Nullstelle haben muss. (Hinweis: Falls einmal auf keiner der beiden Seiten der Gleichung Null steht, bringt man am besten zuerst alles auf eine gemeinsame Seite und definiert erst danach die Funktion!)

Sei also $f : \mathbb{R} \to \mathbb{R}$ definiert durch $f(x) := x^5 + \frac{4}{1+|x|+x^2}$. Jetzt beginnen wir damit, herauszufinden, ob die Funktion alle Voraussetzungen des Zwischenwertsatzes erfüllt.

f ist stetig, da $x \mapsto x^5$, $x \mapsto |x|$ und $x \mapsto 1 + x^2$ stetig sind und weil der Nenner für alle $x \in \mathbb{R}$ positiv ist.

Jetzt brauchen wir noch zwei Stellen a und b mit $f(a) < 0 < f(b)$. Dann gibt es nämlich nach dem Zwischenwertsatz eine Stelle x_0 zwischen a und b, die als Funktionswert $f(x_0) = 0$ hat. Wie findet man solche Stellen a und b?

Durch Probieren: In der komprimierten Musterlösung wurde -2 für a und 0 für b gewählt. (Im Allgemeinen wird man solche Zahlen a und b ausprobieren, für die die Funktionswerte $f(a)$ und $f(b)$ leicht zu berechnen sind. Bei uns ist also 0 eine praktische Wahl.) Es gilt $f(-2) = (-2)^5 + \frac{4}{1+|-2|+(-2)^2} = -32 + \frac{4}{7} < 0$. Passt! Und es gilt $f(0) = 0^5 + \frac{4}{1+|0|+0^2} = 4 > 0$. Passt auch!

Aus dem Zwischenwertsatz folgt nun, dass es eine reelle Zahl $x_0 \in (-2, 0)$ geben *muss*, die als Funktionswert die Zahl 0 hat. Schließlich liegt 0 ja zwischen $f(-2)$ und $f(0)$. Für dieses x_0 gilt also $f(x_0) = 0$. Insbesondere löst x_0 die Gleichung

$$x_0^5 + \frac{4}{1 + |x_0| + x_0^2} = 0$$

und wir sind fertig!

Es gibt demnach eine reelle Zahl, die die Gleichung löst. Welchen Wert diese Zahl hat, war nicht gefragt!

12.2.3 Funktionsgrenzwerte ohne l'Hospital

Bemerkung: Alle drei Teilaufgaben lassen sich mit den Regeln von de l'Hospital relativ leicht lösen. In der Aufgabenstellung wird allerdings explizit gefordert, die Aufgabe ohne die Regeln von de l'Hospital zu lösen. Dabei lernt man nämlich auch etwas über den Umgang mit Grenzwerten, was bei Anwendung der Regeln von de l'Hospital verborgen bliebe.

a) Wir untersuchen den Grenzwert $\lim\limits_{y \to +\infty} \frac{y}{e^y}$. Vom Gefühl her würde man wohl sagen, dass dieser Grenzwert Null ist („e^y wächst ja viel schneller als y"). Nachdem das Gefühl aber auch täuschen kann, müssen wir diese Vermutung nun formal beweisen:

e^y ist seiner Definition nach[4] die Reihe

$$e^y := \sum_{k=0}^{\infty} \frac{y^k}{k!} = 1 + y + \frac{y^2}{2!} + \frac{y^3}{3!} + \dots$$

Insbesondere gilt: $e^y \geq 1 + y + \frac{y^2}{2}$ für $y \geq 0$, da alle Summanden nichtnegativ sind. Damit sind wir aber fertig, denn wir erhalten

$$0 \leq \frac{y}{e^y} \leq \frac{y}{1 + y + \frac{y^2}{2}}$$

für $y \geq 0$ und der rechte Ausdruck geht mit $y \to +\infty$ gegen Null (das müsste man eigentlich noch zeigen. Mittlerweile können wir hier aber schon ein wenig salopper arbeiten – Sie wissen ja, wie man so etwas beweisen würde). Es gilt also

$$\lim_{y \to +\infty} \frac{y}{e^y} = 0$$

nach dem Einschließungskriterium, insbesondere existiert der Grenzwert. Fertig!

b) Hier ist der Grenzwert

$$\lim_{t \searrow 0} t \ln t$$

zu untersuchen. Bei dieser Aufgabe ist es schon schwieriger, richtig zu raten, was das Ergebnis sein wird. t konvergiert mit $t \searrow 0$ natürlich gegen Null, $\ln t$ allerdings gegen $-\infty$. Und über $0 \cdot (-\infty)$ können wir keine Aussage machen. Zur formalen Lösung der Aufgabe braucht man einen kleinen Trick (den man allerdings auch bei anderen Aufgaben gut gebrauchen kann):

Wir *substituieren* $t = e^{-y}$ in der Hoffnung, dass wir Aufgabe a) verwenden können (dort kommt ja auch der Term e^{-y} vor). Was wird dann aber aus dem zweiten Faktor,

[4] Zumindest wird e^x häufig so definiert. Sehen Sie in Ihrer Vorlesungsmitschrift nach, ob das auch bei Ihnen der Fall war bzw. ob das Resultat zumindest als Satz bewiesen wurde!

also aus $\ln t$? Wenn $t = e^{-y}$ ist, dann bewirkt das Anwenden des Logarithmus $\ln t = \ln e^{-y} = -y$. Insgesamt gilt also:

$$t \cdot \ln t = e^{-y} \cdot (-y) = -\frac{y}{e^y}.$$

Das sieht ja erfreulicherweise (bis auf das Minus) schon ganz nach der Funktion aus Aufgabe a) aus. Wir müssen nur noch überlegen, wogegen unsere neu definierte Variable y läuft. t soll laut Voraussetzung von oben gegen 0 laufen. Das ist aber gleichbedeutend damit, dass $-\ln t$ (also y) gegen $+\infty$ läuft. Anstatt den Grenzwert

$$\lim_{t \searrow 0} t \ln t$$

zu berechnen, können wir also auch den Grenzwert

$$\lim_{y \to +\infty} -\frac{y}{e^y}$$

berechnen. Den kennen wir aber aus Aufgabe a) und damit gilt:

$$\lim_{t \searrow 0} t \ln t = -\lim_{y \to +\infty} \frac{y}{e^y} = 0.$$

c) Die Idee in Teil b) war also, den gefragten Grenzwert durch Substitution auf einen schon bekannten Grenzwert zurückzuführen. Das klappt auch in Aufgabe c). Zuerst schreibt man den Term x^x um und erhält:

$$\lim_{x \searrow 0} \frac{x^x - 1}{x \ln x} = \lim_{x \searrow 0} \frac{e^{x \ln x} - 1}{x \ln x}.$$

Dieser Ausdruck ähnelt dem wichtigen Grenzwert

$$\lim_{t \to 0} \frac{e^t - 1}{t} = 1.$$

Das bringt uns auf die Idee der Substitution $t := x \ln x$. Wie in Aufgabe b) müssen wir nun noch überlegen, ob $t = x \ln x$ tatsächlich für $x \searrow 0$ wie gewünscht gegen Null strebt. Diese Information liefert uns aber glücklicherweise Aufgabe b). Wir können also tatsächlich schreiben:

$$\lim_{x \searrow 0} \frac{e^{x \ln x} - 1}{x \ln x} = \lim_{t \to 0} \frac{e^t - 1}{t} = e'(0) = 1.$$

Damit sind wir fertig.

12.2.4 Funktionenfolgen

a) Wir untersuchen die Funktionenfolge f_n mit $f_n(x) = x^n - x^{n+1}$ im Intervall $[0, 1]$ auf punktweise bzw. gleichmäßige Konvergenz.

Punktweise Konvergenz: Wir möchten wissen, wie sich die Funktionenfolge für ein fest gewähltes $x \in [0, 1]$ verhält, wenn n gegen ∞ strebt.

Nachdem sowohl $f_n(0) = 0^n - 0^{n+1} = 0$ als auch $f_n(1) = 1^n - 1^{n+1} = 0$ für alle $n \in \mathbb{N}$ gilt, haben wir schon einmal $\lim_{n\to\infty} f_n(0) = \lim_{n\to\infty} f_n(1) = 0$. An diesen beiden Stellen strebt also die Funktionenfolge punktweise gegen Null.

An allen anderen Stellen, also für $x \in (0,1)$, konvergiert aber sowohl x^n als auch x^{n+1} gegen Null und damit gilt: $\lim_{n\to\infty} f_n(x) = x^n - x^{n+1} = 0$.

Die Funktionenfolge f_n konvergiert folglich für alle $x \in [0,1]$ punktweise gegen $f : [0,1] \to \mathbb{R}$ mit $f \equiv 0$.

Gleichmäßige Konvergenz: Um die gleichmäßige Konvergenz nachzuweisen, verwenden wir nicht ihre Definition, sondern den folgenden Satz 12.1, der üblicherweise in Analysis-Vorlesungen bewiesen wird.

Satz 12.1 Sei D eine Menge, $f_n, f : D \to \mathbb{R}$ seien beschränkte Funktionen. Die Funktionenfolge $\{f_n\}_{n\in\mathbb{N}}$ konvergiert genau dann gleichmäßig gegen f, wenn

$$\lim_{n\to\infty} ||f_n - f||_\infty = 0.$$

$||f_n - f||_\infty$ ist eine Zahlenfolge, nämlich (laut Definition der Supremumsnorm) die Folge

$$\{ \sup_{x\in[0,1]} |f_n(x) - f(x)|\}_{n\in\mathbb{N}}$$

oder eingesetzt:

$$\{ \sup_{x\in[0,1]} (x^n - x^{n+1})\}_{n\in\mathbb{N}}$$

Was ist nun aber das Supremum der Menge $\{x^n - x^{n+1} | x \in [0,1]\}$? Das können wir bestimmen, indem wir das *Maximum* der Menge berechnen (stetige Funktionen nehmen auf kompakten Intervallen ja ein Maximum an):

Wir geben dazu dem Ausdruck $x^n - x^{n+1}$ einen Namen, um besser damit arbeiten zu können. Sei also $h_n(x) := x^n - x^{n+1}$. Wir wollen das Maximum dieser Funktion berechnen, deshalb setzen wir die erste Ableitung gleich Null:

$$h'_n(x) = nx^{n-1} - (n+1)x^n \overset{!}{=} 0 \Leftrightarrow x = \frac{n}{n+1} \text{ oder } x = 0.$$

Jetzt überprüfen wir noch, ob das Maximum tatsächlich an der Stelle $x = \frac{n}{n+1}$ liegt, oder aber am Rand des Intervalls $[0,1]$:

$h_n(0) = h_n(1) = 0$ und $h_n(\frac{n}{n+1}) = \left(\frac{n}{n+1}\right)^n - \left(\frac{n}{n+1}\right)^{n+1} > 0$. Also nimmt h_n das Maximum tatsächlich an der Stelle $x = \frac{n}{n+1}$ an. Insbesondere ist

$$\sup_{x\in[0,1]} (x^n - x^{n+1}) = \left(\frac{n}{n+1}\right)^n - \left(\frac{n}{n+1}\right)^{n+1}.$$

Jetzt müssen wir „nur“ noch davon den Grenzwert berechnen und hoffen, dass er gleich Null ist. Dazu formen wir ein wenig um:

$$\lim_{n\to\infty} \left(\frac{n}{n+1}\right)^n - \left(\frac{n}{n+1}\right)^{n+1} = \lim_{n\to\infty} \left(\frac{n}{n+1}\right)^n \left(1 - \frac{n}{n+1}\right).$$

Diesen Ausdruck schätzen wir jetzt nach oben und unten ab:

$$0 \leq \left(\frac{n}{n+1}\right)^n \left(1 - \frac{n}{n+1}\right) < 1 \cdot \left(1 - \frac{n}{n+1}\right).$$

Die rechte und die linke Seite der Ungleichungskette konvergieren mit $n \to \infty$ gegen Null und damit nach dem Einschließungskriterium auch der Ausdruck in der Mitte.

Die Funktionenfolge f_n konvergiert also auf dem Intervall $[0,1]$ sogar gleichmäßig gegen $f \equiv 0$.

b) Wir untersuchen jetzt die punktweise und gleichmäßige Konvergenz der Funktionenfolge g_n mit $g_n(x) = \frac{x}{\sqrt{x^2+\frac{1}{n}}}$ auf $(0, +\infty)$.

Punktweise Konvergenz: Für fest gewähltes $x \in (0, +\infty)$ gilt:

$$\lim_{n\to\infty} \frac{x}{\sqrt{x^2 + \frac{1}{n}}} = \frac{x}{\sqrt{x^2}} = \frac{x}{x} = 1.$$

Die Funktionenfolge konvergiert also auf $(0, +\infty)$ punktweise gegen die Funktion $g \equiv 1$.

Gleichmäßige Konvergenz: Wir haben die Konvergenz von

$$\sup_{x\in(0,+\infty)} |g(x) - g_n(x)|,$$

also im vorliegenden Fall von

$$\sup_{x\in(0,+\infty)} |1 - \frac{x}{\sqrt{x^2 + \frac{1}{n}}}|$$

zu untersuchen.

An dieser Stelle brauchen wir ein wenig mathematisches Gespür. Wenn eine Funktionenfolge *nicht* gleichmäßig konvergent ist, dann liegt das meistens an Schwierigkeiten am Rand des Definitionsbereiches. Sie kennen vielleicht das Beispiel der Funktionenfolge l_n mit $l_n(x) = x^n$ auf $[0,1]$. Diese ist zwar punktweise, aber nicht gleichmäßig konvergent. An der Stelle $x = 1$ hat die entsprechende Grenzfunktion nämlich eine „Sprungstelle", was gleichmäßige Konvergenz unmöglich macht.

Hier ist das ganz analog. Für $x \searrow 0$ (also wieder am Rand des Definitionsbereiches) gilt nämlich: $\lim_{x\searrow 0} g_n(x) = \frac{0}{\sqrt{0^2+\frac{1}{n}}} = 0$ für alle $n \in \mathbb{N}$. Dies erkennt man auch an einer graphischen Darstellung der Funktionenfolge (siehe Abbildung 12.6).

Wir werden versuchen, eine Nullfolge $x_n \in (0, +\infty)$ so zu konstruieren, dass der Abstand zwischen $g_n(x_n)$ und $g(x_n) = 1$ für alle $n \in \mathbb{N}$ einen bestimmten positiven Wert nicht unterschreitet. Das Supremum dieses Abstandes über alle $x \in (0, +\infty)$ kann dann nämlich nicht wie gewünscht gegen Null konvergieren. Und das machen wir jetzt:

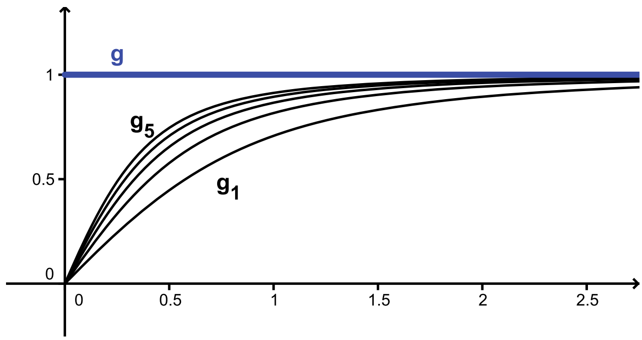

Abbildung 12.6: Graphische Darstellung der Funktionenfolge g_n aus Übung 10.8 b)

Sei $x_n := \frac{1}{\sqrt{n}}$ die angesprochene Nullfolge (warum wir hier diese Nullfolge wählen, werden Sie gleich sehen). Dann gilt nach der Definition des Supremums:

$$\sup_{x\in(0,+\infty)} |1 - \frac{x}{\sqrt{x^2+\frac{1}{n}}}| \geq 1 - \frac{x_n}{\sqrt{x_n^2+\frac{1}{n}}} = 1 - \frac{\frac{1}{\sqrt{n}}}{\sqrt{\frac{1}{n}+\frac{1}{n}}} = 1 - \frac{\frac{1}{\sqrt{n}}}{\sqrt{\frac{2}{n}}} = 1 - \frac{1}{\sqrt{2}}.$$

(Der Ausdruck $\frac{1}{\sqrt{n}}$ lässt sich also praktischerweise kürzen, das liegt eben an der geschickten Wahl der Nullfolge.) Nachdem also

$$\sup_{x\in(0,+\infty)} |1 - \frac{x}{\sqrt{x^2+\frac{1}{n}}}| \geq 1 - \frac{1}{\sqrt{2}}$$

für alle $n \in \mathbb{N}$ ist, kann die Folge

$$||g - g_n||_\infty = \sup_{x\in(0,+\infty)} |1 - \frac{x}{\sqrt{x^2+\frac{1}{n}}}|$$

insbesondere nicht gegen Null konvergieren.

Die Funktionenfolge ist also diesmal nicht gleichmäßig konvergent.

12.3 Themen aus Analysis 2

12.3.1 Grenzwert einer Funktion $f : \mathbb{R}^2 \to \mathbb{R}$

Wir sollen in dieser Aufgabe alle Punkte $(a, b) \in \mathbb{R}^2$ bestimmen, für die der Grenzwert

$$\lim_{(x,y)\to(a,b)} \frac{xy}{x+y}$$

existiert.

Es lohnt sich eine Fallunterscheidung. Warum? Weil viele Punkte $(a, b) \in \mathbb{R}^2$ gar keine großen Schwierigkeiten bereiten und wir diese Punkte zuerst ganz leicht abhandeln können (1. Fall). Danach können wir uns gezielt auf die problematischen Situationen konzentrieren (2. und 3. Fall).

1. Fall: Solange der Nenner des Bruches ungleich Null bleibt, bereitet der Grenzübergang keine Schwierigkeiten. Wir betrachten hier also vorerst den Fall $a + b \neq 0$. Der Grenzwert lässt sich dann einfach berechnen:

$$\lim_{(x,y)\to(a,b)} \frac{xy}{x+y} = \frac{ab}{a+b}.$$

Fertig!

Wir betrachten jetzt jene Fälle, bei denen der Nenner im Grenzübergang Null wird. Davon gibt es wieder zwei prinzipiell verschiedene – nämlich einen, bei dem der Zähler ungleich Null bleibt und einen, bei dem auch der Zähler Null wird.

2. Fall: Wir kümmern uns zunächst um den Fall $a + b = 0$, wobei $a \neq 0$ (also $b = -a$ und $a \neq 0$) ist.

Wollen wir jetzt den obigen Grenzwert berechnen, so müssen wir die Situation $(x, y) \to (a, -a)$ betrachten. Der Zähler xy konvergiert dann gegen $-a^2 \neq 0$, der Nenner $(x + y)$ gegen 0. Also existiert der Grenzwert

$$\lim_{(x,y)\to(a,b)} \frac{xy}{x+y} = \lim_{(x,y)\to(a,-a)} \frac{xy}{x+y}$$

in diesem Fall nicht.

3. Fall: Es bleibt jetzt nur noch eine Situation übrig, die wir durch obige Fälle noch nicht abgedeckt haben. Wir fassen jetzt also noch $a + b = 0$ ins Auge, wobei $a = 0$ (also $a = b = 0$) gilt.

Wir haben es demnach mit dem Grenzwert

$$\lim_{(x,y)\to(0,0)} \frac{xy}{x+y}$$

zu tun. In diesem Fall würden sowohl Nenner als auch Zähler gegen Null konvergieren. Wir können also nicht sofort entscheiden, ob der Grenzwert existiert. Wir müssen nun entweder zwei Nullfolgenpaare finden, entlang derer der Grenzwert unterschiedliche Werte annimmt. Oder aber wir können beweisen, dass der Grenzwert entlang jedes beliebigen Nullfolgenpaares derselbe ist.

Wir versuchen zunächst zwei solche Folgenpaare zu finden: Wählt man z. B. $\{(x_n, y_n)\} = \{(\frac{1}{n}, 0)\}$ und setzt ein, dann gilt

$$\lim_{n\to\infty} \frac{x_n \cdot y_n}{x_n + y_n} = \lim_{n\to\infty} \frac{\frac{1}{n} \cdot 0}{\frac{1}{n} + 0} = 0.$$

Jetzt sollten wir noch ein Folgenpaar finden, entlang dem eine von 0 verschiedene Zahl herauskommt. Das ist hier gar nicht so leicht. Wir hätten gerne, dass die „dominierenden" Nullfolgen in Zähler und Nenner den gleichen Grad haben. Was wir damit meinen, sieht man vielleicht besser in einer Rückschau.

Wir verraten Ihnen daher zunächst, wie man das Folgenpaar geschickt wählen kann und sehen uns im Nachhinein an, warum es funktioniert hat. Wähle also $\{(\tilde{x}_n, \tilde{y}_n)\} = \{(\frac{1}{n}, \frac{1}{n^2} - \frac{1}{n})\}$. Damit folgt

$$\lim_{n\to\infty} \frac{\tilde{x}_n \cdot \tilde{y}_n}{\tilde{x}_n + \tilde{y}_n} = \lim_{n\to\infty} \frac{\frac{1}{n} \cdot (\frac{1}{n^2} - \frac{1}{n})}{\frac{1}{n} + (\frac{1}{n^2} - \frac{1}{n})} = \lim_{n\to\infty} \frac{\frac{1}{n^3} - \frac{1}{n^2}}{\frac{1}{n^2}}.$$

Wir stehen nun vor einer Situation, bei der die „dominierenden" Nullfolgen in Zähler und Nenner den gleichen Grad haben (das ist durch die geschickte Wahl von $\tilde{y}_n$ gelungen, weil sich dadurch die Terme $\frac{1}{n}$ im Nenner aufgehoben haben). Wir erweitern den Bruch jetzt nur noch mit n^2 und sind fertig:

$$\lim_{n\to\infty} \frac{\frac{1}{n^3} - \frac{1}{n^2}}{\frac{1}{n^2}} = \lim_{n\to\infty} \frac{\frac{1}{n} - 1}{1} = -1.$$

Jetzt sehen Sie vielleicht ein, warum wir die Nullfolgen $\frac{1}{n^2}$ als dominierend bezeichnet haben. Die Nullfolge mit noch größerem Exponenten (nämlich $\frac{1}{n^3}$) spielt für die Grenzwertberechnung ja keine Rolle.

Jetzt nochmal zur Wahl von $\tilde{y}_n = \frac{1}{n^2} - \frac{1}{n}$. Warum musste man diesmal eine Differenz zweier Brüche wählen? Hätten wir für $\tilde{y}_n$ eine triviale Nullfolge genommen, die nur aus einem Bruch besteht (damit sind z. B. $\{\frac{1}{n}\}$ oder $\{\frac{1}{n^2}\}$, usw. gemeint), dann hätte die dominierende Nullfolge im Zähler immer größeren Grad gehabt als die Nullfolge im Nenner (weil im Zähler ja multipliziert, im Nenner aber nur addiert wird) und wir hätten wieder den Grenzwert Null erhalten.

So aber sind die beiden gewünschten Folgen gefunden und der Grenzwert existiert auch in diesem dritten Fall nicht.

12.3.2 Differenzierbarkeit einer Funktion $f : \mathbb{R}^3 \to \mathbb{R}^3$

Wir sollen die Differenzierbarkeit der Funktion f im Punkt a mit Hilfe der Definition der Differenzierbarkeit nachweisen.

Falls f im Punkt a differenzierbar ist, können wir nach Definition der Differenzierbarkeit schreiben:

$$f(a + \hat{x}) = f(a) + T(\hat{x}) + R(\hat{x})$$

mit $\lim\limits_{\hat{x}\to 0} \frac{R(\hat{x})}{||\hat{x}||} = 0$. Die lineare Abbildung T ist dann die gesuchte Ableitung $Df(a)$.

Wenn wir es also schaffen, unsere gegebene Funktion f so als Summe zu schreiben, dass wir einen konstanten Term $f(a)$ bekommen, einen linearen Funktionsterm $T(\hat{x})$ und einen Restterm $R(\hat{x})$, für den $\lim\limits_{\hat{x}\to 0} \frac{R(\hat{x})}{||\hat{x}||} = 0$ gilt, dann können wir den Funktionsterm der Ableitung $T(\hat{x})$ direkt ablesen.

Betrachten wir zunächst den Punkt $a = (0, 0, 0)$, definieren $\hat{x} := (x, y, z)$ und bringen

f auf die gewünschte Gestalt:

$$f(a+\hat{x}) = f(\hat{x}) = \begin{pmatrix} xz-y \\ x^2+z \\ x+2y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -y \\ z \\ x+2y \end{pmatrix} + \begin{pmatrix} xz \\ x^2 \\ 0 \end{pmatrix}.$$

Dabei ist der erste Ausdruck der konstante Vektor $f(a)$, der zweite Ausdruck enthält alle linearen Terme und der letzte Ausdruck die Terme höherer Ordnung (also Produkte aus den involvierten Variablen).

Wir haben den Funktionsterm $T(\hat{x})$ damit wahrscheinlich schon gefunden: Wir erwarten nämlich, dass $T(x,y,z) = \begin{pmatrix} -y \\ z \\ x+2y \end{pmatrix} = Df(0,0,0)\hat{x}$ gilt. Dazu müssen wir nur noch prüfen, ob $\lim\limits_{\hat{x}\to 0} \frac{1}{\sqrt{x^2+y^2+z^2}} \cdot \begin{pmatrix} xz \\ x^2 \\ 0 \end{pmatrix} = 0$ ist:

$$\left| \frac{1}{\sqrt{x^2+y^2+z^2}} \cdot \begin{pmatrix} xz \\ x^2 \\ 0 \end{pmatrix} \right| = \sqrt{\frac{x^2z^2+x^4}{x^2+y^2+z^2}} = |x|\sqrt{\frac{x^2+z^2}{x^2+y^2+z^2}} \leq |x| \to 0$$

für $\hat{x} \to 0$. Fertig.

Jetzt analog für den zweiten Punkt $\tilde{a} = (1,1,0)$:

$$\begin{aligned} f(\hat{x}+\tilde{a}) &= f(x+1,y+1,z) = \begin{pmatrix} (x+1)z-(y+1) \\ (x+1)^2+z \\ (x+1)+2(y+1) \end{pmatrix} = \begin{pmatrix} z+xz-1-y \\ 1+2x+x^2+z \\ 3+x+2y \end{pmatrix} \\ &= \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix} + \begin{pmatrix} z-y \\ 2x+z \\ x+2y \end{pmatrix} + \begin{pmatrix} xz \\ x^2 \\ 0 \end{pmatrix}. \end{aligned}$$

Der Restterm ist derselbe wie in der obigen Teilaufgabe. Wir haben also schon nachgewiesen, dass $\lim_{\hat{x}\to 0} \frac{R(\hat{x})}{||\hat{x}||} = 0$ gilt.

Wir erhalten $T(x,y,z) = \begin{pmatrix} z-y \\ 2x+z \\ x+2y \end{pmatrix} = Df(\tilde{a})\hat{x}$ als Funktionsterm der Ableitung und sind fertig.

12.3.3 Mehrdimensionale Kettenregel

a) Wir sollen zunächst die partiellen Ableitungen von f „berechnen". Über f haben wir die Information, dass

$$f(x,y) = g(x^2y, x+y, 2xy^3)$$

ist. f ist also die Verkettung zweier Funktionen. Zunächst wird nämlich der Punkt (x,y) auf $(x^2y, x+y, 2xy^3)$ abgebildet und danach die Funktion g auf diesen Punkt angewandt. Um das etwas besser fassen zu können, definieren wir eine Funktion h, die die erste dieser beiden Abbildungen darstellen soll:

$$h(x,y) := \begin{pmatrix} x^2y \\ x+y \\ 2xy^3 \end{pmatrix} =: \begin{pmatrix} u \\ v \\ w \end{pmatrix},$$

also $h : \mathbb{R}^2 \to \mathbb{R}^3$.

Dann gilt: $f = g \circ h$. Wir können dann mit Hilfe der Kettenregel die Ableitung $Df(x,y)$ ausrechnen und daraus die partiellen Ableitungen ablesen:

$$\begin{aligned}
Df(x,y) &= Dg(h(x,y)) \cdot Dh(x,y) = Dg(u,v,w) \cdot Dh(x,y) \\
&= \begin{pmatrix} \frac{\partial g}{\partial u}(u,v,w) & \frac{\partial g}{\partial v}(u,v,w) & \frac{\partial g}{\partial w}(u,v,w) \end{pmatrix} \cdot \begin{pmatrix} 2xy & x^2 \\ 1 & 1 \\ 2y^3 & 6xy^2 \end{pmatrix} \\
&= \begin{pmatrix} \frac{\partial g}{\partial u} \cdot 2xy + \frac{\partial g}{\partial v} + \frac{\partial g}{\partial w} \cdot 2y^3 & \frac{\partial g}{\partial u} \cdot x^2 + \frac{\partial g}{\partial v} + \frac{\partial g}{\partial w} \cdot 6xy^2 \end{pmatrix}.
\end{aligned}$$

Man liest ab: $\frac{\partial f}{\partial x}(x,y) = \frac{\partial g}{\partial u}(u,v,w) \cdot 2xy + \frac{\partial g}{\partial v}(u,v,w) \cdot 1 + \frac{\partial g}{\partial w}(u,v,w) \cdot 2y^3$

und $\frac{\partial f}{\partial y}(x,y) = \frac{\partial g}{\partial u}(u,v,w) \cdot x^2 + \frac{\partial g}{\partial v}(u,v,w) \cdot 1 + \frac{\partial g}{\partial w}(u,v,w) \cdot 6xy^2$.

b) Konkret sollen wir jetzt $Df(-1,1) = f'(-1,1)$ berechnen. Wir haben dazu

$$Dg(1,0,-2) = g'(1,0,-2) = (0,-1,1)$$

vorgegeben.

Damit wir das Resultat aus a) verwenden können, müsste laut Kettenregel $f(-1,1) = g(1,0,-2)$ gelten. Das ist aber der Fall, wie man durch Einsetzen leicht sieht.

Mit Aufgabe a) folgt jetzt

$$\begin{aligned}
Df(-1,1) &= \begin{pmatrix} 0 \cdot (-2) + (-1) \cdot 1 + 1 \cdot 2 & 0 \cdot 1 + (-1) \cdot 1 + 1 \cdot (-6) \end{pmatrix} \\
&= \begin{pmatrix} 1 & -7 \end{pmatrix}.
\end{aligned}$$

12.3.4 Lokale Umkehrbarkeit

Diese Aufgabe öffnet den Blick auf die Lösbarkeit von Gleichungssystemen und zeigt, wie dazu der Satz 7.2 über die lokale Umkehrbarkeit genutzt werden kann.

Wir sollen in dieser Aufgabe zeigen, dass sich das Gleichungssystem

$$\begin{aligned}
e^{2x+y} - \cos(xy) &= s \\
e^x - \cos(x+y) &= t
\end{aligned} \tag{12.1}$$

in einer Umgebung des Punktes $(s,t) = (0,0)$ lösen lässt.

Explizites Berechnen einer Lösung für ein gegebenes Wertepaar (s,t) ist leider nicht möglich (versuchen Sie es ruhig einmal). Sonst wären wir rasch fertig!

Wir werden stattdessen den Satz über die lokale Umkehrbarkeit einer Funktion verwenden, wenngleich die Situation derzeit nicht danach aussieht. Wir haben ja schließlich (noch) gar keine Funktion vor uns! Das werden wir aber jetzt ändern und definieren naheliegenderweise eine Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ durch:

$$f(x,y) := \begin{pmatrix} e^{2x+y} - \cos(xy) \\ e^x - \cos(x+y) \end{pmatrix}.$$

Um den Satz anwenden zu können brauchen wir zunächst einmal einen Punkt (x,y), der auf den gewünschten Punkt $(s,t) = (0,0)$ abgebildet wird – also eine Lösung des Gleichungssystems (12.1) für $(s,t) = (0,0)$. Aufgrund der komplizierten Gestalt des Gleichungssystems versucht man diese Lösung am besten nicht zu *berechnen*, sondern zu *sehen* bzw. zu *erraten*.

Der Punkt $(x,y) = (0,0)$ löst das Gleichungssystem (12.1) für $(s,t) = (0,0)$. Oder anders ausgedrückt: $f(0,0) = (0,0)$. Die Frage ist nun, ob man auch zu Wertepaaren in der Nähe des Punktes $(s,t) = (0,0)$ eine Lösung (x,y) angeben kann. Wir suchen also zu vorgegebenem Bildpunkt (s,t) den zugehörigen Punkt (x,y). Und das hört sich schon sehr nach Invertierbarkeit an!

Wir prüfen also die Voraussetzungen des Satzes 7.2. Offensichtlich ist f als Komposition stetig differenzierbarer Funktionen stetig differenzierbar und es gilt $f(0,0) = (0,0)$. Es ist also nur noch zu überprüfen, ob $\det Df(0,0) \neq 0$ ist:

$$Df(x,y) = \begin{pmatrix} e^{2x+y} \cdot 2 + \sin(xy) \cdot y & e^{2x+y} + \sin(xy) \cdot x \\ e^x + \sin(x+y) & \sin(x+y) \end{pmatrix},$$

$$\det Df(0,0) = \det \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} = -1 \neq 0.$$

Nach dem Satz über lokale Umkehrbarkeit existieren offene Umgebungen U von $(x,y) = (0,0)$ und V von $(s,t) = (0,0)$, sodass $f : U \to V$ bijektiv ist. Insbesondere existiert die Umkehrfunktion $f^{-1} : V \to U$ und damit zu jedem Wertepaar $(s,t) \in V$ ein Punkt $(x,y) \in U$ mit $f(x,y) = (s,t)$, also eine (in U sogar eindeutig bestimmte) Lösung des Gleichungssystems (12.1). Wir haben damit wie gewünscht eine offene Umgebung von $(s,t) = (0,0)$ gefunden, in der das Gleichungssystem lösbar ist.

12.4 Themen aus der Linearen Algebra 1

12.4.1 Gauß-Jordan-Algorithmus

Wir stellen zunächst die Koeffizientenmatrix des Gleichungssystems auf:

$$\left(\begin{array}{ccccc|c} 0 & 4 & 3 & 1 & 5 & 1 \\ 1 & 3 & 2 & 4 & 2 & 1 \\ 3 & 1 & 2 & 1 & 0 & 1 \\ 2 & 2 & 3 & -2 & 3 & 1 \end{array}\right).$$

Nach dem Gauß-Jordan-Algorithmus müssen wir nun an Position (1,1) [sprich: in der 1. Zeile und 1. Spalte; linke obere Ecke] eine Zahl ungleich Null schaffen und zwar durch Tauschen zweier Zeilen, falls nicht sowieso schon eine Zahl ungleich Null an der Position steht. Das geht natürlich nur, wenn überhaupt in irgendeiner Zeile der 1. Spalte eine Zahl ungleich Null steht. (Ist dies einmal nicht der Fall, dann muss man mit der 2. Spalte weitermachen.)

Was heißt das für unseren Fall? Wir haben an Position (1,1) eine Null stehen. In der 2. Zeile der 1. Spalte aber eine Eins, also eine Zahl ungleich Null, deshalb vertauschen wir nun die 1. und die 2. Zeile:

$$\left(\begin{array}{ccccc|c} 1 & 3 & 2 & 4 & 2 & 1 \\ 0 & 4 & 3 & 1 & 5 & 1 \\ 3 & 1 & 2 & 1 & 0 & 1 \\ 2 & 2 & 3 & -2 & 3 & 1 \end{array}\right).$$

Als Nächstes muss man an Position (1,1) eine Eins schaffen, das heißt, man muss die Zeile durch die Zahl teilen, die an dieser Position steht. Bei uns steht da aber schon eine Eins, also brauchen wir nichts zu machen.

Jetzt zieht man die erste Zeile a_{21}-mal von der 2. Zeile ab (a_{21} sei die Zahl, die in der 2. Zeile und 1. Spalte steht). Und genauso macht man das für alle anderen Zeilen auch.

Für uns heißt das: Wir ziehen 0-mal die 1. Zeile von der 2. Zeile ab. Wir ziehen 3-mal die 1. Zeile von der 3. Zeile ab. Wir ziehen 2-mal die 1. Zeile von der 4. Zeile ab. Und erhalten:

$$\left(\begin{array}{ccccc|c} 1 & 3 & 2 & 4 & 2 & 1 \\ 0 & 4 & 3 & 1 & 5 & 1 \\ 0 & -8 & -4 & -11 & -6 & -2 \\ 0 & -4 & -1 & -10 & -1 & -1 \end{array}\right).$$

So erhält man nämlich in der 1. Spalte lauter Nullen (in der 1. Zeile natürlich nicht!).

Gut, weiter geht's. Wir wiederholen das gleiche Spiel mit der Position (2,2). Steht dort eine Zahl ungleich Null? Ja, die Vier. Dann müssen wir jetzt die 2. Zeile durch 4 teilen, um dort eine Eins zu erhalten:

$$\left(\begin{array}{ccccc|c} 1 & 3 & 2 & 4 & 2 & 1 \\ 0 & 1 & \frac{3}{4} & \frac{1}{4} & \frac{5}{4} & \frac{1}{4} \\ 0 & -8 & -4 & -11 & -6 & -2 \\ 0 & -4 & -1 & -10 & -1 & -1 \end{array}\right).$$

Und jetzt ziehen wir wieder die entsprechenden Vielfachen der 2. Zeile von allen anderen Zeilen ab, sodass dort in der 2. Spalte nur noch Nullen stehen. (Hier also: 3-mal 2. Zeile von der 1. Zeile abziehen; 8-mal 2. Zeile zur 3. Zeile addieren; 4-mal 2.

Zeile zur 4. Zeile addieren.)

$$\left(\begin{array}{ccccc|c} 1 & 0 & -\frac{1}{4} & \frac{13}{4} & -\frac{7}{4} & \frac{1}{4} \\ 0 & 1 & \frac{3}{4} & \frac{1}{4} & \frac{5}{4} & \frac{1}{4} \\ 0 & 0 & 2 & -9 & 4 & 0 \\ 0 & 0 & 2 & -9 & 4 & 0 \end{array}\right)$$

Weiter geht's mit Position (3,3). Hier steht eine Zwei, also müssen wir die 3. Zeile durch 2 teilen:

$$\left(\begin{array}{ccccc|c} 1 & 0 & -\frac{1}{4} & \frac{13}{4} & -\frac{7}{4} & \frac{1}{4} \\ 0 & 1 & \frac{3}{4} & \frac{1}{4} & \frac{5}{4} & \frac{1}{4} \\ 0 & 0 & 1 & -\frac{9}{2} & 2 & 0 \\ 0 & 0 & 2 & -9 & 4 & 0 \end{array}\right).$$

Jetzt wird sie wieder mit entsprechenden Vielfachen von den anderen Zeilen abgezogen ($\frac{1}{4}$-mal zur 1. Zeile addiert; $\frac{3}{4}$-mal von der 2. abgezogen; 2-mal von der 4. abgezogen):

$$\left(\begin{array}{ccccc|c} 1 & 0 & 0 & \frac{17}{8} & -\frac{5}{4} & \frac{1}{4} \\ 0 & 1 & 0 & \frac{29}{8} & -\frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 1 & -\frac{9}{2} & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Und jetzt sind wir auch schon fertig, da die Koeffizientenmatrix in Treppennormalform vorliegt. Nun können wir die Lösungsmenge einfach daraus ablesen:

x_1, x_2, x_3 lassen sich mit Hilfe von x_4 und x_5 darstellen:

$$\begin{aligned} x_1 =& \frac{1}{4} - \frac{17}{8} x_4 + \frac{5}{4} x_5 \\ x_2 =& \frac{1}{4} - \frac{29}{8} x_4 + \frac{1}{4} x_5 \\ x_3 =& 0 + \frac{9}{2} x_4 - 2x_5. \end{aligned}$$

Für x_4 und x_5 darf man also beliebige Werte einsetzen. Damit ergibt sich die Lösungsmenge:

$$\mathbb{L} = \{ \begin{pmatrix} \frac{1}{4} - \frac{17}{8}s + \frac{5}{4}t \\ \frac{1}{4} - \frac{29}{8}s + \frac{1}{4}t \\ \frac{9}{2}s - 2t \\ s \\ t \end{pmatrix} \mid s, t \in \mathbb{R} \}.$$

12.4.2 Basisergänzung

Vorüberlegung: Damit die beiden Vektoren zu einer Basis ergänzt werden können, müssen sie selbst im Vektorraum liegen und linear unabhängig sein. Im Vektorraum liegen sie klarerweise und linear unabhängig sind sie auch, da für beliebige $\lambda, \mu \in \mathbb{Q}$ aus

$$\lambda \cdot \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} + \mu \cdot \begin{pmatrix} -2 \\ 0 \\ 5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

schon direkt $\lambda = 0$ und $5\mu = 0$, also auch $\mu = 0$ folgt.

Außerdem wissen wir, dass jede Basis von $\mathbb{Q}^3$ aus *drei* Vektoren besteht. Das bedeutet für unser Beispiel, dass die beiden Vektoren durch *einen* weiteren Vektor zu einer Basis ergänzt werden müssen.

Strategien:

1. Wir erraten einen Vektor v, der die beiden Vektoren zu einer Basis ergänzt, und zeigen auf gewohnte Weise, dass die Vektoren

   $$w_1,\ w_2,\ v$$

   tatsächlich eine Basis des $\mathbb{Q}^3$ bilden.

2. Wir machen uns den Austauschsatz von Steinitz und dessen Beweis zu Nutze und tauschen die beiden Vektoren gegen zwei Basisvektoren einer uns schon bekannten Basis des $\mathbb{Q}^3$ aus.

Strategie 1 wird für dieses Beispiel noch ziemlich gut funktionieren. Müsste man dagegen z. B. mehr als einen Vektor ergänzen, wird diese Strategie unzuverlässig. Deshalb wollen wir uns hier an Strategie 2 halten, da sie eine Vorgehensweise liefert, die für alle Vektorräume funktioniert, von denen man schon eine Basis kennt.

Wir überlegen uns zunächst, wie wir den Satz benutzen können. Dazu schauen wir ihn uns noch einmal an:

Satz 12.2 (Austauschsatz von Steinitz) Seien V ein K-Vektorraum (K ein Körper), $\{v_1, \ldots, v_n\}$ eine Basis von V und $w_1, \ldots, w_i$ ($i \geq 0$) linear unabhängige Vektoren aus V.

Dann existiert eine Indexmenge $I \subseteq \{1, \ldots, n\}$ mit $\#I = i$, sodass die Vektoren $w_1, \ldots, w_i$ zusammen mit allen Vektoren v_j mit $j \in \{1, \ldots, n\} \backslash I$ eine Basis von V bilden.

Interpretation des Satzes: Man nimmt i geeignete Vektoren aus der Basis $\{v_1, \ldots, v_n\}$ heraus. Die übrigen Vektoren bilden dann zusammen mit den Vektoren $w_1, \ldots, w_i$ eine neue Basis von V. Man kann also geeignet gewählte Vektoren der gegebenen Basis gegen die Vektoren $w_1, \ldots, w_i$ austauschen.

Wir können den Satz verwenden, wenn wir eine bekannte Basis des $\mathbb{Q}^3$ nehmen und jetzt zwei der Vektoren gegen w_1 und w_2 austauschen. Wir können dafür zum Beispiel die Standardbasis benutzen:

$$\{\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix}\}.$$

Bezogen auf den Satz haben wir also folgende Vektoren gegeben:

$$w_1=\begin{pmatrix}3\\1\\0\end{pmatrix},w_2=\begin{pmatrix}-2\\0\\5\end{pmatrix}\text{ und }v_1=\begin{pmatrix}1\\0\\0\end{pmatrix},v_2=\begin{pmatrix}0\\1\\0\end{pmatrix},v_3=\begin{pmatrix}0\\0\\1\end{pmatrix}.$$

Demnach ist $i = 2$ und wir müssen eine Indexmenge $I \subseteq \{1,2,3\}$ finden, sodass gilt: w_1, w_2 zusammen mit allen Vektoren v_j mit $j \in \{1,2,3\} \setminus I$ bilden eine Basis. (Dass es eine solche Indexmenge gibt, liefert uns der Satz.)

Das bedeutet aber nichts anderes, als dass wir zwei der drei Vektoren v_1, v_2, v_3 durch w_1 und w_2 ersetzen. Nun müssen wir noch herausbekommen, *welche* zwei der drei Vektoren wir dazu nehmen müssen. Das liefert uns der Satz nicht direkt, aber ein Blick in den Beweis zeigt uns den Weg:

Beweis des Austauschsatzes: Per Induktion nach i.

Induktionsanfang für $i = 0$: Trivial, da dann keine Vektoren ausgetauscht werden und die Ursprungsbasis $v_1, \ldots, v_n$ stehen bleibt.

Für $i > 0$ vollziehen wir einen Induktionsschritt von $i-1$ nach i. Das heißt, wenn wir $i-1$ linear unabhängige Vektoren austauschen können, dann können wir auch i linear unabhängige Vektoren austauschen.

Nach Induktionsvoraussetzung gibt es eine Indexmenge $I' \subseteq \{1, \ldots, n\}$ mit $\#I' = i-1$, sodass $w_1, \ldots, w_{i-1}$ zusammen mit allen v_j mit $j \in \{1, \ldots, n\} \backslash I'$ eine Basis von V bilden.

Da $w_1, \ldots, w_{i-1}$ zusammen mit allen v_j mit $j \in \{1, \ldots, n\} \backslash I'$ eine Basis bilden, können wir den Vektor w_i als Linearkombination dieser Basisvektoren schreiben:

$$w_i = \sum_{j=1}^{i-1} a_j w_j + \sum_{\substack{j=1\\ j \notin I'}}^{n} a'_j v_j. \tag{12.2}$$

Da die Vektoren $w_1, \ldots, w_i$ linear unabhängig sind, kann man w_i nicht als Linearkombination der Vektoren $w_1, \ldots, w_{i-1}$ schreiben. Das heißt, in der Darstellung (12.2) muss mindestens ein Koeffizient a'_j von 0 verschieden sein. Es existiert also ein $j_0 \in \{1, \ldots, n\} \backslash I'$ mit $a'_{j_0} \neq 0$.

Setzt man $I := I' \cup \{j_0\}$, dann bilden $w_1, \ldots, w_i$ zusammen mit allen v_j mit $j \notin I$ eine Basis von V.

Der Beweis dieser letzten Aussage ist für unsere Aufgabe nicht mehr von Belang.

Was bedeutet dieser Beweis für die vorliegende Aufgabe? Wir wenden den Induktionsschritt mehrfach an – erst von 0 nach 1, dann von 1 nach 2. Das heißt, wir gehen in zwei Schritten vor:

1. Schreibe w_1 als Linearkombination von v_1, v_2 und v_3. Überprüfe, welche Koeffizienten in der Linearkombination ungleich 0 sind (einen muss es laut Beweis des Satzes geben) und tausche dann den Vektor w_1 gegen den entsprechenden Vektor v_i aus. Wir erhalten die neue Basis $\{w_1, v_{i_1}, v_{i_2}\}$ $(i_1, i_2 \neq i)$.
2. Wir wiederholen Schritt 1 mit dem Vektor w_2 und der neuen Basis $\{w_1, v_{i_1}, v_{i_2}\}$.

Bemerkung: Wie man sieht, lässt sich diese Vorgehensweise beliebig oft wiederholen. Sie führt also auch zum Ziel, wenn man mehr als zwei Vektoren austauschen muss.

Ausführung:

1. Wir schreiben w_1 als Linearkombination von v_1, v_2, v_3, denn in der Induktionsvoraussetzung ist die Indexmenge $I' = \emptyset$ und die Vektoren v_1, v_2, v_3 bilden eine Basis:

$$\begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} = 3 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

   Demnach dürfen wir w_1 gegen v_1 oder v_2 tauschen, nicht aber gegen v_3, da dort in der Linearkombination der Koeffizient 0 steht. Wir tauschen hier gegen v_1.

   Das bedeutet in der Sprache des Satzes, dass $a'_1 = 3$, $a'_2 = 1$, $a'_3 = 0$ ist und wir $j_0 = 1$ wählen. Die neue Indexmenge wird wie folgt gebildet:

$$I = I' \cup \{j_0\} = \emptyset \cup \{1\} = \{1\}.$$

   Damit folgt aus diesem Induktionsschritt, dass w_1 zusammen mit allen v_j mit $j \in \{1,2,3\} \backslash I$ eine Basis von $\mathbb{Q}^3$ bildet. Oder anders ausgedrückt: $\{w_1, v_2, v_3\}$ ist Basis von $\mathbb{Q}^3$.

2. Wir machen nun den Induktionsschritt von 1 nach 2. Dabei wird die Indexmenge I aus dem 1. Schritt zur Indexmenge I' und wir wissen laut Induktionsvoraussetzung (und aus Schritt 1), dass $\{w_1, v_2, v_3\}$ eine Basis ist. Wir müssen nun w_2 als Linearkombination dieser drei Vektoren schreiben (das geht, weil sie eine Basis bilden):

$$\begin{pmatrix} -2 \\ 0 \\ 5 \end{pmatrix} = -\frac{2}{3} \cdot \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} + \frac{2}{3} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 5 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

   Da alle drei Koeffizienten ungleich 0 sind, dürfen wir alle drei Vektoren durch w_2 ersetzen. Aber aufgepasst: Natürlich ist es nicht sinnvoll, w_2 gegen w_1 zu tauschen, da wir ja eine Basis wollen, die beide Vektoren enthält! (Dass mindestens einer der anderen Vektoren einen Koeffizienten ungleich 0 haben muss, liefert uns wieder der Beweis des Satzes.) Wir können also zum Beispiel w_2 gegen v_3 tauschen und erhalten die neue Basis $\{w_1, w_2, v_2\}$.

   In der „Satzsprache" haben wir nun also $j_0 = 3$ gewählt, damit die Indexmenge $I = I' \cup \{j_0\} = \{1,3\}$ erhalten und wissen, dass w_1, w_2 zusammen mit allen v_j mit $j \in \{1,2,3\} \setminus I$ eine Basis bilden. Es gibt aber nur noch ein $j \in \{1,2,3\}$, das nicht in I liegt – nämlich $j = 2$. Und damit ergibt sich die neue Basis $\{w_1, w_2, v_2\}$.

Fazit: Wir haben die Vektoren w_1, w_2 durch den Vektor v_2 zu einer Basis von $\mathbb{Q}^3$ ergänzt. Dazu haben wir die Aussage und den konstruktiven Beweis des Austauschsatzes verwendet.

Bemerkung: Es gibt unendlich viele Vektoren, die w_1 und w_2 zu einer Basis des $\mathbb{Q}^3$ ergänzen. Man kann jeden Vektor nehmen, der nicht in der von w_1 und w_2 aufgespannten Ebene liegt. In unserem Beispiel haben wir die Standardbasis des $\mathbb{Q}^3$ gewählt. Würden wir eine andere Ausgangsbasis wählen, würden wir meistens auch einen anderen Ergänzungsvektor erhalten.

12.4.3 Invertierbarkeit von Matrizen

Strategie: Wir versuchen die Matrix A durch Zeilenumformungen in die Einheitsmatrix E_3 zu überführen. Die gleichen Umformungsschritte führen wir parallel an E_3 durch. Schaffen wir es, A in die Einheitsmatrix zu überführen, dann ist A invertierbar (denn es gilt dann $\mathrm{rg}(A) = 3$). Die inverse Matrix erhalten wir gerade als die parallel umgeformte Einheitsmatrix E_3.

Sollte A nicht invertierbar sein, wird es uns nicht gelingen, die Matrix in die Einheitsmatrix zu überführen. Wir werden stattdessen durch die Umformungen eine Matrix erhalten, bei der eine Nullzeile auftaucht. Diese Matrix hat dann keinen vollen Rang. Da sich der Rang durch die Zeilenumformungen nicht ändert, kann auch A keinen vollen Rang haben. Und damit ist A dann nicht invertierbar.

Wir erkennen also durch die Umformungen, die wir zur Bestimmung der inversen Matrix ohnehin machen müssten, ob die Matrix überhaupt invertierbar ist. Das heißt, es ist ökonomisch, gleich mit den Umformungen zu starten.

Durchführung: Wir schreiben die Matrizen A und E_3 nebeneinander:

$$\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & a & 0 & 1 & 0 \\ 1 & a & 1 & 0 & 0 & 1 \end{array}.$$

Jetzt formen wir um. Wir wollen links die Einheitsmatrix erhalten. Deswegen ist es sinnvoll, zunächst die 1. Zeile von der 2. Zeile und von der 3. Zeile abzuziehen:

$$\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & a-1 & -1 & 1 & 0 \\ 0 & a-1 & 0 & -1 & 0 & 1 \end{array}.$$

Hieran erkennen wir aber sofort, dass die Matrix für $a = 1$ nicht invertierbar ist. Denn ist $a = 1$, dann stehen in der 2. und 3. Zeile nur Nullen. Die Matrix hat dann Rang 1 und ist somit nicht invertierbar.

Für $a \neq 1$ können wir weiter umformen. Wir teilen die 2. und die 3. Zeile beide durch $a - 1$ (das dürfen wir, da ja nun $a - 1 \neq 0$ gilt):

$$\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -\frac{1}{a-1} & \frac{1}{a-1} & 0 \\ 0 & 1 & 0 & -\frac{1}{a-1} & 0 & \frac{1}{a-1} \end{array}.$$

Wir vertauschen die 2. und 3. Zeile:

$$\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -\frac{1}{a-1} & 0 & \frac{1}{a-1} \\ 0 & 0 & 1 & -\frac{1}{a-1} & \frac{1}{a-1} & 0 \end{array}.$$

Wir ziehen die 2. Zeile von der 1. Zeile ab:

$$\begin{array}{ccc|ccc} 1 & 0 & 1 & 1+\frac{1}{a-1} & 0 & -\frac{1}{a-1} \\ 0 & 1 & 0 & -\frac{1}{a-1} & 0 & \frac{1}{a-1} \\ 0 & 0 & 1 & -\frac{1}{a-1} & \frac{1}{a-1} & 0 \end{array}.$$

Und schließlich ziehen wir noch die 3. Zeile von der 1. Zeile ab:

$$\begin{array}{ccc|ccc} 1 & 0 & 0 & 1+\frac{2}{a-1} & -\frac{1}{a-1} & -\frac{1}{a-1} \\ 0 & 1 & 0 & -\frac{1}{a-1} & 0 & \frac{1}{a-1} \\ 0 & 0 & 1 & -\frac{1}{a-1} & \frac{1}{a-1} & 0 \end{array}.$$

Rechts steht nun die inverse Matrix von A:

$$A^{-1} = \begin{pmatrix} \frac{a+1}{a-1} & -\frac{1}{a-1} & -\frac{1}{a-1} \\ -\frac{1}{a-1} & 0 & \frac{1}{a-1} \\ -\frac{1}{a-1} & \frac{1}{a-1} & 0 \end{pmatrix} = \frac{1}{a-1} \begin{pmatrix} a+1 & -1 & -1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}.$$

12.4.4 Kommutierende Matrizen

Wir müssen also zeigen, dass nichtverschwindende Einträge in der Matrix A nur auf der Diagonalen auftreten. Außerdem müssen die Einträge auf der Diagonalen alle gleich sein. Da für jedes $X \in M_{n\times n}(K)$ $AX = XA$ gelten muss, wählen wir ein beliebiges und schauen, was wir dadurch für die Einträge der Matrix A erhalten. Am einfachsten ist es natürlich, sich X als eine Matrix mit möglichst vielen Nullen zu wählen, da dann das Matrizenprodukt einfach wird. Wir probieren das mal für die Matrix

$$X = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

aus. Wenn wir A als Matrix mit Einträgen a_{ij} schreiben ($A = (a_{ij})$), ist

$$AX = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{2n} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & 0 & \cdots & 0 \end{pmatrix}$$

und

$$XA = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{2n} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Da die beiden Matrizen nach Voraussetzung gleich sein müssen, folgt $a_{11} = a_{11}$ und $a_{12} = \ldots = a_{1n} = 0$ und $a_{21} = \ldots = a_{n1} = 0$.

Also hat A schon mal die Form

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Jetzt lassen wir die Eins in X auf der Diagonalen wandern. Dadurch erhalten wir ähnliche Ergebnisse. Das wollen wir nun formalisieren. Wir definieren $X = (x_{ij}) \in M_{n\times n}(K)$ wie folgt: Sei $i \in \{1, \ldots, n\}$ beliebig, aber fest gewählt und sei $x_{ii} = 1$ und $x_{lj} = 0$ für alle $(l, j) \neq (i, i)$. Dann gilt[5]:

$$(AX)_{lj} = \sum_{k=1}^{n} a_{lk} x_{kj} = \begin{cases} a_{li} & \text{für } j = i \\ 0 & \text{für } j \neq i \end{cases} \tag{12.3}$$

und

$$(XA)_{lj} = \sum_{k=1}^{n} x_{lk} a_{kj} = \begin{cases} a_{ij} & \text{für } l = i \\ 0 & \text{für } l \neq i \end{cases}. \tag{12.4}$$

Das heißt, die Matrizen haben die Form:

$$AX = \begin{array}{c} i\text{-te Spalte} \\ \downarrow \\ \begin{pmatrix} 0 & \cdots & 0 & a_{1i} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & a_{2i} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{ni} & 0 & \cdots & 0 \end{pmatrix} \end{array}$$

und

$$XA = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \\ a_{i1} & \cdots & a_{in} \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \\ \end{matrix}$$

[5] Mit $(AX)_{lj}$ ist der Eintrag der Matrix AX in Zeile l und Spalte j bezeichnet.

Da die beiden Matrizen aber wieder gleich sein müssen, gilt $a_{ii} = a_{ii}$, $a_{ij} = 0$ und $a_{li} = 0$ für alle $l, j \neq i$.

Kann man das auch aus den Gleichungen (12.3) und (12.4) ableiten? Klar, dafür müssen wir vier Fälle unterscheiden. Seien c_{lj} die Einträge der Matrix $(AX)_{lj} = (XA)_{lj}$, dann gilt:

1. Fall: $j = i, l = i$

$c_{lj} = a_{li} = a_{ii}$ und $c_{lj} = a_{ij} = a_{ii}$. Das liefert keine Information für A.

2. Fall: $j = i, l \neq i$

$c_{lj} = a_{li}$ und $c_{lj} = 0$, also $a_{lj} = 0$.

3. Fall: $j \neq i, l = i$

$c_{lj} = 0$ und $c_{lj} = a_{ij}$, also $a_{ij} = 0$.

4. Fall: $j \neq i, l \neq i$

$c_{lj} = 0$ und $c_{lj} = 0$, also keine Information für A.

Weil wir i aber beliebig gewählt hatten, heißt das, dass wirklich nur auf der Diagonalen Einträge ungleich Null stehen können. Mittlerweile wissen wir also, dass die Matrix A nur die Form

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

haben kann.

Es bleibt nun noch zu zeigen, dass die Einträge auf der Diagonalen alle gleich sind. Auch hier wollen wir das wieder mit möglichst wenig Aufwand, also einem „einfachen“ X, erreichen. Bisher haben wir die Matrizen betrachtet, die nur auf der Diagonalen *eine* Eins und sonst nur Nullen als Einträge hatten. Vielleicht reicht es ja für den nächsten Schritt diejenigen Matrizen zu betrachten, die auch nur eine Eins und sonst Nullen haben, in denen die Eins jedoch nicht auf der Diagonalen steht. Schauen wir uns das doch erst einmal genauer für

$$X = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

an. Dann ist

$$AX = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} 0 & a_{11} & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$XA = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \cdot \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix} = \begin{pmatrix} 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

und somit gilt $a_{11} = a_{22}$.

Für die anderen Fälle führen wir das wieder formal durch: Sei $X \in M_{n\times n}(K)$ so gewählt, dass für ein beliebiges, aber festes $i \in \{1, \ldots, n\}$ gilt: $x_{1i} = 1$ und $x_{lj} = 0$ für alle $(l, j) \neq (1, i)$. Dann ist:

$$(AX)_{lj} = \sum_{k=1}^{n} a_{lk}x_{kj} = \begin{cases} a_{l1} & \text{für } j = i \\ 0 & \text{für } j \neq i \end{cases},$$

$$(XA)_{lj} = \sum_{k=1}^{n} x_{lk}a_{kj} = \begin{cases} a_{ij} & \text{für } l = 1 \\ 0 & \text{für } l \neq 1 \end{cases}.$$

Da aber a_{lj} nur ungleich 0 ist, wenn $l = j$ gilt, folgt:

$$(AX)_{lj} = \begin{cases} a_{11} & \text{für } (l, j) = (1, i) \\ 0 & \text{für } (l, j) \neq (1, i) \end{cases},$$

und

$$(XA)_{lj} = \begin{cases} a_{ii} & \text{für } (l, j) = (1, i) \\ 0 & \text{für } (l, j) \neq (1, i) \end{cases}.$$

Da die Einträge nun wieder nach Voraussetzung gleich sind, gilt für $(l, j) = (1, i)$: $c_{lj} = a_{11} = a_{ii}$. Nun hatten wir i aber beliebig gewählt, damit gilt $a_{11} = \ldots = a_{nn}$. Setzen wir nun noch $a := a_{11}$, so folgt die Behauptung:

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{11} \end{pmatrix} = \begin{pmatrix} a & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a \end{pmatrix} = a \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$
$$= a \cdot E_n.$$

12.5 Themen aus der Linearen Algebra 2

12.5.1 Eigenwerte, Eigenräume, charakteristisches Polynom und Minimalpolynom

Da es einfacher ist, die geforderten Objekte für Matrizen zu bestimmen, stellt man als Erstes die zugehörige Abbildungsmatrix auf. Dazu benötigt man natürlich eine Basis des Vektorraums. Die Menge

$$B := \{x^0, x^1, x^2, \ldots, x^n\}$$

bildet gerade eine Basis von V. Um die Abbildungsmatrix aufzustellen, müssen wir wie gewohnt vorgehen und die Bilder der Basisvektoren als Linearkombination der Basisvektoren darstellen:

$$\begin{aligned}
A(x^0) &= 0,\\
A(x^1) &= 1 \cdot x^0,\\
A(x^2) &= 2 \cdot x^1,\\
A(x^3) &= 3 \cdot x^2,\\
&\vdots\\
A(x^n) &= n \cdot x^{n-1}.
\end{aligned}$$

Daraus erhält man dann die Darstellungsmatrix bzgl. der Basis B:

$$M_B^B(A) = \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 3 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 0 & n \\ 0 & \cdots & \cdots & 0 & 0 & 0 \end{pmatrix}.$$

Mit Hilfe dieser Matrix können wir das charakteristische Polynom von A bestimmen (die charakteristischen Polynome der Darstellungsmatrix und der Abbildung sind nämlich gleich!). Wegen der einfachen Form der Matrix kann man die Eigenwerte und das charakteristische Polynom direkt ablesen:

$$\mathrm{charpol}_A(x) = \mathrm{charpol}_{M_B^B(A)}(x) = \det(xE_n - M_B^B(A)) = x^{n+1},$$

da es sich um eine $(n+1) \times (n+1)$ obere *Dreiecksmatrix* handelt. $\mathrm{charpol}_A$ hat nur 0 als Nullstelle und damit ist 0 der einzige Eigenwert.

Der Eigenraum ist dann auch leicht abzulesen, denn x^0 liegt im Kern ($A(x^0) = 0$) und $\dim(\ker(A)) = \dim(V) - \mathrm{rg}(A) = (n+1) - n = 1$. Damit erzeugt x^0 schon den Kern:

$$V(0, A) = \ker(A - 0id_V) = \ker(A) = \langle x^0 \rangle.$$

Im Übrigen folgt dieses Ergebnis auch aus direkten Überlegungen, welche Elemente von A auf das Nullpolynom abgebildet werden.

Wir müssen uns nun noch überlegen, wie das Minimalpolynom aussieht. Da das Minimalpolynom das charakteristische Polynom teilt, muss

$$\mathrm{minpol}_A(x) = x^i$$

für ein $i \in \{0, 1, \ldots, x^{n+1}\}$ gelten. Es ist also zu überlegen, ob schon eine Potenz kleiner als $n+1$ die Nullabbildung ergibt ($i = 0$ kann dann natürlich nicht sein). Das ist aber genau dann der Fall, wenn alle Basisvektoren durch A^i auf Null (die Null ist hier das Null*polynom*) abgebildet werden, also insbesondere auch x^n. Da nun aber

$$\begin{aligned}
&A(x^n) = n \cdot x^{n-1} \neq 0,\\
&A^2(x^n) = n \cdot (n-1) \cdot x^{n-2} \neq 0,\\
&\vdots\\
&A^n(x^n) = n \cdot (n-1) \cdot \ldots 2 \cdot 1 \cdot x^0 \neq 0,\\
&A^{n+1}(x^n) = 0
\end{aligned}$$

gilt, muss das Minimalpolynom schon gleich dem charakteristischen Polynom sein. Also $\text{minpol}_A(x) = \text{charpol}_A(x) = x^{n+1}$.

12.5.2 Diagonalisierbarkeit

Eine Voraussetzung für die Diagonalisierbarkeit einer Matrix ist, dass das charakteristische Polynom vollständig in Linearfaktoren zerfällt. Allerdings ist diese Bedingung nicht hinreichend. Es gelten aber folgende Äquivalenzen: Eine Matrix ist genau dann diagonalisierbar, ...

1. ... wenn das Minimalpolynom nur einfache Nullstellen besitzt (also aus lauter verschiedenen Linearfaktoren besteht).
2. ... wenn sich der zugehörige Vektorraum als direkte Summe der Eigenräume schreiben lässt.
3. ... wenn die Dimension der Eigenräume jeweils mit der Vielfachheit des entsprechenden Eigenwerts als Nullstelle des charakteristischen Polynoms übereinstimmt.

Welche dieser drei Äquivalenzen (es lassen sich sogar noch mehr formulieren) man zum Beweis benutzt, ist egal.

Wir fangen damit an, die Diagonalisierbarkeit von A zu überprüfen: Es fällt auf, dass A aus zwei Blöcken besteht, nämlich aus $A_1 = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ und $A_2 = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$.

Nun gilt aber für solche Blöcke:

$$A \text{ diagonalisierbar } \Leftrightarrow A_1 \text{ und } A_2 \text{ sind diagonalisierbar.}$$

Kennt man diese Aussage nicht, kann man auch direkt mit der Matrix A arbeiten. Das ist natürlich deutlich aufwendiger, da wir es dann nicht mehr mit (2×2)-Matrizen sondern mit einer (4×4)-Matrix zu tun haben. Die einzelnen Schritte verlaufen aber genauso, wie wir sie hier gleich für die Matrix A_1 durchführen werden.

A_1 hat nun als einzigen Eigenwert 2, denn:

$$\text{charpol}_{A_1}(x) = \det(xE_2 - A_1) = \begin{vmatrix} x-2 & -1 \\ 0 & x-2 \end{vmatrix} = (x-2)^2$$

hat nur 2 als Nullstelle und die Nullstellen des charakteristischen Polynoms sind gerade die Eigenwerte.

Gleichzeitig sehen wir aber auch, dass das charakteristische Polynom in Linearfaktoren zerfällt. Die notwendige Bedingung für die Diagonalisierbarkeit ist also gegeben.

Wir müssen an dieser Stelle nun noch weitere Überlegungen anstellen.

Für das Minimalpolynom gibt es noch zwei Möglichkeiten (jeder Linearfaktor des charakteristischen Polynoms ist auch Linearfaktor des Minimalpolynoms und das Minimalpolynom muss das charakteristische Polynom teilen): $x-2$ oder $(x-2)^2$. Im ersten Fall wäre nach 1. A_1 diagonalisierbar, im zweiten Fall nicht.

Wäre $x-2$ das Minimalpolynom, so müsste aber A_1 Nullstelle des Polynoms sein. Dies ist aber nicht der Fall, denn $A_1 - 2E_2 = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Also ist $A_1 - 2E_2$ nicht die Nullmatrix und A_1 somit nicht Nullstelle des Polynoms $x-2$. Dieses kann damit nicht das Minimalpolynom sein. Somit ist $(x-2)^2$ das Minimalpolynom, welches 2 als doppelte Nullstelle besitzt. Daraus folgt, dass A_1 und damit auch A nicht diagonalisierbar ist.

Kennt man das Minimalpolynom nicht, kann man alternativ auch den Eigenraum (oder dessen Dimension) zum einzigen Eigenwert 2 berechnen:

$$V(2, A_1) = \ker(A_1 - 2E_2) = \ker \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rangle.$$

Man sieht, dass die Dimension des Eigenraumes 1 ist, die Vielfachheit von 2 als Nullstelle des charakteristischen Polynoms ist hingegen 2. Somit ist A_1 nicht diagonalisierbar.

Alle Berechnungen und Überlegungen, die wir gemacht haben, waren unabhängig davon, ob wir den Körper $\mathbb{Q}$ oder $\mathbb{C}$ betrachten. Deshalb ist A tatsächlich in beiden Fällen nicht diagonalisierbar.

Überprüfen wir nun die Diagonalisierbarkeit von B:

B lässt sich leider nicht in kleinere Blöcke zerlegen, deshalb müssen wir die ganze Matrix betrachten. Auch hier schauen wir uns als Erstes das charakteristische Polynom von B an:

$$\begin{aligned}
\mathrm{charpol}_B(x) = \det(xE_4 - B) &= \begin{vmatrix} x & -1 & 0 & 0 \\ 0 & x & 0 & -1 \\ -1 & 0 & x & 0 \\ 0 & 0 & -1 & x \end{vmatrix} \\
&\overset{\text{Lapl. 1. Zeile}}{=} x \cdot \begin{vmatrix} x & 0 & -1 \\ 0 & x & 0 \\ 0 & -1 & x \end{vmatrix} + 1 \cdot \begin{vmatrix} 0 & 0 & -1 \\ -1 & x & 0 \\ 0 & -1 & x \end{vmatrix} \\
&\overset{\text{Sarrus}}{=} x^4 - 1 = (x^2-1)(x^2+1) = (x-1)(x+1)(x^2+1).
\end{aligned}$$

Über $\mathbb{Q}$ hat x^2+1 keine Nullstellen, über $\mathbb{C}$ die Nullstellen i und $-i$. Damit zerfällt das charakteristische Polynom über $\mathbb{Q}$ nicht in Linearfaktoren und B ist somit über $\mathbb{Q}$ nicht diagonalisierbar.

Über $\mathbb{C}$ zerfällt das charakteristische Polynom in lauter verschiedene Linearfaktoren. Diese Linearfaktoren tauchen dann aber auch im Minimalpolynom von B auf und zwar auch maximal einmal, weil das Minimalpolynom das charakteristische Polynom teilen muss. Demnach zerfällt auch das Minimalpolynom von B in lauter verschiedene Linearfaktoren. Das bedeutet aber, dass B in diesem Fall diagonalisierbar ist.

Ohne Minimalpolynom kann man wie folgt argumentieren:

Wir haben die Eigenwerte von B: $1, -1, i, -i$. Sie sind alle einfache Nullstellen des charakteristischen Polynoms. Wenn wir uns nun noch überlegen können, dass die Eigenräume jeweils auch Dimension 1 haben, sind wir fertig.

Das könnten wir natürlich für jeden Eigenwert einzeln machen, wir können die vier Fälle aber auch zusammen abhandeln:

Sei a einer dieser Eigenwerte, dann gilt:

$$V(a,B) = \ker(B - aE_4) = \ker \begin{pmatrix} -a & 1 & 0 & 0 \\ 0 & -a & 0 & 1 \\ 1 & 0 & -a & 0 \\ 0 & 0 & 1 & -a \end{pmatrix}$$

$$\overset{\text{3.\&1.Z. vertauscht}}{=} \ker \begin{pmatrix} 1 & 0 & -a & 0 \\ 0 & -a & 0 & 1 \\ -a & 1 & 0 & 0 \\ 0 & 0 & 1 & -a \end{pmatrix}$$

$$\overset{(a\text{-mal 1.Z})+3.\text{Z}}{=} \ker \begin{pmatrix} 1 & 0 & -a & 0 \\ 0 & -a & 0 & 1 \\ 0 & 1 & -a^2 & 0 \\ 0 & 0 & 1 & -a \end{pmatrix}$$

$$\overset{\text{2.\& 3.Z. vertauscht}}{=} \ker \begin{pmatrix} 1 & 0 & -a & 0 \\ 0 & 1 & -a^2 & 0 \\ 0 & -a & 0 & 1 \\ 0 & 0 & 1 & -a \end{pmatrix}$$

$$\overset{(a\text{-mal 2.Z.})+3.\text{Z.}}{=} \ker \begin{pmatrix} 1 & 0 & -a & 0 \\ 0 & 1 & -a^2 & 0 \\ 0 & 0 & -a^3 & 1 \\ 0 & 0 & 1 & -a \end{pmatrix}$$

$$\overset{\text{3.\& 4.Z. vertauscht}}{=} \ker \begin{pmatrix} 1 & 0 & -a & 0 \\ 0 & 1 & -a^2 & 0 \\ 0 & 0 & 1 & -a \\ 0 & 0 & -a^3 & 1 \end{pmatrix}$$

$$\overset{(a^3\text{-mal 3.Z.})+4.\text{ Z.}}{=} \ker \begin{pmatrix} 1 & 0 & -a & 0 \\ 0 & 1 & -a^2 & 0 \\ 0 & 0 & 1 & -a \\ 0 & 0 & 0 & 1-a^4 \end{pmatrix}.$$

Die ersten drei Zeilen der Matrix sind linear unabhängig. Da nun aber a^4 für alle Eigenwerte jeweils 1 ist, stehen in der letzten Zeile nur Nullen. Also hat wirklich jeder Eigenraum die Dimension 1.

Mit etwas Hintergrundwissen kann man die Berechnung der Eigenräume umgehen: Jeder Eigenraum hat mindestens die Dimension 1, weil zu jedem Eigenwert ein Eigenvektor gehört, ein Eigenvektor nie der Nullvektor ist und ein Eigenvektor immer im Eigenraum liegt. Also enhält jeder Eigenraum einen vom Nullvektor verschiedenen Vektor und muss somit eine Dimension größer oder gleich 1 haben.

Da einerseits die Eigenräume und damit auch die Summe der Eigenräume immer Unterräume des Vektorraums sind und andererseits die verschiedenen Eigenräume auch immer nur den Nullvektor gemeinsam haben, ist die Summe der Dimensionen der Eigenräume gleich der Dimension der Summe der Eigenräume (direkte Summe!). Und die ist wiederum kleiner als die Dimension des Vektorraums. Formal:

$$\sum_{i=1}^{k} \dim(V(\lambda_i, A)) = \dim(\bigoplus_{i=1}^{k} V(\lambda_i, A)) \leq \dim V.$$

Also folgt in diesem Beispiel:

$$\begin{aligned} 4 &\leq \dim(V(1,B)) + \dim(V(-1,B)) + \dim(V(i,B)) + \dim(V(-i,B)) \\ &= \dim[V(1,B) \oplus V(-1,B) \oplus V(i,B) \oplus V(-i,B)] \\ &\leq 4, \end{aligned}$$

woraus $\dim(V(a,B)) = 1$ folgt für $a \in \{1, -1, i, -i\}$.

Bemerkung: Diese Argumentation funktioniert im Übrigen ganz allgemein und liefert die Aussage „Zerfällt das charakteristische Polynom einer Matrix in lauter verschiedene Linearfaktoren, so ist die Matrix bereits diagonalisierbar“. (Die Umkehrung dieser Aussage gilt im Allgemeinen nicht!)

12.5.3 Linearformen

Für diese Aufgabe brauchen wir einige Eigenschaften von Vektorräumen, Unterräumen und Faktorräumen. Nämlich:

1. Ist $U \subseteq V$ ein Unterraum eines endlich-dimensionalen K-Vektorraums V, dann gilt:

   a) $\dim(U) \leq dim(V)$.

   b) Ist $\dim(U) = \dim(V)$, folgt $U = V$.

   c) Es existiert ein Unterraum $U' \subseteq V$ mit $V = U \oplus U'$.

2. Ist $B = \{v_i \mid i \in I\}$ eine Basis von V und $B' = \{v_j \mid j \in J\}$ eine in B enthaltene Basis von U, dann ist $B'' = \{v_i + U \mid i \in I \backslash J\}$ eine Basis von V/U.

   Außerdem:

   Ist $B' = \{u_i \mid i \in I\}$ eine Basis von U und $B'' = \{v_j + U \mid j \in J\}$ eine Basis von V/U, dann ist $B = \{v_j, u_i \mid i \in I, j \in J\}$ eine Basis von V.

3. Ist $A \in \mathrm{Hom}_K(V, W)$, so ist $\ker(A)$ ein Unterraum von V und $\mathrm{im}(A)$ ein Unterraum von W.
4. $V / \ker(A) \cong \mathrm{im}(A)$ (*Homomorphiesatz*)

Behalten wir das für die Aufgabe im Hinterkopf. Wir sollen für zwei Linearformen eines K-Vektorraums V ($f, g \in V^*$) zeigen, dass ein $\lambda \in K$ existiert mit $g = \lambda \cdot f$, falls für alle $v \in V$ mit $f(v) = 0$ stets schon $g(v) = 0$ folgt.

Nehmen wir uns also beliebige $f, g \in V^*$ mit dieser Eigenschaft und zeigen für sie die Behauptung.

Aus der Voraussetzung folgt schon: $\ker(f) \subseteq \ker(g) \subseteq V$. Außerdem wissen wir, dass $\ker(f)$ und $\ker(g)$ Unterräume von V sind (vgl. 3.).

Wir unterscheiden zwei Fälle:

1. Fall: $\ker(g) = V$.

Dann ist $g(v) = 0$ für alle $v \in V$. Wählen wir dann $\lambda = 0 \in K$, sind wir fertig, denn $(\lambda f)(v) = 0f(v) = 0 = g(v)$ für alle $v \in V$ und somit ist $0 \cdot f = g$.

Oder kurz: Da $g(v) = 0$ für alle $v \in V$ ist, folgt $g = 0$ und damit natürlich auch $g = 0 \cdot f$.

2. Fall: $\ker(g) \subsetneq V$.

Aufgabe ist es nun, ein $\lambda \in K$ zu finden, sodass für alle $v \in V$ schon $g(v) = \lambda f(v)$ gilt.

Dazu können wir zunächst folgende Überlegungen anstellen: Ist $v \in \ker(f)$, also $f(v) = 0$, dann ist auch $g(v) = 0$. Daraus folgt aber trivialerweise: $g(v) = \lambda f(v)$ (sogar für jedes $\lambda \in K$).

Das hilft uns also erst mal nicht bei der Suche nach einem passenden λ, aber es stößt uns auf ein Problem: Nehmen wir mal an, wir haben ein $v \in \ker(g)$, das nicht in $\ker(f)$ liegt, dann müsste für das gesuchte λ gelten:

$$0 = g(v) = \lambda f(v).$$

Da $f(v) \in K\backslash\{0\}$ ist, würde dann schon $\lambda = 0$ folgen. Dann wäre aber $g = 0$ und damit $\ker(g) = V$, was einen Widerspruch zur Voraussetzung darstellt. Also muss

$$\ker(g) \backslash \ker(f) = \emptyset$$

gelten. Das ist aber äquivalent zur Aussage

$$\ker(g) = \ker(f),$$

da $\ker(f) \subseteq \ker(g)$ gilt. Das muss also gelten, außerdem wird uns diese Aussage helfen, ein λ zu finden. Also beweisen wir das zunächst:

Dabei hilft uns der Homomorphiesatz weiter. Dazu machen wir uns als Erstes Gedanken über die Bilder der beiden Abbildungen. Das ist hier aber nicht schwer. f und g bilden ja in den Körper K ab, also gilt schon mal

$$\mathrm{im}(f), \mathrm{im}(g) \subseteq K$$

(insb. folgt mit 1.a)): $\dim(\mathrm{im}(f)), \dim(\mathrm{im}(g)) \leq \dim(K) = 1$). Das heißt, entweder gilt $\mathrm{im}(f) = K$ oder $\mathrm{im}(f) = \{0\}$ (analog natürlich für g). Dass Letzteres nicht gelten kann, überlegt man sich wie folgt:

Da $\ker(g) \subsetneq V$ ist, existiert ein $v \in V$ mit $g(v) \neq 0$. Also ist $g(v) \in \mathrm{im}(g)$ ein Element ungleich 0 und somit folgt die Behauptung.

Da mit $g(v) \neq 0$ auch $f(v) \neq 0$ folgt (Aussagenlogik: $[A \Rightarrow B] \Leftrightarrow [\neg B \Rightarrow \neg A]$), gilt dann auch $\mathrm{im}(f) = K$.

Nun kommen wir zum Homomorphiesatz. Dieser liefert uns angewandt auf f *und* g:

$$V/\ker(f) \cong \mathrm{im}(f) = K = \mathrm{im}(g) \cong V/\ker(g).$$

Daraus folgt aber $\ker(f) \cong \ker(g)$. Da $\ker(f) \subseteq \ker(g)$ und $\dim(V/\ker(f)) = \dim(V/\ker(g)) < \infty$ gilt, folgt schon

$$\ker(f) = \ker(g).$$

Der Homomorphiesatz liefert uns aber noch mehr: Setzen wir

$$U := \ker(f) = \ker(g),$$

dann wissen wir $V/U \cong K$, also gilt $\dim_K(V/U) = \dim_K(K) = 1$. Das heißt V/U wird von *einem* Vektor erzeugt. Also ist $B'' := \{v + U\}$ für ein $v \notin U$ eine Basis von V/U. Ist B' wiederum eine Basis von U, dann wissen wir aus 2., dass $B = \{v\} \cup B'$ eine Basis von V ist. Somit folgt aber:

$$V = \langle v \rangle \oplus \langle B' \rangle = \langle v \rangle \oplus U.$$

So weit, so gut, wozu haben wir das jetzt eigentlich gemacht? Wir wollen schließlich ein λ finden, sodass $g(v) = \lambda f(v)$ für alle $v \in V$ ist.

Für $u \in U = \ker(f) = \ker(g)$ gilt ja:

$$g(u) = 0 = \lambda \cdot 0 = \lambda \cdot f(u). \tag{12.5}$$

Das liefert uns also keine Bedingung an λ.

Da für v

$$g(v) = \lambda f(v) \tag{12.6}$$

gelten soll und sowohl $g(v)$ als auch $f(v)$ ungleich 0 sind (nun im Falle $v \notin U$!), muss dann $\lambda = \frac{g(v)}{f(v)}$ sein.

Jetzt wissen wir also, wie wir λ wählen müssen. Nun müssen wir nur noch zeigen, dass dieses λ auch schon wirklich das gesuchte ist.

Um das zu beweisen, nehmen wir uns ein beliebiges w in V. Dann existieren aber, da $V = U \oplus \langle v \rangle$ ist, $u \in U$ und $\mu_1, \mu_2 \in K$ mit

$$w = \mu_1 u + \mu_2 v.$$

Es folgt schließlich:

$$\begin{aligned} g(w) &= \mu_1 g(u) + \mu_2 g(v) \\ &\overset{(12.5),(12.6)}{=} \mu_1(\lambda \cdot f(u)) + \mu_2(\lambda \cdot f(v)) \\ &= \lambda \cdot f(\mu_1 u + \mu_2 v) \\ &= \lambda \cdot f(w) \\ &= (\lambda f)(w). \end{aligned}$$

12.5.4 Skalarprodukt

Folgende Eigenschaften und Tricks könnten uns beim Bearbeiten der Aufgabe behilflich sein:

1. Man wendet die Voraussetzung auf $v + w$, $v - w$, $\lambda v + \mu w$ oder Ähnliches an. f ist ja bilinear und A linear.
2. Oft ist es auch nützlich zu wissen, dass $f(v, v) \geq 0$ ist, was man dann z. B. auf $f(v + w, v + w)$ oder $f(\lambda v, \lambda v)$ anwenden kann.
3. Die Linearität des SKPs wird ausgenutzt.
4. Um $v = w$ zu zeigen, kann man $f(v - w, v - w) = 0$ beweisen. Denn daraus folgt nach Definition des SKPs: $v - w = 0$.
5. Um zu zeigen, dass das SKP in $\mathbb{C}$ gleich 0 ist, kann man zeigen, dass sowohl Realteil wie auch Imaginärteil des SKPs gleich 0 sind. Dafür multipliziert man einmal mit λ^2 und einmal mit $i\lambda \cdot (\bar{i}\lambda)$ und wendet dann die Voraussetzungen an.

Kommen wir nun aber zur Aufgabe. Wir haben einen K-Vektorraum V ($K = \mathbb{R}, \mathbb{C}$) mit SKP f und einen Endomorphismus $A \in \mathrm{Hom}_K(V, V)$ mit der Eigenschaft

$$f(A(v), v) = 0 \text{ für alle } v \in V$$

gegeben.

Zu a): Wir sollen beweisen: Ist $K = \mathbb{C}$, dann ist $A = 0$. Das heißt, wir müssen $A(v) = 0$ für alle $v \in V$ zeigen! Die erste Idee (vgl. 4.) ist, dass wir

$$f(A(v), A(v)) = 0$$

beweisen, weil daraus die Behauptung folgt. Der Trick bei diesem Beweis ist, die Voraussetzung auf $v + w$ anzuwenden (vgl. 1.). Das geht, da $v + w$ ein Vektor aus V

ist! Mit dieser Idee erhalten wir:

$$\begin{aligned}
0 &= f(A(v+w), v+w) \overset{A \text{ lin.}}{=} f(A(v)+A(w), v+w) \\
&\overset{f \text{ lin. in 1.Komp.}}{=} f(A(v), v+w) + f(A(w), v+w) \\
&\overset{f \text{ semilin. in 2.Komp.}}{=} 1f(A(v), v) + 1f(A(v), w) + 1f(A(w), v) + 1f(A(w), w) \\
&\overset{\text{Vor.}}{=} f(A(v), w) + f(A(w), v).
\end{aligned}$$

Jetzt wenden wir die Idee noch auf $v + iw$ an:

$$\begin{aligned}
0 &= f(A(v+iw), v+iw) = f(A(v), v+iw) + if(A(w), v+iw) \\
&= f(A(v), v) - if(A(v), w) + if(A(w), v) - i^2 f(A(w), w) \\
&\overset{\text{Vor.}}{=} -if(A(v), w) + if(A(w), v).
\end{aligned}$$

Wir erhalten also die beiden Gleichungen:

$$\begin{aligned}
f(A(v), w) + f(A(w), v) &= 0 \\
-if(A(v), w) + if(A(w), v) &= 0.
\end{aligned}$$

Dividieren wir die zweite Gleichung durch $-i$, so bekommen wir die Gleichungen:

$$\begin{aligned}
&f(A(v), w) + f(A(w), v) = 0 \\
&f(A(v), w) - f(A(w), v) = 0.
\end{aligned}$$

Aus diesen beiden Gleichungen erhalten wir also durch Addition Informationen über $f(A(v), w)$:

$$f(A(v), w) = 0,$$

und zwar für alle v und w aus V! Da aber $A(v) \in V$ ist (A ist *Endo*morphismus!), dürfen wir insbesondere $w = A(v)$ setzen und erhalten:

$$f(A(v), A(v)) = 0.$$

Dies war genau unsere erste Idee und somit ist die Aussage bewiesen.

Zu b): In a) wurde gezeigt, dass in komplexen Vektorräumen nur die Nullabbildung die Voraussetzung erfüllt. Hier sollen wir nun zeigen, dass es in reellen Vektorräumen durchaus solche Abbildungen gibt, die nicht die Nullabbildung sind.

Um so eine Abbildung zu finden, muss man sich klar machen, was von A eigentlich verlangt wird. Da steht ja nichts anderes als: Das Bild jedes Vektors unter A muss orthogonal zu diesem Vektor sein.

Wir nehmen den $\mathbb{R}^2$ als Vektorraum. Dort können wir uns recht einfach etwas unter der Orthogonalität von Vektoren vorstellen. Zwei Vektoren sind doch genau dann orthogonal (bzgl. des Standardskalarprodukts/Euklidischen SKPs), wenn sie einen 90°-Winkel einschließen. Das heißt aber doch gerade, dass die Drehung um 90° einen Vektor auf einen zu sich orthogonalen Vektor abbildet. Also ist die Drehung um 90°

gerade so eine gesuchte Abbildung. Die Vorstellung ist klar, nur beweisen müssen wir das nun noch.

Ist A die Drehung um $90°$ und $B := \{e_1, e_2\}$ die Standardbasis des $\mathbb{R}^2$ (es handelt sich bei B insbesondere um eine Orthonormalbasis!), dann brauchen wir uns nur zu überlegen, wohin die Einheitsvektoren abgebildet werden, denn dadurch ist die Abbildung schon eindeutig bestimmt.

Wenn man den Vektor e_1 um $90°$ dreht (entgegen dem Uhrzeigersinn), landet man gerade bei e_2, d. h.

$$A(e_1) = e_2.$$

Wenn man den Vektor e_2 um $90°$ dreht landet man bei $-e_1$, also ist

$$A(e_2) = -e_1.$$

Und es gilt dann natürlich:

$$\begin{aligned}
&f(Ae_1, e_1) = f(e_2, e_1) = 0\\
&f(Ae_2, e_2) = f(-e_1, e_2) = 0.
\end{aligned}$$

Außerdem folgt für alle $v \in V$ (schreibe $v = \lambda e_1 + \mu e_2$; $\lambda, \mu \in \mathbb{R}$):

$$\begin{aligned}
f(A(v), v) &= f(A(\lambda e_1 + \mu e_2), \lambda e_1 + \mu e_2) = f(\lambda Ae_1 + \mu Ae_2, \lambda e_1 + \mu e_2)\\
&= \lambda^2 f(Ae_1, e_1) + \lambda\mu f(Ae_1, e_2) + \mu\lambda f(Ae_2, e_1) + \mu^2 f(Ae_2, e_2)\\
&= \lambda^2 \cdot 0 + \lambda\mu f(e_2, e_2) + \mu\lambda f(-e_1, e_1) + \mu^2 \cdot 0\\
&= \lambda\mu \cdot 1 - \lambda\mu \cdot 1\\
&= 0.
\end{aligned}$$

Klarerweise gilt auch $A \neq 0$, da ja z. B. $A(e_1) = e_2 \neq 0$ ist.

Machen Sie sich auch klar, warum man dieses Beispiel nicht auf den $\mathbb{C}^2$ übertragen kann! Wo geht es schief?

Sachverzeichnis

C. Ableitinger, A. Herrmann, *Lernen aus Musterlösungen zur Analysis und Linearen Algebra*,
DOI 10.1007/978-3-658-02353-9, © Springer Fachmedien Wiesbaden 2013

Symbolverzeichnis

$\mathbb{N}$	Menge der natürlichen Zahlen $\{1, 2, 3, \dots\}$
$\emptyset$	leere Menge
$\#M$	Anzahl der Elemente der Menge M
$\mathbb{L}$	Lösungsmenge
$n!$	Fakultät
$\binom{n}{k}$	Binomialkoeffizient
$\wedge$	und
$\vee$	oder
δ_{ij}	Kroneckersymbol
$\overset{!}{=}$	wird gleichgesetzt
$\overset{?}{=}$	noch zu beweisende Gleichheit
$\equiv$	konstant gleich
$\cong$	isomorph
$0.\bar{9}$	Periode
$\bar{z}$	komplex konjugiert
$[x]$	nächstkleinere ganze Zahl zu x (Gaußklammer)
$\ln(x)$	natürlicher Logarithmus
$\{x_n\}_{n\in\mathbb{N}}$ bzw. $\{x_n\}$ bzw. x_n	Folge
$\limsup\limits_{n\to\infty} x_n$	Limes superior
$\liminf\limits_{n\to\infty} x_n$	Limes inferior
$f'(x)$ bzw. $Df(x)$	erste Ableitung von f (Jacobi-Matrix)
$f''(x)$ bzw. $D^2f(x)$	zweite Ableitung von f (Hesse-Matrix)
$f^{(n)}(x)$	n-te Ableitung von f
$\dfrac{\partial^2 f}{\partial y \partial x}$	zweite partielle Ableitung, wobei zunächst nach x und dann nach y abgeleitet wird
$T_{n,f,x_0}(h)$	Taylorpolynom n-ten Grades zur Funktion f mit Entwicklungspunkt x_0
$\|\|f\|\|_\infty$	Supremumsnorm von f
$\lim\limits_{x\searrow x_0} f(x)$	rechtsseitiger Grenzwert
$\lim\limits_{x\nearrow x_0} f(x)$	linksseitiger Grenzwert
f^{-1}	Umkehrfunktion zu f
0_V	Nullvektor des Vektorraumes V
$\oplus$	direkte Summe
$\langle v_1, \dots, v_m \rangle$	Erzeugnis/Lineare Hülle der Vektoren $v_1, \dots, v_m$
$M_{m\times n}(K)$	Menge der $(m \times n)$-Matrizen über K.

$\mathrm{rg}(A)$	Rang der Matrix A
E_n	$(n \times n)$-Einheitsmatrix
A^T	transponierte Matrix zu A
A^{-1}	inverse Matrix zu A
$\mathrm{Hom}_K(V, W)$	Menge der linearen Abbildungen von V nach W über dem Körper K
$\ker(f)$	Kern der Abbildung f
$\mathrm{im}(f)$	Bild der Abbildung f
$M_B^C(f)$	Darstellungs-/Abbildungsmatrix von $f : V \to W$ bzgl. der Basen B von V und C von W.
$V(\lambda, A)$	Eigenraum von A zum Eigenwert λ
$V^\infty(\lambda, A)$	verallgemeinerter Eigenraum
$\mathrm{charpol}_A(x)$	charakteristisches Polynom von A
$\mathrm{minpol}_A(x)$	Minimalpolynom von A
$\mathrm{JNF}(A)$	Jordansche Normalform von A
$\mathrm{rJNF}(A)$	rationale Jordannormalform von A
V/U	Faktorraum (Quotientenraum) von V nach U
$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$	Permutationsabbildung ($i \mapsto \sigma(i)$)
V^*	Dualraum zum Vektorraum V
B^*	duale Basis zur Basis B
f^*	duale Abbildung zu f
kgV	kleinstes gemeinsames Vielfaches
$K[x]$	Menge der Polynome über dem Körper K
f^{ad}	adjungierte Abbildung zu f